فهرست مطالب

		صفحه
اعضای کمیته‌های علمی وبرگزاری		ک
پیشگفتار		ل

سخنرانی های جلسه افتتاحیه

پیام وزیر فرهنگ وآموزش عالی		۱
سخنرانی ریاست دانشگاه صنعتی اصفهان		۴
سخنرانی ریاست مرکز آمار ایران		۹
سخنرانی دبیر کمیته علمی کنفرانس		۱۶

مقالات فارسی

کاربرد مدلهای عمومی خطی در تحلیل عوامل خطر موثر بر آتشی	علیرضا ابدی و ...	۱۹
سرشماری عمومی ساختمان	ابوالفضل ابوالقاسمی قمی	۳۹
معرفی تئوری های مرگ ومیر	علی اکبر احمدی	۴۵
بررسی خطای نمونه‌ای در طرح آمارگیری کشاورزی ۱۳۶۶ وتاثیر طبقه‌بندی بعد از آمارگیری در کاهش خطا	محمدصادق احمدی دانش آشتیانی	۴۶
نقش آمار در اندازه‌گیری سطح تکنولوژی	علی اصغر آزادگان	۴۷
زنجیره‌های مارکف مرتبه P	مجتبی اسحقی	۵۷

		صفحه
کوتاهترین بازه‌های اطمینان برای خانواده‌توزیعهای که‌تکیه‌گاه‌آنها به‌پارامتروابسته‌است	مجیداسدی‌و...	۷۱
روش بوت استرپ	ایلنر ابرانمنش	۷۲
UMVUE برای P(X<Y)	قاسم‌تارمست	۸۱
بررسی‌آماری وضع‌هباتیت‌B	اکبرحسن زاده‌و...	۹۱
روشهای عملی اجرای طرحهای آمارکشاورزی	محمدباقرحسین پورمقدم	۹۲
سرشماری‌ونمونه‌گیری بصورت توام	محمدباقرحقیقی انارکی	۱۰۲
محاسبه‌شاخص نابرابری توزیع‌کالا (روغنهای نباتی) به‌تفکیک مناطق شهری و روستایی	طاهره‌دفتری وطن پرست	۱۱۱
آماردربخش کشاورزی، مشکلات وراه‌حل‌ها	علیرضادهقان	۱۲۵
بررسی اجمالی وضعیت آمارهای ثبتی کشور	عبدالکریم‌دی پیر	۱۲۶
تعدیل شاخص ازدحام‌بااستفاده‌ازضریب نابرابری جینی برای استانهای‌کشور	فریده‌دیباشی	۱۳۶
ضریب همبستگی داخلی	توفیق رجائی	۱۴۸
پرسشنامه، سنگ زیربنای آمارگیریها و سرشماریها	پرویزرضائی‌تهرانی	۱۵۴
اطلاع‌رسانی سالنامه‌آماری کشور	سیدعلیرضارفیع زاده	۱۵۸
فرهنگ آماری کلمات قرآن کریم	محمودروحانی	۱۷۱
وجوه‌تشابه‌واافتراق سرشماریهای عمومی نفوس ومسکن کشور	ابوالقاسم‌سادات دربندی	۱۷۷
منابع خطادرآمارگیریها	محمدباقرسخاوت	۱۸۹
مدیریت داده‌هادرمقیاس بزرگ	جمشیدسراوندی راد	۲۰۵

استنباط آماری از واریانس در حالت بزرگ نمونه‌ای	بهرام سلیمانی	صفحه ۲۱۶
تعیین مدل توابع تقاضای گازطبیعی در بخش های صنعتی-تجاری و خانگی در ایران	محمدرضا سینائی	۲۲۶
خطی مشی حنضاب صف M/G/1 با فراغتهای سرویس دهنده	غلامحسین شاهکار	۲۳۷
آینده آمار در ایران	امیر شاهپور شاهین	۲۵۳
بررسی اثر عوامل اقتصادی، اجتماعی بر سطح باروری	سیدحسن شمس	۲۶۵
خطا و کنترل خطا در استخراج	حمیدرضا صاحب	۲۶۶
عوامل موثر در ارتقای کیفیت وارائه بهنگام آمار	جعفر صادقی نامور	۲۸۱
مشکلات مربوط به تهیه چارچوب آماری	صدیقه صالحی	۲۹۱
آمارگیری جاری جمعیت سال ۱۳۶۵ و استخراج ماشینی	عبدالحسین صدقیانی	۲۹۹
احتمال پیشامدهای فازی	سیدمحمود ظاهری و ...	۳۴۰
نقش مراکز اطلاعاتی در پژوهش های آماری	پروین ظه ...	۳۲۹
کاربرد آمار و اطلاعات در مدیریت و اقتصاد از طریق طراحی و اجرای سیستم اطلاعات مدیریت	جعفر عسگری	۳۴۲
تجزیه و تحلیل الگوی مصرف خانوارهای شهری در استان تهران	علی عسگری	۳۵۲
کاربرد نقشه در آمار	نجم‌الدین علامه	۳۵۴

		صفحه
مروری برنمونه‌گیری سیستماتیک	علی عمیدی	۳۶۵
منابع آماری	شهلاغفرانی	۳۶۶
چگونگی آمارخدمات درایران	فرزانه‌فرزین	۳۷۲
شاخص های قیمت	محمدحسین فرشودی	۳۹۶
نقش تعاریف ومفاهیم‌درطرحهای آماری	کبری فندی	۴۲۴
مدل سازی وشبیه‌سازی تغییرات ارتفاع‌سطح آب دریای خزر...	اکبرگلدسته	۴۳۰
کاربردآزمونهای آماری دربررسی میزان امنیت دستگاههای رمزکننده مخابراتی	اکبرگلدسته‌و...	۴۳۱
تعیین عوامل موثردرمدت زمان شیردهی‌به‌طریق آنالیزمسیری	حسین محجوب و...	۴۳۲
کدوکدگذاری درآمار	علی محرابی	۴۳۳
روش تهیه‌طرحهای فنی خاصه طرح فنی هزینه‌ودرآمدخانواده	علی اکبرمظهری	۴۴۴
شناخت ویژگی های جمعیت معلول بااستفاده‌ازروشهای آماری	سیدرضاسیفی	۴۵۷
مسائل آموزش آماربرای دانشجویان علوم‌کشاورزی	مسعودمنصوری	۴۷۰
نظارت وکنترل کیفی درطرحهای بزرگ آماری	حسین منطقی	۴۷۱
سیستمهای خبره‌استقرائی وکاربرد آن درآمار	سیدباقرمیراشرفی و...	۴۸۲
الگوی پیشنهادی‌درموردتجهیزیافته‌های سرشماری های جمعیتی درزمینه‌موالید	محمدحسین نجاتیان	۴۹۳

		صفحه
تهیه‌جدول عمرجمعیت کشوردرسال ۱۳۶۵	طهنورالهی	۵۰۸
برآوردرفرآیندهای میانگین متحرک با استفاده‌ازروش کمترین مربعات	حسینعلی نیرومند	۵۳۷
بررسی اثرات عوامل اجتماعی واقتصادی مرباروری و ...	فاطمه‌هرندی	۵۵۰
معرفی طرح آمارگیری ازقیمت فروش محصولات وهزینه‌خدمات کشاورزی	فرخ یونسی	۵۶۸

چکیده پوسترها

سرشماریهای جمعیتی کشوراطلاعات مسکن	نسرین افتخاری	۵۸۶
طرح شناسایی نیازهاوامکانات آماری کشور	اکبرافشار	۵۸۷
آمارگیری ازمرغداریهای صنعتی کشور	جلال افضل زاده‌ناشینی	۵۸۸
توزیع شغلی ومیزان تحصیلات زنان شاغل درسرشماری ۱۳۶۵	فرشته‌امیرفریار	۵۸۹
آمارگیری ازفعالیت اقتصادی خانوار	بهرام‌میراحمدی	۵۹۰
برآوردافزوده‌حمل ونقل جاده‌ای مسافر دراقتصادملی‌باستفاده‌ازتکنیکهای‌آماری	محمدرضاانصاری	۵۹۱
آمارگیری ازمنابع‌آب	فریده‌باقری	۵۹۲
مقایسه‌تعدادآبادیهای خالی‌ازسکنه‌کشور درنوسرشماری وعلت افزایش آن	محمدتربتی سراف	۵۹۳
ساختارسنی جمعیت درجامعه‌عشایری ایران	فضل الله‌توکلی	۵۹۴

		صفحه
ارائه‌تجربیات مربوط به‌تهیه‌اطلس عشایری وتهیه‌پیش نویس اطلس سرشماری عمومی نفوس ومسکن وهمکاری درتهیه‌اطلس ملی	علی‌اصغرتیمافره ...	۵۹۵
اشتغال زنان درایران	سهرانگیزچنگیزی آشتیانی	۵۹۶
شاخصهای آموزش عالی درایران، ۱۳۶۹ـ۱۳۶۱	صغری حاج سیدجوادی	۵۹۷
بررسی روندمزدوحقوق ومقایسه‌آن با بهره‌وری کاردرکارگاههای بزرگ صنعتی درسالهای مختلف	وجیه‌خلیلی	۵۹۸
روش جمع آوری آماردرفعالیت پرورش کرم ابریشم	غلامعلی خیری	۵۹۹
روندمصرف انرژی درکشورطی سالهای اخیر وبررسی مصرف انرژی درصنعت سیمان	محمودنرودی	۶۰۰
بررسی آماری صنعت چرمایران	پرویزرحمانی	۶۰۱
بررسی آگاهی، بینش وعملکردزنان همسردار ۱۵ـ۴۹ ساله‌روستایی ایران نسبت به‌برنامه تنظیم‌خانواده(۱۳۶۹)	منصورضایی	۶۰۲
بررسی صنعت فرش ایران	سعیدزارعدوست	۶۰۳
طرح آمارگیری ازمحل خدمات بیمارستانی	ثریازاهد	۶۰۴
ویژگی شاغلان بخش صنعت درکشور	علیرضاسحرخیز	۶۰۵
بررسی آماری نیروی انسانی شاغل دربخش ساختمان کشور	اسکندرسروش	۶۰۶
بررسی آماری، مشخصات بیماران مسموم بستری شده‌دربخش مسمومین	محمودرحیمی ـ بهرام‌سلیمانی	۶۰۷
شاخصهای کمی‌مسکن	فاطمه‌سمیعی	۶۰۸
تحلیل علتهای مرگ ومیردرایران‌براساس اطلاعات ثبتی سالهای ۶۵ تا ۶۹	حجت ا... سیف الهی تنکران	۶۰۹

		صفحه
بررسی آماری حمل ونقل درایران	علی‌محمدشاه‌محمدی	۶۱۶
طرح آمارگیری ازکارگاههای بازرگانی و تجزیه‌وتحلیل نتایج آن	نسرین شعبانی	۶۱۱
بررسی راههای روستایی کشور	منوچهرصالحی	۶۱۲
کاربردسنجش ازدوردرآمارگیری کشاورزی ومنابع‌طبیعی	نصرت الله‌صالحی	۶۱۳
جمعیت وآثاراقتصادی	احمدصفایی نیک	۶۱۴
شناسایی وضعیت فعالیت آماری سازمانهای تحت پوشش سازمان اموراداری واستخدامی کشور	علی اصغرصفایی‌نیک	۶۱۵
طرح آمارگیری ازخدمات آموزشی	پروین طوفانیان	۶۱۶
بررسی آماری موسسات تحقیقاتی کشور	مولودکامکار	۶۱۷
آماردرایران	مجیدرضاکتابدار	۶۱۸
بررسی کوتاه‌درموردساختارجمعیتی اقلیتهای مذهبی ایران براساس سرشماری سالهای ۱۳۳۵، ۱۳۴۵،۱۳۵۵و ۱۳۶۵	حسن کیانی خوزستانی	۶۲۰
آمارگیری کشاورزی‌درایران	سیداحمدمحدث	۶۲۱
نقش وپایگاه‌اجتماعی- اقتصادی زن درجوامع پیشرفته‌ودرحال توسعه‌باتکیه‌برکاروتحصیل	مهنازمحسن پورسی‌سی	۶۲۲
طراحی نظامهای آمارSISبه‌عنوان نمونه، نظام‌آماری شرکت سهامی شیلات ایران	علی اوسط ظفری	۶۲۳
مقایسه‌نسبت سوادجامعه عشایری، روستایی‌وشهری کشوروبررسی علت آن ازدیدگاه‌آمار	میراحسان معاون اسلامی	۶۲۴
بررسی وضعیت مرگ ومیردرکشوربرحسب جنس	مسعودملک پور	۶۲۵

		صفحه
نظری به‌سالخوردگی جمعیت	احمدعلی منصوریان	۶۱۶
مشابرکوچنده‌درگذشته، حال وآینده	محمدصادق‌نامی منش	۶۲۷
ضروری برافزایش ورشدجمعیت شهرهای کشور طی دوره‌سی ساله(۱۳۶۵ ـ ۱۳۳۵)	محمداسماعیل ناصحی	۶۲۸
بررسی وضع‌زناشویی بانگاهی به‌وضع‌سواد و فعالیت	مجیدضروری	۶۲۹
مصرف نهایی دولت وتشکیل سرمایه‌ثابت ناخالص دولت‌درساختمان وماشین آلات و تجهیزات	همانیکنام	۶۳۰
آگاهی، بینش وعملکردزنان ۴۹ ـ ۱۵ ساله همسردارشهری ایران نسبت به‌برنامه تنظیم‌خانواده	قاسم‌بادکاروحسین ملک افضلی	۶۳۱
تجزیه‌وتحلیل داده‌های آماری نمرات دانش آموزان دردرسهای علوم‌ریاضی	هدایت یاسائی مهیدی	۶۳۲
بررسی آمارانرژی (برق)	افسانه‌یزدجی	۶۳۴

اعضای کمیته‌های علمی وبرگزاری نخستین کنفرانس آمار ایران

کمیته‌علمی : (به‌انتخاب شورای اجرایی انجمن آمار ایران)

جوادبهبودیان	استاددانشگاه‌شیراز
احمدپارسیان	دانشیاردانشگاه‌شیراز
مرتضی جمشیدیان	استادیاردانشگاه‌صنعتی اصفهان
محمدنگابی	استادیاردانشگاه‌صنعتی اصفهان
علی‌رجالی	دانشیاردانشگاه‌صنعتی اصفهان
علی‌زینل همدانی	استادیاردانشگاه‌صنعتی اصفهان
سلطان محمدصدوقی الوندی	دانشیاردانشگاه‌شیراز

کمیته‌برگزاری : (اعضای هیئت علمی دانشکده‌ریاضی دانشگاه‌صنعتی اصفهان)

سعید پولادساز

مرتضی جمشیدیان

احمد عقانی

فروزان خرم‌پژوه

محمد نگابی

علی رجالی

غلامعلی رئیسی اردلی

علی زینل همدانی

حمید رضا مرزبان

پیشگفتار

نخستین کنفرانس آمار ایران که پیگیری امور اجرایی آن از فروردین ماه سال ۱۳۷۰ آغاز شده بود بسته همت اعضاء هیئت علمی دانشکده ریاضی دانشگاه صنعتی اصفهان با همکاری هیئت رئیسه محترم دانشگاه ، انجمن آمار ایران و برخی از موسسات دیگر و بویژه مرکز آمار ایران در روز سه شنبه پنجم خرداد ماه هزار و سیصد و هفتاد و یک با حضور بیش از پانصد نفر از صاحبنظران و کارشناسان آمار با پیام آقای دکتر معین وزیر محترم فرهنگ و آموزش عالی افتتاح گردید. در مراسم افتتاحیه آقای دکتر سقائیان نژاد ریاست محترم دانشگاه صنعتی اصفهان ، آقای دکتر جمشیدی ریاست محترم مرکز آمار ایران و آقای دکتر بهبودیان رئیس کمیته علمی کنفرانس سخنانی ایراد نمودند. این کنفرانس پس از سه روز فعالیت با ارائه مجموعا" یکصد و پنجاه مقاله و پوستر و یک کارگاه محاسبات آماری به کار خود پایان داد.

از جمله فعالیت های جانبی کنفرانس ، می توان تشکیل نخستین مجمع عمومی انجمن آمار ایران ، تشکیل یک میزگرد تحت عنوان " جایگاه تدریس آمار در مدارس کشور" و برگزاری نمایشگاه بین المللی کتب تخصصی آمار با بیش از دو هزار و چهارصد عنوان و همچنین نمایشگاه نرم افزارهای آماری نام برد. مرکز آمار ایران نیز نمایشگاهی از فعالیت های آماری ایران در زمینه های مختلف تهیه دیده بود.

سمینارهای ویژه دبیران محترم آموزش و پرورش و همچنین یک بازدید علمی از فعالیت های آماری سازمان برنامه و بودجه اصفهان نیز یکی از فعالیت های جانبی این کنفرانس بود.

لازم است از همکاری نهادها و سازمانهائی که در برگزاری این کنفرانس حمایت نموده اند و بویژه از مساعدت محترم ریاست جمهوری ، هیئت رئیسه دانشگاه صنعتی اصفهان ، وزارت فرهنگ و آموزش عالی ، وزارت امور خارجه ، وزارت پست و تلگراف ، مرکز آمار ایران ، سازمان برنامه و بودجه ، بانک مرکزی ایران ، استانداری اصفهان ، شهرداری اصفهان ، سازمان برنامه و بودجه اصفهان ، ذوب آهن اصفهان ، مرکز تحقیقات فیزیک نظری و ریاضی ، شرکت داده گاوی ، شرکت مهندسی پالس الکترونیک ، دفتر هواپیمایی جمهوری اسلامی شعبه اصفهان ، اداره کل فرهنگ و ارشاد اسلامی اصفهان و همچنین موسسات انتشاراتی

که‌دربرگزاری نمایشگاه‌کتاب مارایاری داده‌اندتشکروسپاسگزاری نمائیم.

به‌دنبال انتشارمجموعه‌مقالات مدعوین نخستین کنفرانس آمار، این مجموعه‌شامل سخنرانی هـــمـــای افتتاحیه، مقالات شرکت کنندگان شامل دوبخش فارسی وانگلیسی وچکیده‌پوسترهای ارائه‌شده‌می‌باشد.

لازم‌به‌توضیح است که‌مقالات چاپ شده‌داوری نشده‌ومسئولیت مطالب آن بعهده‌نویسندگان می‌باشد، در ضمن بعضی ازمقالات به‌دلیل عدم‌دریافت بموقع‌متن کامل ویاکیفیت نامطلوب فقط بصورت چکیده‌چاپ گردیده‌است .

بدیهی است برای دریافت اصل مقالات وپوسترهاکه‌بصورت چکیده‌دراین مجموعه‌به‌چاپ رسیـــده‌است می‌توان بانویسنده‌مکاتبه‌نمود.

امیدواست که‌باانجام‌این کارتوانسته‌باشیم‌گامی درجهت شناسائی ، پیشرفت وپیشبردعلم‌آماردرکشـــور برداشته‌باشیم.

کمیته‌های برگزاری وعلمی نخستین کنفرانس آمارایران

سخنرانی‌های جلسهٔ افتتاحیه

پیام وزیر فرهنگ وآموزش عالی

" . . . وکلُّ شَيءٍ عندَهُ بمقدار، عالِمُ الغَيبِ والشَّهادَةِ الْكبيرُالمُتعال "

وهرچیزرانزداومقداری معین است ، دانای نهان وآشکار، آن خدای بزرگ متعال

(سوره رعد ۱۰ و ۱۱)

بادرودبه‌روان پاک رهبرکبیرومعماربزرگ انقلاب اسلامی حضرت امام‌خمینی(ره) وادای احتـــرام به‌حضورمقام‌معظم‌رهبری، تشکیل نخستین کنفرانس آمارایران رابه‌عنوان گامی ارزنده‌وضروری بـــرای تبیین اهمیت علم‌آمار، ارزیابی‌نتایج جدیدترین دستآوردهای علمی جهانی درزمینه‌های آمارواحتمـال ونقش آن دربرنامه‌ریزی توسعه‌کشورهابه‌فال نیک می‌گیرم‌وازاهتمام‌ومساعی‌همه‌دست اندرکاران ایـــن مهم‌تقدیروتشکرمی‌کنم.

میهن اسلامی ماکه‌سابقه‌ای دیرین وجایگاهی ارزنده‌درفرهنگ وتمدن مشرق زمین دارد، درروندبازیافت هویت وشکوه‌تاریخی خویش بایدمهدپرپائی اینگونه‌مجامع‌علمی‌باشدومحققان ودانشوران متعهد وآگاه‌بتوانندازاین طریق به‌بررسی همه‌جانبه‌اولویتها ونیازهای علمی وتوسعه‌فرهنگی، اجتماعـــی و اقتصادی کشوربپردازند.

دگرگونیها وپیشرفتهای وسیع وسریع علمی وفنی وضرورت دستیابی بهره‌هاوزمانهای نوین آمـــــوزشی وپژوهشی ببهره‌مندی به‌موقع وشایسته‌ازعلم‌وتکنولوژی جدیدازپدیده‌های جدی است که‌چشمانـــــداز تمدن‌آفرین دوران نسل دانشگاهی ماگشوده‌است .

آغازعصرفرامدنی وگستردگی وشتاب نظام‌اطلاع رسانی وارتباطات ، نیازهاوضرورت های جدیـــدی‌را مطرح می‌سازد. تکنولوژی پیچیده‌جدیدکه‌می‌کوشد " دانایی" رابه‌عنوان منبع تازه‌ای ازقدرت ، مطـــرح

سازدوبرداشت علمی صنعتی وفرهنگی رابه‌مدابزارهاوامواج فزاینده‌اطلاعات به‌طورمداوم دگرگـــون نمایدجهان سوم‌راباتصمیم‌گیری های دشوارتری مواجه‌کرده‌است . آنچه‌بیش ازهرچیزموردنیازهمـــــه جوامع درحال توسعه‌است توجه‌به‌جوهره‌سلطه‌طلب وغارتگرنظام‌مسلط جهانی وکشف موقعیت وهویت خودبه‌عنوان گام‌اول توسعه‌است . واقعیت اینست که‌علت وعامل اساسی توسعه‌نیافتگی جهان ســـوم بیش ازآنچه‌سخت افزاری وفنی باشدمشکلی فرهنگی است . باورداشتن نیرو وتوان جوامع خـــــودی و وجودپیشینه‌ومنابع ارزنده‌علمی وفرهنگی درمیان مسلمانان ، هاضمه‌قوی وسالم‌برای جـــــذب و بهره‌مندی ازعلوم‌وتکنولوژی جدیدمی‌سازد. جوامع مسلمان دراین مرحله‌تاریخی بیش ازهرزمـــــان برای شکوفایی وباروری قدرت وخلاقیت فکری خویش نیازمندتوجه‌جدی وعملی به‌امربرنامه‌ریـــــزی آموزش علمی وفنی، تحکیم‌زیربناهای توسعه‌علمی وفرهنگی وگسترش وتعمیق ارتباطات تحقیقاتـــــی وآموزشی‌هستندودراین میان علم‌آماروجمع آوری، نگهداری ، تجزیه‌وتحلیل واستفاده‌بهینـــــــه از اطلاعات جایگاه‌ویژه‌ای رابه‌خوداختصاص می‌دهد.

وزارت فرهنگ وآموزش عالی براساس مسئولیت های قانونی ورسالت فرهنگی وانقلابی نسبت بـــــه توسعه وتقویت آموزش وتحقیق ، تامین نیروی انسانی متخصص وکارآمدوارتقای سطح علمی کشـــور، اینگونه‌مجامع علمی فرخنده‌ومبارک راحمایت وتشویق می‌نمایدونکات ذیل راموردتاکیدقرارمی‌دهد:

۱ـ برنامه‌ریزی توسعه نیروی انسانی درابعادوسیع توباتوجه‌به‌کمبودشدیدمتخصصان آماردرسطـــوح مختلف کارشناسی، کارشناسی ارشدودکترا .

۲ـ توجه‌به‌آمارکاربردی بمنظوراستفاده‌سایررشته‌های دانشگاهی ونیزتحلیل پدیده‌هاومعضلات اجتماعی بعهده‌آمارجهت استفاده‌مسئولان ودستگاههای اجرایی .

۳ـ توجه‌ویژه‌به‌آزمایشگاههای آماروتهیه‌امکانات وتجهیزات موردنیاز.

۴ـ مسئولیت خطیراستادان آماروانفورماتیک برای توسعه‌وتقویت این رشته‌هاباتوجه‌به‌اهمیـــــــت روزافزون علم‌انفورماتیک درعصرحاضر.

۵ـ تلاش بیشتراعضای محترم‌هیات علمی درزمینه‌گسترش دوره‌های تحصیلات تکمیلی وتحقیقات .

۶ـ بازنگری محتوای برنامه‌های آموزشی باتوجه‌به‌پیشرفت های علمی روزونیازهای جامعه.

۷ـ اهتمام بیشترنسبت به‌گسترش دامنه‌فعالیت های انجمن های تخصصی درزمینه‌آماروانفورماتیک .

۸ـ توسعه‌ارتباط میان مراکزعلمی وتحقیقاتی وآماری کشوروارتباط باسایرمراکزبین المللی بمنظور بهره‌مندی هرچه‌بیشترازدستآوردهای نوین علمی وفنی .

۹ـ ضرورت استفاده‌بیشترازروشها وتکنیکهای جدیدآماری دربرنامه‌ریزی توسعه‌وسازندگی کشور.

سخنرانی ریاست دانشگاه صنعتی اصفهان

" یرفع الله الذین آمنومنکم والذین اوتوالعلم درجات "

مقدم میهمانان عزیز، اساتید ارجمند، دانشمندان وپژوهشگران داخلی وخارجی ، پیشکسوتان جامعه آمارکشور، مسئولین وکارشناسان آمارسازمانهای آماری کشور، دبیران آمارسراسرکشور، ایرانیان مقیم خارج که برای این کنفرانس به ایران مسافرت نموده اند، هیات انجمن اسلامی علوم آمارکه ازپاکستان آمده اندودانشجویان ومسئولین ومقامات محترم استان راازجانب خود، هیات رئیسه ودانشگاهیان به دانشگاه صنعتی اصفهان خیرمقدم عرض نموده ازصمیم قلب به همه شماها خوش آمدمی گویم. ازاینکه چنددقیقه ای به بیانه عرض خیرمقدم وقت گرانبهای کنفرانس رامی گیرم پیشاپیش پوزش می طلبم.

هنوزبیش ازدوماه ازسال جدیدنگذشته وماشاهدسه نشست دررشته های مختلف علوم پایه دردانشگاه صنعتی اصفهان بوده ایم که دونشست آن مربوط به دانشکده ریاضی ماست . نشست اولی بمدت دوهفته ازتاریخ پانزدهم الی سیام اردیبهشت ماه تحت عنوان کارگاه بین المللی هندسه وجبر باکمک مرکز بین المللی فیزیک نظری تازه درهفته گذشته بپایان رسیدوبرای علوم ریاضی کشوردوره بسیار موفق وپربرکتی بود، درتاریخ شانزدهم اردیبهشت ما شاهدبرگزاری اولین گردهمائی کاتالیزور وکاربرد آن توسط دانشکده شیمی وباهمکاری مرکزتحقیقات نفت بودیم که آنهم درجای خوداز اهمیت قابل توجهی برخوردار بود. وامروزما شاهدبرگزاری نخستین کنفرانس آمارایران توسط دانشکده ریاضی وانجمن نوپنیاد آمارایران هستیم.

بنده انگیزه اصلی برگزاری این نشستها راچیزی جزسپاسگزاری ازنقش زنده، رشدوتوسعه علوم ونیازمند " تکنولوژی نمی دانم. تازه این نشستها نیست که ما دردانشگاه صنعتی اصفهان داشته ایم، آمارا بگونه

کنفرانسها درسطح کشورزیادوبرقراری آنها برای مجامع علمی وجامعه ما بسیار نویدبخش است وخوش بمن گویدمجددا" یک نهضت علمی درجامعه ما شروع شده که پیام آورنوراست وروشنایی ونابودکننده جهل و تاریکی .

شاید محبتهای من دراین رابطه کمی خوشبینانه واحتمالا" توام با احساسات باشد، ولیکن وقتی انسان با یک نظربه تاریخ این مملکت ، افتخارات علمی گذشته آن رابررسی می نمایدوازنظردیگربامیدی واقعبینانه عقب افتادگیمان ازعلم وتکنولوژی جهان امروزرا می بیند، آنوقت این گونه فعالیتها برایش معنای خاصی دارد، گویا روزنه امیدی دروجودش شعله ورمی گردد، برای روشن شدن مطلب اجازه میدهید شمه ای ازاین افتخارات رابگونه ای فهرست وارمتذکرشوم:

کشورما دارای فرهنگ غنی وسرشارازمعنویت وسعی وتلاش بوده ، مگرکلمه الگوریسم (Algorism) منتسب ازالخوارزمی بمعنی محاسبه ارقام ازنام نامی محمدبن موسی خوارزمی نابغه ایرانی که درریاضی، نجوم وجغرافیا سرآمددوران بوده گرفته شده است ؟ مگرزمانی علم ومعرفت وصنعت ماهمراه باآوازه ابوعلی سینا ازدروازه های روم شرقی عبورنکرده وبازارهای جهانی آن روزگارتا چین رافتح ننمود؟ ودیگرانی، شیخ بهایی ها ، ابوریحان بیرونی ها ، خواجه نصیرالدین طوسیها وعمرخیامها و . . . . و تازه اینها نخستین دانشمندان ایرانی هستندوبنده هیچ ذکری ازدانشمندان دیگرکشورهای اسلامی نکرده ام ! تنها کافی است سری به موزه ها وکتابخانه های بزرگ دنیازده شودتا آثاروگنجینه های تفکرایرانی اسلامی برای ما هما بیشتر روشن شود، بنده به هیچ وجه قصدندارم تفکرکهنه گرائی وگذشته گرائی راتبلیغ کنم چون معتقدهستم که بایدبا واقعیات روبروشد.

اخیرا" نیکسون رئیس جمهوری اسبق آمریکاکتابی دارتباط با ایران نوشته که ضمن بررسی وتائیدتاریخ پرافتخارکشورما وموفقیتهای دوران بعدازانقلاب به پتانسیل های عظیم نهفته این ملت اشاره نموده وضمن هشداربه رغم خود دستورالعمل هائی رابرای عیاد حاکمه فعلی آمریکاتدوین نموده است .

اینها را من نمی گویم که تسکینی باشدبرای گرفتاریهایمان وبعبارتی این صحبتها برای رنگ کسرمن نیست ، چونکه همانطوریکه گفتم بایدواقع بین بودوبپذیرفت که درقلمروعلم وتکنولوژی بسیار عقب هستیم، وشایدبا سرعتی فراترازسرعت نورباید گام برداشت تابه قافله علم جهانی برسیم، اما همه صحبت

من این است که ما هیچ چیزی از مظاهر انسانی را در خیطه علم و تکنولوژی از دیگران کم نداشته و نداریم و اگر امروز از آن عقب مانده ایم وضعیتی است که برای ما بوجود آورده اند. اما حال که ایران اسلامی پرچمدار اوزش والای انسانی لقب یافته است انتظار میرود فواصلی را که از نظر علمی و تکنولوژی برای ما ایجاد کرده اند را با قاطعیت و توکل و اراده از میان برداریم. اینجاست که باید برای رسیدن به این اهداف خط مشی جدیدی برای خود ترسیم نمائیم:

الف - باید علما و دانشمندان و پژوهشگران با توجه به مسئولیت خطیر خود فارغ از هرگونه خودخواهی و حساسیت همدیگر را پیدا نمایند و در جهت رفع معضلات علمی و عقب افتادگیها چاره جویی نمایند.

ب - باید جایگاه خاص علم و عالم در جامعه از طرف مردم و مسئولین مشخص شود و در جهت تقویت آنها از هیچ کوششی فروگذاری نگردد.

ج - باید حداقل امکانات اولیه جهت آنها فراهم شود تا آنها بتوانند مسئولیت خطیر خود را انجام دهند.

د - باید دولو اینکه مجبور باشیم از نان شبمان بزنیم و برکوشش و تلاشمان بیفزائیم تا از علم و تکنولوژی معاصر غافل نمانیم و در یک کلام به صراحت به شما می گویم که مسئولان عزیز مملکت باید آن قدر به علم و تکنولوژی ارج بگذارند و از آن حمایت کنند و بی پرده و تعارف به مردم بگویند چند ده سال باید کمتر خورد تا صاحب علم و تکنولوژی بشویم.

حال با این اوصاف جایگاه این کنفرانسها بیشتر معانی خاص خود را پیدا نموده و بار مسئولیت ما را سنگین تر می کند. کشور ما دوران بازسازی خود را سپری می نماید و در همه ابعاد احتیاج به برنامه ریزی می باشد و سنگ زیربنای برنامه ریزی ، آمار و اطلاعات است . امروزه آمار همانند یک زبان و بعنوان یک پایه در تمام علوم مطرح است ، مخصوصا " حضور و وجود کامپیوتر در عصر حاضر و نرم افزارهای مختلف آماری بیش از پیش بر اهمیت این علم افزوده است . خوشبختانه در کنار کنفرانس کارگاهی در زمینه کاربرد کامپیوتر برگزار گردیده است که هدف آن آشنا سازی جامعه آماری کشور با بسته های نرم افزاری کامپیوتری مورد نیاز در علوم مختلف آماری می باشد.

نمایشگاه نرم افزار با بیش از ۱۲ نرم افزار آماری ، و نمایشگاه کتب داخلی و خارجی از جمله برنامه های دیگر این کنفرانس است .

دانشگاه صنعتی اصفهان بعنوان یکی از قطبهای علمی کشور و در عین حال جوان همواره کوشیده است که زمینه ساز رشد و توسعه همه زمینه های علم و تکنولوژی و مخصوصا " علوم پایه باشد. ما در اینجا علاوه بر تقریبا " تمام رشته های مهندسی و کشاورزی دارای سه دانشکده در علوم پایه، ریاضی ، فیزیک و شیمی هستیم که دانشکده ریاضی ما در این میان دارای جایگاه مخصوصی است . فعالیتهای پژوهشی دانشکده مخصوصا " در جهت عمومی نمودن علوم ریاضی بسیار چشمگیر و از آنجمله می توان برگزاری کارگاه هندسه و جبر، آزمونهای کارشناسی ارشد، بنیانگزاری مسابقات ریاضی تحت عنوان المپیاد ریاضی، چاپ نشریات پیک ریاضی و همکاری در انتشار فرهنگ واندیشه ریاضی را نام برد. فعالیتهای آموزشی دانشکده نیز بسیار قابل توجه و دارای کیفیت خاص خود می باشد. من لازم است در اینجا آماری را که در همین هفته از طریق سازمان سنجش آموزش کشور وابسته به وزارت فرهنگ و آموزش عالی بدستم رسیده و در هفته گذشته در سمینار معاونین آموزشی دانشگاههای کشور مورد بحث و بررسی قرار گرفته است را به اطلاع شما ها برسانم، این آمار مربوط به آزمون ورودی به دوره های کارشناسی ارشد سال تحصیلی ۷۰ ـ ۶۹ می باشد. تقریبا " در تمام رشته ها نیکه در دانشگاه برقرار است ، میانگین نمره آزمون دانشجویان دانشگاه صنعتی اصفهان در میان سایر دانشگاههای کشورها جائز رتبه اول و یا رتبه دوم و یا سوم بوده است . این در شرایطی است که کیفیت ورودیهای دانشگاههای تهران به هیچ وجه قابل مقایسه با ورودیهای سایر دانشگاههای شهرستانی نبوده و نیست . یعنی اینکه اگر معدل ورودیها نیز مد نظر قرار بگیرند و امکانات و شرایط دانشگاههای شهرستان نیز مورد توجه واقع شود جایگاه حقیقی این رتبه ها بیشتر مشخص خواهد شد، من فقط در اینجا به آمار دانشکده ریاضی بسنده می کنم. میانگین نمره آزمون ورودیه کارشناسی ارشد رشته ریاضی در میان دانشگاههای کشور در سال ۷۰ـ۶۹ رتبه دوم را کسب کرده که در تبه اول متعلق به یکی از دانشگاههای تهران است . این حاکی از فعالیت مطلوب آموزش در دانشگاه و در دانشکده ریاضی ما است . همه چیز خوشبختانه آماده است که دانشکده ریاضی در سال آینده دوره دکتری خود را راه اندازی نماید و این جزء برنامه های دانشکده و دانشگاه می باشد.

من مطمئن هستم که این کنفرانس انشاء ... با موفقیت به اهداف خود خواهد رسید و برکات آن شامل رشد و توسعه همه علوم و همچنین علوم ریاضی خواهد شد. ثمرات این کنفرانس هم اکنون شامل حال دانشگاه صنعتی اصفهان شده است . علاوه بر توفیق میزبانی شما عزیزان که از جهت معنویت شامل برکات الهی

است . بهمراه برگزاری این کنفرانس دانشگاه از طریق مرکز تحقیقات فیزیک و ریاضی وابسته به وزارت فرهنگ و آموزش عالی مجهز به سیستم E-mail گردیده است . دانشگاه صنعتی اصفهان اولین دانشگاهی است که از این طریق با مجامع علمی در اروپا متصل می‌گردد. و مطمئنا " این امکان می‌تواند از نظر مبادلات علمی و پژوهشی و حمایت از دوره‌های عالی در داخل دانشگاه بسیار موثر واقع شود. از طریق E-mail همچنین دانشگاه می‌تواند از وقات فراغت سوپر کامپیوترها نیز بهره مند گردد.

و دیگر برکات این کنفرانس برای دانشگاه آشنایی اعضای محترم کادر هیات علمی دانشگاهها و ایرانیان مقیم خارج با دانشگاه صنعتی اصفهان و پتانسیل های آن میباشد که از همین جا بنده آنها را دعوت به همکاری می‌نمایم.

ارگانهای مختلفی ما را یاری دادند تا این کنفرانس بتواند برگزار شود و مسلما " اگر همکاری این ارگانها نمی‌بود برقراری این کنفرانس غیر ممکن می‌بود. در راس این ارگانها همکاری تنگاتنگ انجمن آمار ایران را می‌توان نام برد. اجازه دهید که بنده نیز ذکر نام ارگانهای دیگر بواسطه تعدد آنها خودداری نمایم من از همه این ارگانها از صمیم قلب تشکر و قدردانی می‌نمایم. در اینجا جا دارد از افرادیکه در دانشگاه و دانشکده ریاضی شب و روز زحمت کشیدند تا این کنفرانس برقرار شود که در راس آن رئیس محترم دانشکده و همکارانشان را می‌توان نام برد و همچنین روابط عمومی و امور بین الملل نیز تشکر و قدردانی نمود. تشکر خاص متعلق به مسئول اجرایی کنفرانس برادر ارجمند، جناب آقای دکتر رجالی است که حقیقتا " از همه چیز خود در برگزاری این سمینار مایه گذاشته‌اند انشاء . . . مورد قبول درگاه حق واقع شود. در پایان امید است که این کنفرانس در عمل بتواند قدم بلندی در جهت شکوفائی استعدادها برداشته و منشاء اثر خدمات پر ارزشی برای کشور ما واقع گردد. انشاء . . .

سخنرانی رئیس مرکزآمارایران

" کارآیی آمارهای تولیدشده درکشوروسیزان کاربردآن "

قبل ازآغازسخنرانی، لازم‌می‌دانم‌مسرت فراوان خودرااز برگزاری نخستین کنفرانس آمارایران در حضور وزیرمحترم‌فرهنگ وآموزش عالی، اساتیدودبیران ودانش پژوهان ومحققان ارجمندومیهمانان خارجی جمهوری اسلامی‌ایران به‌ویژه عزیزانی دانشمندو والامقام‌ازدانشگاه‌صنعتی اصفهان ونیز برای تشکیل اولین مجمع‌عمومی انجمن آمارایران که‌آرزوی چندین ساله‌این خدستگزارجامعه‌آماری کشور بوده‌است ، اعلام‌دارم‌وازهمه‌برگزارکنندگان سمینارتشکروقدردانی نمایم. امیدوارم‌راهی‌که‌امروزگشوده شده‌است هرگزبسته‌نشودوبااستعانت ازقدرت لایزال الهی وهمت همه‌دست اندرکاران ، به‌اهداف متعالی خودکه‌همانا‌ارتقای‌فعالیتهای علمی آماری چه‌به‌شکل کاربردی وچه‌به‌شکل نظری است درکوتاهترین زمان ممکن نائل آییم‌و وارد دوران شکوفایی وسازندگی شویم.

دراینجالازم‌می‌دانم‌به‌عنوان مسئول مرکزآمارایران که‌عمده‌تولیدکننده‌آماردرکشوراست مراتب تشکر خودرااز همه‌همکاران خوبمان درواحدهای‌آماری بانک مرکزی جمهوری اسلامی‌ایران، وزارت کشاورزی ، وزارت صنایع سنگین ، وزارت آموزش وپرورش ، وزارت فرهنگ وآموزش عالی و وزارت بهداشت ودرمان وآموزش پزشکی به‌خاطرکمک های فنی وتخصصی‌شان به‌جامعه‌آماری کشوراعلام دارم‌وآرزوکنم‌که‌هریک به‌اقتضای نوع فعالیت خوداز جنبه‌های فنی، اجرایی ومالی، دراین اقدام‌ارزنده دانشگاه‌صنعتی اصفهان مشارکت فعال داشته‌باشند.

می‌دانیم‌که‌در‌ده‌آخرقرن بیستم‌باعصرانفجاراطلاعات نامگذاری کرده‌اندوبه‌حق نیزاین نامگذاری‌شایسته این زمان است . اکنون درهردهسال، کل اطلاعات جامعه‌بشری دوبرابرمی‌شودواین خود به‌معنی شتاب فزاینده‌به‌طرف پیشرفت وتوسعه‌چرخهای عظیم‌تکنولوژی باشیروی فکری وابزارمهم‌اطلاعاتی آن چنان

باشتاب بیشتری درآمده است که اگر چنین امری در کار آن نباشیم به سرعت از مافاصله گرفته و ما را در پشت سر خواهد نهاد، یعنی کما کنون با آن مواجهیم، منبع بدآماری است که در انتهای زنجیره اطلاعات قرار می گیرد.

این زنجیره را بدشکل زیر طرح می نماییم:

Intelligent → Wisdom → Knowledge

Information → Data → (Statistics)

در اینجا، بدون اینکه به جزئیات بپردازیم واژه مفوله Statistics که به تعبیری آمار است می شویم. در این کلام مبحثهای نظری را به دانشگاه و اساتید ارجمند سپرده و وارد بحث تولید آمار می شویم.

از مسائل بسیار مهمی که در تولید آمار در کشور اسلامی ایران مطرح است و جادارد که در این مجال و مقال طرح شود، یکی چگونگی و نحوه تولید آمار در کشور است و دیگری، کاربرد آن، به تعبیر دیگر دو مطلب اساسی در زمینه تولید آمار در کشور مطرح است، یکی میزان کارآیی آمار تولید شده از دیدگاه سازمانی (Efficiency of Production) و دیگری میزان کاربرد آن توسط استفاده کنندگان است - یعنی (User Oriented) است یا نه؟ باید تعیین کرد که آیا آمار تولیدی کشور در جهت تامین نیاز مصرف کننده

در مبحث اول، یعنی تعیین کارآیی آمار تولید شده مناسب است به مطالعاتی که در سال ۱۳۶۲، تحت عنوان آمارگیری از امکانات آماری کشور، توسط مرکز آمار ایران انجام شد اشاره شود و سپس نتیجه گیری لازم به عمل آید.

بررسی نتایج طرح فوق نشان می دهد که از مجموع ۲۱۴۰ سازمان عمده عمومی و خصوصی مورد بررسی ۱۶۳۰ سازمان (۷۶ درصد) دارای فعالیت تبدیل آمار بوده و در ۵۱۰ سازمان یعنی ۲۴ درصد این فعالیت صورت نمی گرفته است. قابل توجه است که از مجموع سازمانهای مورد بررسی، ۹۲/۴ درصد دارای مدیریت عمومی و ۷/۶ درصد دارای مدیریت خصوصی بوده اند.

در همین جا باید متذکر شد که در طرح مزبور، به هر گونه فعالیتی که منجر به تولید آمار می شود چه در قالب یک مجموعه به عنوان واحد آماری و چه در قالب یک فرد به طور انقطاعی یا مستمر، تولید آمار و فعالیت آماری

گفته‌می‌شود. بنابراین ، جای تعجب داردکه‌بااین تعریف، در ۲۴ درصدسازمانهای موردبررسی ، آمار تولیدنمی‌شود. درجایی که‌هرگونه‌فعالیت اقتصادی- اجتماعی به‌طورمستقیم‌یاغیرمستقیم‌منجربه‌تولید آمارمی‌شود، ضرورت دارداطلاعات تولیدشده‌به‌طورمجردیاترکیبی همراه‌بااطلاعات برونزای سیستم جهت ادامه‌حیات وپویایی‌فعالیت درمسیرصحیح خودمورداستفاده‌قرارگیرد. با سازمانهایی اعم از دولتی وغیردولتی‌مواجه‌می‌شویم‌که‌درآنها آمارتولیدنمی‌شود. تعدادابن سازمانها نیزکم‌نیست، ازهر چهارسازمان ، فقط سه‌سازمان به‌تولیدآمارا شتغال دارندکه‌جای تاسف زیاددارد.

ازکل سازمانهای تهیه‌کننده‌آمار، ۸۳۳ سازمان یعنی ۵۱ درصددارای واحدتولیدآمارهستندو ۷۹۷ سازمان (۴۹ درصد) دیگربدون داشتن واحدتولیدآماربه‌تهیه‌وجمع آوری اطلاعات آماری اقدام‌می‌کنند.

به‌تعبیردیگر، فقط نصف سازمانهای موردبررسی، دریک قالب سازمان یافته‌مشغول جمع آوری آمار واطلاعات هستندوبقیه‌بااستفاده‌ازنیروی انسانی وبدون آنکه‌ازامکانات سازمان یافته‌خاصی استفاده کنند، دست به‌جمع آوری آماروا طلاعات می‌زنند. فقدان یک قالب سازمان یافته‌برای تولیدآمار، باعث می‌شودکه‌تهیه‌استانداردهاوتعاریف ومفاهیم‌یکسان ، روندجمع آوری اطلاعات درطول زمان وبالاخره‌انجام‌مقایسه‌های بین سازمانی دچارمشکلات فراوانی شود. این امربه‌نوبه‌خودموجب می‌شودکه تدوین وارائه‌آمارهای این سازمانهاودریک مجموعه‌مانندسالنامه‌آماری غیرممکن شود. درسازمانهایی که‌واحدآماری خاصی‌برای تولیدآماروجودندارد، به‌مجردی که‌فردمسئول جمع آوری آمارتغییرکندتمام فعالیتهای آماری درآن سازمان ، تحت تاثیراین تغییر، قرارخواهدگرفت .

در ۱۶۲۵ سازمانی‌که‌فعالیت تهیه‌آماردرآنها انجام‌می‌شود، درمجموع ، ۹۰۷۳ نفردرزمینه‌تولیدآمار فعالیت می‌کنند. بدین ترتیب ، به‌طورمتوسط درهرسازمان تهیه‌کننده‌آمار، ۵/۶ نفربه‌کارتهیه‌آمار اشتغال دارندکه ۴۳/۲ درصدآنان فقط دراستان تهران مشغول انجام‌وظیفه‌بوده‌اند.

اکنون بنیست دراین زمینه‌محاسبه‌ای‌انجام‌شود.

فرض کنیدکه‌هرماه‌به‌طورمتوسط ، برای هریک ازافرادی که‌درتولیدآمارمشغول به‌کارهستند، سازمان مربوط مبلغ ۱۲۰۰۰۰ ریال که‌شامل حقوق ومزایاواضافه‌کاروکمکهای نقدی ازقبیل حق التحقیق وعیدی ودیگرپرداختهای پرسنلی است ، هزینه‌کند

۹۰۰۰۰ + ۳۰۰۰۰ + ۲۰۰۰۰ = ۱۴۰۰۰۰

سایر اضافه‌کار حقوق
پرداختها

اگراین سازمان درهرماه‌برای سایرامورشامل هزینه‌ماموریت ، خارج اتومبیل ، هزینه‌آموزش ، هزینه ملزومات ولوازم‌اداری، هزینه‌آب وبرق ونگهداری ساختمان وبالاخره‌هزینه‌پردازش اطلاعات وانتشار نتایج هرفعالیت آماری، تقریبا" ۱۲۰۰۰۰ ریال به‌ازای هرنفرهزینه‌کند، بافرض اینکه‌هیچ گونه فعالیت عمده‌آماری کمابیش‌به‌پرسنل اضافی‌داشته‌باشد، انجام‌دهنده‌درسال، حدود

هزینه‌پرسنلی	۱۵۲۴۱۶۴۰۰۰۰
هزینه Overhead	۱۱۹۷۶۳۶۰۰۰۰
	۲۷۲۱۹۰۰۰۰۰۰

بیست وهفت میلیارد ودویست ونوزده‌میلیون ریال برای تولیدآمارهای جاری سازمانهای مختلف کشورهزینه‌می‌شود.

بی‌مناسبت نیست اشاره‌کنیم‌که‌در ژاپن با ۱۲۵ میلیون نفرجمعیت (بیش از ۲ برابرجمعیت فعلی ایران) که‌یک کشورکاملا" پیشرفته‌درتولیدآمارمحسوب‌می‌شود، پرسنل موجوددرای تولیدآمارهای جاری‌برابر ۱۲۷۰۰ نفروهزینه‌تولیدآماردرهرسال برابر ۴۶ میلیاردین است .

نتایج بررسی ۱۶۲۵ سازمان تهیه‌کننده‌آمار واطلاعات نشان می‌دهدکه‌تعداد ۵۴۶ سازمان درطول سالهای ۱۳۶۲-۶۶ درمجموع، ۱۸۵۶ فعالیت آماری انجام‌داده‌اند، ازتعداد ۵۴۶ سازمان ، ۳۰۱ سازمان (۵۵ درصد)، وضعیت آماری خودرا خاتمه‌یافته‌گزارش کرده‌اند. ۱۶۰ سازمان (۲۷/۵ درصد) فقط دارای فعالیت آماری دردست انجام‌بوده‌و ۹۵ سازمان (۱۷/۴ درصد) هردونوع‌فعالیت خاتمه‌یافته‌وفعالیت دردست انجام‌را داشته‌اند.

ازتعداد ۱۸۵۶ فعالیت آماری مزبور، ۷۲ درصدطول دوره‌موردگزارش خاتمه‌یافته‌ونتایج آن در اختیارا ستفاده‌کنندگان آماروا طلاعات قرارگرفته‌است و ۲۸ درصدآنها درپایان سال ۱۳۶۶ همچنان

دردست اجرا بوده‌است .

ازمجموع فعالیتهای آماری خاتمه‌یافته‌سازمانها ، حدود ۶۰ درصدمربوط به‌فعالیتهای آماری مستمر وحدود۴۰ درصدمربوط به‌فعالیتهای آماری موردی بوده‌است .

بدین ترتیب ، ازکل فعالیتهای آماری که‌درطول سالهای ۱۳۶۲-۶۶ انجام‌شده‌است ، فقط نتایج ۱۳۲۹ فعالیت آماری ارائه‌ودراختیارمتقاضیان احتمالی قرارگرفته‌است . حال می‌توان هزینه فعالیتهای آماری خاتمه‌یافته‌رابرای دوره ۵ ساله‌فوق محاسبه‌کرد. کل هزینه ۵ ساله‌تولیدآمار ، حدود ۱۳۶۱۰۰۰۰۰۰۰۰ ریال است . بنابراین، برای هرفعالیت آماری خاتمه‌یافته، حدود ۱۰۰میلیون ریال هزینه‌شده‌است . این رقم‌بسیاربالاست ، زیرابرای انجام‌یک سرشماری عمومی جمعیت دریک جامعه ۵۰ میلیون نفری ، هزینه‌استخدام ۳۰ هزارنفرماموروبه‌کارگیری ۷۰۰۰ اتومبیل ورانده حدود ۷۵۰۰۰۰۰۰۰۰ ریال برآوردمی‌شودکه‌معادل هزینه‌اجرای ۷۵ طرح آماری درکشوراست . بدین ترتیب ، می‌توان به‌صراحت اعلام داشت که‌هرچندکاربردآماردرایران گسترده‌وسیعی ندارد، اماهزینه تولیدآن بسیاربالاست وبه‌عنوان یک کالای لوکس قلمدادمی‌شود.

نکته‌قابل توجه‌دیگردراین بررسی، میزان تحصیلات کارکنانی است که‌درتولیدآمارفعالیت دارند.

بررسی مدارک تحصیلی کارکنان شاغل درتهیه‌آمارنشان می‌دهدکه ۲۲/۱ درصدآنان دارای تحصیلات لیسانس وبالاتر، ۷/۹ درصددارای تحصیلات فوق دیپلم، ۵۴/۳ درصددارای مدارک دیپلم و ۱۵/۶ درصد دارای تحصیلات کمترازدیپلم‌هستند. ازمجموع کارکنان شاغل درتهیه‌آمار، ۵۶ نفردارای مدرک تحصیلی دکترا هستند. بدین ترتیب ملاحظه‌می‌شودکه‌تولیدآماردرکشوربا هزینه‌بسیارانجام‌می‌شود. حدود ۷۰ درصدازکسانی که‌به‌تولیدآماراشتغال دارند، دارای تحصیلات دیپلم‌وکمترهستندوآماری راتولید می‌کنندکه‌ازسطح‌نازلی برخوردارست . این امرنیزجای تعجب دارد، زیرااگرآماروتولیدآماریک مقوله علمی وتخصصی است چرابایدتوسط این گروه‌وباصرف هزینه‌بسیارودرنهایت باکیفیت نازل تولیدشود.

نکته‌قابل ذکراین است که‌بیشترین میزان فعالیتهای تولیدآماربه‌صنعت وپس ازآن به‌جمعیت ونیروی انسانی اختصاص داشته‌است . به‌عبارت دیگر، ازکل فعالیتهای آماری ۲۳/۵ درصددرزمینه صنعت، ۱۶/۸ درصددرزمینه‌جمعیت ونیروی انسانی، ۹/۷ درصددرزمینه‌آموزش وپرورش، ۸/۴ درصد

درزمینه‌بهداشت ودرمان وخدمات رفاهی، ۷/۱ درصددرزمینه‌کشاورزی و ۵/۹ درصددرزمینه‌فعالیتهای بازرگانی بوده‌است .

بدین ترتیب می‌توان گفت که‌توزیع فعالیتهای آماری بین فعالیتهای مختلف اقتصادی‌واجتماعی نیزیک توزیع متناسب به‌شمارنمی‌رود. این موضوع‌را می‌توان بابررسی دیگری به‌روشنی نشان داد.

دریک بررسی که‌درسالهای ۱۳۶۹ و ۱۳۷۰ به‌عمل آمد، وضعیت استفاده‌ازآماروااطلاعات درواحدهای برنامه‌ریزی واجرایی کشورمورد‌مطالعه‌قرارگرفت . براساس این مطالعه، شاخصهایی تهیه‌شدکه‌نسوع ومیزان‌استفاده‌ازاطلاعات آماری تولیدشده‌راشخص می‌کندبا استفاده‌ازاین شاخصها می‌توان مشخص کرد که‌فعالیت تولیدآمارتاچه‌حدی بانیازهای موجودوآتی استفاده‌کنندگان ازآمارانطباق دارد.

این بررسی تحت عنوان " طرح شناسایی اطلاعات موردنیازارزیابی برنامه‌اول اقتصادی، اجتماعی، وفرهنگی جمهوری اسلامی‌ایران" انجام‌شده‌است . از ۱۲۳۴ اطلاع موردنیازاعلام شده‌توسط استفاده کنندگان، ۱۱۰۶ اطلاع یعنی ۸۳ درصدآن تولیدمی‌شود، ولی درعین حال، ۷ درصداطلاعات آماری تولید شده‌ازنظربوشش جغرافیایی، اشکالات‌چارچوبی ، ابهامات درتعاریف ومفاهیم، عدم‌رعایت طبقه بندی‌های استاندارد، استفاده‌نامناسب ازروشهای نمونه‌گیری وبالاخره‌مسایل مربوط‌به‌امورا جرایی ویردازش اطلاعات ، دارای اشکال بوده‌وبنابراظهاراستفاده‌کنندگان، موردتایید‌قرارنگرفته‌است .

بدین ترتیب ، ۱۰۳۵ اطلاع موجودیعنی ۷۷ درصدآماری که‌دراختیاراستفاده‌کنندگان قرارداشته‌است کم‌وبیش موردتایید‌آنان بوده‌است وبقیه‌آمارارائه‌شده‌یادچارمسئله‌بوده‌یاهیچ گونه‌کاربردی نداشته‌است

همچنین قابل ذکراست که‌از ۱۱۰۶ موردااطلاع تولیدشده، دورمزمانی ۶۵ درصدازاطلاعات بسا زمان آماری مطلوب انطباق داشته‌است و۳۵ درصدبقیه‌دارای زمانهای آماری غیرمطلوب بوده‌است .

همچنین ۷۶ درصدازاطلاعات تولیدشده، غیرازسطوح جغرافیایی موردانطباق داشته‌وسطح جغرافیایی ۲۴ درصدازاطلاعات استفاده‌شده، غیرازسطوح جغرافیایی موردنیازبوده‌است . یعنی اطلاعات آماری - موردقبول نیزخودازنظرسطح جغرافیایی وزمان آماری‌چاراشکال است . به‌ویژه‌آنکه‌طبق بررسی فوق، اطلاعاتی که‌ازنظرکیفیت ، غیرقابل قبول تشخیص داده‌شده‌است به‌طورعمده‌دربخش صنایع‌است . یعنی درست همان بخشی که‌بیشترین فعالیتهای آماری رابه‌خودا‌ختصاص داده‌است .

درمجموع‌می‌توان گفت که‌هرچنداطلاعات وآمارموجود، برای ارزیابی فعالیتهای آماری چندان کافی و وافی به‌مقصودنیست ولی تولیدآماردرکشوردچارمشکلات جدی‌وفراوانی است که‌می‌بایست هرچه زودترموردتوجه‌قرارگیرد. دراین زمینه، همکاری وهمفکری سازمانهای عمده‌تولیدکننده‌آمارمانند مرکزآمارایران ومجامع‌علمی مانندد‌انشگاهها‌ضرورتی تام‌دارد .

خوشبختانه، بدون هیچ گونه‌مداهنه‌وخصیت سازمانی‌بایدبه‌عرض برسانم‌که‌درچندسال اخیر، همکاران ارجمندمادرمرکزآمارایران کوششهایی فراوانی‌به‌عمل آورده‌اندکه‌تولیدآماررا‌هم‌صرفا" درجهت تامین نیازهای مصرف کننده (User Oriented) قراردهندوهم‌هزینه‌آن را مقرون به‌صرفه (Cost effective) کنند. هرچنداین کوششها‌با‌امکانات محدودموجودهماهنگی لازم‌راندارد‌ولی ازآنجاکه‌مسئولان اجرایی کشور، جایگاه‌آماررا‌کم‌وبیش شناخته‌وبه‌دنبال دستیابی به‌اطلاعات صحیح هماهنگ ، به‌هنگام‌وقابل استفاده‌هستند، امیدمی‌رودآینده‌درخشانتری درزمینه‌فعالیتهای آماری درکشور به‌وجودآید.

درپایان ضمن آرزوی موفقیت برای کلیه‌همکاران محترم‌ومدعوین گرامی امیدوارم‌که‌این سمینار منشاء‌خیرات وبرکات زیادی‌برای‌دست اندارکاران فعالیتهای آماری کشورباشدوموجب شودکه‌درتولید آماراز‌سطوح علمی‌وتخصصی لازم‌بهره‌گیری شودوبدین طریق، آمارتولیدی قابلیت استفاده‌بیشتری‌یابد.

سخنرانی دبیرکمیته‌علمی کنفرانس

شرکت کنندگان گرامی

بسیارخوشوقتم‌که‌ازطرف کمیته‌علمی نخستین کنفرانس آماروایران وازطرف انجمن آماروایران مقدم همه‌شما ، به‌ویژه‌میهمانان بین المللی، را به‌شهرتاریخی وفرهنگی اصفهان گرامی دارم.

درکشورمااین نخستین کنفرانس نیست که‌درآن موضوع‌آمارمطرح می‌گردد. ازنخستین کنفرانس ریاضی کشوردرشیرازتابیست وسومین آن درباختران ، همواره‌درکنارسخنرانیهای‌ریاضی سخنرانیهای آماروهمراه‌شده‌اند. پس این کنفرانس برای چیست ؟ آماردرعصرمابک دیسیپلین مهم علمی باکاربرد فراوان دربیشتررشته‌هامی‌باشد. توانایی برای درک بحثهای آماری برای هرپژوهشگرمدیراهمیت دارد. هرفردتحصیل کرده‌تاحدودی‌نیازبه‌سوادآماری دارد. آماربرای هرسازمانی به‌منظوربرنامه‌ریزی، آینده - نگری ، برخوردبا‌امورتصادفی‌وارزیابی ریسک ضروری است . این کنفرانس برای تاکیدبراهمیت آمار، برای بحث روی‌مطالعات آماری، برای بالابردن کاربردآماردرپژوهش ، آموزش ، بهداشت ، صنعت ، کشاورزی ، بازاریابی واقتصاد، وسازندگی می‌باشد.

درتمام‌کشورهای جهان هنوزسوادآماری آنطورکه‌لازم‌می‌باشدگسترش پیدانکرده‌است . بااینکه‌درسه دهه‌اخیرآماروکامپیوتردوش به‌دوش هم‌باسرعتی عجیب ما راغافلگیرکرده‌اند، هنوزمردم عادی وباسواد می‌پرسند "آماروچیست ؟" ، درکشورمابرای مردم‌عادی کلمه‌آمارفورا" اذهان‌آماروبابت احوال راتداعی می‌کند. این یک تنانی بجامی‌باشد، زیراآماردرحقیقت ازاعدادوارقام‌استفاده‌می‌کند، ولی اندیشه‌آماری برای خودمستقل می‌باشد. به‌عنوان مثال فیشرانگلیسی، نابغه‌آمارقرن بیستم، باانتشارهای آماری خودواندکی ریاضی

تحولی بزرگ درپژوهشهای کشاورزی وژن شناسی ابداع نمود. کارهای او ودیگران درچنددهه‌گذشته‌منطقی واقعی آمار راشرح می‌دهد. آماربمعنی اعدادوارقام‌مربوط به‌زندگی وجهان (تحت عنوان داده‌ها) وبرداشت معقول ازآنها (به‌کمک اندیشه‌های آماری) برای حل مسائل زندگی.

این همه‌کتاب ومجله‌ومقاله‌درزمینه‌آمار، اهمیت وپیشرفت آمار رادرعصرمابیان می‌دارد. کشورما هم، چه‌پیش ازانقلاب اسلامی چه‌بعدازآن، دراین راه‌دارای سهمی بسزااست. این کنفرانس خود نشانگررشدورسائی آماردرمیهن ماوزحمات آماردانان ایرانی درگذشته‌وحال می‌باشد. ناگفته‌نماندکه‌در طی جنگ تحمیلی باهمه‌دشواریها، آماردانان ایرانی‌برای تنظیم‌برنامه‌های درسی وبرای پژوهش درسطح جهانی فعال بوده‌اند. مرکزآمارایران دوبارسرشماری کرده‌است وفعالانه‌روی داده‌های جمعیتی مطالعه‌کرده‌ومی‌کند. ده‌هاکتاب آمارتالیف وترجمه‌شده‌اند. ده‌ها مقاله‌علمی درمجلات بین المللی توسط پژوهشگران آمارمنتشرشده‌اند. لااقل در ۶ دانشگاه‌برنامه‌های کارشناسی یاکارشناسی ارشدیادکترا درحال اجرامی‌باشند. بسیاری ازدانشجویان ماباعلاقه‌فراوان دراین رشته‌کارمی‌کنندومغزهای آماری این کشورخواهندشد. امیداست که‌ازوجودآنها به‌نحواحسن استفاده‌شودوامکانات لازم‌برای آموزش وپژوهش آنها هرچه‌بیشترفراهم‌آورند. خدانکندکه‌هرگزدراثربی‌توجهی، این مغزهافرارکرده‌جذب کشورهای دیگرشوند.

مایلم‌به‌مقامات محترم‌دانشگاه‌صنعتی اصفهان، به‌ویژه‌به‌رئیس محترم‌دانشگاه‌ورئیس دانشکده ریاضی، که‌بانی نخستین کنفرانس‌آمارکشورماشده‌اندصمیمانه‌تبریک گفته‌وازآنهاتشکرنمایم. برگزاری یک کنفرانس علمی همه‌جا، به‌ویژه‌درشرایط سخت بعدازجنگ، کاری دشواراست، واقعاً" مستلزم دانایی، توانایی، ازخودگذشتگی، تلاش فراوان، وهمت عالی‌می‌باشد. خوشوقتانه‌کمیته‌برگزاری کنفرانس صاحب این صفات می‌باشند. بنابراین وظیفه‌خودمی‌دانم، که‌به‌خاطرفداکاری وتلاش بی‌پایان درعرض ۱۲ ماه‌گذشته، ازاعضای‌کمیته‌برگزاری صمیمانه‌تشکرنمایم.

ازحمایت بی‌دریغ مرکزآمارایران، ازمساعدت‌فراوان وزارت فرهنگ وآموزش عالی ایران، واز همکاری مرکزتحقیقات فیزیک نظری وریاضیات سپاسگزاریم.

جادارد‌صراحتاً" ازتمام‌مقامات محترم‌شهراصفهان، به‌ویژه‌استانداری، شهرداری، ذوب آهن،

سازمان برنامه‌وبودجه، بخاطرکمکهای ارزنده‌صمیمانه‌قدردانی نمایم.

امیدداست که‌این نخستین کنفرانس یک پیروزی چشمگیری برای علم‌آماردرکشورما باشد. ابا مسرت خوش وپربارباید برای آماردانان ارجمندازشهرهای ایران وازخارج آرزومی‌کنم. این یک فرصت بزرگ می‌باشد تابایکدیگرآشناشویدوبتوانیدعقیده‌نااریب خودرادرباره‌این کنفرانس ابرازدارید.

کاربرد مدلهای عمومی خطی در تحلیل عوامل خطر مؤثر بر آنمی

علیرضا ابدی* - کاظم محمد

محمدرضا زالی - محمدرضا مسجدی

وزارت بهداشت و دانشگاه تربیت مدرس

خلاصه :

در این تحقیق به بررسی وضعیت بیماری آنمی و تعیین عوامل خطرمؤثر بر آن در شهر تهران پرداخته می شود. اطلاعات مورد نیاز از طرح بررسی سلامت و بیماری که زیر نظر معاونت پژوهشی وزارت بهداشت ودرمان انجام پذیرفته می باشد.

برای انجام این تحقیق هموگلوبین خون ۲۲۰۱ نفر از افراد ساکن شهر تهران اندازه گیری شد که با توجه به آن افراد سالم و بیمار تشخیص داده شدند.

از افراد مورد بررسی ۵۵/۸٪ دارای بیماری آنمی بودند که براساس تفکیک جنس ۴۴/۵٪ از آنها مرد و ۵۶/۸٪ از آنها زن بودند و براساس تفکیک سنی در گروه سنی ۱۰-۱ سال ۴۳/۲٪ و در گروه سنی ۱۴-۱۰ سال ۵۵/۲٪ و در گروه سنی ۱۵-۶۴ سال ۵۶/۲٪ بیمار وجود داشت. براساس تفکیک محل سکونت ۶۷٪ در شمال شهر و ۴۳/۸٪ در مرکز شهر و ۵۶/۲٪ در جنوب شهر ساکن بودند.

در بررسی وضعیت تغذیه و وضعیت رفاهی مشخص گردید که در شهر تهران تغذیه و بیماری و وضعیت رفاهی با آنمی ارتباط معنی داری دارند.

در بررسی منطقه محل سکونت مشخص شد که متوسط هموگلوبین خون افراد ساکن در مرکز شهر تهران بیشتر از سایر نقاط جنوب و شمال تهران می باشد و این مسئله تا حد زیادی هوا یا نوع مواد غذایی را تحت الشعاع قرار می دهد که بررسی بیشتر این مسئله ضروری به نظر می رسد.

به‌جهت تعیین عوامل خطرموثر بر بروز آنمی با کمک برنامه‌های کامپیوتری SPSS و GLIM از مدلهای لگ خطی و لجیت استفاده شد که در بررسی متغیرهای سن و محل سکونت و بیماری آنمی ، اثر متقابل بین سن و محل سکونت و اثر جنس روی بیماری نشان داده .

---

* : استادیار دانشگاه تهران ** : معاون پژوهشی وزارت بهداشت ،درمان و آموزش پزشکی ***: استادیار دانشگاه علوم پزشکی ایران

مقدمه :

درنوشته‌های قرن هفدهم‌بیماری سبزشرح‌داده‌شده‌است که‌بوسیله‌ٔ گیاهان با آهن داروئی مداوا می شده‌است . بیماری سبزنامها ولیه‌بیماری کم‌خونی می باشد .

طبق تعریف بیماران مبتلابه‌کم‌خونی به‌دلیل کاهش قابل توجه‌درتوده‌ٔ گلبول قرمزدچارکمبودظرفیت حمل اکسیژن بوسیله‌ٔ خون می گردندوچون حجم خون دریک سطح ثابت حفظ می گرددد‌رنتیجه‌کم‌خونی مستلزم‌کاهش درغلظت گلبولهای قرمزیا هموگلوبین خون است . مهمترین عمل گلبول قرمزانتقال ورساندن اکسیژن ازریه‌ها به‌تمام‌سلولهای بدن وبازگرداندن موادزائد حاصل ازمصرف اکسیژن درسلولها یعنی گازکربنیک به‌ریه‌ها می باشد .

روش بررسی :

جمع آوری اطلاعات به‌روش نمونه‌گیری خوشه‌ای می باشدکه‌هرخوشه‌شامل هفت‌خانوار است . جامعه‌ٔ موردبررسی جمعیت ۶۹-۲ ساله‌مقیم‌شهران را تشکیل می دهد .

جدول تعریف بیماری آنمی بر مبنای میزان هموگلوبین (gr/100ml)

بر اساس تعریف W.H.O

2 سال تا 5 سال	11
6 سال تا 14 سال	12
آقایان	13
خانمهای حامله	11
خانمهای غیر حامله	12

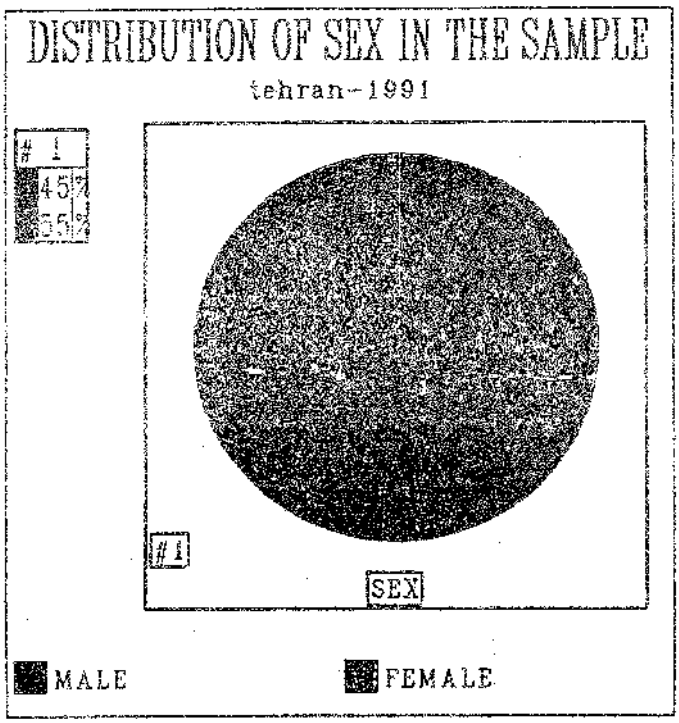
درمورد تشخیص بیماری از روی میزان هموگلوبین مطالعات متعددی انجام شده و گزارشات مختلفی وجود دارد. سازمان بهداشت جهانی در سال ۱۹۸۹ طی مقاله‌ای کمترین حد طبیعی هموگلوبین را در گروههای سنی و به تفکیک جنسی گزارش نموده و ذکر می‌نماید که پائین تر از این مقدار را باید کم‌خونی اطلاق نمود.

با انجام آزمونهای آماری مشخص شد که این تقسیم‌بندی سنی و جنسی دقیقاً با جامعه مورد بررسی انطباق ندارد زیرا بین میانگین هموگلوبین خون افراد در گروههای سنی ذکر شده تفاوت معنی داری دیده شد و تنها در گروه سنی ۶۹-۱۵ سال بین میانگین هموگلوبین خون دو جنس زن و مرد تفاوت معنی دار دیده شد.

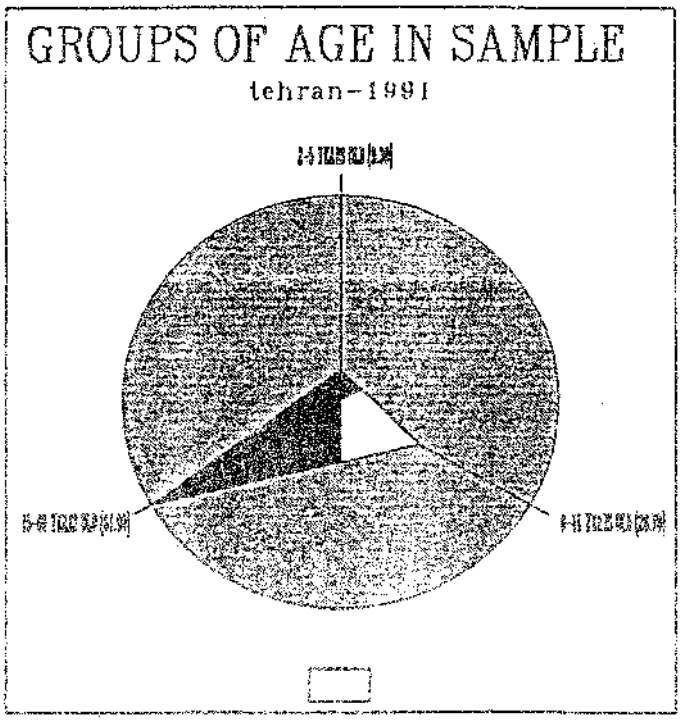
نتایج: توزیع سن و جنس افراد مورد بررسی در نمودار شماره ۱ و شماره ۲ نشان داده شده است. ۵۰/۳٪ از افراد مورد بررسی مرد و ۵۵٪ از آنها زن بودند. ۹/۳٪ در گروه سنی ۵-۲ سال و ۲۶/۷٪ در گروه سنی ۱۴-۶ سال و ۶۴٪ در گروه سنی ۶۹-۱۵ سال بودند.

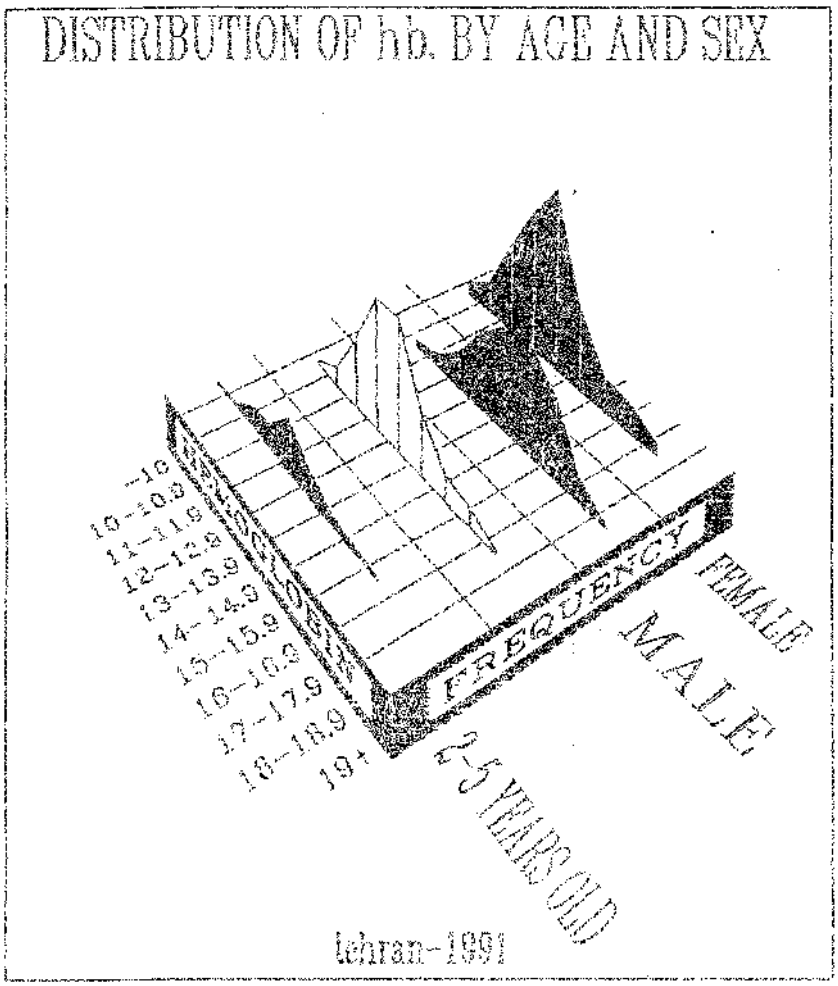
توزیع هموگلوبین خون افراد مورد بررسی به تفکیک سن و جنس در نمودار شماره ۳ نشان داده شده است.

نمودار شماره ۱ :



نمودار شماره ۲ :




DISTRIBUTION OF hb. BY AGE AND SEX
FREQUENCY
FEMALE
MALE
AGE GROUPS
-10
10-10.9
11-11.9
12-12.9
13-13.9
14-14.9
15-15.9
16-16.9
17-17.9
18-18.9
19+
25 YEARS OLD
Tehran-1991

یکی از مشکلاتی که در اکثر تحقیقات اجتماعی وجود دارد چگونگی اطلاع پیدا کردن از وضعیت اقتصادی افراد مورد بررسی است . در پاسخ به سوال میزان درآمد غالبا " افراد یا مقدار واقعی را نمی گویند و یا فقط یکی از درآمدهای اصلی را می گویند و کل درآمد خود را ذکر نمی کنند . بنابراین با این سوال نمی توان از وضعیت اقتصادی افراد برداشت درستی داشت . در مورد شغل هم به همین صورت است که یا افراد از ذکر آن خودداری می ورزند و یا مشاغل کلی را ذکر می کنند که از روی آن نمی توان وضعیت اقتصادی را مورد سنجش قرار داد . مثلا" در شغل آزاد یا کارمندکه در هر دو گروه ممکن است فردی با وضعیت بسیار خوب اقتصادی باشد یا ممکن است با وضعیت بسیار بد اقتصادی باشد .

برای رفع این مشکل تدبیر جدیدی بکار گرفته شد که با تقسیم سطح زیر بنای منزل (سطح سقف) بر تعداد افراد خانواده سطح زیربنای سرانه بدست آمد که نشان می دهد به هر فرد از افراد خانوار چه مقدار رفاه می رسد و نشانگر وضعیت رفاهی افراد است . در این طرح نیز برای ما وضعیت رفاهی افراد مهم بود که در وضع سلامت یا بیماری فرد می تواند موثر باشد .

برای گروه بندی متغیر وضعیت رفاهی اگر به هر فرد خانواده کمتر از ده متر فضا برسد آنرا جزء گروه وضع رفاهی بد قرار دادیم . اگر به هر نفر بین ده تا بیست متر مربع فضا برسد جزء گروه ضعیف قرار دادیم . اگر به هر نفر بین بیست تا چهل متر مربع فضا برسد جزء گروه متوسط قرار دادیم و اگر به هر نفر بیشتر از ۴۰ متر مربع فضا برسد جزء گروه خوب قرار دادیم .

در مورد وضعیت رفاهی در گروه وضع اقتصادی بد ۶/۱٪ بیمار و در گروه وضع اقتصادی ضعیف ۵/۹٪ بیمار و در گروه وضع اقتصادی متوسط ۵/۳٪ بیمار و در گروه وضع اقتصادی خوب ۶٪ بیمار مشاهده گردید که ملاحظه می شود نسبت بیمار در هر گروه تقریبا " یکسان است و نتایج آزمونهای آماری نیز تفاوت معناداری را نشان نداد . در نتیجه وضع رفاهی تاثیر بر بیماری نداشت .

نمودار شماره ۲ وضعیت رفاهی بیماران در سه منطقه شمال، مرکز و جنوب شهر را نشان می دهد.

در مورد نوع تغذیه فراوانی مصرف سی نوع ماده غذایی در دو روز را بررسی نمودیم که این سی نوع ماده غذایی را به سه گروه مواد پروتئینی، مواد غذایی اصلی گیاهی و سبزیجات و میوه تقسیم بندی نمودیم. در مورد مصرف مواد یکه کمبود استفاده از آنها در بروز آنمی مؤثر ذکر می شود مثل گوشت قرمز و سبزیجات و برگی و بسیاری مواد پروتئینی تفاوت معناداری بین متوسط مصرف در گروه بیماروسالم دیده نشد.

در سه منطقه شمال و مرکز و جنوب شهر در مورد متوسط مصرف نوزده نوع مواد غذایی تفاوت معنادار دیده شد در حالیکه در مقایسه متوسط مصرف در دو گروه بیماروسالم فقط در مورد مصرف یک نوع ماده غذایی تفاوت معنادار دیده شد که نشان می دهد بیماری آنمی در تهران به خاطر سوء تغذیه نمی باشد و مردم مناطق جنوب شهر تهران که مصرف متوسط غذایی در آنجا پائین تر بوده حداقل سوء دلازم که به بیماری مبتلا نشوند را دریافت داشته اند.

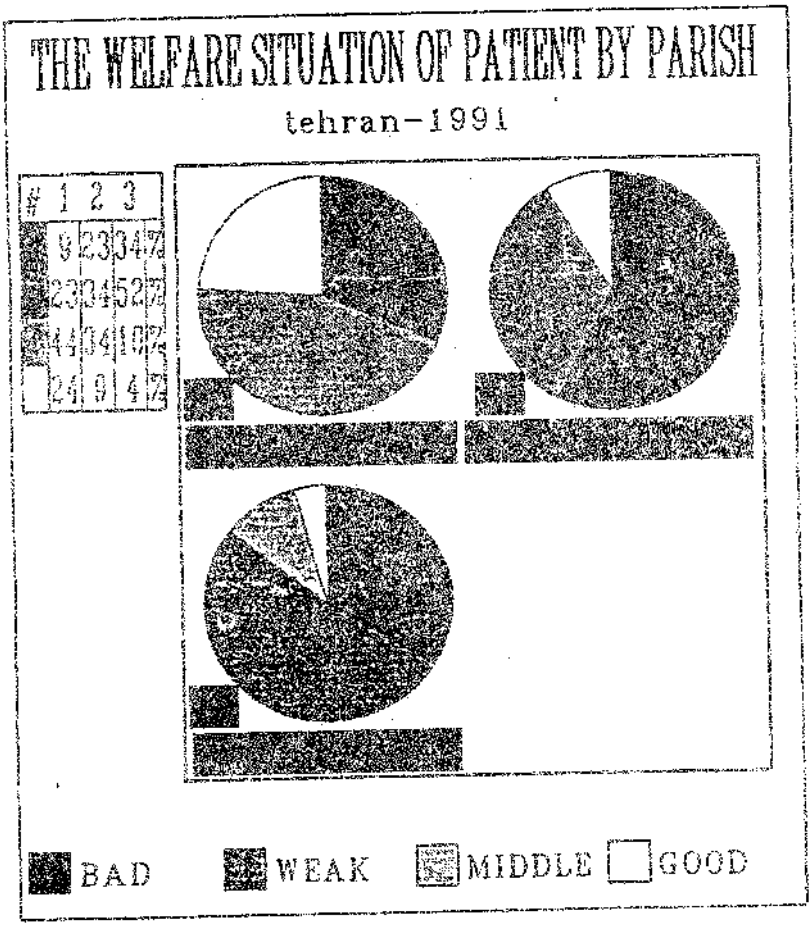
از سوی دیگر چون سوء تغذیه ارتباط مستقیم با وضعیت رفاهی افراد دارد بی از اینکه دیدیم وضع رفاهی در بروز بیماری تأثیری نداشته انتظار هم چنین بود که وضع تغذیه نیز در بروز بیماری تأثیری نداشته باشد.

جدول شماره ۱ : وضعیت بیماری آنمی بر اساس وضعیت رفاهی

بیماری آنمی		وضعیت رفاهی			
		بد	ضعیف	متوسط	خوب
افراد سالم	تعداد	816	1412	1185	453
	درصد	93.3	94.1	94.7	94
افراد بیمار	تعداد	53	88	66	29
	درصد	6.1	5.9	5.3	6

(شهر تهران - 1370)

نمودار شماره ۲ :



به‌جهت تعیین عوامل خطرمؤثر برآتمی به‌بررسی نوع ارتباط بین سن و جنس و محل سکونت وبیماری آتمی پرداخته‌ایم که‌خروجی برنامه‌های کامپیوتری وتحلیل آنها ارائه‌شده‌است .

مدل شماره : ۱

بررسی مدل لگ خطی بین بیماری آتمی و محل سکونت و جنسس و سن

با استفاده از برنامه SPSS

Tests that K-way and higher order effects are zero.

K	DF	L.R. Chisq	Prob	Pearson Chisq	Prob	Iteration
۴	۴	۷/۱۳۴	/۱۳۹۵	۵/۹۳۲	/۲۰۳۵	۲
۳	۱۶	۲۳/۳۶۰	/۱۰۳۴	۱۹/۰۵۳	/۲۶۵۹	۳
۲	۲۹	۹۳/۳۰۳	/۰۰۰۰	۸۳/۱۹۷	/۰۰۰۰	۴
۱	۳۵	۶۲۴۱/۲۱۳	/۰۰۰۰	۷۱۸۸/۲۲۳	/۰۰۰۰	۰

Tests of PARTIAL associations.

Effect Name	DF	Partial Chisq	Prob	Iter
Y١.X١.X٢	٢	٣/٧٥٨	/٣٢٥٨	٣
Y١.X١.ANH	٢	١١/٢٢٣	/٠٣٣٣	٣
Y١.X٢.ANH	٢	/٣٩٤	/٧٨١٣	٣
X١.X٢.ANH	٢	/٣٥١	/٧٩٨٣	٣
Y١.X١	٢	١٨/٢٩٣	/٠٠١٠	٣
Y١.X٢	٢	/٧٢٣	/٣٩٤٧	٣
X١.X٢	٢	٢٠/١٣٣	/٠٠٠٠	٣
Y١.ANH	٢	١٢/٢٢٩	/٠٠٠٨	٣
X١.ANH	٢	٧/٠١٧	/٠٢٩٩	٣
X٢.ANH	١	٩/٣٨٧	/٠٠١٩	٣
Y١	٢	١١٢/٧٥١	/٠٠٠٠	٣
X١	٢	٢٠١٢/٢١٢	/٠٠٠٠	٣
X٢	١	٣٨/٥١١	/٠٠٠٠	٣
ANH	١	٣٦٧٢/٣٢٥	/٠٠٠٠	٣

Backward Elimination for DESIGN ١ with generating class

Y١.X١.X٢.ANH

The final model has generating class

Y١.X١.ANH
X٢.ANH
X١.X٢

تحلیل مدل شماره ۱ :

در بررسی نوع ارتباط بین بیماری آنمی و سن وجنس ومحل سکونت با توجه به نتایج برنامه SPSS آزمون فرض صفربودن اثر مرتبه k وبالاتر نشان می‌دهد که فرض صفربودن اثر مرتبه سه وبالاتر پذیرفته می شود .

باید توجه داشت که چون این آزمون معنی داربودن تمام جملات مرتبه سه وبالاتر را تواما "بررسی می نماید ممکن است درعین معنی دارنبودن اثر مرتبه سه وبالاتر یکی از جملات جزء مرتبه سه در آزمون جملات جزء معنی دار گردد که در اینجا می بینیم جمله $Y_1.X_1.AN.$ معنی دار شده است . در این موارد سطوح معناداری نیز غالبا " تفاوت چندانی با هم ندارند . مقدار P برای پذیرش صفربودن اثر مرتبه ۲ وبالاتر ۱/۰ است ومقدار P برای معنی داری اثر جمله $X_1.Y_1.ANM$ ۰/۰۲۴ است .

مدل نهایی تولیدشده توسط کلاس $Y_1.X_1.ANM$ نشانگر همبستگی بین بیماری وسن ومحل سکونت است و $X_2.ANM$ نشانگر همبستگی بین جنس وبیماری است و $X_1.X_2$ نشانگر همبستگی بین سن وجنس است . مدل مناسب برای برازش به اطلاعات با توجه به سلسله مراتبی بودن مدل بصورت زیر نمایش داده می شود :

$$\text{Log } m_{ijk} = U+U_{ANM}(i)+U_{X1}(j)+U_{X_2}(k)+U_{y_1}(l)+U_{X1.X_2}(jk)+U_{X_2.ANM}(ik)$$
$$+U_{X_1.y_1}(jl)+U_{X_1.ANM}(ij)+U_{y_1.ANM}(il)+U_{X_1.y_1.ANM}(ijl)$$
$$i = 1,2 \qquad j = 1,2$$
$$k = 1,2 \qquad l = 1,2,3$$

در این مدل اثربیماری واثرسن واثرجنس واثرمحل سکونت واثرمتقابل بین سن وجنس واثرمتقابل بین جنس وبیماری واثرمتقابل بین سن ومحل واثر متقابل بین سن وبیماری واثرمتقابل بین محل وبیماری واثرمتقابل بین سن ومحل وبیماری وجوددارند .

پس از اینکه مقادیر مشاهده شده ومقادیر مورد انتظار روبا قیمانده ها وباقیمانده های استاندارشده ، تست نیکوئی برازش نسبت درستنمایی X^2 انجام شد که $P = 0/542$ نشان از پذیرش فرض مناسب بودن برازش مدل دارد .

مدل شماره : ۲

بررسی مدل تجمیعی بین بیماری آنمی و محل سکونت و جنس و سن

با استفاده از برنامه GLIM .

terms = ۱ + Y۱ + X۱ + X۲ + Y۱/X۱

X۱	Y۱	X۲	۱	۲
۱	۱	۱	./.۱۹۹۳۶۵۱۶	./.۳۰۲۵۷۹۵۱
		۲	./...	./...
		۳	./.۴۴۳۶۲۲.۷	./.۶۴۵.۳۵۸۲
	۲	۱	./.۶۲۵۷۲۵۹۸	./.۹۲۹۱۸۷۲۹
		۲	./.۱۵۵۳۲.۸۹	./.۲۳۶۵۸۱۸۲
		۳	./.۲۵۲۹۹۹۳۲	./.۶۷۸۷۵۷۹۵
۲	۱	۱	./.۵۷۳۱۱۳۹۲	./.۸۵۳۳۸۲۵۷
		۲	./.۳۸۳۶.۸۱۵	./.۵۷۶۸۷۴۳۲
		۳	./.۵۱۲۹۹۷۷۷	./.۷۶۶۲۵۷۲۹

unit	observed	out of	fitted	residual
۱	۴	۶۴	۱/۲۷۵	۱/۰۱۲
۲	۰	۵۷	۱/۷۲۸	-۱/۲۲۲
۳	۰	۴۱	۰/۰۰۰	-۰/۰۰۷
۴	۰	۵۰	۰/۰۰۰	-۰/۰۰۹
۵	۳	۸۶	۳/۸۱۵	-۰/۲۲۷
۶	۷	۹۳	۶/۱۸۵	۰/۲۲۹
۷	۸	۱۵۶	۹/۳۶۱	-۰/۵۸۲
۸	۱۷	۱۶۲	۱۵/۲۲۹	۰/۲۲۴
۹	۳	۱۲۲	۲/۲۲۸	۰/۵۱۲
۱۰	۳	۱۵۹	۳/۷۲۲	-۰/۲۲۷
۱۱	۱۲	۲۵۲	۱۱/۵۰۶	۰/۷۵۲
۱۲	۱۲	۲۲۲	۱۶/۲۹۲	-۰/۶۲۶
۱۳	۱۸	۲۷۶	۱۹/۸۲۰	-۰/۲۲۲
۱۴	۳۱	۲۵۹	۲۹/۱۷۰	۰/۲۰۶
۱۵	۱۲	۲۵۹	۱۲/۷۷۲	-۰/۲۸۷
۱۶	۲۹	۲۷۲	۲۷/۲۲۸	۰/۲۵۰
۱۷	۲۲	۲۲۵	۲۱/۸۰۲	۰/۲۶۲
۱۸	۲۷	۶۲۶	۲۸/۱۹۸	-۰/۱۸۰

	x^2 / x^1	1	2	
x^1	1	0.2.2874501	0.0000015429	
	2	0.0000016.68	0.0000015429	
	3	0.0.2222222.7	0.0.6660.25842	
	1	0.0.6660.25842	0.0.9291AV229	
	2	0.0.1866222.A1	0.0.2226A1A12	
	3	0.0.2822222.2	0.0.6VAV6V960	
	1	0.0.6922112.2	0.0.AD2222A7	
	2	0.0.2822.A16	0.0.6V6AV222	
	3	0.0.6122222.2	0.0.V6662V229	

دربررسی مدل نجیت بین بیماری آنمی وسن وجنس ومحل سکونت ابتدا برآورد با رامترها برای مدل نهایی $1 + y_1 + X_1 + X_2 + y_1.X_1$ ارائه شده است . این عبارت نشان می دهدکه درمدل اثرمحل سکونت وسن وجنس واثرمتقابل بین محل سکونت وسن روی بیماری آنمی وجوددارد .چون درمدل بیماری بصورت متغیر وابسته است لذا تاثیرسن برجنس دراین مدل واردنشده ولی درمدل SPSS این اثرنیزوجوددارد . با این توضیح مشخص می شودکه هماهنگی لازم بین جوابهای مدل SPSS با مدل GLIM وجوددارد .

پس از ارائه مقادیرمشاهده شده ومقادیربرازش شده وباقیمانده ها جدول ریسک ارائه شده است . با توجه به جدول فوق می توان خطرنسبی ابتلابه بیماری را درهر گروه سنی وجنسی برحسب محل سکونت محاسبه نمود .

برای مثال با تقسیم مقدارخطرزنان شهل شهردرگروه سنی ۱۴-۶ برمقدار خطرزنان شهل شهردرگروه ۵-۲ سال مقدار ۳/۰۷ بدست می آیدکه نشان می دهددر زنان شهل شهرخطررفتن ازگروه سنی ۵-۲ سال به ۱۴-۶ سال ۳/۰۷ برابرافزایش خواهدیافت .

لازم به ذکر است که دراین مدل ازبین متغیرهای سن وجنس ومحل اثر متقابل بین سن ومحل روی بیماری وجوددارد ولی اثرمتقابل بین سن وجنس واثر متقابل بین جنس ومحل روی بیماری وجودندارد . برای نشان دادن این مطلب ازجدول ریسک می توان استفاده نمود .

بعنوان مثال برای بررسی اثرمتقابل بین سن وجنس روی بیماری اگر درمحل ثابت خطر ابتلا به بیماری دریک جنس ازسنی به سن دیگربرابرمقدارخطر جنس دیگرازهمان سن به سن دیگرباشد یا نسبت این دومقدارخطرحدودیک باشد

نتیجه می گیریم اثر متقابل بین سن وجنس روی بیماری وجودندارد . اما اگر این نسبت مخالف یک شد نشانگر وجود اثر متقابل بین سن وجنس روی بیماری است

برای نشان دادن تاثیر متقابل بین سن ومحل روی بیماری ملاحظه می شود مثلا" خطر ابتلا به بیماری از مردان ۵-۲ ساله به مردان ۶-۱۴ ساله در شمال شهر به ۲/۱۴ برابر افزایش می یابد و همین خطر در جنوب شهر به ۱/۵۲ برابر افزایش می یابد و نسبت این دو مقدار ۲/۰۷ است که از یک فاصله دارد . در نتیجه اثر متقابل بین سن و محل روی بیماری وجود خواهد داشت .

برای نشان دادن عدم تاثیر متقابل بین سن وجنس روی بیماری ملاحظه می شود که خطر ابتلا به بیماری از مردان ۱۴-۶ ساله به مردان ۶۹-۱۵ ساله در مرکز شهر به مقدار ۲/۴۶ برابر افزایش می یابد همین خطر برای زنان ۱۴-۶ ساله به زنان ۶۹-۱۵ ساله در مرکز شهر به مقدار ۲/۴۴ برابر افزایش می یابد و نسبت این دو مقدار ۱/۰۱ است که نشانگر عدم تاثیر متقابل بین سن وجنس روی بیماری است .

برای نشان دادن عدم تاثیر متقابل بین جنس ومحل روی بیماری ملاحظه می شود خطر ابتلا به بیماری در گروه سنی ۶۹-۱۵ ساله شمال شهر از مردان به زنان ۱/۴۹ برابر افزایش می یابد و همین خطر در مرکز شهر به ۱/۵۰۳ برابر افزایش می یابد . نسبت این دو مقدار ۱/۰۰۹ است که نشانگر عدم تاثیر متقابل بین جنس ومحل روی بیماری است .

بحث :

درمورد محل سکونت نیز به نتایج مهمی دست یافتیم . تهران را به منطقه شمال و مرکز و جنوب تقسیم بندی نمودیم . متوسط هموگلوبین افراد ساکن در مرکز شهر در هر سه گروه سن تفاوت معناداری با متوسط هموگلوبین مردم شمال و جنوب شهر نشان داده و از آنها بالاتر بود . بنظر می رسد عامل آلودگی هوا با توجه به دلایل زیر می تواند توجیه منطقی برای این موضوع باشد :

آلودگی هوای مرکز شهر به اکسیدکربن بیشتر از مناطق شمال و جنوب شهر است زیرا اکسیدکربن بیشتر در اثر احتراق ناقص ماشینهای بنزین سوز تولید می شود که در منطقه مرکز شهر نیز تردد ماشینهای بنزین سوز بیشتر از مناطق دیگر است . میل ترکیبی هموگلوبین خون با اکسیدکربن دویست برابر بیشتر از میل ترکیبی آن با اکسیژن است و مقدار کمی اکسیدکربن از هوای استنشاق شده می تواند مقدار متنابهی از هموگلوبین را به کربکسی هموگلوبین تبدیل نماید .

این باند تشکیل شده دیگر نمی تواند عمل انتقال اکسیژن به بافتهای بدن را به انجام رساند از سوی دیگر مانع جدا شدن مولکولهای اکسیژن از هموگلوبین می شود و با کاهش فشار نسبی گاز اکسیژن در خون نیروی محرکه انتشار اکسیژن در بافتهای بدن را نیز کاهش می دهد . درمجموع باعث کاهش تحویل اکسیژن به بافتهای بدن می گردد و به همین خاطر بدن برای رفع این نقیصه به تولید بیشتر هموگلوبین میپردازد و بنظر میرسد به همین دلیل متوسط هموگلوبین افراد مرکز شهر از دیگر مناطق بالاتر می باشد .

بطور کلی در مورد بیست منطقه تهران متوسط هموگلوبین افراد ساکن در منطقه ۹ تهران نیز در هر سه گروه سنی ذکر شده از متوسط هموگلوبین افراد دیگر مناطق بالاتر بود که ظاهراً به خاطر وجود فرودگاه مهرآباد در این منطقه است که

تردد بسیار زیاد هوا بیماران و بروز زدرا رتفاع کم باعث آن شده است .

در تعریف W.H.O برای بیماری توجه به ارتفاع محل سکونت از سطح دریا شده است که هر مقدار ارتفاع بیشتر شود باید مقدار تعیین هموگلوبین برای بیماری نیز افزایش یابد . بنظر میرسد که لازم است در تعریف بیماری آنمی از روی میزان هموگلوبین به موضوع آلودگی CO هوا نیز توجه شود . زیرا در مناطق آلوده تر بصورت طبیعی هموگلوبین افراد بالاتر خواهد بود . جدول شماره ۲ .

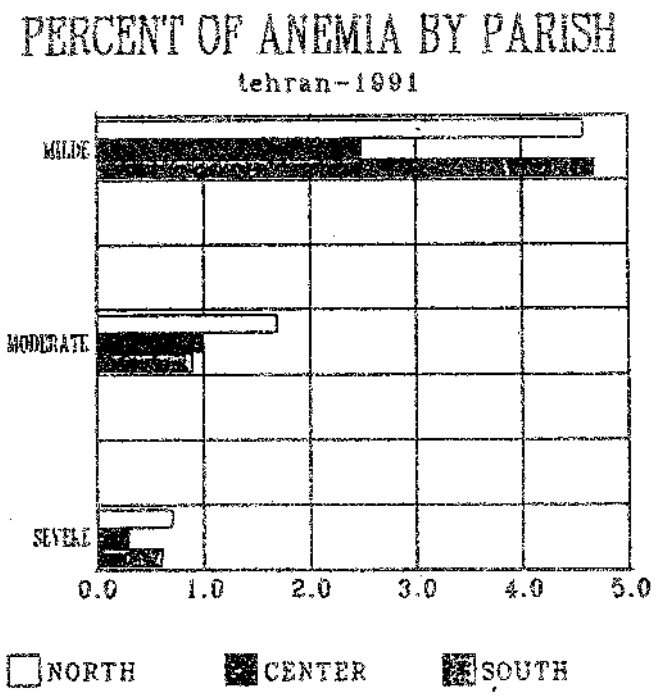
لازم به ذکر است که چون نوع این مطالعه مقطعی بوده است و از این نوع مطالعه نمی توان در مورد علتها نتیجه گیری قطعی نمود لذا در ادامه این تحقیق طرحی در معاونت پژوهشی وزارت بهداشت و درمان در حال بررسی است که طی آن هموگلوبین خون ۴۰۰ نفر از زنان خانه دار به همراه میزان اکسیدکربن محل سکونت آنها اندازه گیری می شود تا چگونگی ارتباط میان اکسیدکربن و هموگلوبین خون روشنتر گردد .

جدول شماره ۲: بررسی تفاوت میانگین هموگلوبین بین سه منطقه محل سکونت به تفکیک گروه سنی
علامت (*) نشانگر تفاوت معنی دار بین دو محل سکونت در سطح 5% می باشد

	2 - 5			6 - 14			15 - 69		
	شمال شهر	مرکز شهر	جنوب شهر	شمال شهر	مرکز شهر	جنوب شهر	شمال شهر	مرکز شهر	جنوب شهر
شمال شهر									
مرکز شهر	*		*	*		*	*		*
جنوب شهر									
SIG - F	0.0205			0.0000			0.0009		

(شهر تهران - 1370)

نمودار شماره ۵ :



فهرست منابع :

۱) بیماریهای خون ـ طب داخلی هاریسون ـ ۱۹۹۱ ـ ترجمه آقای دکتر شبر وانی

۲) خون چیست ـ نوشته آقای فرخ انصاری و احمدعابدینی ـ انتشارات آذر

۳) بهداشت ـ نوشته آقای دکترنیلفروشان ، آقای دکترضرابی ، آقای دکترمیرفتاحی

1)"Nutritional Anemias " Technical report world health organization - 1968

2)"Generalized linear models" Mc.Cullagh - Nelder' 1984

3)"Hematology"-Basic principles and practice Ronald - hoffman

4)"Text book of Paediatrics "-John-o.forfar Gavin-Ciarncll.

سرشماری عمومی ساختمان

ابوالفضل ابوالفتحی قمی

مرکز آمار ایران

مقدمه

علیرغم نقش نسبی حائز اهمیت بخش ساختمان و فعالیتهای ساختمانی در ایجاد اشتغال، سرمایه گذاری، تولید ناخالص داخلی، جریان مبادلات بین بخشهای اقتصادی، تعدیل نوسانات اقتصادی، گسترش زمینه مناسب برای انجام سایر فعالیتهای اقتصادی و ... متاسفانه اطلاعات آماری اندکی در ارتباط با آن گردآوری می‌شود و به طور نسبی می‌توان گفت که از جمله حوزه هایی از نظام اقتصادی کشور است که با کمبود جدی اطلاعات آماری روبه رواست.

بر پایه آخرین آمارهای در دسترس، تعداد شاغلان فعالیتهای ساختمانی کشور در سال ۱۳۶۵ بالغ بر ۱۲۰۷ هزار نفر بوده است که نزدیک به ۱۲ درصد کل اشتغال کشور است. همچنین در سال ۱۳۶۸، کل سرمایه گذاری انجام شده در ساختمان در سطح کشور حدود ۲۶۰۵ میلیارد ریال بوده است که بیش از ۸۱ درصد کل تشکیل سرمایه ثابت ناخالص کشور است. ضمناً در همین سال، در نتیجه انجام فعالیتهای ساختمانی حدود ۱۲۳۵ میلیارد ریال ارزش افزوده ایجاد شده است که نزدیک به ۵ درصد کل تولید ناخالص داخلی بدون نفت کشور است. از کل تشکیل سرمایه ثابت در ساختمان حدود ۶۵ درصد توسط بخش خصوصی و ۳۵ درصد توسط بخش عمومی انجام شده است. از این میزان تشکیل سرمایه حدود ۶۰ درصد در ساختمانهای مسکونی و ۴۰ درصد در سایر ساختمانهاست.

منابع آمار ساختمان

در زمینه منابع اطلاع آماری و طرحهای آمارگیری در زمینه ساختمان و فعالیتهای ساختمانی می‌توان به موارد زیر اشاره کرد:

- در هر سرشماری عمومی نفوس و مسکن، اطلاعات آماری مربوط به عرضه و مصالح ساختمانی عمده به کار رفته در واحدهای مسکونی و تعداد اطاق آنها جمع آوری می‌شود که تنها در ارتباط با قسمتی از مصادیق عینی مفهوم ساختمان، یعنی واحدهای مسکونی تحت تصرف خانوارهای ساکن آنهاست. به بیان دیگر، این سرشماریها، حتی مشخصات کلیه واحدهای مسکونی را نیز جمع آوری نمی‌کنند.

ـ بانک مرکزی ایران، همه ساله ، از طریق انجام یک طرح نمونه ای ، آمارهای مربوط به فعالیتهای ساختمانی بخش خصوصی در مناطق شهری کشور درارتباط با ساختمانهای جدید راجمع‌آوری می‌کند .

ـ دربرخی سازمانها وادارات و موسسات (مانند وزارت راه ، راه آهن ، سازمان بنادر وکشتیرانی ، سازمان هواپیمایی و فرودگاهها و ...) اطلاعات ثبتی معینی که درارتباط باعملکرد آن سازمانهاست ، جمع آوری می‌شود . برخی از این اطلاعات به بخش ساختمان و فعالیتهای ساختمانی مربوط می‌شود .

به‌این ترتیب ملاحظه می‌شود که اطلاعات آماری موجود در سازمانها و موسسات و نیزآمارهای جمع آوری شده‌از طریق طرحهای آماری درزمینه ساختمان و فعالیتهای ساختمانی موردی بوده و در چارچوب یک ساختار جامع ، هماهنگ ، بهنگام و از پیش اندیشیده شده در راستای تحقق اهداف‌مشخص قرار ندارد . لذا مرکز آمارایران در راستای انجام وظایف خود اقدام به تهیه مجموعه طرحهایی‌دراین زمینه کرده‌است که یکی از آنها طرح " سرشماری عمومی ساختمان " است .

هدف از اجرای این طرح گردآوری آمارو اطلاعات مربوط به مشخصات وویژگیهای ساختمانهای موجود مناطق و روستایی استانهای کشور است . باتوجه به‌این که مفهوم ساختمان به معنای وسیع آن دارای گستره بسیار وسیعی است . لذا دراین طرح تنها به ساختمان به مفهوم خاص آن توجه می‌شود که محتوی آن به شرح زیر است :

تعریف مفهوم " ساختمان "

با توجه به این که " ساختمان " مفهوم و مقوله اساسی و بنیادینی است که طرح سرشماری عمومی ساختمان براساس آن تهیه شده است و واحد آماری آن را تشکیل می‌دهد، لذا بایدویژگیهای آن شناخته شود تابتوان ساختمانها را ازیکدیگر ونیز از دیگرپدیده های جهان واقعیت تمیز داد .

طبق تعاریف بین‌المللی که مورد قبول و عمل سازمانهای آماری کشور بوده و ضمنا" در بررسیها و محاسبات اقتصادی نیز از آن پیروی می‌شود، ساختمان به مفهوم وسیع آن شامل زمینهای آماده سازی شده ، بناهای مسکونی ، بناهای غیر مسکونی ، راه ، جاده ، راه آهن ، بندرو اسکله ، شبکه

آب و برق و گاز وفاضلاب ، فرودگاه ، سد و ... هرگونه کالبد فیزیکی نظیر آنها می‌شود که درنتیجه کارانسان برروی زمین ایجاد می‌شود .

به وضوح دیده می‌شود که این مجموعه ، هرچند از نظر ویژگیشان به عنوان ساختمان دریک مقوله طبقه بندی می‌شود، از نظر سایر ویژگیهای خود دارای خصوصیات منحصر بفردی هستند . لذا برای جمع آوری اطلاعات مربوط به این ویژگیها لازم است طرحهای آماری ـ اطلاعاتی متعدد ومتنوعی تهیه واجراشود . تعداد و ساختار این طرحها وابسته به اهدافی است که برجمع آوری آمارهای ساختمان مترتب است . همین امر یکی از علتی است که موجب نپرداختن جدی وکار آمد به آمارگیریهای ساختمان درگذشته شده است .

در طرح سرشماری عمومی ساختمان ، به مجموعه بناهای مسکونی و غیر مسکونی و عرصه مرتبط با آنها ، ساختمان گفته شده است . بنابراین دیده می‌شود که این طرح نیزتنها بخشی از ساختمان راتحت پوشش قرارمی‌دهد که همانا ساختمان به مفهوم خاص آن است . علت نیز آن بوده است که اطلاعات مربوط به سایر ساختمانها را از طرق دیگری (که از نظر نیروی انسانی، زمان ، منابع مالی و ... به هزینه های کمتری احتیاج دارد) می‌توان جمع آوری کرد و گذشته از آن ، میزان اهمیت وموارد استفاده از ویژگیهای گروههای مختلف ساختمانها در برنامه ریزیها و طرح ریزیهای شهری ومنطقه ای متفاوت است . بنابراین طرحها و روشهای دیگری برای جمع آوری آمارو اطلاعات سایر ساختمانها در نظر گرفته شده است .

درطرح حاضر، ساختمان به عنوان فضایی متشکل از یک یا چند بناو عرصه متعلق و مرتبط با آن در نظر گرفته است . بنابراین ساختمان فضایی است محصور، ثابت ، مستقل که تمام یا قسمتی از آن مسقف بوده و برای سکونت یا فعالیت اقتصادی ساخته شده است ، مشروط به آن که مساحت قسمت مسقف آن از ۲ متر مربع کمتر نباشد .

دریک سرشماری ، نحوه اجرای کار باید به گونه‌ای طراحی و پیش بینی شود که هیچ یک از واحدهای آماری مورد نظر از قلم نیفتد و یا بیش از یکبار به‌حساب نیایند . این امردرسرشماری عمومی ساختمان از طریق انتخاب دومرحله‌ای واحد آماری مدنظر قرارگرفته است .

دراین سرشماری ، باتوجه به ارتباط بین ساختمان و قطعه زمین ، درمرحله اول قطعه زمین به عنوان واحد آماری انتخاب شده است و سپس درمرحله دوم ساختمان به عنوان واحدآماری مدنظر قرار گرفته است . بنابراین اطلاع ازمفهوم قطعه زمین دراین طرح به شناخت بیشتر آن کمک می‌کند .

مفهوم قطعه زمین ونقش آن

باتوجه به هدف طرح که شمارش کلیه ساختمانها و گردآوری اطلاعات و آمار مربوط به ویژگیهای آن است ، لذا به گونه‌ای باید بتوان آن را به عنوان واحد آماری طرح مشخص ساخت . در این راستا، باتوجه به واقعیات عینی ، قطعه زمین به عنوان واحدآماری مرحله اول طرح سرشماری عمومی ساختمان مدنظر قرارگرفت .

قطعه زمین سطحی است که توسط معبر ، مانع طبیعی یا مصنوعی ، دیوار، پرچین و نظایر آنها از اطراف خود محدود شده است .

بنابراین ، قطعه زمین به صور زیر دیده می‌شود :

- زمین اشغال شده توسط یک یاچند بنا (که همان ساختمان باشد)
- زمین اشغال نشده توسط یک یاچند بنا که می‌تواند محصور یا غیر محصور باشد .

در طرح سرشماری عمومی ساختمان ، کلیه قطعه زمینهای واقع درمحدوده شهرها و آبادیهای کشور فهرست می‌شوند و ویژگیهای چارچوبی آنها پرسش می‌شود . باتوجه به این که قسمت اعظم قطعه زمینهای فهرست شده را قطعه زمینهای اشغال شده توسط یک یاچند بنا (= ساختمان) تشکیل داده و در ضمن ویژگیهای مورد نظر برای ساختمانها به میزان معنابهی بیشتر از ویژگیهای سایر انواع قطعه زمینهاست و در پرسشنامه طرح ، این ویژگیها مورد پرسش قرارمی‌گیرد، لذاساختمان به عنوان واحد آماری مرحله دوم مد نظر قرار گرفته است . لذا محتوی اصلی طرح را جمع آوری آمارهای مربوط به ساختمانهای کشور تشکیل می‌دهد و به همین دلیل به عنوان سرشماری عمومی ساختمان شناخته شده است .

اقلام آماری مورد نظر در طرح

پس از مشخص کردن هرساختمان و ایجاد تمایز بین آن و سایر ساختمانها، اطلاعات و اقلام آماری زیر در مورد هرساختمان گردآوری می‌شود که مجموعه آنها تصویر دقیق وروشنی از ویژگیها و مشخصات کالبدی ، سکونتی ، اقتصادی و اجتماعی ساختمانهای کشور را به دست می‌دهد .

- وضعیت ساختمان از نظر فعالیتهای ساختمانی (تمام شده ، نیمه تمام در دست ساخت ، نیمه تمام متوقف)
- مساحت زمین و زیربنای ساختمان .
- مصالح به کار رفته در بنای ساختمان و نوع اسکلت آن
- تعداد طبقات ساختمان
- وضعیت روکار و نوع روکار ساختمان
- نوع بام و مصالح به کار رفتهٔ براندودکردن بام ساختمان .
- سازنده و مالک ساختمان
- سال ساخت ساختمان
- تاسیسات و تجهیزات ساختمان
- نوع فعالیت اقتصادی در ساختمان و تعداد کارکنان و مساحت زیربنای مرتبط با آنها
- مشخصات قسمت مسکونی ساختمان (شامل زیربنا، تعداد اطاق ، تعداد واحد مسکونی ، تعداد خانوار و جمعیت ساکن آنها و ...)

مفهوم ومحتوی و نیز محدوده واجزای هریک از این اقلام وابسته به ملاحظات عینی و هدفهای طرح است که خود بحثهای مفصلی را می‌طلبد که برای ایجاز کلام از آن درمی‌گذریم .

روش اجرا و پوشش

این طرح به صورت مراجعه مستقیم و حضوری مامور آمارگیر به کلیه واحدهای آماری و مشاهده آن و نیز مصاحبه با فرد مطلع انجام می‌شود وکلیه قطعه زمینها و ساختمانهای واقع در محدوده طرح (شهر یا آبادی) را پوشش می‌دهد . برای اجرای این طرح نیز تمهیداتی در نظر گرفته شده است که

متناسب باهدفهای طرح و وضعیت عینی ساختمانهای مورد مراجعه است .

طرح مورد اشاره به‌صورت آزمایشی درسال ۱۳۶۹ در مسجد سلیمان به‌اجرا درآمد که نتایج آن منتشر و مورد استفاده برنامه ریزان قرار گرفته است . نتایج حاصل از اجرای طرح نشان دهنده قابلیت عملی قابل قبول روشها و مفاهیم نظری و عملی پیش بینی شده در آن است که به دلیل کوتاهی وقت از آن می‌گذریم و علاقمندان را به مطالعه مجموعه مدارک و اسناد این طرح دعوت می‌کنیم .

معرفی تئوریهای مرگ‌ومیر

علی‌اکبر احمدی

مرکز آمار ایران

۱ـ مقدمه

برای معرفی تئوریهای ریاضی و آماری مرگ‌ومیر باید طرحهای مربوط به توصیف ویژگیها و قوانین عمومی مرگ‌ومیر و همچنین روشهای محاسبه را شناخت. مطالعه یک مدل مرگ‌ومیر میتواند ما را در روند مرگ‌ومیر در گذشته وهمچنین بدست آوردن مدلسی برای پیش بینی وقایع آینده یاری دهد. مدلهای ارائه شده غالبا" همبستگی بین سن ومرگ‌ومیر را نشان میدهند و تا وقتی در این مدلها سن به عنوان یک متغییر مطرح میشود، تئوریهای مرگ و میر به عنوان یکی از وجوه تئوری سن (aging) در نظر گرفته می‌شود.

شاید اولین مدل ریاضی مرگ‌ومیر مدل گامپرتزاست (Gompertz) که درسال ۱۸۲۵ میلادی منتشر شد. این مدل عمدتا" رابطه مرگ‌را با سن نشان میدهد. با توجه به این که فرض گامپرتز برا ینکه اصل استوار بوده که قدرت انسان با بالا رفتن سن در مقابله بامرگ کاهش می یابد. این مدل یک تابع نمایی است.

مدل گامپرتز در سال ۱۸۶۰ توسط ماه هام (Makeham) کاملترشد و بالاخره پرکز (Perks) مدل گامپرتز ومک هام را کامل نموده ومدل بهتری را ارائه نمود.

ارا ئه مدلهای ریاضی مرگ‌ومیر بر اساس عوامل ژنتیکی یکی دیگر از مباحث تئو ریهای مذکور است. به عبارت دیگر علاوه بر دخالت متغیرهای سن اگر عوامل ژنتیکی رانیز در نظر بگیریم، اینگونه مدلها پدید میآیند. و بالاخره ارائه مدلهای استوکاستیک، حد مطلوبها رائه مدلهای مرگ‌ومیر است. زیرا منطبق با مشاهدات آماری درزمینه مرگ‌ومیر و عوامل آن میباشد.

در این مقاله که هدف معرفی تئوریهای مرگ‌ومیر است ابتدا به یک سری قرار دادهها و شاخصهایی که

معمولا " در جداول مرگومیـــر وجود دارد اشاره خواهد شـــد .

۲-جداول مرگومیر ،نمادهای قراردادی ومحاسبه احتمال مرگومیـــر

معمولا " درجداول مربوط به مرگومیر ازعلایـــم ونمادهای زیر استفاده می شـــود .

l_x ..... تعداد افـــــرادی کـــه در ســـن x سالگـــی زنده باشنـــد .

d_x ..... تعداد افرادی که درانتروال سنی $[\ x\ ,\ x+1\ ]$ می میرنـــد .

q_x ..... احتمال این که فردی x سالش باشد وبه سن $(x+1)$ نرســـد .

p_x ..... احتمال این که فردی x سالش باشدوتا $(x+1)$ سالگی هم زنده باشد .

${}_nq_x$ ..... احتمال اینکه فردی x سالش باشدو قبل از رسیدن به سن $(x+n)$ می میرد .

${}_np_x$ ..... احتمال این‌که فردی x سالش باشد وتا $(x+n)$ سالگی هم زنده باشـــد .

${}_{n|m}q_x$ احتمال این که فردی x سالش باشد در فاصله n سال اول زنده باشدو در فاصله m سال دوم بمیـــرد .

${}_nd_x$ ..... تعداد افرادی که در انتروال سنی $[\ x\ ,\ x+n\ ]$ میمیرنـــد .

رابطه بین نمادهای فوق الذکـــر به شرح زیر میباشـــد .

$$d_x = l_x - l_{(x+1)} \quad , \quad {}_nd_x = l_x - l_{(x+n)}$$

$$q_x = \frac{d_x}{l_x} \quad , \quad p_x = \frac{l_{x+1}}{l_x} \quad , \quad q_x + p_x = 1$$

$${}_nq_x = \frac{l_x - l_{(x+n)}}{l_x} \quad , \quad {}_np_x = \frac{l_{(x+n)}}{l_x}$$

$${}_{n|m}q_x = \frac{l_{(x+n)} - l_{(x+n+m)}}{l_x}$$

$${}_{n|m}q_x = {}_np_x \cdot {}_mq_{(x+n)}$$

با استفاده از روابط مذکور احتمالهای مختلف مرگ و میر را می‌توان محاسبه کرد. به چند مثال زیر توجه کنید.

مثال ۱ـ احتمال این که فرد ۲ ساله قبل از این که به پنج سالگی بمیرد چقدر است.

$$_3q_2 = \frac{l_2 - l_5}{l_2} = 1 - \frac{l_5}{l_2} = 1 - {}_3p_2$$

مثال ۲ـ احتمال این که فردی ۱۰ ساله باشد و تا ۲۰ سالگی هم زنده باشد چقدر است.

$$_{10}p_{10} = \frac{l_{20}}{l_{10}} = 1 - {}_{10}q_{10}$$

مثال ۳ـ احتمال این که دو مرد یکی ۵ ساله و دیگری ۲۵ ساله، ۵ سال دیگر هم زنده باشند چقدر است.

$$_5p_{10} \cdot {}_5p_{25} = \frac{l_{10+5}}{l_{10}} \cdot \frac{l_{25+5}}{l_{25}} = \frac{l_{15}}{l_{10}} \cdot \frac{l_{30}}{l_{25}}$$

مثال ۴ـ احتمال این که فردی در فاصله n سال اول زنده باشد و در فاصله m سال دوم بمیرد.

$$_{n|m}q_x = \frac{l_{(x+n)} - l_{(x+n+m)}}{l_x}$$

x $x+n$ $x+n+m$

۳ـ مدلهای قاطع (۱)

مدل گامپرتز عمدتاً رابطه مرگ را باسن نشان میدهد زیرا فرض گامپرتز بر این اصل استوار بوده است که قدرت انسان با بالا رفتن سن در مقابله با مرگ کاهش می‌یابد.

قبل از معرفی مدل گامپرتز، به محاسبه قدرت مرگ میسر (تابع مرگ و میر) میپردازیم.

در قسمتهای قبل گفتیم که برای محاسبه احتمال این که فردی x ساله باشد و قبل از رسیدن به سن $x+n$ بمیرد، برابر $_nq_x$ است که عبارتست از:

$$_nq_x = \frac{l_{(x)} - l_{(x+n)}}{l_{(x)}}$$

---

(۱) Deterministic Models

حال اگر بخواهیم این احتمال را در فاصله زمانی h بدست آوریم یعنی در انتروال سنی $[x \text{ و } x+h]$

بنا بر این :

$$\lim_{h\to 0} \frac{{}_h q_x}{h} = -\frac{1}{l_x} \lim_{h\to 0} \frac{l_x - l_{x+h}}{h} = -\frac{1}{l_x} \cdot \frac{dl_x}{dx}$$

$$\boxed{\mu(x) = \lim_{h\to 0} \frac{{}_h q_x}{h} = -\frac{1}{l(x)} \cdot \frac{dl(x)}{dx}} \quad (1)$$

تابع $\mu(x)$ ، مرگ و میر را در هر لحظه زمان مشخص میکند . با استفاده از رابطه (۱) داریم :

$$\mu(x)\,dx = -\frac{dl(x)}{l(x)} \Rightarrow \int \mu(x)dx = -\ln l(x)$$

$$\boxed{l(x) = e^{-\int \mu(x)\,dx}} \quad (2)$$

رابطه (۲) تابع حیات را برحسب تابع مرگ و میر نشان می دهد .

$$\ln l(x) = -\int \mu(x)dx \Rightarrow -\int_0^y \mu(x)dx = \left[\ln l(x)\right]_0^y = \ln l(y) - \ln l(0)$$

$$-\int_0^y \mu(x)dx = \ln \frac{l(y)}{l(0)} \Rightarrow \frac{l(x)}{l(0)} = e^{-\int_0^x \mu(x)dx}$$

$$\Rightarrow \boxed{l(x) = l(0)\, e^{-\int_0^x \mu(x)dx}} \quad (3)$$

رابطه (۳) تابع حیات را بر حسب تابع مرگ و میر در فاصله زمانی $[x \text{ و } 0]$ سالگی نشان می دهد و بالاخره تعداد افرادی که در فاصله زمانی $[x \text{ و } x+n]$ می میرند از رابطه (۴) بدست میاید .

$$n\,d_x = l(x) - l(x+n) = \int_x^{x+n} \mu(x)\, l(x)\, dx$$

با توجه به رابطه (۲) ملاحظه می شود که تابع $l(x)$ یک تابع نمایی نزولی است که با توجه به شناخت تابع مرگ و میر ($\mu(x)$) و تابع حیات ($l(x)$) به معرفی مدلهای مرگ و میر

می‌پردازیم . مدل مرگ و میر گامپرتز در سال ۱۸۲۵ به صورت $f(x)=Bc^x$ ارائه شده است .
که در آن B میزان مقاومت در برابر مرگ و میر و c نرخ مرگ و میر است . B و c پاراستر هایی
هستند که با استفاده از مشاهدات موجود در جداول مرگ و میر بدست میآیند . مدل گامپرتز برای سنین
بچگی و پیری مناسب نبوده ، بلکه برای گروههای سنی مناسب است که یک عامل خارجی مستقل از سن
در ارتباط با مرگ و میر آنها وجود دارد . این عوامل می‌توانند در طبیعت به طور تصادفی و یا موردی
از ضرب و جرح و امراض عفونی بوجود آمده باشد . با توجه به این عوامل مک هام (Makeham)
در سال ۱۸۶۰ یک عامل ثابت را به مدل گامپرتز اضافه کرد و مدل $f(x)=A+Bc^x$ را ارائه
نمود . اگر چه فرمول " مک هام " بیشتر از فرمول " گامپرتز " با مشاهدات مطابقت دارد (بویژه
برای سنین جوانی) ولی در هر دو مدل میزان مرگ و میر زیادی را در سنین آخر زندگی بدست می دهند ،
زیرا در بعضی از موارد مشاهده شده است که نرخ مرگ و میر کندتر از "روند نمایی" مدل گامپرتز
افزایش یافته و حتی همتراز هم میشود . برای غلبه بر این مشکل " پرکس " (Perks) فرمول
گامپرتز و مک هام را به صورت زیر ارائه نمود .

$$f(x)=\frac{A+Bc^x}{1+Dc^x}$$

که در آن مخرج برای کم کردن نرخ مرگ و میر در سنین بالا موثر می باشد .

۴ - تئوریهای ژنتیکی (۱)

بدیهی است که خصوصیتهای مختلف افراد مربوط به اختلافات ژنتیکی آنهاست . به نظر
میرسد که طول عمر نیز جزء این خصوصیات باشد . بنا براین میتوانیم این فرض را در نظر بگیریم
که طول عمر یک نفر بدون در نظر گرفتن "تاثیرات محیط" کاملا" بستگی به ساخت ژنتیکی او دارد . در این
مورد مثالی توسط : آردی کلارک " (R . D . Clark) پیشنهاد شده است . کلارک انسانها "
مساله خود را به برآورد نرخ مرگ و میر با توجه به سن مطرح کرد ، به طوری که این مرگ و میر ممکن است
نتیجه عوامل محیطی نظیر تصادف یا بیماری باشد . چنین مرگ و میر هایی را کلارک به عنوان " پیش بینی
طبیعی پایان زندگی " نامیده و تمامی مرگ و میر های دیگر را تحت عنوان " مرگ های پیری " که ریشه
ژنتیکی دارند طبقه بندی نمود . او فرض کرد که چنین مرگهایی را نمی‌توان به تصویق انداخت . برای برآورد
" مرگ و میرهای بر اثر پیری "کلارک ابتدا گروههای سنی متوالی را مورد مطالعه قرار داد و نسبت مرگ

(۱) Genetic theories

و میر تمامی افـــراد یک‌گروه (بیشتر بر اثـــر بیماریهای قبلی بوده انـــد) رابدست آورد . کلارك تشخیص داد که بعضـــی از مرگ هـــا در گروه " مرگ بر اثـــر پیـــری" نیستند . کلارك از این نسبتها در تعیین مدل " مرگ و میر بر اثر پیری " (عوامل ژنتیکی) استفاده کرد و با بکار بردن آنها جـــدول امید زندگی را برای مردان طی سالهای ۱۹۳۲ - ۱۹۳۰ بنـــا کـــرد .

یک مدل ریاضـــی برای طول عمر بر اساس عوامل ژنتیکی که قابـــل کاربرد برای زندگـــی تمام پستانداران می باشـــد . در سال ۱۹۵۹ توسط زیـــلارد (Szilard) پیشنهاد شد . پایه واساس فرض " زیلارد " این است که : سن در هنگام مرگ منحصرا " توسط عوامل ژنتیکی فرد تعیین می شـــود . به این معنـــی که تعدادی از ژنها ی توارث در این مورد دخالت می‌کنند و بقیه ژنها که برای سلامتـــی باقی می ماننـــد ممکن است به طور تصادفـــی صدمه ببینند .

۵ - مدلهای احتمالـــی (۱)

مدلهای احتمالی به‌طور ایده ال و در حد قابل قبولـــی قابلیت کاربرد در جمعیتی همگن را داراســـت . یعنی افرادی که بنیان جمعیت را میسازنـــد اساسا " از نقطه نظر خصوصیات شبیه همهستند . چنین جمعیتی از این نـــوع را ، فقط در صورتی می‌توان در نظر گرفت که عده آنها زیاد باشـــد (حجم جامعه زیاد باشد) . زیـــرا که تغییرات هر خصوصیت در رابطه با میانگین اش " کوچك " خواهـــد بود . از طرف دیگر مدل احتمالی وجـــود نوسانات آماری در مشاهـــدات پدیده ها را می پذیرد ، به عبارت دیگر مدل احتمالی یک متغیر ، توزیـــع احتمالات را در بطـــن خود داراست . یعنی وقتـــی که افراد با خصوصیات متفاوتـــی توصیف می‌شـــوند این مـــدل نسبتا " پیچیده می شـــود .

" بیرد " در جهت ادامه تحقیقات خود تصمیم گرفت که یک علت خاص مرگ و میر را جدا کرده و مـــدل منحنـــی مرگ و میر این علت را بررسی کنـــد . بدین منظور ، مرگ و میـــر ناشی از سرطـــان ریـــه که در انگلستان بیـــش از حد شیوع کرده بود برگزید . در ابتداء فرض کرد که مرگ و میر در سن x در طـــول سال t با فرمول زیر بیـــان میشـــود .

$$\mu^{S}_{x,t} = K \cdot H_{t-x} \cdot \Phi_t \cdot F_x$$

که در آن H_{t-x} تابعـــی است فقط از سال تولـــد و Φ_t تابعـــی وابسته در سال مورد مطالعـــه و F_x تابعی از سن می باشد . تابـــع H_{t-x} توسط " بیرد " ثابـــت در نظر گرفته شده اســـت ،

(۱) Stochastic Models ۴۵/۶

زیرا که ترکیب ژنتیک یک جمعیت در طول زمان بکندی تغییر می‌کند، Φ_x ممکن است ناپیوسته باشد زیرا که مرگ و میر در سالها توسط اپیدمی ممکن است ناپیوسته باشد و باتوجه به شرایط اقلیمی و محیط بوجود آیند و تابع F_x توسط " بیرد " بعنوان مرگ و میر خالص در نظر گرفته شده است .

بالاخره " بیرد " (Beard) بر اساس عوامل ژنتیکی و بکارگیری مدلهای احتمالی و مدلهای قاطع مدل زیر را معرفی کرد .

$$\mu(x) = \frac{\int \eta(s)\,\mu^{s}(x)\cdot \exp\left(-\int_0^x \mu^{s}(k)\,dk\right) ds}{\int \eta(s)\,\exp\left(-\int_0^x \mu^{s}(k)\,dk\right) ds}$$

در رابطه فوق $\mu(x)$ میانگین وزنی $\mu^{s}(x)$ است و "بیرد" نشان داد که در موارد خاصی که $\eta(s)$ به عنوان یک تابع گاما توزیع شده باشد و $\mu^{s}(x)$ به صورت فرم " ملک هام " باشد، $\mu(x)$ از یک توزیع مجانبی تبعیت می‌کند . نسبت $\eta(s)$ به‌عنوان یک شاخص عوامل ژنتیکی جمعیت می‌باشد .

در سال ۱۹۶۰ مدلی معادل مدل " بیرد " توسط ساشر (Sacher) پیشنهاد شده بدنبال ارائه مدلهای استوکاستیک وتکمیل نسبی آنها ، یک تئوری مرگ و میر توسط استرحلر (Strehler) و میلدوند (Mildvand) پیشنهاد شد . در این مدل " اتلاف انرژی مورد لزوم برای بازسازی هر زیر سیستم ارگانیک " مورد نظر بوده است . این آقایان قوه زیست هر زیرسیستم را به صورت " ماکزیمم انرژی که برای بازگرداندن آن به حالت تعادل لازم است " بیان نمودند . حال باتوجه به اینکه " نوسانات تقاضا برای مصرف انرژی اساسا " بستگی به نوسانات محیط انرژی زای (محیط داخلی ارگانیزم) ارگانیزم دارد " بنابراین یک مدل استوکاستیک در تئوریهای مرگ و میر ایجاد خواهد شد . ساشر و تروکو (Trocco) بدنبال معرفی تئوریهای مرگ و میر بر اساس تلفیق متناسب مدلهای قاطع ، ژنتیکی و به تبع آن ساخت ، مدلهای احتمالی ، ایده اشان را به صورت زیر مطرح کرده اند که میانگین مقادیر متغیرهای فیزیولوژیکی اساسی ، نقطه ای در فضای η بعدی است به طوری که در دور این نقطه یک سطح چند بعدی محاط شده است ، مادامی که نقطه ای مصرف نیازهای فیزیولوژیکی یک ارگانیزم ، در داخل منطقه محاط شده ، حرکت کند این ارگانیزم زنده باقی می‌ماند واگر خارج این سطح چند بعدی محاط شده باشد ، ارگانیزم

می‌میرد، ساشر و شروکسوا این حد را "Lethal bound" نامیدند.

در خاتمه دست یابی به تئوریهای مرگ و میر به طوری که از هر نظر جنبه کاربری آنها مطرح باشد هدف نهایی آمارشناسان و متخصصین ژنتیک است. بنابراین در جداول مرگ و میر استاندارد که معمولا" شاخصهایی اندازه گیری می‌شوند که بر اساس مشاهدات جمعیتی استوار است باید شاخصهای مختلف ژنتیکی هم منظور شود تا مدلهای احتمالی به صورت کاربردی مورد استفاده قرار گیرد. ولی از آنجایی که ساخت شاخصهای عوامل مختلف ژنتیکی کاربسیار دشواری است معمولا" در جداول مرگ و میر صحبتی از آنها بمیان نمی آید.

در این مقاله تئوریهای مرگ و میر به‌طور خلاصه ازابتدا، تاکنون مطرح شد، مدلهای ارائه شده بیشتر قالب ریاضی و یا احتمالی محض دارد وکمتر به صورت عملی میتوان ازآنها استفاده کرد. امیدوارم که‌مطرح شدن تئوریهای مرگ و میر باعث شود دانشجویان دوره‌های کارشناسی ارشد و دکترا در رشته‌های آمار، جمعیت شناسی، ژنتیک ونظایر آنها، بتواند بر روی مدلهای فوق ویامدلهای ارائه شده دیگر به صورت عملی وکاربردی فعالیت کرده و مدل مرگ و میر را برای جامعه ایران و گروههای سنی مختلف باعوامل ژنتیکی اقلیمی خاص کشور ارائه دهند.

منابع مورد استفاده :

(1) Techniques of Population Analysis by Barclay
(2) Introduction to Demography by Spiegelman

بررسی خطای نمونه‌ای در طرح آمارگیری کشاورزی ۱۳۶۶ و تأثیر طبقه‌بندی بعد از آمارگیری در کاهش خطا

محمدصادق احمدی‌دانش‌آشتیانی

مرکز آمار ایران

۵شنبه ۱۵:۳۰ تالار ۲

خطای نسبی در نتایج این طرح بیش از حدّ مورد انتظار بوده است و بررسی واریانس برآوردها نشان می‌دهد که سهم واریانس بین PSU بیش از ۹۵درصد واریانس برآورد است. به منظور کاهش میزان خطا، در این بررسی سعی شده است از نتایج سرشماری عمومی کشاورزی ۱۳۶۷ و برخی اطّلاعات کمکی دیگر به منظور طبقه‌بندی مجدّد آبادیها (PSUها) کمک گرفته شود و برآوردهای جدید براین‌اساس محاسبه شود. به این ترتیب تاکنون ۵ نوع طبقه‌بندی مجدّد انجام شده و خطای نسبی برآوردها با خطای نسبی در طرح اولیّه (بدون طبقه‌بندی مجدّد) مقایسه شده است. چون هیچیک از این روشها تأثیر قابل ملاحظه در کاهش خطای نسبی نداشته‌اند هیچکدام برای استفاده عملی توصیه نمی‌شوند. با توجّه به اینکه تعداد اقلام آماری پرسش‌شده بسیار زیاد و متنوّع هستند، نتیجه‌گیری می‌شود که این اقلام به چند گروه تقسیم شوند و برای هرگروه طبقه‌بندی براساس صفت مناسب همان انجام شود.

نقش آمار در اندازه‌گیری سطح تکنولوژی

علی‌اصغر آزادگان

مرکز آمار ایران

با پیشرفت دانش بشر به‌تدریج داده‌های کیفی به کمی‌تبدیل شده و نتیجتا " برای آمار و اطلاعات استفاده های جدیدتری مطرح می‌شود . یکی ازآخرین کاربردهای آمارواطلاعات ،بهره‌گیری‌از آن دراندازه‌گیری سطح تکنولوژی در فعالیتهای‌تولیدی می‌باشد . قبل ازتشریح کاربردآمارواطلاعات دراندازه گیری سطح تکنولوژی ضرورت داردمختصری درخصوص تکنولوژی بیان شود . اولین سئوالی که مطرح می‌شوداین است که به طورکلی تکنولوژی چیست ؟ تاکنون برای تکنولوژی تعاریف گوناگونی ارائه شده‌است و هر محققی آن رااززاویه خاصی‌مورد بررسی قراردادده وبرای آن تعریفی ارائه کرده‌است . اگر تعاریف عنوان شده برای تکنولوژی جمع آوری شودمحققا " کتاب قطوری‌را تشکیل خواهد داد . از میان تعاریف گوناگون ارائه شده‌برای تکنولوژی ، یکی از معتبرترین آنها به شرح ذیل است :

دریک سیستم تولیدی ، تکنولوژی‌عامل تبدیل داده به ستانده دریک فضای ملی است . برای‌اندازه‌گیری تکنولوژی لازم است ابتدا عوامل آن شناسایی شود . از نظرکلی تکنولوژی دارای چهارعامل " ماشین آلات وتجهیزات (TECHNOWARE) " ، " توانایی‌های‌انسانی (HUMANWARE) " ، " دانش واطلاعات فنی (INFORWARE) " و " سازماندهی‌ومدیریت (ORGANWARE) " است .

نمودار ۱ ـ عوامل تشکیل دهنده تکنولوژی را نشان می‌دهد .

الف ـ عامل ماشین آلات و تجهیزات (TECHNOWARE) ماشین‌آلات و تجهیزات قلب هرسیستم تولیدی است . این ابزار عمدتا " از طریق توانایی ما و با استفاده از اطلاعاتی که

در طول زمان اندوخته می‌شود توسعه می‌یابد. درجات پیچیدگی این عامل طیف وسیعی را تشکیل می‌دهد که از ابزار دستی مانند آچار آغاز و به روبات ختم می‌شود. درجات پیچیدگی این عامل به شرح ذیل است :

- ماشین آلات و تجهیزات دستی
- ماشین آلات و تجهیزات برقی
- ماشین آلات و تجهیزات فنی عمومی
- ماشین آلات و تجهیزات فنی تخصصی
- ماشین آلات و تجهیزات فنی کامپیوتری
- ماشین آلات و تجهیزات فنی تمام کامپیوتری

برای اندازه گیری سهم این عامل در تکنولوژی به هریک از درجات پیچیدگی امتیازی نسبت داده شده است و برحسب اینکه تکنولوژی مورد استفاده در فرآیند تولید از این عامل به چه میزانی از پیچیدگی استفاده می‌کند امتیاز آن درجه از پیچیدگی به آن عامل تعلق می‌گیرد.

ب ـ توانایی های انسانی (HUMANWARE)، توانایی‌های انسانی در انجام عملیات صنعتی نقش کلیدی دارد این توانایی‌ها است که تجهیزات را سود بخش می‌نماید. برای توانایی‌های انسانی نیز درجات پیچیدگی گوناگونی پیش بینی شده است که از توانایی کار با ماشین شروع و به توانایی در نوع آوری ختم می‌شود. در حال حاضر درجات پیچیدگی که برای توانایی‌های انسانی در نظر گرفته شده است به قرار ذیل است :

- توانائی کار با ماشین
- توانایی نصب و راه اندازی ماشین
- توانایی تعمیر
- توانایی تولید قطعات یدکی
- توانایی بهبود کمیت و کیفیت

ـ توانایی در نوآوری

همانگونه که قبلاً " اشاره رفت به هر یک از درجات پیچیدگی امتیازی اختصاص داده شـــده است و بر حسب اینکه در آینده تولید از این عامل در چه سطحی بهره می‌گیرد امتیاز آن سطح در محاسبات منظور می‌شود.

ج ـ دانش فنی و اطلاعات (INFORWARE) ، اطلاعات نمودار تجمع دانش بشری است . امروز میزان دانش موجود به نحو مستمری در حال رشد است و بنابر این اطلاعات موجـــود نیازمند به روز شدن می‌باشد و اگر این عمل انجام نشود استفاده بهینه از ماشین آلات و تجهیزات و توانایی های انسانی ممکن نخواهد شد. درجات پیچیدگی دانش فنی و اطلاعات به شرح ذیل است :

ـ اطلاعات آشنا کننده

ـ اطلاعات توضیح دهنده

ـ اطلاعات مشخص کننده

ـ اطلاعات بهره گیرانه

ـ اطلاعات تشریح کننده

ـ اطلاعات تصمیم دهنده

ـ اطلاعات ارزیابی کننده

د ـ سازماندهی و مدیریت (ORGAWARE) ، سازماندهی و مدیریت بهره‌گیری از ماشیـــن آلات و تجهیزات ، توانایی های انسانی و دانش فنی و اطلاعات را در هر فعالیت صنعتـــی هماهنگ می‌کند تا منابع طبیعی به نحو مطلوب به مواد تولید شده تبدیل شود . اگر چـــه نوع سازماندهی و مدیریت مورد نیاز به میزان پیچیدگی تجهیزات ، توانایی های انسانـــی و دانش فنی و اطلاعات بستگی دارد اما در مواردی می‌تواند به عنوان عامل تعیین کننـــده سه عامل دیگر نقش داشته باشد . برای این عامل نیز مانند سه عامل دیگر درجات پیچیدگـــی

تعیین شده است که به شرح ذیل ارائه می‌شود:

ـ مرحله تلاش

ـ مرحله وابستگی

ـ مرحله شهود

ـ مرحله حمایت

ـ مرحله تثبیت

ـ مرحله کامیابی

ـ مرحله رهبری

برای اندازه گیری سطح تکنولوژی در یک فرآیند تولید " ضریب کمک فنی (TCC)[۲] " تعریف شده است که این ضریب از طریق مدل زیر که یک مدل کاپ داگلاس است محاسبه می‌شود.

$$TCC = T^{\beta_T} . H^{\beta_h} . I^{\beta_i} . O^{\beta_o}$$

که در آن :

TCC ، ضریب کمک فنی

T ، ماشین آلات و تجهیزات

H ، توانایی‌های انسانی

I ، دانش فنی و اطلاعات

O ، سازماندهی و مدیریت

می‌باشد.

بر اساس اینکه در فرآیند تولید از هر یک از عوامل چهارگانه تکنولوژی در چه سطحی از درجات پیچیدگی استفاده می‌شود به T ، H ، I و O امتیازی تعلق می‌گیرد و این امتیازات پس از استاندارد شدن ، در مدل منظور می‌شود.

۲) TECHNOLOGY CONTRIBUTION COEFICIENT

محاسبه ضریب کمک فنی برای تولید کالا در یک واحد یا یک کشور نقطه‌شان می‌دهد که در فرآیند تولید کالا چه سطحی از تکنولوژی مورد استفاده قرارگرفته است و همانگونه که در مدل [illegible] مشاهده می‌شود مقدار تولید در ضریب کمک فنی نقشی ایفا نمی‌کند - به عنوان مثال ضریب کمک فنی نقطه‌شان می‌دهد که کارخانه الف در تولید نخ نایلون در مقایسه با کارخانه ب از تکنولوژی پیشرفته‌تری استفاده می‌کند حال آنکه ممکن است تولید کارخانه الف هزار تن و تولید کارخانه ب یک میلیون تن در سال باشد . در جهت غلبه دادن به مقدار تولید در اندازه گیری سطح تکنولوژی تعریف دیگری تحت عنوان " محتوی تکنولوژی افزوده ($TCA^{۳}$) " مطرح می‌شود .

محتوی تکنولوژی افزوده از رابطه زیر حاصل می‌شود :

$$Tca = \lambda \cdot Tcc \cdot VA$$

که در آن :

λ ، ضریب فضای مثلی تکنولوژی است که در آن فضا عمل تبدیل انجام می‌گیرد .

VA ، ارزش افزوده ایجاد شده می‌باشد .

روش اندازه گیری سطح تکنولوژی در تعدادی واحد تولیدی به مرحله عمل در آمده است که از آن جمله می‌توان صنایع تولید الیاف مصنوعی و صنایع تولید آهن و فولاد را نام برد . در اجرای این روش با توجه به ویژگی های صنایع تولیدی در کشور تغییراتی در پارامترهای موجود در هر یک از مدل به شرح ذیل ایجاد شده است :

۱ - توانایی های انسانی ، برای عامل نیروی انسانی صفات خاصی در نظر گرفته شده است و وجود یا عدم این صفات در چهار گروه کارگران ، تکنسینها ، مدیران و کارکنان تحقیق و توسعه مورد بررسی قرار گرفته است ، شرح صفات تعیین شده به قرار ذیل است :

- ابتکار .

۳) TECHNOLOGY CONTENT ADDED

درحالت کلی ممکن است تساوی اول بدون برقراری تساوی دوم

درست باشد درا ینصورت زنجیررا نا همگن گوئیم .

ماتریس تغییروضعیت :

۲-۱- تعریف :

فرض کنیم فضای وضعیت متناهی ودارای K عضوباشد. تعداد

$i_0,\dots i_{p-1}Tj_1,\dots j_p$ رابرای احتمال مشروط زیرا نتخاب میکنیم :

$$(1.2)\quad P(X_n=i_0,\dots,X_{n-p+1}=i_{n-p+1} \mid X_{n-1}=j_1,\dots,X_{n-p}=j_p)$$

$$(i_k,j_k=1,\dots,K)$$

$$\begin{cases} P(X_n=i_1 \mid X_{n-1}=i_1,\dots,X_{n-p}=i_p) & \text{if } i_m=j_m \quad m=1,\dots,p-1 \\ 0 \quad \text{otherwise} \end{cases}$$

$i_0Ti_1,\dots,i_p$ ماتریسی مربعی ازمرتبه K^p خواهد

بود بطوریکه درسطراول K نخستین درایه غیرصفروبقیه درایه ها دقیقا

صفرهستند درسطردوم ازدرایه K+1 ام تادرایه 2K ام غیرصفروبقیه

صفرهستندو.... نظیروضعیت فوق ازسطر K+1 ام تاسطر 2K ام تکرار

میشود.....

۳-۱ ـ مثال : اگردریک زنجیرمارکف مرتبه 2 فضای وضعیت $\{0,1,2\}$

باشدآنگاه ماتریس تعریف شده در 2-1 بصورت زیرخواهدبود :

$$\begin{bmatrix} 0T00 & 0T01 & 0T02 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0T10 & 0T11 & 0T12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0T20 & 0T21 & 0T22 \\ 1T00 & 1T01 & 1T02 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1T10 & 1T11 & 1T12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1T20 & 1T21 & 1T22 \\ 2T00 & 2T01 & 2T02 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2T10 & 2T11 & 2T12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2T20 & 2T21 & 2T22 \end{bmatrix}$$

چنانکه ملاحظه میشودمجموع عناصرهرستون عبارتست از :

$$(1.3) \quad 1 = \sum_{i_0} P(X_n = i_0 \mid X_{n-1} = i_1, X_{n-2} = i_2)$$

این نتیجه درحالت کلی برای T نیزصحیح است .

۳-۱ ـ قضیه : اگر T ماتریس مذکوردر 1.2 و $Q_0 = P_{x_0, \ldots, x_{p-1}}$

ماتریس توزیع اولیه زنجیرباشدکه درآن

$$P_{x_0, \ldots, x_{p-1}} = P(X_0 = x_0, \ldots, X_{p-1} = x_{p-1})$$

آنگاه $Q\ P_{x_1,\ldots,x_p}$ عبارتست از $TQ_{\cdot}$.

برهان : $P(X_1=x_1,\ldots,X_p=x_p)=\sum_{X_0} P(X_0=x_0,\ldots,X_p=x_p) =$

$\sum_{X_0} P(X_p=x_p \mid X_{p-1}=x_{p-1},\ldots,X_1=x_1,X_0=x_0)P(X_0=x_0,\ldots,X_{p-1}=x_{p-1})$

$\sum_{X_0} x_p Tx_{p-1},\ldots,x_0 Px_0,\ldots,x_{p-1}$

$\Rightarrow \left[P(X_1=x_1,\ldots,X_p=x_p)\right] = \left[\sum_{X_0} x_p Tx_{p-1},\ldots,x_0 Px_0,\ldots,x_{p-1}\right]$

۵-۱ - نتیجه : Q_n ماتریس توزیع توام $X_n,\ldots,X_{n+p}$

از $T^n Q_0$ بدست میآید.

برهان : با استفاده از خاصیت شرکت پذیری ضرب معمولی ماتریسها نتیجه میشود.

فرم ادغام شده ماتریس انتقال collapsed form

۵-۲ - تعریف : ماتریس C از مرتبه $K\times K^p$ را فرم ادغام شده ماتریس T گوئیم هرگاه C متشکل از اعضای ناصفر T باشدکه در 1.2 به شکل $Ti_1,\ldots,i_p$ نشان داده شد.

با توجه به این تعریف شکل ادغام شده ماتریس انتقال T مذکور در مثال عبارتست از:

$$\begin{bmatrix} 0T00 & 0T01 & 0T02 & 0T10 & 0T11 & 0T12 & 0T20 & 0T21 & 0T22 \\ 1T00 & 1T01 & 1T02 & 1T10 & 1T11 & 1T12 & 1T20 & 1T21 & 1T22 \\ 2T00 & 2T01 & 2T02 & 2T10 & 2T11 & 2T12 & 2T20 & 2T21 & 2T22 \end{bmatrix}$$

۲-۱- کاربردفرم ادغام شده ماتریس انتقال

اگرتابع توزیع احتمال حاشیه‌ای X_n را بصورت K عنصری برداری $P_n = P(X_n = i)$ تعریف کنیم آنگاه $P_n = CQ_{n-1}$

$$P_n = \sum_{i_1} \cdots \sum_{i_p} P(X_n = i_1, \ldots, X_{n-p} = i_p)$$

واضح است که T را میتوان از C بدست آورد و برعکس . همچنین اگر $1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

آنگاه $1'C = 1'$

****

قضیه زیرکمک میکندکه تمام تئوری مربوط به زنجیرمارکف مرتبه اول را درزنجیرهای مارکف مرتبه p بکاربریم :

۲-۱- قضیه : هرزنجیرمارکف مرتبه p قابل تحویل به زنجیرمارکف مرتبه اول است .

اثبات : فرض کنیم $\{X_t : t \in I\}$ یک زنجیر مارکف مرتبه p با فضای وضعیت S باشد (S متناهی) Y_{n-1} را وضعیت فرایند در مراحل $n-p$ ام و ... و $n-1$ ام تعریف می کنیم . بنابراین $Y_{n-1}=(X_{n-p},\dots,X_{n-1})$ در اینصورت $\{Y_t : t \in I\}$ فرایندی است با فضای وضعیت S' زیرمجموعه‌ای از $S^k=S\times\dots\times S$ و داریم :

$$P(Y_n=y_n \mid Y_{n-1}=y_{n-1}, Y_{n-2}=y_{n-2}) = P(X_n=x_n,\dots,X_{n-p+1}=x_{n-p+1},$$

$$X_{n-1}=x_{n-1},\dots,X_{n-p}=x_{n-p},X_{n-2}=x_{n-2},\dots,X_{n-p-1}=x_{n-p-1}) =$$

$$P(X_n=x_n,\dots,X_{n-p+1}=x_{n-p+1} \mid X_{n-1}=x_{n-1},X_{n-2}=x_{n-2},\dots,$$

$$X_{n-p-1}=x_{n-p-1}) = P(X_n=x_n \mid X_{n-1}=x_{n-1},\dots,X_{n-p-1}=x_{n-p-1})$$

که بنا به خاصیت مارکف مرتبه p در $\{X_t\}$

$$= P(X_n=x_n \mid X_{n-1}=x_{n-1},\dots,X_{n-p}=x_{n-p}) = P(X_n=x_n,\dots,$$

$$X_{n-p+1}=x_{n-p+1} \mid X_{n-1}=x_{n-1},\dots,X_{n-p}=x_{n-p}) =$$

$$P(Y_n=y_n \mid Y_{n-1}=y_{n-1})$$

پس :

$$P(Y_n=y_n \mid Y_{n-1}=y_{n-1}, Y_{n-2}=y_{n-2}) = P(Y_n=y_n \mid Y_{n-1}=y_{n-1})$$

که نشان میدهد $\{Y_t\}$ یک زنجیر مارکف مرتبه اول است .

قضیه زیر ارتباط بین ماتریس انتقال در دو فرایند مذکور را تعیین می کند .

۱۰۹ قضیه : اگر $\{Y_t\}$ زنجیر مارکف مرتبه اول ، بدست آمده از زنجیر مارکف مرتبه p $\{X_t\}$ باشد آنگاه ماتریس تغییر وضعیت در دو فرایند یکسان است .

اثبات :

$$y_n T y_{n-1} = P(Y_n = y_n \mid Y_{n-1} = y_{n-1}) =$$

$$P(X_n = x_n, \cdots, X_{n-p+1} = x_{n-p+1} \mid X_{n-1} = x_{n-1}, \cdots, X_{n-p} = x_{n-p}) =$$

$$P(X_n = x_n \mid X_{n-1} = x_{n-1}, \cdots, X_{n-p} = x_{n-p}) = x_n T x_{n-1} \cdots x_{n-p}$$

یعنی

$$T_{Y_t} = y_n T y_{n-1} = x_n T x_{n-1} \cdots x_{n-p} = T_{X_t}$$

که در آن T_{X_t} و T_{Y_t} بترتیب ماتریسهای تغییر وضعیت مربوط به دو زنجیر $\{X_t\}$ و $\{Y_t\}$ است .

شرط تحویلنا پذیری $\{Y_t\}$ برحسب $\{X_t\}$

قضیه :

۱۱۰ : اگر در $\{X_t\}$ از یک وضعیت مفروض بتوان در یک مرحله به هر وضعیت دلخواه در S رسید آنگاه $\{Y_t\}$ تحویلنا پذیر است .

اثبات : مطابق فرض داریم $i, j \in S$ $\forall\ i, j: P_{ij} > 0$

برای آنکه $\{Y_t\}$ ساده نشدنی باشد می بایست برای هر m و n ای k ای بیابیم بشکلی : $P^k(y_m, y_n) > 0$ ثابت می کنیم این شرط برقرار است :

فرض کنیم $y_m=\underline{z}=(z_1,\ldots,z_p)$ و $y_m=\underline{t}=(t_1,\ldots,t_p)$

همچنین فرض کنیم p و q کوچکترین عددی باشند که $z_q=t_1,\ldots,z_p=t_{p-q}$

$P(y_n,y_m)$ یعنی وقوع بشکل زیر :

$$\underbrace{z_1\ z_2\ \ldots\ z_q\ \ldots\ z_p}_{t_1\ \ldots\ \ldots\ t_p}$$

در مطابق مفروضات قضیه احتمال آن برابر است با :

$$P_{z_1}P_{z_1,z_2}\cdots P_{z_{q-1},t_1}P_{t_1,t_2}\cdots P_{t_{p-1},t_p}$$

چون تمام عناصر عبارت اخیر طبق فرض قضیه مثبت هستند پس حاصلضرب آنها

هم مثبت است و در نتیجه $P^q(y_m,y_n)>0$ پس $K=q$

در صورتیکه q با شرایط فوق موجود نباشد آنگاه : $P^p(y_n,y_m)=P_{z_1}P_{z_1,z_2}\cdots P_{z_{p-1},z_p}\cdot P_{z_p,t_1}\cdot P_{t_1,t_2}\cdots P_{t_{p-1},t_p}>0$

قضیه حدی در مدل انورگرسیو زنجیر مارکف مرتبه : p

۱۱-۱ قضیه : فرض کنید $\{X_t : t\in I\}$ بصورت زیر تعریف شده باشد :

$$P(X_n=j_0 \mid X_{n-1}=j_1,\ldots,X_{n-p}=j_p)=\sum_{i=1}^{p}\lambda_i q^{(i)}_{j_0 j_1} \qquad (1.5)$$

که در آن $\sum_{i=1}^{p}\lambda_i=1$ و $Q=[q_{jk}]$ یک ماتریس نامنفی $K\times K$

که جمع ستونهایش یک است . بطوریکه $0<\sum_{i=1}^{p}\lambda_i q^{(i)}_{jk_i}<1$

و $\Pi=(\pi_1,\ldots,\pi_k)$ بردار گونه‌ای باشد که $\pi_i>0$ $(i=1,\ldots,k)$

و $\sum_{i=1}^{k}\pi_i=1$ و $Q\Pi=\Pi$ آنگاه ،

$$\lim_{n\to\infty}P(X_n=j \mid X_p=i_p,\ldots,X_1=i_1)=\pi_j \qquad (i_1,\ldots,i_p,j=1,\ldots,k)$$

اثبات : فرض کنیم T ماتریس انتقال مربعی از مرتبه K^p مذکور

در T^{p-1} باشد هرسطر T با یک p تایی بصورت $i_0,\ldots,i_{p-1}$

ارتباط دارد بقسمیکه i_0 باکندی بسیار تغییرمیکند، i_1 باکندی کمتر

و..... وبطورمشابه ستونهای T با $j_1,\ldots,j_p$ مرتبط اندکه به همان

ترتیب مرتب شده اند.

ابتدا نشان میدهیم که T ارگودیک است (تحویل ناپذیر

وناد وره‌ای) . تمام حالتهای آنها هم مرتبطند پس T تحویلناپذیر

است واز میان اعضای قطری T ، k تا ناصفرند آنهایی که مربوطند

به $i_0=i_1=\ldots=i_{p-1}=j_1=\ldots=j_p=1 \quad (i=1,\ldots,k)$

که طبق فرض $q_{11}>0$ پس این k حالت نامتناوب هستند. اما T

تحویلناپذیراست پس تمام حالتها نامتناوبند وبدلیل متناهی بودن T

توزیع ایستای ξ را داردکه دران $T\xi=\xi$ یا عناصر

$\xi_{i_1\ldots i_p}=\lim P(X_{n-1}=i_1,\ldots,X_{t-p}=i_p)$ صدق میکنند. توجه آرایشی که دران

i_1 بسیارکند تغییرمیکند و..... فرض کنیم $w=(w_1,\ldots,w_k)$

توزیع حالتهای یک بعدی مربوطه با تقریب‌خوبی ξ فرم انتظام خسته

که یک ماتریس $K\times K^p$ است واضح است که درحالت کلی $C\xi=w \quad (A\cdot1)$

همچنین برای حل (۱-۵) $C=\sum_{j=1}^{p}\lambda_j U_j$ که دران

$$A_{j,k}=\begin{cases}0 & \text{if } j=k\\ -1 & \text{if } j\neq k\end{cases} \qquad \text{و} \qquad U_j=A_{j,1}\times\ldots\times A_{j,1}$$

و $\times$ علامت ضرب مستقیم ماتریسها از راست است یعنی $A\times B=\left[a_{ij}B\right]$

حال C را به روش دیگری محاسبه میکنیم . عنصر r ام $U_j\xi$ برابر است با

$$\sum_{i_1\cdots i_p} q_{ri_j}\,\xi_{i_1\cdots i_p} = \sum_{i} q_{ri_j} \sum_{\substack{i_h;h=1,\ldots,p\\ h\neq j}} \xi_{i_1\cdots i_p}$$

$$= \sum_{ij} q_{ri_j}\, w_{ij}$$

که عنصر r ام QW است بنابراین

$$(A.2)\qquad C\xi=\sum_{j=1}^{p} q_{\lambda_j}(QW)=QW\sum_{j=1}^{p} q_{\lambda_i}=QW$$

مقایسه A.1 و A.2 نشان میدهد $QW=W$ که جمل منحصر بفرد آن توسط Cox و Miller در (1965p124) عبارتست از $W=\pi$

معرفی محکهایی برای تخمین مرتبه زنجیرمارکف . برآوردهای AIC و BIC

AIC ; Akaike's Information Criterion

BIC ; Bayesian Information Criterion

برآورد AIC فقط برای زنجیرهای مارکف با فضای وضعیت متناهی است که فرض میشود ساده ، ندنی و نا متناوب (ارگودیک) هستند وطبق یک فرض ثانوی مرتبه زنجیر مارکف مفروض درحالیکه ناشناخته است از عدد ثابتی مانند m بیشتر نیست .

با داشتن n مشاهده از زنجیر $\{X_t\}$ تابع درست نمایی برای زنجیر مارکف مرتبه p عبارتست از :

$$L(X,p)=\prod_{i_1 \cdots i_p} (p_{i_1 \cdots i_{p+1}})^{n_{i_1 \cdots i_{p+1}}}$$

که در آن $n_{i_1 \cdots i_{p+1}}$ نشاندهنده شمار انتقالهایی بصورت $i_1 \to \cdots \to i_{p+1}$ است که در نمونه رخ داده ماکزیمم درستنمایی دراز ای برآورد ML از $p_{i_1 \cdots i_{p+1}}$ بدست می آید Billingsley(1961a) نشان داده است برآورد ماکزیمم نسبت درستنمایی برای $p_{i_1 \cdots i_{p+1}}$ عبارتست از $\hat{p}_{i_1 \cdots i_{p+1}}=\frac{n_{i_1 \cdots i_{p+1}}}{n_{i_1 \cdots i_p}}$ که در آن $n_{i_1 \cdots i_p}=\sum_{i_{p+1}} n_{i_1 \cdots i_{p+1}}$

بنابراین ماکزیمم تابع درستنمایی دراز ای این برآورد عبارت

است $\prod_{i_1}^{n_1}\cdots\prod_{i_{p+1}}^{n_{p+1}}(\hat{p}_{i_1\cdots i_{p+1}})^{n_{i_1}\cdots i_{p+1}}$ $H_p(X)=$ برای استفاده از قضیه

نیمن پیرسن نسبت ماکزیمم تابع درستنمایی $\lambda_{p,m}$ را تشکیل میدهیم :

$$\lambda_{p,m}=\frac{H_p(\underline{X})}{H_m(\underline{X})}$$

که برای آزمون $H_0:p_T=p$ در برابر $H_1:p_T=m$ (p_T مرتبه واقعی زنجیر است) تشکیل داده میشود. اگر

$$\eta_{p,m}=-2\mathrm{Ln}\lambda_{p,m}$$

$$\eta_{p,m}\sim\chi^2_{(k^m-k^p)(k-1)}$$

برآوردکننده AIC برای رتبه زنجیر مارکف بصورت زیر تعریف میشود :

۱۲-۱ تعریف : برآورد بعنی انتخاب میشود که :

$$AIC(p_{AIC})=\min_{0\le p<m}AIC(p)$$ که در آن

$$AIC(p)=\eta_{p,m}-2(k^m-k^p)(k-1) \qquad (1.6)$$

مشاهده میشود که 1.6 آماره نسبت درستنمایی است که بوسیله تابع مجازات دوبرابر درجه آزادی اصلاح می گردد. برای این روش نداشتن بهینگی نشان داده میشود. Kats بوسیله قضیه زیر نشان میدهد که این روش با احتمالی مثبت سازگار نیست

۱۲-۱- قضیه : فرض کنید $(j=0,\ldots,k-1)$ Z_j متغیرهای تصادفی

مستقل با توزیع $\chi^2_{k^j(k-1)^2}$ باشند. قرار می دهیم :

$$S_p=\sum_{j=p}^{K-1} Z_j - 2k^j(k-1)^2 \qquad p=0,\ldots,k-1$$

آنگاه :

$$\lim P\left\{\hat{p}_{AIC}=p\right\} = \begin{cases} 0 & \text{if } 0 \leqslant p < p_T \\ P(S_p=\min S_j \quad (p_T \leqslant j < m) & \text{if } p_T \leqslant p < m \end{cases}$$

برای یافتن برآوردکننده سازگار با استفاده از روش بیز برآوردکننده سازگار زیر بدست می آید .

۱۴-۱ تعریف : برآوردکننده BIC برای مرتبه زنجیر مارکف بگونه‌ای انتخاب می شود که

$$BIC(\hat{p}_{BIC})=\min_{0 \leqslant p < m} BIC(p)$$

که در آن

$$1.7 \quad BIC(p)=\eta_{p,m}-(K^m-K^p)(K-1)\frac{\ln n}{n}$$

چنانکه ملاحظه میشود تابع مجازات برآوردماره نسبت درستنمایی عبارتست از درجه آزادی ضریب لگاریتم طبیعی تعداد دنباله . مقایسه دو رابطه 1.6 ، 1.7 نشان میدهد که برای $n \geqslant 8$

$$\hat{p}_{BIC} \leqslant \hat{p}_{AIC}$$

۱۵-۱ قضیه : p_{BIC} برآوردکننده سازگار مرتبه واقعی زنجیر مارکف (p_T) است و یا

$$\lim p(\hat{p}_{BIC}=p_T)=1$$

اثبات در Katz .

References

1. RAFTERY (1985) Amodei for high order Markov chain
j.r. Statist. soc.B(£ 1985) 47 No.3pp.528-539.

2.PEGRAM.1960 Ans autoregressive model for multilag Markovc chain
j.r. Appl.prob. 17 (1960) No.17 p.p.350-362

3.KATS (1981) Onsome criteria for estimating the order of Markov chain.
TECHNOMETRICS. vol.23 No.3 August 1981 pp. 243-249

4.COX and MILLER (1965) The theory of stochastic proce sses

5.BILLINGSLEY (1961a) Statistical methods in Markov chain
Ann Math . Statist. 32 . pp .12-40

کوتاهترین بازه‌های اطمینان برای خانواده توزیعهایی که تکیه‌گاه آنها به پارامتر وابسته است

مجید اسدی*- ناصر رضا ارقامی

دانشگاه فردوسی مشهد

شنبه ۱۵:۰۰ تالار ۶

فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال مفروض: $f(x,\theta)=\frac{g(x)}{h(\theta)} \quad a(\theta)\le x\le b(\theta)$ باشد. HUZURBZAR (1955) نشان داد که کمیت $Q=\frac{h(T)}{h(\theta)}$ (T آماره بسنده است) یک کمیت محوری است و براساس آن کوتاهترین بازه اطمینان $(1-\alpha)100\%$ را برای θ را بدست آورد. Ferentinos (1990) خانواده توزیع:

$$f(x,\theta)=\frac{g(x)}{h(\theta_1,\theta_2)} \quad a(\theta_1,\theta_2)<x<b(\theta_1,\theta_2)$$

را مورد بررسی قرار داد و کوتاهترین بازه اطمینان برای $h(\theta_1,\theta_2)$ و θ_1 و θ_2 را بدست آورد. در این مقاله نشان می‌دهیم که کوتاهترین بازه اطمینان $(1-\alpha)100\%$ و بازه اطمینان مبتنی بر درستنمایی در خانواده‌های بالا یکسانند و در حالتی که متغیر تصادفی X از نوع گسسته باشد و دارای تابع احتمال $f_\theta(x)=\frac{g(x)}{h(\theta)}x\in\{d,\dots,\theta\}$ کوتاهترین بازه اطمینان برای θ در سطح $(1-\alpha)100\%$ را بدست می‌آوریم.

روش بوت استرپ (Boot Strap Method)

انیس ایرانمنش

دانشگاه فردوسی مشهد

فرض کنید بخواهیم درمورد برآوردگری مثل $\hat{\theta}$ ویژگیهایی از قبیل واریانس، اریبی و یا توزیع آن را بدست آوریم. درحالتی که توزیع جامعه معلوم است، معمولاً می توان با شبیه گیری های متعدد از جامعه ویژگیهای موردنظر را بدست آورد:

$$\begin{cases} X_1, X_2, \cdots, X_n \longrightarrow \hat{\theta}_1 \\ X_1, X_2, \cdots, X_n \longrightarrow \hat{\theta}_2 \\ \vdots \\ X_1, X_2, \cdots, X_n \longrightarrow \hat{\theta}_B \end{cases}$$

$$Bias(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_b - \theta$$

$$Var(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B} (\hat{\theta}_b - E(\hat{\theta}))^2 \cong \frac{1}{B}\sum_{b=1}^{B} (\hat{\theta}_b - \hat{\theta}^{\cdot})^2$$

$$\hat{\theta}^{\cdot} = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_b$$

هم چنین با داشتن مشاهدات متعدد می توان توزیع $\hat{\theta}$ را به شکل زیر بدست آورد:

$$F_{\hat{\theta}}(t) = p(\hat{\theta} \leq t) = \frac{\#\{\hat{\theta}_b \leq t\}}{B}$$

واضح است درحالتی که جامعه معلوم است بنا برقانون اعداد بزرگ وقتی $B \to \infty$ آنگاه:

پارامترها	برآوردگرها
$F_{\hat{\theta}}(t)$	$\frac{\#\{\hat{\theta}_b \le t\}}{B}$

درعمل توزیع جامعه دراختیا رمانیست وفقط نمونه‌ای ازجامعهٔ اصلی در دست است دراین حالت با نمونه‌گیری های مکرربا جایگذاری ازنمونهٔ موجود ویژگیهای موردنظربرآوردگررا برآوردمی کنیم . این روش به‌روش بوت‌استرپ موسوم است وتوسط Efron (1979) ابداع گردیده است .

نمونه تصادفی موجود $X_1, X_2, \dots, X_n$

$$\begin{cases} X_1^*, X_2^*, \dots, X_n^* \longrightarrow \hat{\theta}_1^* \\ X_1^*, X_2^*, \dots, X_n^* \longrightarrow \hat{\theta}_2^* \\ \vdots \\ X_1^*, X_2^*, \dots, X_n^* \longrightarrow \hat{\theta}_B^* \end{cases}$$

$$\widehat{Bias}(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}^* - \hat{\theta}$$

$$\widehat{Var}(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B} (\hat{\theta}_b^* - \hat{\theta}^*)^2$$

$$\hat{\theta}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_b^*$$

هم چنین با داشتن مشاهدات متعدداز نمونهٔ موجودکه به جای جامعه مورد استفاده قرارگرفته است میتوان توزیع $\hat{\theta}$ را برآوردکرد ، که توزیع بوتاسترپ $\hat{\theta}$ نامیده میشود :

$$\widehat{F_{\hat{\theta}}}(t) = P(\hat{\theta}^* \le t) = \frac{\#\{\hat{\theta}_b^* \le t\}}{B}$$

درحالتی که توزیع جامعه مجهول است ونمونه گیری ازنمونهٔ ابتدایی

انجام شده است در واقع از $\hat{F}$ نمونه گرفته ایم که $\hat{F}$ تابع توزیع تجربی نمونه است، از طرفی $\hat{F}$ برآوردگر درستنمایی ماکزیمم نا پارامتری F است و بنا به خاصیت پایایی برآوردگر درستنمایی ماکزیمم وقتی $B \to \infty$ آنگاه :

برآوردگرهای درستنمایی ماکزیمم $\longrightarrow$ برآوردگرهای بوت استرپ

امروزه چون محاسبات با کامپیوتر انجام می پذیرد در عمل می توان B را بسزرگ اختیار کرد تا نتیجه مطلوب حاصل گردد.

مثال ۱ـ فرض کنید واریانس میانه یک جامعه مجهول مورد نظر باشد، به روش بوت استرپ $\hat{\theta}_1^*, \hat{\theta}_2^*, \hat{\theta}_3^*, \dots, \hat{\theta}_B^*$ را که میانه‌های نمونه‌های بوت استرپ هستند بدست می آوریم و

$$var(\hat{\theta}^*) = \widehat{var}(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B}(\hat{\theta}_b^* - \hat{\theta}^{*\cdot})^2$$

را محاسبه می کنیم. $var(\hat{\theta}^*)$ برآورد بوت استرپ واریانس میانه نامیده می شود.

مثال ۲ـ می دانیم ضریب همبستگی نمونه برای ضریب همبستگی جامعه یک برآوردگر اریب است و گاهی لازم است مقدار این اریبی را بدانیم، با استفاده از روش B.S. می توان مقدار این اریبی را برآورد کرد. فرض کنید ρ ضریب همبستگی جامعه و r ضریب همبستگی نمونه اولیه با شد و نمونه (X_1, Y_1) و (X_2, Y_2) و ... و (X_n, Y_n) را در اختیار داشته باشیم، نمونه‌های بوت استرپ را بدست آورده، داریم :

$$r_1^* \longrightarrow (X_1^*, Y_1^*), (X_2^*, Y_2^*), \dots, (X_n^*, Y_n^*)$$

$$r_2^* \longrightarrow (X_1^*, Y_1^*), (X_2^*, Y_2^*), \dots, (X_n^*, Y_n^*)$$

$$\vdots$$

$$r_B^* \longrightarrow (X_1^*, Y_1^*), (X_2^*, Y_2^*), \dots, (X_n^*, Y_n^*)$$

که $r_1^*, r_2^*, \ldots, r_B^*$ ضریب همبستگی های نمونه‌های بوت استرپ هستند.

$$\widehat{Bias}(r) = Bias(r^*) = \frac{1}{B}\sum_{b=1}^{B} r_b^* - r$$

حال فرض کنید بخواهیم برای یک پارامتر بازهٔ اطمینان بیابیم. روشی که efron برای بازهٔ اطمینان پیشنهاد کرده است بدین ترتیب است که ابتدا توزیع بوت استرپ $\hat{\theta}$ را بدست می آوریم، سپس تعریف می کنیم:

$\hat{\theta}_{\alpha}$ = چندک متناظر با احتمال α در توزیع بوت استرپ $\hat{\theta}$

$\hat{\theta}_{1-\alpha}$ = " " " $1-\alpha$ " " "

efron بازهٔ (۱) $[\hat{\theta}_{\alpha}, \hat{\theta}_{1-\alpha}]$ را بازهٔ اطمینان بوت استرپ $1-2\alpha$ برای θ معرفی می‌کند. اما وی در مقالاتش در مورد توجیه این عمل ودلیل منطقی به کاربردن این بازهٔ شرح کافی نداده است. ذیلاً سعی خواهیم کرد استفاده از (۱) را برای بازهٔ اطمینان θ توجیه کنیم.

فرض کنید $X_1, X_2, \ldots, X_n$ نمونه‌ای از جامعهٔ موردنظر با تابع چگالی احتمال f_θ باشد. فرض کنید $\hat{\theta}$ برآوردگر درستنمایی ماکزیمم θ و $\sigma_{\hat{\theta}}$ تابع توزیع $\hat{\theta}$ باشد. هم چنین فرض کنید تابعی اکیداً "صعودی" g وثابت τ وجود دارد به نحوی که توزیع $Z = \tau[g(\hat{\theta}) - g(\theta)]$ یک توزیع متقارن (متقارن نسبت به مبدأ) با تابع توزیع N باشد. البته چون برآوردگرهای درستنمایی ماکزیمم به طور مجانبی نرمال هستند فرض فوق (لااقل به طور تقریبی) زیاد بی جا نیست. از آن گذشته خیلی از توزیع‌های نامتقارن با یک تبدیل مناسب به یک توزیع متقارن تبدیل می شوند. مثلاً r (ضریب همبستگی یک نمونه از توزیع نرمال دومتغیره) دارای توزیعی با چولگی زیاد است ولی توزیع

$\left[\frac{1}{2}\ln\frac{1+\rho}{1-\rho} - \frac{1}{2}\ln\frac{1+\beta}{1-\beta}\right]\sqrt{n-3}$ تقریبا " نرمال استاندارد است .

قضیه : تحت شرایط فوق یک حداطمینان بالای α-ابرای θ عبارتست از $\hat{\theta}$ یعنی چندک متناظربا احتمال $1-\alpha$ درتوزیع $G_{\theta}^{-1}(1-\alpha)$ یک حداطمینان بالای $1-\alpha$ برای θ است . برای اثبات قضیه فوق لم زیررا ارائه میکنیم .

لم : فرض کنید $T=\hat{\theta}$ برآوردگر θ و $G_{\theta}(t)$ تابع توزیع T باشد و $G_{\theta}(t)$ نسبت به θ اکیدا " نزولی باشدآنگاه اگرجوابمعادله $G_{\theta}(T)=\alpha$ (که درآن θ مجهول محسوب می شود) را با $\bar{\theta}$ نشان دهیم ، $\bar{\theta}$ یک حداطمینان بالادرسطح α-ابرای θ است .

اثبات - فرض کنید $K_t(\theta)$ همان $G_{\theta}(t)$ باشدکه درآن θ متغیرمحسوب شودیعنی $G_{\theta}(t)=K_t(\theta)$ و $K_t(\theta)$ اکیدا " نزولی است . (معادله $G_{\theta}(T)=\alpha$ حتما " جواب داردچون G نسبت به θ اکیدا " نزولی است پس یک بیک است وهمواره جواب دارد .)

حال معادله $G_{\theta}(t)=\alpha$ رادرنظرگرفته وازطرفین آن K_t^{-1} می گیریم :

$$K_t^{-1}\left(G_{\theta}(t)\right)=K_t^{-1}(\alpha)$$

$$K_t^{-1}\left(K_t(\theta)\right)=K_t^{-1}(\alpha) \quad \Longrightarrow \quad \theta=K_t^{-1}(\alpha)=\bar{\theta}=K_t^{-1}(\alpha)$$

ازطرفی توزیع متغیرتصادفی $G_{\theta}(T)$ درفاصله $(0,1)$ یکنواخت است .

(چون اگر X یک متغیرتصادفی باتابع توزیع F باشدآنگاه $F(X)$ دارای توزیع یکنواخت درفاصله $(0,1)$ است .)

لـــذا :

$$p(G_\theta(T) > \alpha) = 1-\alpha \Rightarrow p(K_T(\theta) > \alpha) = 1-\alpha$$

چون $K_T(\theta)$ اکیدا" نزولی است ، $K_T^{-1}(\alpha)$ نیز اکیدا" نزولـــــی است پس :

$$p(\theta < K_T^{-1}(\alpha)) = 1-\alpha \Rightarrow p(\theta < \bar{\theta}) = 1-\alpha$$

اثبات قضیـــه : تحت شرایط قضیه چون $Z = \tau[g(\hat{\theta}) - g(\theta)]$ داریم :

$$K_{\hat{\theta}}(\theta) = G_\theta(\hat{\theta}) = H\{\tau(g(\hat{\theta}) - g(\theta))\} \qquad (*)$$

وتعریف می کنیم : $z_\alpha = H^{-1}(\alpha)$

دراینجا $K_{\hat{\theta}}(\theta) = G_\theta(\hat{\theta})$ یعنی $K_{\hat{\theta}}(\theta)$ همان $G_\theta(\hat{\theta})$ است که متغیرآن θ درنظرگرفته می شود. حال اگـــــرازطرفیـــــن $G_\theta(\hat{\theta}) = \alpha$ ، H^{-1} بگیریم باتوجه به رابطه (*) داریم:

$$\tau[g(\hat{\theta}) - g(\theta)] = z_\alpha \longrightarrow \begin{cases} g(\hat{\theta}) = \frac{z_\alpha}{\tau} + g(\theta) & (*1) \\ g(\theta) = g(\hat{\theta}) - \frac{z_\alpha}{\tau} & (*2) \end{cases}$$

اگرازطرفیـــن معادله $G_\theta(\hat{\theta}) = \alpha$ ، G^{-1} بگیریم وباتوجه به‌رابطه (*1) داریم :

$$G_\theta^{-1}(\alpha) = \hat{\theta} = g^{-1}\{g(\theta) + \frac{z_\alpha}{\tau}\}$$

اگرازطرفین‌معادله $K_{\hat{\theta}}(\theta) = \alpha$ ، K^{-1} بگیریم وباتوجه‌به‌رابطه (*2)

۲* داریم :

$$K^{-1}_{\hat{\theta}}(\alpha) = \theta = g^{-1}\left\{ g(\hat{\theta}) - \frac{z_\alpha}{c} \right\}$$

ازطرفی چون $z_{1-\alpha} = -z_\alpha$ داریم :

$$K^{-1}_{\hat{\theta}}(\alpha) = \theta = g^{-1}\left\{ g(\hat{\theta}) + \frac{z_{1-\alpha}}{c} \right\}$$

لـــذا :
$$K^{-1}_{\hat{\theta}}(\alpha) = G^{-1}_{\hat{\theta}}(1-\alpha)$$

ازطرفی طبق تعریف $\bar{\theta}$ داریم : $\bar{\theta} = K^{-1}_{\hat{\theta}}(\alpha)$ لذا :

$$\bar{\theta} = G^{-1}_{\theta}(1-\alpha)$$

یعنی $\bar{\theta}$ چندک متناظربا احتمال $1-\alpha$ درتوزیع $\hat{\theta}$ است .

$$P(\theta < \bar{\theta}) = 1-\alpha$$

$$P(\theta > \underline{\theta}) = 1-\alpha$$

$$P(\underline{\theta} < \theta < \bar{\theta}) = 1 - P[(\theta < \underline{\theta}) \cup (\theta > \bar{\theta})] = 1-\alpha-\alpha = 1-2\alpha$$

یعنی $(\underline{\theta}, \bar{\theta})$ یک بازهٔ اطمینان $1-2\alpha$ برای θ است . حـــال اگربه جای $G_\theta(t)$ از $G_{\hat{\theta}}(t)$ استفاده کنیم ، $\bar{\theta}$ بـــر آورد می شودا ما توزیع $G_{\hat{\theta}}$ همان توزیع بوتاسترپ $\hat{\theta}$ است ، یعنی $\bar{\theta}$ به وسیلهٔ چندک متناظربا احتمال $1-\alpha$ درتوزیع بوتاسترپ $\hat{\theta}$ و $\underline{\theta}$ بـــه وسیلهٔ چندک متناظربا احتمال α درتوزیع بوتاسترپ $\hat{\theta}$ برآوردمی شود ، این برآوردگرها را با $\hat{\bar{\theta}}$ و $\hat{\underline{\theta}}$ نشان میدهیم . لذا بازهٔ $(\hat{\underline{\theta}}, \hat{\bar{\theta}})$ یک بازهٔ اطمینان $1-2\alpha$ تقریبی برای θ است .

دربعضی موارددفرض متقارن بودن توزیع Z نسبت به مبدأ " ممکن است صحت نداشته باشددراین گونه موارد efron بازهٔ اطمینان بوت استرپ بـــا اصلاح اریبی را پیشنهاد می کند . بازهٔ اطمینان B.C. یااصلاح اریبی به

شرح زیر است :

فرض کنید $\tau\left[g(\hat{\theta})-g(\theta)\right]+z_0$ دارای توزیع متقارن H باشد

با اثباتی مشابه قضیه قبل بدست می آوریم که :

چندک متناظر با احتمال $H(z_{1-\alpha}+2z_0)$ در توزیع بوت استرپ $\hat{\theta}^{*}-\hat{\theta}$

چندک متناظر با احتمال $H(z_{\alpha}+2z_0)$ در توزیع بوت استرپ $\hat{\theta}^{*}-\hat{\theta}$

در این حالت لازم است H و z_0 را بدانیم ، efron فرض کرده است :

توزیع نرمال استاندارد $= H$

$$z_0 = H(G_0(\hat{\theta})) = \Phi^{-1}(G_0(\hat{\theta}))$$

لذا بازه $\left[\hat{G}^{-1}(\Phi(2z_0-z_\alpha)) \; , \; \hat{G}^{-1}(\Phi(2z_0+z_\alpha))\right]$

بازه اطمینان تقریبی $1-2\alpha$ برای θ از روش بوت استرپ اصلاح شده است .

مـــــراجـــــــع

1 Efron , B. (1979) The Jackknife the bootstrap and other resampling plans .

2 Efron , B . (1979) Book strap methods : another Look at the Jackknife

Annals of statistics 1979 , Vol;7,No. 1,1- 26 .

3 Efron ,B. (1985) Bootstrap confidence intervals for a class of parametric problems.

Biometrica , 72, 1,pp.45- 58

4 THomas J.Diciccio Joseph Romano A review of Bootstrap confidence intervals.

J.R.Statist. SOC . B (1988) 50 , No. 3,pp. 338 - 350

UMVUE برای $P(X < Y)$

قاسم تارمست

دانشگاه اهواز

مفهوم عبارت $P(X<Y)$ در تحقیقات صنعتی کاربرد فراوان دارد. قیمت سوخت انرژی هسته‌ای بسیار خوب با قیمت سوخت مولدهای قدرت الکتریکی رقابت می‌کند و انرژی هسته‌ای یک انرژی قابل استفاده و مناسب و جایگزین مناسبی برای سوختهای فسیلی در تولید گرمای مورد نیاز در صنایع میباشد. افزایش استفاده این منبع تولید انرژی بایستی با بیشترین درجه ایمنی همراه باشد. مفهوم قابلیت اعتماد معیار مناسبی جهت تومیف ایمنی مراکز قدرت و انرژی هسته‌ای در جهان است. قابلیت اعتماد یک سیستم یعنی احتمال اینکه آن سیستم در یک زمان معین تحت شرایط داده شده بطور موفقیت آمیز در سرویس باشد و بکار خود ادامه دهد. همچنین قابلیت اعتماد یک سیستم احتمال اینستکه قدرتهای مولفه‌های یک سیستم بیشتر از فشارهای وارده بر مولفه‌های متناظر سیستم مذکور باشد. برای هر دو مفهوم قابلیت اعتماد یک سیستم، عبارت $P(X<Y)$ مورد استفاده و بوسیله آن سیستم مورد بررسی و ارزیابی قرار میگیرد.

اگر $X=[x,x,\ldots,x]$ و x یک زمان معین و $Y=\{Y_1,Y_2,\ldots,Y_n\}$ و Y_i بیانگر طول عمر زامین مولفه سیستم مورد بررسی باشد

عبارت $P[X<Y]$ احتمال اینستکه سیستم مذکور تحت فشارهای وارده با موفقیت بکار خود ادامه دهـــــد.

بعلت اهمیت عبارت $P(X<Y)$ لازمست آنرا برآورد کنیم. توزیعهای مهم آماری را برای X , Y در حالتهای مختلف در نظر میگیریم و در صورت امکان UMVUE , MLE برای عبارت $P(X<Y)$ را بدست میاوریم.

الف فرض کنید $x=t$ یک زمان معین و Y_i طول عمر یک مولفه دارای توزیع نمایی با پارامتر A باشد. قابلیت اعتماد این سیستم عبارتست از:

$$R(t)=P(Y>t)=\exp(-At) \quad , \quad A>0$$

UMVUE برای $R(t)$ عبارتست از:

$$\hat{R}(t)=\left[\frac{\sum_{i=1}^{n}y_i-t}{\sum_{i=1}^{n}y_i}\right]^{n-1} \qquad \sum_{i=1}^{n}y_i>t$$

MLE برای $R(t)$ عبارتست از:

$$\hat{R}_m(t)=\exp(-t/\bar{y}) \quad , \quad \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

ب فرض کنید $x=t$ یک زمان معین و Y طول عمر یک مولفه دارای توزیع وایبل با مقدار معلوم پارامتر α باشد. قابلیت اعتماد این سیستم عبارتست از:

$$R(t)=P(Y>t)=\exp[-(bt)^{a}], \quad t, a, b >0$$

UMVUE برای R(t) عبارتست از :

$$\hat{R}(t)=\left[\frac{\sum_{i=1}^{n} y_i^{a}-t}{\sum_{i=1}^{n} y_i^{a}}\right]^{n-1} \quad , \quad \sum_{i=1}^{n} y_i^{a}>t$$

MLE برای R(t) عبارتست از :

$$R_m(t)=\exp\left(-\frac{n}{\sum_{i=1}^{n} y_i^{a}}\, t^{a}\right)$$

۳ـ حالت (۲) را در نظر گیرید و فرض کنید Y دارای توزیع گاما به شرح زیر باشد :

$$f(y)=\frac{a^{m}}{\Gamma(m)} y^{m-1} e^{-ay} I_{(0,\infty)}(y) \qquad a\ ,\ m>0$$

که در آن m عددیست مثبت و صحیح. قابلیت اعتماد سیستم مورد بررسی عبارتست از :

$$R(t)=P(Y>t)=\sum_{j=0}^{m-1} \frac{e^{-at}(at)^{j}}{j!}$$

اگر m معلوم باشد، MLE برای R(t) عبارتست از :

$$R_m(t)=\sum_{j=1}^{m-1} \frac{e^{-\frac{nm}{\bar{y}}t}\cdot\left(\frac{m}{\bar{y}}t\right)^{j}}{j!}$$

۴ ـ فرض کنید x=t یک زمان معین و Y طول عمر یک مولفه دارای توزیع $N(\mu,\sigma^2)$ باشد. قابلیت اعتماد این سیستم عبارتست از :

$$R(t)=F\left(\frac{\mu-t}{\sigma}\right) \quad , \quad F(a)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{u^2}{2}}\, du$$

UMVUE برای R(t) عبارتست از :

$$\hat{R}(t)= F\left(\frac{\bar{y}-t}{\sqrt{\frac{n-1}{n}}\cdot\sigma}\right)$$

MLE برای R(t) عبارتست از :

$$\hat{R}m(t)= F\left(\frac{\bar{y}-t}{\sigma}\right)$$ σ معلوم است.

$$\hat{R}m(t)= F\left(\frac{\bar{y}-t}{s}\right)$$ σ نامعلوم است.

۵ ـ فرض کنید یک سیستم با P مولفه مستقل و $X=[t,t,\ldots,t]$ و t یک زمان معین و Y_i طول عمر iامین مولفه با توزیع نمائی باشد. قابلیت اعتماد این سیستم عبارتست از :

$$R(t)=P(Y>X)=\prod_{i=1}^{P} P(Yi>t)=\exp(-t\sum_{i=1}^{P} Ai)$$

UMVUE برای R(t) عبارتست از :

$$\hat{R}u(t)=\prod_{\alpha=1}^{P}\left[\frac{\sum_{i=1}^{n} y_{i\alpha}-t}{\sum_{i=1}^{n} y_{i\alpha}}\right]^{n-t}$$

MLE برای R(t) عبارتست از :

$$\hat{R}m(t)=\prod_{\alpha=1}^{P}\exp\left[\frac{-nt}{\sum_{i=1}^{n} y_{i\alpha}}\right]$$

۶ـ فرض کنید X , Y متغیرهای تصادفی به ترتیب دارای توزیع نمائی با پارامترهای A و B باشند. قابلیت اعتماد یک مولفه با قدرت Y و تحت فشار X عبارت است از :

$$R=P(Y>X)=A/(A+B)$$

اگر $X_1, X_2, \ldots, X_{n_1}$ و $Y_1, Y_2, \ldots, Y_{n_2}$ به ترتیب دو نمونه تصادفی مستقل به ترتیب روی X , Y باشند،

$$T_x = \sum_{i=1}^{n_1} X_i \qquad T_y = \sum_{i=1}^{n_2} Y_i$$

UMVUE برای R عبارتست از :

$$\hat{R} = \sum_{j=0}^{n_1-1} (-1)^j \frac{(n_1-1)!\,(n_2-1)!}{(n_1-1+j)!\,(n_2-1-j)!} \left(\frac{T_x}{T_y}\right)^j \quad , T_x \le T_y$$

$$\hat{R} = 1 - \sum_{j=0}^{n_1-1} (-1)^j \frac{(n_1-1)!\,(n_2-1)!}{(n_2-1+j)!\,(n_1-1-j)!} \left(\frac{T_y}{T_x}\right)^j \quad , T_x > T_y$$

MLE برای R عبارتست از :

$$\hat{R}_m = \frac{\bar{Y}}{\bar{Y}+\bar{X}}$$

A و B نامعلوم هستند.

۷- فرض کنید X,Y دارای توزیع نمائی مشترک زیر باشند.

$$\bar{F}(x,y) = P(X>x, Y>y) = \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12} \mathrm{Max}(x,y)]$$

$$R = \Gamma(X<Y) = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_{12}}$$

اگر $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_{12}$ به ترتیب MLE برای $\lambda_1, \lambda_2, \lambda_{12}$ باشند، MLE برای R عبارتست از :

$$\hat{R}_m = \frac{\hat{\lambda}_1}{\hat{\lambda}_1+\hat{\lambda}_2+\hat{\lambda}_{12}}$$

اگر $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$ یک نمونه تصادفی زوجی روی X و Y باشد بطوری که برای n_1 تا از این زوجها $X_i>Y_i$ و برای n_2 تا از این زوجها $X_i<Y_i$ و $n=n_1+n_2$ باشد،

ـ تمایل به رشد و ارتقا

ـ تمایل به کار دسته جمعی

ـ تمایل به جدی بودن

ـ میل به قبول خطر

ـ وقت شناسی

۲ ـ دانش فنی و اطلاعات ، برای این عامل صفاتی در نظر گرفته شده است . وجود یا عـــدم این صفات در هر واحد مورد ارزیابی قرار گرفته است شرح صفات این عامل به‌قرارزیراست :

ـ قابل دسترسی بودن

ـ وسعت ارتباط

۲ ـ سازماندهی و مدیریت ، برای این عامل هفت پارامتر به شرح ذیل منظور شده است :

ـ سود دهی

ـ درصد بهره برداری از ظرفیت

ـ آینده نگری

ـ استقلال

ـ سازماندهی برای بهبود مهندسی

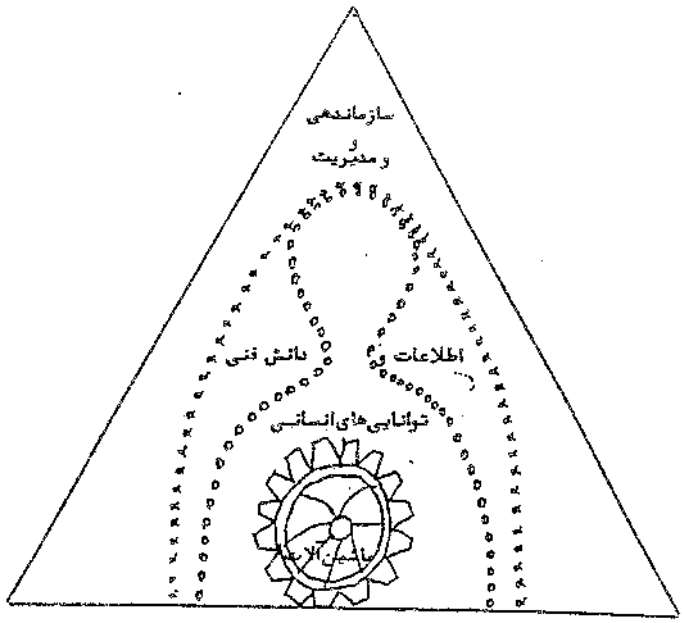
ـ نوسازی

ـ درصد هزینه تحقیقات به فروش

بیان روش کمی کردن هریك از عوامل به جهت طولانی بودن از حوصله این مقاله خارج است وازاین رو نقطه به ادامه در نمونه از نتایج کار انجام شده در صنایع تولید الیاف مصنوعی و تولید آهـن و فولاد بسنده می‌شود . نمودار ۱ سطح تکنولوژی به کار رفته در صنایع تولید آهن و فولاد را در کشور

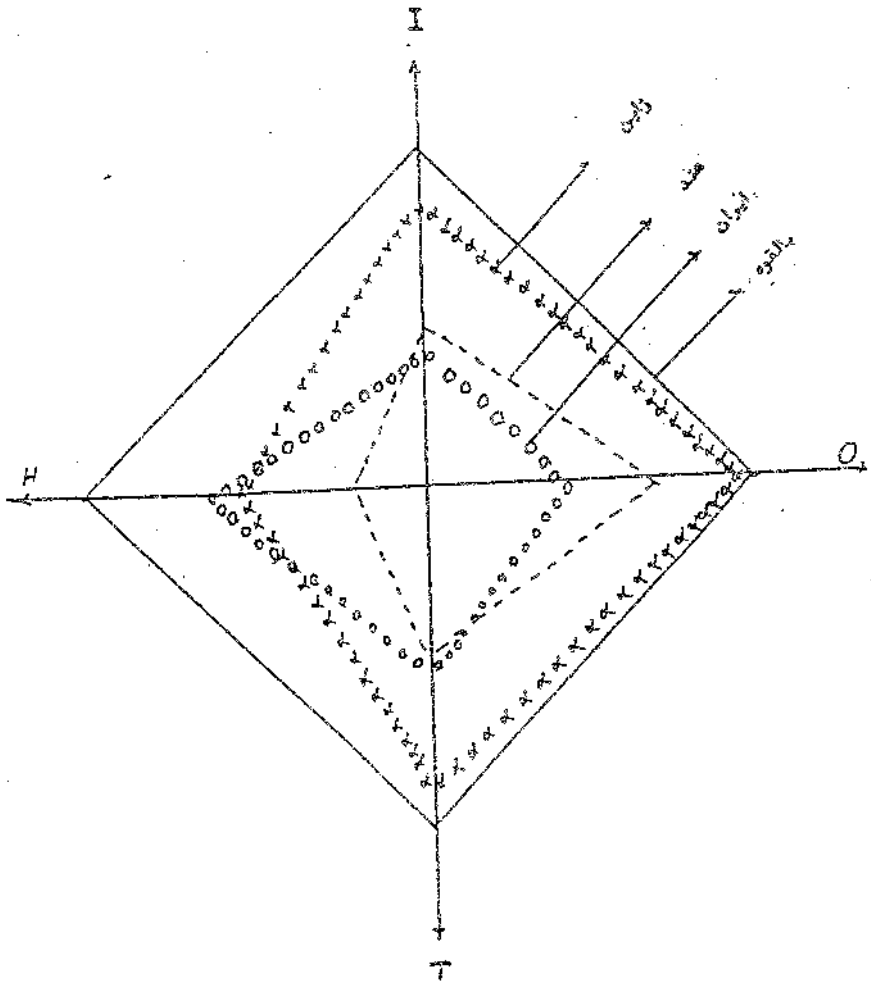
ایران ، هند و ژاپن از یک سو و فاصله آن را تا بالاترین سطح تکنولوژی موجود نشان می دهد . حال آنکه نمودار اندازه‌گیری سطح تکنولوژی مورد استفاده در تولید الیاف مصنوعی را در کشورهای ایران و امریکا مورد مقایسه قرار داده و نشان می‌دهد که تا چه اندازه این تکنولوژی‌ها از سطح بالقوه آن پائین تراست .

نمودار ۱ ـ عوامل تشکیل دهنده محتوی تکنولوژی

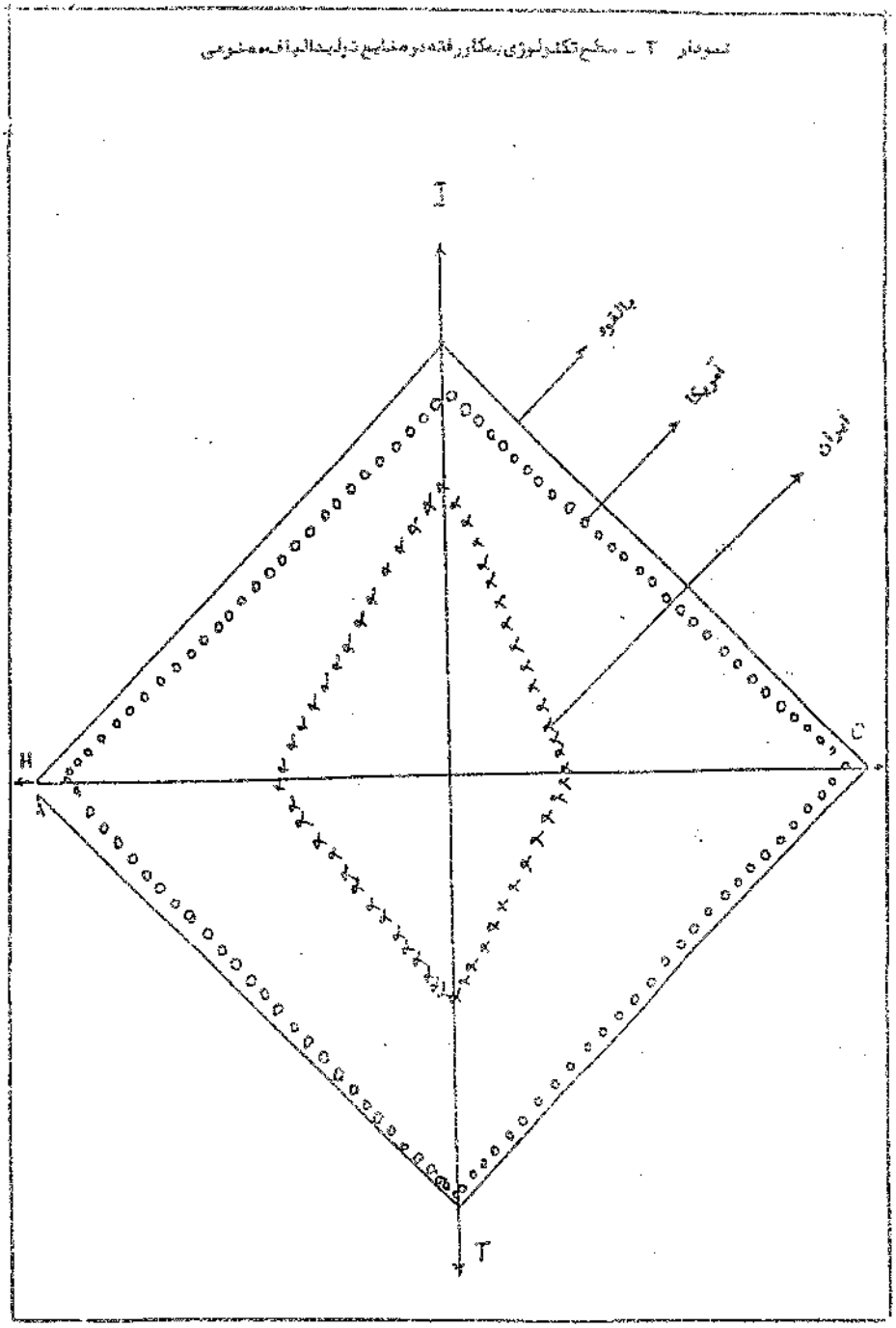


محتوی سازمانی تکنولوژی	ORGAWARE
محتوی اطلاعات و دانش فنی تکنولوژی	INFORWARE
محتوی توانایی‌های انسانی تکنولوژی	HUMANWARE
محتوی ماشین آلات تکنولوژی	TECHNOWARE

نمودار ۲ ـ سطح تکنولوژی به‌کار رفته در صنایع تولید آهن و فـــــولاد



نمودار ۲ ـ سطح تکنولوژی بهکار رفته در صنایع تولید الیاف مصنوعی



زنجیرهای مارکف مرتبهٔ p

مجتبی اسحقی

دانشگاه آزاد اسلامی

خلاصه :

ابتدا زنجیرهای مارکف مرتبه p وفرم ماتریس انتقال مربوط به آن دردوشکل معمولی وادغام شده معرفی می شود. سپس قضیه‌ای- ثابت میشودکه طبق آن هرزنجیرمارکف مرتبه p به‌یک زنجیرمارکف مرتبه اول قابل تحویل است. بعدازآن شرطی روی زنجیرمرتبه p گذاشته میشودکه درصورت برقراری آن زنجیرمرتبه اول بدست‌آمده تحویل ناپذیر خواهدشد. قضیه حدی درمدل اتورگرسیوزنجیرمارکف اثبات میشودونهایتا یک محک تخمین مرتبه زنجیرمارکف وگسترش بیزی آن معرفی میشود.

۱-۱ - تعریف :

فرآیندتصادفی $\{X_t : t\in T\}$ که مقادیرش را درمجموعه شمارای S اختیارمیکندزنجیرمارکف ازمرتبه p گوئیم هرگاه :

$$(1.1)\quad P(X_n=i_0 \mid X_{n-1}=i_1,\dots,X_{n-p}=i_p,X_{n-p-1}=i_{p+1},\dots)=P(X_n=i_0 \mid X_{n-1}=i_1,\dots,X_{n-p}=i_p)=P_{i_p,\dots,i_1}$$

$$\hat{\lambda}_1 = \frac{n_2}{\sum_{i=1}^{n_2} x_{i2}} = \frac{1}{\bar{x}_2} \qquad \hat{\lambda}_2 = \frac{n_1}{\sum_{i=1}^{n_1} y_{i1}} = \frac{1}{\bar{y}_1}$$

$$\hat{\lambda}_{12} = \frac{1}{\bar{y}_2} - \frac{1}{\bar{y}_1} \qquad \text{یا} \qquad \hat{\lambda}_{12} = \frac{1}{\bar{x}_1} - \frac{1}{\bar{x}_2}$$

مثال : فرض کنید یک نمونه تصادفی با اندازه $n=10$ از یک جامعه نمائی دو متغیره وابسته به شرح زیر بدست آمده است :

Xi : ۲/۱ ۰/۸ ۱/۳ ۰/۹ ۰/۷ ۰/۴ ۱/۲ ۲/۵ ۰/۵ ۱

Yi : ۱/۲ ۲ ۱/۸ ۱/۱ ۰/۸ ۲/۱ ۲ ۰/۸ ۲ ۲

$$\sum_{i=1}^{n_2} x_{i2} = ۲ + ۰/۸ + ۲/۱ + ۰/۸ + ۱/۱ + ۱/۸ + ۲ + ۱/۲ = ۱۱/۸$$

$$\sum_{i=1}^{n_2} x_{i2} = ۶/۸ \qquad \sum_{i=1}^{n_1} x_{i1} = ۴/۶ \qquad \sum_{i=1}^{n_1} y_{i1} = ۴$$

$$\hat{\lambda}_1 = ۱/۱۸$$

$$\hat{\lambda}_2 = ۰/۵$$

$$\hat{\lambda}_{12} = \frac{1}{\bar{y}_2} - \frac{1}{\bar{y}_1} = ۰/۶۸ - ۰/۵ = ۰/۱۸$$

$$Rm = \frac{۱/۱۸}{۱/۱۸ + ۱/۵ + ۱/۱۸} = ۰/۶۳$$

۸ — X و Y دارای توزیع دو متغیره نرمال هستند و به ترتیب بیانگر فشار وارده و قدرت یک مولفه می‌باشند . قابلیت اعتماد این سیستم عبارتست از :

$$R = P(X<Y) = F\left(\frac{\mu_2 - \mu_1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \right)$$

اگر کلیه پارامترهای توزیع مورد نظر نامعلوم باشند ، MLE برای R عبارتست از :

$$\hat{Rm} = F\left(\frac{\bar{y} - \bar{x}}{\sqrt{S_1^2 + S_2^2 - 2\rho S_1 S_2}} \right)$$

۹- فرض کنید یک سیستم با p مولفه دارای قدرتهای $Y_1, Y_2, \ldots, Y_p$ تحت تاثیر فشار X قرار میگیرد بطوریکه :

$$[1 - F(x;\theta_1)]^{\theta_2} = [1 - G(x;\theta_2)]^{\theta_1} \qquad (1)$$

تا وقتیکه k تا از Yiها بیشتر از X هستند این سیستم با موفقیت بکار خود ادامه میدهد. اگر آماره‌های مرتب Y را بصورت $Y_{(1)}, Y_{(2)}, \ldots, Y_{(p)}$ در نظر بگیریم قابلیت اعتماد سیستم مورد نظر عبارتست از :

$$R_{X,p} = P[Y_{(p-k+1)} > X] = 1 - \frac{\Gamma(k+\lambda)\cdot\Gamma(p+1)}{\Gamma(p+1+\lambda)\cdot\Gamma(k)} \quad , \quad \lambda = \frac{\theta_2}{\theta_1} \qquad (2)$$

مثال: فرض کنید شدت جریان الکتریکی X که از یک سیستم با ۳ مولفه عبور میکند دارای توزیع نمائی $f(x) = \frac{1}{5} e^{-\frac{x}{5}}$ باشد، Y_1, Y_2, Y_3 دارای توزیع نمائی $f(y) = \frac{1}{8} e^{-\frac{y}{8}}$ میباشند که در آن Y_1, Y_2, Y_3 بیانگر قابلیت تحمل شدتهای جریان الکتریکی ۳ مولفه سیستم مورد بررسی میباشند و اگر حداقل یک مولفه از این ۳ مولفه دارای قابلیت تحمل شدتهای جریان بیشتری از X باشد، سیستم بخوبی کار می کند.

بنابراین احتمال مناسب کار کردن این سیستم عبارتست از :

$$R_{1,3} = P[Y_{(3)} > X] = 1 - \frac{\Gamma(1+\frac{5}{8})\Gamma(4)}{\Gamma(4+\frac{5}{8})\Gamma(1)} \simeq \%86$$

۱۰- فرض کنید یک سیستم دارای یک مولفه با قدرت Y باشد و تحت تاثیر فشارهای $X_1, X_2, \ldots, X_p$ قرار میگیرند. Xiها و Y از یکدیگر مستقل هستند و Xiها دارای توزیع یکسان $F(x;\theta_1)$ و Y دارای توزیع تجمعی $F(y;\theta_2)$ میباشند بطوریکه $F(y;\theta_2)$ و $F(x;\theta_1)$ در رابطه (۱) صدق میکنند.

اگر حداقل kتا از Xiها کمتر از Y باشند، سیستم با موفقیت کار میکند. بنابراین احتمال مناسب کارکردن این سیستم عبارتست از :

$$R_{K,P} = \frac{\Gamma(P+1)\cdot\Gamma(P+\frac{1}{\lambda}+1-K)}{\Gamma(P+1-K)\cdot\Gamma(P+\frac{1}{\lambda}+1)} \qquad (۳)$$

در روابط (۲) و (۳) تنها λ ممکنست نامعلوم باشد. اگر برآورد حداکثر درستنمائی λ را در روابط فوق جایگزین کنیم، $\hat{Rm}$ بدست میاید.

مثال : فرض کنید $Y_1, Y_2, \ldots, Y_p$ و X مستقل و دارای توزیع نمائی و $p=۴$ و $k=2$ و $\sum_{۱}^{۱۰} Xi = ۲۰$ و $\sum_{۱}^{۱۰}\sum_{\alpha}^{۴} Yi\alpha = ۴۰$ باشد، برآورد حداکثر درستنمائی λ عبارتست از :

$$\hat{\lambda} = \frac{\hat{\theta}_۲}{\hat{\theta}_۱} = \frac{۱۰\sum_{۱}^{۱۰}\sum_{\alpha}^{۴} Y_{i\alpha}}{۴۰\sum_{۱}^{۱۰} X_{i}} = \frac{۱۰\times۴۰}{۴۰\times۲۰} = ./۵$$

MLE برای رابطه (۲) عبارتست از :

$$\hat{Rm} = ۱ - \frac{\Gamma(۲+./۵)\,\Gamma(۵)}{\Gamma(۵+./۵)\,\Gamma(۲)} = ./۳۹$$

مثال : فرض کنید $X_1, X_2, \ldots, X_p$ و Y مستقل و دارای توزیع نمائی و $p=۴$ و $k=2$ و $\sum_{۱}^{۱۰}\sum_{\alpha}^{۴} Xi\alpha = ۴۰$ و $\sum_{۱}^{۱۰} Yi = ۲۰$ باشد، برآورد حداکثر درستنمائی λ عبارتست از :

$$\hat{\lambda} = \frac{\hat{\theta}_۲}{\hat{\theta}_۱} = \frac{۴۰\sum_{۱}^{۱۰} Y_{i}}{۱۰\sum_{۱}^{۱۰}\sum_{\alpha} X_{i\alpha}} = \frac{۴۰\times۲۰}{۲۰} = ۲$$

MLE برای رابطه (۳) عبارتست از :

$$\hat{Rm} = \frac{\Gamma(۵)\,\Gamma(۵+\frac{۱}{۲}-۲)}{\Gamma(۵+\frac{۱}{۲})\,\Gamma(۵-۲)} = \frac{\Gamma(۵)\Gamma(۳/۵)}{\Gamma(۵/۵)\Gamma(۳)} = ./۲۵$$

مثال : فرض کنید $X_1, X_2, \ldots, X_p, X$ و Y مستقل و دارای توزیع وایبل به شرح زیر باشند .

$$F(x;\theta_1) = 1 - e^{-\frac{x^{\alpha}}{\theta_1}}$$

$$F(y;\theta_2) = 1 - e^{-\frac{y^{\alpha}}{\theta_2}} \qquad (۴)$$

این دو توزیع در رابطه (۱) صدق می‌کنند . قابلیت اعتماد این سیستم عبارتست از :

$$R_{k,p} = \frac{\Gamma(p+1)\,\Gamma(p+\frac{1}{\lambda}+1-k)}{\Gamma(p+1-k)\,\Gamma(p+\frac{1}{\lambda}+1)}$$

فرض کنید $p=۴$ و $k=۳$ و $\sum_{i=1}^{10}\sum_{j=1}^{4} x_{ij}^{\alpha}=۴۰$ و $\sum_{i=1}^{10} y_i^{\alpha}=۲۰$ و $\alpha=۲$ و $n=۱۰$ باشد .

$$\hat{\lambda} = \frac{\hat{\theta}_2}{\hat{\theta}_1} = \frac{۴۰\sum_{i=1}^{10} y_i^{۲}}{۱۰\sum_{i=1}^{10}\sum_{j=1}^{4} x_{ij}^{۲}} = \frac{۴\times۲۰}{۴۰} = ۲$$

MLE برای R عبارتست از :

$$\hat{R}_m = \frac{\Gamma(۵)\,\Gamma(۵+۱/۲-۳)}{\Gamma(۵-۳)\,\Gamma(۵+۱/۲)} = \frac{\Gamma(۵)\,\Gamma(۳.۵)}{\Gamma(۳)\,\Gamma(۵.۵)} = R_{k,p} = ./۲۵$$

۱۱ـ فرض کنید X یک بردار تصادفی p مولفه‌ای و Y نیز یک بردار p مولفه‌ای و به ترتیب بیانگر فشارهای وارده بر و قدرتهای مولفه‌های یک سیستم سری pمولفه‌ای باشند .در اینحالت محاسبه $P(X<Y)$ بسیار مشکل و دراکثر حالات غیرممکن میباشد و حدود بالا و پائین برای عبارت $P(X<Y)$ محاسبه میشوند .مرجع (۸) حدود بالا و پائین برای عبارت $p(x<y)$ را مورد بحث قرار داده است.

۱۲ـ بارلو و پیروشن در مرجع (۴) یک سیستم سری معادل برای هر سیستم چند مولفه‌ای را بیان کرده اند . در نتیجه می‌توان عبارت $P(X<Y)$ را برای ارزیابی احتمال مناسب کارکردن هر سیستم pمولفه‌ای مورد استفاده قرار داد .

مراجع

1. BASU, A.P.(1981), THE ESTIMATION OF P(X<Y) FOR DISTRIBUTIONS USEFUL IN LIFE TESTING. NAVAL RESEARCH LOGISTICS QUARTERLY. 28,383-392.
2. BHATTACHARYYA, G.L. AND JOHNSON, R.(1975). STRESS-STRENGTH MODEL FOR RELIALILITY.
3. BHATTACHARYYA, G.L. AND JOHNSON, R.(1974). ESTIMATION OF RELIALILITY IN A MULTICOMPONENT STRESS -STRNGTH MODEL, J.AMER.STATIST. ASSOC.,69, 966-97.
4. TONG, H.(1974). ANOTE ON THE ESTIMATION OF P(Y<X) IN EXPONENTIAL CASE. TECHNOMETRICS 16 ,625.
5. TONG, Y.L. (1980). PROBABILITY INEQUALITIES IN MULTIVARIATE DISTRIBUTIONS.ACADEMIC PRESS, NEW YORK.
6. R.E.BARLOW AND F.PROSCHAN (1981). STATISTICAL THEORY OF RELIABILITY AND LIFE TESTING PROBABILITY MODELS. MCARDLE PRESS, INC.
7. RELIABILITY OF NUCLEAR POWER PLANTS (1975), PROCEEDINGS OF A SYMPOSIUM. INTERNATIONAL ATOMIC ENERGY AGENCY.
8. G. TARMAST (1983). RELIABILITY OF COMPLEX SYSTEN, PHD DISSERTATION. UNIVERSITY OF MISSOURI-COLUMBIA.

بررسی آماری وضع هپاتیت B

اکبر حسن‌زاده*- کاظم محمد

دانشگاه علوم پزشکی تهران

۵شنبه ۱۶:۳۰ تالار ۵

روشن شده است که HBsAg (آنتی ژن سطحی هپاتیت B) یکی از مهمترین اجزای ویروس هپاتیت B بوده و عفونت‌زایی ویروس را به آن نسبت می‌دهند. در این تحقیق به بررسی وضعیت شیوع HBsAg در استان آذربایجان شرقی در رابطه با متغیرهای مختلف پرداخته شده است.

نمونه‌گیری به‌روش خوشه‌ای انجام شده و تعداد افرادی که خون آنها از نظر HBsAg مورد بررسی قرار گرفته است ۳۴۰۹ نفر می‌باشند.

جهت تعیین شیوع HBsAg در رابطه با عوامل مختلف ابتدا به کمک برنامهٔ SPSS به انجام آزمونهای آماری مقدماتی همچون χ^2 پرداخته شده است. سپس براساس از ۲۲ متغیر مستقل مورد بررسی متغیرهای «محل سکونت»، «سن»، «جنس»، «وضعیت اقتصادی»، «مورد اتکل در مقدوم»، «نوع شغل آب آشامیدنی و آدنوپاتی زیر بغل» که به استثنای سن در سطح اشتباه ۵% = α با متغیر وابسته HBsAg همبستگی نشان داده‌اند برای تجزیه و تحلیل دقیق و جامعتر انتخاب شدند و ارتباط این متغیرها با شیوع HBsAg با استفاده از «مدل لجستیک حتی‌الامکان یکجا مورد بررسی قرار گرفته که در این راستا از نرم‌افزار GLIM استفاده شده است.

روشهای عملی اجرای طرحهای آمار کشاورزی

محمدباقر حسین‌پورمقدم

ادارهٔ کل آمار و اطّلاعات کشاورزی

یکی ازعوامل توسعه اقتصادی بهبود وگسترش کشاورزی است لازمه این بهبود شناخت عوامل موثرآن میباشدکه جهت این شناخت نیازمبرم به آگاهی ودردست داشتن آماروا طلاعات دقیق به عنوان ابزاری لازم درشناخت وحل مسائل نه فقطکشاورزی بلکه درزمینه‌های برنامه‌ریزی تصمیم‌گیری ،کنترل ونمایش واقعیتهای موجود میباشیم .

آنچه مسلم است نیل به این اهداف مستلزم تهیه واجرای طرحهای مختلف مبتنی برعلم آمارمیباشد بدین لحاظ مسئولین بخش کشاورزی درکشورخواستارایجادتشکیلات مناسب آماری با توجه به بافت جامعه کشاورزی می‌باشند تاازاین طریق بتوانندواقعیات راچون آینه‌ای پیش‌رو داشته باشند . دراین راستامتولیان آمارکشاورزی باهمه کمبودهاوتنگناهااقدام به تهیه طرحهای مختلف آماری نموده‌اندولی بایستی اذعان داشت که اقدامات انجام یافته به لحاظ کیفیت وکمیت نیازهای این بخش راکه متجاوزازدوهزارقلم آمارمیباشدتامین‌نکرده‌ونمیکندواین‌بیانگراین‌حقیقت است که باید سرلوحه تامین آنهااقدامات اساسی وزیربنائی انجام گیرد .

باتوجه به مقدمه بالاضمن برشمردن مختصری ازنقاط قوت وضعف طرحهای جمع آوری آمارهای کشاورزی به بحث اصلی می‌پردازیم .

بطورکلی سه روش برای جمع آوری واستخراج آماردربخش کشاورزی مورداستفاده قرار میگیرد .

۱ـ آمارهای ثبتی .

۲ـ آمارهای تخمینی .

۳ـ آمارهای استنباطی .

۱ـ آمارهای ثبتی :

به آمارهایی اطلاق میشود که ازدفاترومنابع ثبت شده استخراج میگردد این گونه آمارها بعلت ماهیت‌شان ازدقت بالایی برخوردارند وبسهولت قابل دسترسی می‌باشند .

آماری ازادوات کشاورزی توزیع شده دروزارت کشاورزی برای یك دوره معینی درنظراست بامراجعه به بنگاه توسعه ماشین آلات وادوات کشاورزی وثبت داده‌های موجوداین نیازبرطرف خواهدشد ویاآماربارندگی ایستگاههای هواشناسی .

این نکته ضروری است که توجه شود درمواردی بااینکه آمارثبتی موجود می‌باشد ولی به لحاظ فرصت مورد نظرقابل استفاده نخواهد بود . فرض کنیم میزان تولید محصول چغندرقند بسر حسب استان ویاشهرستان مدنظرباشد دروهله اول بنظرمیرسد که این آمارازدفاترکارخانجات قند استان قابل وصول است ولی باید توجه داشت که بعلت حمل ونقل‌های بین استانی چنین امری میسرنیست .

۲ـ آمارهای تخمینی :

گرچه دامنه عمل این نوع آمارگسترده‌ترازآمارهای ثبتی می‌باشد ولی بعلت ماهیت آن که برخواسته ازحدس وکارشناسی می‌باشد نه علم و واقعیت ازدقت بالایی برخوردارنبوده ونمیتوان با قطعیت درموردآن قضاوت نمود، تصمیم‌گیری کرد ولی درمواردی به علت نیازهای فوری آماری ویا کمبودهای پرسنلی وامکاناتی مجبوربه استفاده ازاین روش می‌باشیم .

فرض کنیم آمارمحصولات کشاورزی محدوده شهری شهراصفهان مد نظرباشد، باتوجه به اینکه وسعت کشت درمحدوده شهری کم می‌باشد وگذشته ازآن چنانچه ازآماراستنباطی استفاده شود داشتن کلیه بهره برداران محدوده شهری اصفهان بعنوان چارچوب برای نمونه گیری یا سرشماری ضروری می‌باشدواین امرمستلزم فهرست برداری ازکلیه مناطق شهری می‌باشد که این خودایجاب می‌کند هزینه، پرسنل، امکانات ووقت زیادی صرف گردد لذا دراینگونه موارد باامکانات موجودآماری دربخش کشاورزی که شرح آن خواهدآمدناچاربه استفاده ازآمارهای تخمینی که بوسیله کارشناسان خبره ازطریق بازدیدهای محلی وگفتگوبامطلعین محلی اخذ میگردد می‌باشیم .

۳ـ آمارهای استنباطی :

آمارهای استنباطی یکی ازعمده‌ترین آمارهای مورداستفاده دربخش کشاورزی است این آمارهامبتنی به روشهای علمی بوده ، بنابراین بایدازاصول وقواعدی پیروی نماید که خودمستلزم تهیه وتدوین طرح فنی است . برای این منظورتعریف مفاهیم زیرضروری است .

- اهداف طرح : اولین مسئله‌ای که به نظر میرسد تعیین اهداف کلی وجزئی طرح میباشد درطرحهای آمارگیری اهداف کلی میتوانند برآورد سطح زیرکشت، میزان تولید، هزینه تولید درواحد سطح وواحد وزن واهداف جزئی محاسبه متوسط عملکرد درواحدسطح بتفکیک وسعت اراضی، تعیین میزان خسارات، میزان نهاده‌های مصرفی و ..... باشند.

- پیوشش طرح : طرح آمارگیری محصولات سالانه کشاورزی بتفکیک گندم وجو سایر محصولات دوکلیه استانهای کشوردوقالب دوطرح جداگانه اجرامیشود.

- جامعه آماری : عبارتست ازکلیه مناطق روستائی وشهری کشور.

چارچوب آماری : دراین آمارگیری ازفریم سرشماری کشاورزی سال ۱۳۶۷ مرکزآمار ایران استفاده شده که اطلاعات مربوط به سطح زیرکشت محصولات سالانه بهره‌برداران درهرآبادی به تفکیک گندم وجووسایرمحصولات مد نظرقرارگرفته شده‌است.

- روش آمارگیری : آمارگیری بصورت نمونه‌گیری به روش طبقه‌بندی دومرحله‌ای میباشد که آبادی بعنوان واحدآماری مرحله اول وبهره‌برداربعنوان واحدآماری درمرحله دوم انتخاب میشوند.

- سطح برآورد : پارامترهای مورد نظربراهداف طرح درسطح شهرستان برآوردمیگردد.

- نحوه طبقه‌بندی : طبقه بندی آبادیهایی که براساس چارچوب مورداستفاده دارای‌کشت محصولات سالانه مورد نظرطرح میباشند درهرمیان ازشهرستانهابراساس سطح زیرکشت آنهـــا بتفکیک گندم وجو سایرمحصولات وبااستفاده ازروش دالینوس انجام میگیرد.

نحوه طبقه بندی به روش دالینوس :

برای اینکارابتداء کلیه آبادیهای هرشهرستان رابراساس سطح زیرکشت محصولات مورد نظرطرح بصورت صعودی مرتب میکنیم پس ازانجام اینکارطبقه‌بندی اولیه راباقرجه مساوی طبق جدول زیرانجام میدهیم .

حدود طبقات	فراوانی درهر طبقه	جذرفراوانی طبقات	فراوانی مطلق جذرفراوانیهای طبقات
$a + h$	f_1	$\sqrt{f_1}$	$\sqrt{f_1} = F_1$
$a + 2h$	f_2	$\sqrt{f_2}$	$\sqrt{f_1} + \sqrt{f_2} = F_2$
⋮	⋮	⋮	⋮
$a + ih$	fi	$\sqrt{fi}$	$\sqrt{f_1} + \sqrt{f_2} + \ldots + \sqrt{fi} = Fi$
⋮	⋮	⋮	⋮
$a + kh$	f_k	$\sqrt{f_k}$	$\sum_{i=1}^{k} \sqrt{fi} = F$

که درآن

a : سطح زیرکشت محصولات مورد نظرطرح درآبادی است که کمترین سطح زیرکشت‌را درشهرستان دارامی‌باشد .

h : فاصله طبقات اولیه

k : تعداد طبقات اولیه

fi : فراوانی طبقه i ام (تعداد آبادی که سطح زیرکشت آنها از $(i-1)h+a$ بیشتر واز $a+ih$ کمترمی‌باشد)

Fi : مجموع جذرفراوانیهای طبقه i ام وطبقات ماقبل آن

F : فراوانی مطلق جذرفراوانیهای کلیه طبقات

پس ازطبقه‌بندی اولیه، طبقه بندی ثانویه (اصلی) بطریق زیرانجام میگردد.

ابتداء مقدار D رااز فرمول زیربدست می‌آوریم :

$$D = F/L$$

که درآن L حداکثرتعداد طبقات مورد نظردرشهرستان می‌باشد.

برای بدست آوردن مرزطبقات ثانویه بادرنظرگرفتن مقدار D وستون آخرجدول (Fiها) بـــه طریق زیرعمل میکنیم :

۱ـمقدار D رابا F_1 مقایسه کرده چنانچه $D < F_1$ باشد یك واحداز L کم کـــرده‌و مجددا" مقدارجدید D راحاسبه می‌نمائیم واین عمل راآنقدرادامه داده تا $D \geq F_1$ گردد.

۲ـمقدار D رابا Fiها (ستون آخرجدول) مقایسه کرده ودوفراوانی مطلق F_i و F_{i+1} راکه مقدارفوق بین آنهاواقع شده مشخص می‌نمائیم قدرمطلق تفاوت D با Fi و (Fi+1) رامشخص کرده مرزطبقه اول درنظرمی‌گیریم، برای بدست آوردن مرزطبقه دوم مقدار D رابا Fi هامقایسه کـــرده ومانند حالت قبل مرزطبقه دوم رامشخص می‌نمائیم وبه همین ترتیب عمل راادامه داده تاحـــدود کلیه طبقات مشخص گردد.

ـتعداد آبادی نمونه درهرشهرستان : باتوجه به خطای برآورد قابل قبول که ۵ درصد کل سطح زیرکشت درهرشهرستان می‌باشد تعداد آبادیهای نمونه لازم ازفرمول زیرمحاسبه میشود.

$$n = \frac{\left(\sum_{h=1}^{K} N_h \cdot S_{fi}\right)^2}{\left(0.05 \sum_{h=1}^{K} \sum_{i=1}^{N_h} x_{hi}\right)^2}$$

مشروط براینکه تعدادآبادیهای نمونه درهرشهرستان از ۵ درصد کل آن تجاوزنکند چنانچه ایـــن تعداد بیش از ۵ درصدباشد ازفرمول زیراستفاده خواهد شد.

$$n = \frac{\left(\sum_{h=1}^{K} N_h \cdot S_{fi}\right)^2}{\left(0.05 \sum_{h=1}^{K} \sum_{i=1}^{N_h} x_{hi}\right)^2 + \sum_{h=1}^{K} N_h \cdot S_h^2}$$

پس ازمشخص شدن تعداد کل نمونه درهرشهرستان برای بدست آوردن تعداد نمونه درهرطبقه

ازفرمول زیراستفاده میشود.

$$n_\ell = \frac{N_\ell \cdot S_\ell}{\sum_{h=1}^{k} N_h \cdot S_h} \times n$$

درفرمولهای بالا داریم:

n : تعداد کل آبادیهای نمونه درهرشهرستان.

n_L : تعداد آبادیهای نمونه درطبقه L ام.

N_L : تعداد کل آبادیهای طبقه L ام.

S_L : انحراف معیارسطح زیرکشت محصولات مورد نظر طرح درآبادیهای طبقه L ام.

S_L^2 : واریانس سطح زیرکشت محصولات مورد نظر طرح درآبادیهای طبقه L ام.

x_{hi} : سطح زیرکشت محصولات مورد نظر طرح درآبادی i ام ازطبقه h ام.

k : تعداد طبقات درهرشهرستان.

لازم به توضیح است درشهرستانهائیکه بعضی ازآبادیهای آن براساس تعریف چارچوب در زمان آمارگیری فاقدکشت محصولات سالانه (گندم وجو ، سایرمحصولات) بوده‌اند درقالب طبقه‌ای جداگانه منظوروحداکثر ۵ درصدآبادیهای فوق بعنوان نمونه بصورت تصادفی ساده انتخاب گردیده تاتغییرات بوجودآمده درفعالیتهای کشاورزی آنهامشخص گردد . پس ازمشخص شدن تعداد نمونه لازم درهرطبقه، آبادیهای نمونه به روش سیستماتیک دایره‌ای انتخاب میگردند . دلیل استفاده‌از روش فوق اینست که خطای برآورد درنمونه‌گیری به روش سیستماتیک دایره‌ای خصوصـــــــا" در چارچوب‌هائی که واحدهای آماری مرتب شده باشند کمترازتصادفی ساده می‌باشد پس ازانتخـــاب آبادیهای نمونه جهت بررسی دقت نمونه گیری دوتست زیرانجام میگردد .

۱ـ توسط نمونه‌های انتخابی، سطح زیرکشت درشهرستان برآورد میشود واین برآوردباکـل سطح زیرکشت شهرستان براساس اطلاعات موجود درچارچوب مقایسه میگردد .

۲ـ خطای برآورد باآماره T درسطح تشخیص ۵ درصدآزمون میشود، بادرنظرگرفتن نتیجه بندهای ۱ و۲ نمونه‌های انتخابی تایید وباطبقه بندی مجدد عمل می‌آید .

پس ازانتخاب آبادیهای نمونه جهت بدست آوردن واحدهای آماری درمرحله‌دوم‌فهرست‌ـ برداری ازآبادیهای مذکوربه طریق ده گردشی انجام تابهره‌برداران ساکن آبادی مشخص شوند . در پایان فهرست برداری اسامی بهره‌برداران غیرساکن شهری لیست وپس‌ازمشخص‌شدن‌بهره‌برداران ساکن وغیرساکن شهری درهرآبادی باتوجه به وسعت سطح زیرکشت بطورصعودی مرتب وبسته دو طبقه کوچك وبزرگ تقسیم میشوند، تعداد بهره برداران کوچك نمونه درهرآبادی نمونه ده درصـــد کل بهره‌برداران کوچك آن بادرنظرگرفتن حداقل۲ بهره‌برداركه بصورت سیستماتیك دایـــره‌ای انجام میگیرد میباشد، همچنین درهرآبادی نمونه کلیه بهره‌برداران بزرگ مورد‌پرسش واقـــــع میشوند . میدانیم درجامعه‌هایی که واحدهای آماری مرتب شده باشند نمونه‌گیری بســـــــه روش سیستماتیك دایره‌ای ازدقت بیشتری نسبت به تصادفی ساده برخوردارمی‌باشد لذابه همین‌دلیل بهره‌برداران براساس سطح زیرکشت مرتب ونمونه گیری انجام میگردد .

پس ازانتخاب بهره برداران نمونه پرسشنامه‌های اطلاعاتی ازطریق پرسش ازبهـــره ـ برداران تکمیل میگردد، برای اینکارابتدا کلاسهای آموزشی جهت پرسنل اجرائی برگزاروسپس اکیپ های اجرایی متشکل ازآمارگیران ، بازبین‌هاوناظرین‌فنی عملیات صحرایی راشروع کرده پس ازتکمیل پرسشنامه توسط آمارگیران بازبین هائی که به همین منظورآموزش دیده‌اند کلیـه باربینی پرسشنامه‌هارا انجام میدهندتا اشکالات احتمالی مشخص وبامراجعه مجدد به بهـــره ـ

برداران موثق گردد . پس ازتکمیل پرسشنامه‌های اطلاعاتی دراستان به مرکزارسال ومجسمه‌ها " بررسی ، کنترل وکدگزاری میگردد وجهت استخراج ماشینی به قسمت پانچ فرستاده میشود پس ازپانچ اطلاعات لیستهای کنترلی ازکامپیوترگرفته میشودتا اشتباهات پانچی توسط ادیت چشمی برطرف گردد نهایتا " اطلاعات پرسشنامه‌های هرشهرستان برای کنترل نهائی توسط ادیت ماشینی درسطح یک هکتارتبدیل وباتوجه به پراکندگی آن برای هرآیتم اطلاعاتی درشهرستان وبا استفاده ازروش بررسی داده‌های آماری ازطریق نقاط میانه، چارکها وفریبال داده‌های دورویرت درسری مذکورمشخص وماجایگزینی میانه داده‌ها بجای داده‌های دورویرت اصلاح نهائی صورت میگردد .

فرمولهای برآورد :

برآورد سطح زیرکشت، میزان تولیدو ... هرمحصول درسطح شهرستان ازفرمول زیرحاصل میگردد .

$$S = \sum_{h=1}^{k} \frac{N_h}{n_h} \sum_{j=1}^{n_h} \left(\frac{M_{hj}}{m_{hj}} \sum_{i=1}^{m_{hj}} \sum_{p=1}^{s_{hji}} \frac{x_{hjip}}{f_p} + \sum_{p=1}^{r_{hj}} \sum_{q=1}^{t_{hjp}} \frac{y_{hjpq}}{d_q} \right)$$

که درآن

N_h تعدادکل آبادیهای طبقه h ام.

n_h تعدادآبادیهای نمونه طبقه h ام.

k تعدادطبقات.

M_{hj} تعدادکل بهره‌برداران کوچك آبادی j ام درطبقه hام.

m_{hj} تعداد بهره برداران نمونه کوچك آبادی j ام درطبقه h ام.

s_{hji} تعداد بهره برداریهای مستقل بهره بردارکوچك i ام درآبادی j ام ازطبقه hام.

x_{hjip} اندازه صفت مورد نظربهره برداری p ام ازبهره‌برداركوچك i ام درآبادی j ام از طبقه h ام.

f_p تعداد بهره برداران بهره برداری p ام.

r_{hj} تعدادبهره‌برداران بزرگ آبادی j ام ازطبقه h ام.

t_{hjp} تعدادبهره‌برداریهای بهره‌برداربزرگ i ام درآبادی j ام ازطبقه h ام.

y_{hjpq} اندازه صفت مورد نظربهره‌برداری q ام ازبهره‌برداربزرگ i ام درآبادی j ام از طبقه h ام.

d_q تعداد بهره برداران بهره برداری q ام.

هرچند تلاش شده ومیشود روشهای علمی وعملی پرتوان وشناخته شده درسطح جهان جهت تهیه وتدوین طرحهای آماری مورد استفاده قرارگیرد ولی عوامل متعددی درمواردی نتایج طرحها رازیرسوال میبرد که بایستی روشهای موثرتری به مورداجراگذاشته شودتعدادی ازعوامل فوق بشرح زیرمیباشد.

بافت جامعه موردبررسی(تنوع وپراکندگی کشت) واجرای طرحهاچندمنظوره که نمیتواند جوابگوی نیازهای آماری دربخش کشاورزی باشد لذابایستی حداقل برای محصولات اساسی هرکدام طرح جداگانه‌ای تدوین واجراگردد تابرآوردداطلاعاتی ازدقت بیشتری برخوردار گردد.

مراجعه دربلک نوبت به پاسخگویان، بدلیل اینکه فاصله کاشت تازمان برداشت اغلب محصولات طولانی میباشد بنابراین اگراطلاعاتی درمورد کاشت درزمان پایان برداشت مدنظر باشدازدقت بالایی برخوردارنخواهدبود.

مشکلات ونارسائیهای موجود: اهم این مشکلات ونارسائیهارابشرح زیرمیتوان برشمرد.

- مهمترین عامل دراخذاطلاعات دقیق پاسخگومیباشد که متاسفانه دربیشترمواردبعلت کمبود فرهنگ آماری درجامعه‌مسائل صنفی پاسخگویان دارای اریبی میباشند.
- عدم وجود شبکه اجرایی ونظارتی قوی ومستمروسیاست مشخص آماری.
- عدم شناخت لازم ازآمارتوسط اکثرافرادجامعه مخصوصا " مسئولین.
- عدم انطباق تشکیلات آماری متناسب بانیازهای آن.
- کمبود توان علمی وعملی پرسنل درنظام آماری موجود.
- عدم وجودچارچوبهای آماری مطمئن جهت تدوین واجرای طرحهای نمونه‌ای.
- وجود تبعیضهای شغلی بین پرسنل آماری بادیگران.
- کمبود شدید پرسنل وامکانات درواحدهای اجرائی دراستانهاومرکز.
- ثابت نبودن آمارگیران درمراکزاخذاطلاعات وعدم ارتباط سریع ومنسجم آنهابامرکز.

اگرچه تلاش میشود بهترین طرحهای آماری منطبق باصول شناخته شده تهیه وتدوین گردد همواره عوامل نامبرده یا لاازتوانی که بالقوه وبالفعل درنظام آماری دیده میشودبطور محسوسی کاسته که بایستی بارفع این عوامل وایجاد وگسترش شبکه انفورماتیک جهت

تسریع در ارسال اطلاعات بدست آمده گامی اساسی برداشته شود تا وضعیت نامطلوب کنونی به یک وضعیت نسبتا " مطلوب تبدیل گردد .

به این امید که در آینده به مسئله آمار و نقش ارزنده آن در زمینه‌های مختلف آنگونه برخورد شود که شایسته است .

سرشماری و نمونه‌گیری به صورت توأم

محمدباقر حقیقی انارکی

مرکز آمار ایران

۱- مقدمه :

افزایش سریع نیازهای اطلاعاتی جوامع و پیچیدگی‌های آن از یک طرف و معایب سرشماریها به لحاظ هزینه زیاد، تدارکات مشکل و عدم امکان دستیابی به برخی از اطلاعات در سرشماریها از طرف دیگر، آمارشناسان و دست اندرکاران سازمانهای آماری را بر آن داشت تا برای رفع این مشکل از تکنیک‌های آمارگیری نمونه‌ای استفاده بعمل آورند. اگرچه استفاده از نمونه‌گیری در سرشماریها در سایر زمینه‌ها مانند طراحی و آزمون روشهای سرشماری، کنترل و ارزیابی نتایج سرشماری و انتشار نتایج زودرس در کشور ما سابقه طولانی دارد ولی تکنیک نمونه‌گیری بهمراه سرشماری تنها در طرح جاری جمعیت سال ۱۳۷۰، مورد استفاده قرار گرفته است. در طرح جاری جمعیت، اطلاعات کلی شامل سن، جنس، وضع اقامت، وضع مهاجرت، محل تولید، تحصیل و سواد، وضع فعالیت و وضع زناشوئی از تمام افراد کشور سئوال شد و اطلاعات تفصیلی شامل پنج بخش مهاجرت، تحصیلات، اشتغال و بیکاری، باروری و مرگ و میر تنها از ده درصد خانوارهای کشور اخذ گردید.

با توجه به شرایط فعلی کشور و نیاز روز افزون به اطلاعات در زمینه‌های مختلف انتظار میرود در سرشماریهای آینده از تکنیک نمونه‌گیری به همراه سرشماری استفاده بعمل آید و ضمن جمع آوری اطلاعات تفصیلی و پیچیده مورد نیاز کشور در هزینه سرشماریها نیز صرفه جوئی موثر بعمل آورد.

۲ـ روشهای مختلف جمع‌آوری آمار و اطلاعات

صرف‌نظر از آمارهای ثبتی که در اثر جریان گردش یک فعالیت تولید می‌شود و نیز جدا از اطلاعاتی که از آزمایشهای انجام‌شده در زمینه علوم مختلف حاصل می‌شود، امروز، در اکثر کشورها آمار و اطلاعات به‌طور عمده با استفاده از دو روش زیر به‌دست می‌آید.

۱ـ۲ـ سرشماریها

۲ـ۲ـ آمارگیریهای نمونه‌ای

لازم به‌تذکر است که امروزه ما هواره‌های اطلاعاتی نیز بخشی از اطلاعات مورد نیاز جوامع را تولید می‌کنند ولی به‌دلیل عدم ارتباط با موضوع مورد نظر، از بحث راجع به آن خودداری می‌شود.

ذیلاً " مزایا و معایب دو روش فوق‌الذکر به‌جهت ارتباط با موضوع مورد بحث به طور خلاصه شرح داده می‌شود.

۳ـ مزایا و معایب سرشماریها و آمارگیریهای نمونه‌ای

در یک دیدگاه کلی می‌توان مزایای سرشماری را به عنوان معایب نمونه‌گیری تلقی کرد و برعکس معایب سرشماری را به‌عنوان مزایای نمونه‌گیری در نظر گرفت. بنابراین با مشخص شدن مزایا و معایب یکی از دو روش جمع‌آوری آمار و اطلاعات یعنی سرشماری و نمونه‌گیری، مزایا و معایب روش دیگر مشخص می‌شود.

۱ـ۳ـ مزایای سرشماریها

الف ـ سرشماریها امکان تهیه چارچوبهای آماری مختلف جهت انجام آمارگیریهای نمونه‌ای را به دست می‌دهند.

ب ـ سرشماری امکان ارائه اطلاعات در سطوح کوچک جغرافیائی را فراهم می‌سازند.

ج ـ درسرشماریها خطای نمونه‌گیری وجود ندارد . بنابراین اطلاعات حاصله از سرشماریها مستند و قابل مقایسه میباشند .

۲ـ۲ ـ معایب سرشماریها

الف ـ سرشماریها معمولا " باهزینه زیاد انجام می‌شوند و به تدارک امکانات ونیروی انسانی زیادی نیاز دارند . بنابراین نمیتوان آنرا در فواصل زمانی نزدیک به اجرا درآورد . از اینرو دستیابی به اطلاعاتی از جامعه که به سرعت تغییر می‌کند از طریق سرشماری تقریبا " غیرممکن میباشد .

ب ـ معمولا " به خاطر حجم زیاد کار در سرشماریها ، امکان نظارت دقیق وجود ندارد . بنا براین میزان اشتباهات غیرنمونه‌ای در سرشماریها معمولا " بیشتر از آمارگیریهای نمونه‌ای است .

ج ـ درسرشماریها اخذ برخی از اطلاعات که ماهیت پیچیده‌ای دارند امکان پذیر نیست . مانند اخذ اطلاعات درزمینه باروری زنان ، درآمد و . . .

د ـ به علت حجم زیاد اطلاعات جمع‌آوری شده در سرشماریها ، استخراج این حجم از اطلاعات به زمان زیادی نیاز دارد و ازاین نظر استفاده از اطلاعات سرشماریها معمولا " هنگامی میسر می‌شود که اطلاعات بهنگام بودن خود را از دست داده است .

لازم به ذکر است که مزایا و معایب سرشماری که درفوق مطرح شده است دریک شرایط عادی وبا درنظر گرفتن عوامل امکانات میتواند مطرح باشد . درغیراین صورت بعضی چنانچه عوامل امکانات را درنظر نگیریم معمولا " سرشماری ایده‌آلترین شیوه جمع‌آوری آمار و اطلاعات میباشد .

۴ـ اساس ترکیب نمونه‌گیری و سرشماری

پایه و مبنای کاربرد نمونه‌گیری به همراه سرشماری استفاده از نقاط قوت هر دو روش جمع‌آوری آماریعنی سرشماریها و نمونه‌گیریها است . بدین معنی که در کاربرد نمونه‌گیری به همراه سرشماری چنان ترکیبی از دو روش به وجود میاوریم که در مجموع نقاط قوت هر دو روش را در بر داشته باشد . برای ساخت چنین ترکیبی مطالعات ذیل ضروری است .

برای شروع چنین مطالعه‌ای ابتدا باید در هر سرشماری این سئوال مطرح شود که " آیا انجام یک آمارگیری نمونه‌ای به همراه سرشماری ضرورت دارد یا نه ؟ "

برای رسیدن به جواب سئوال فوق باید در مورد هر یک از اقلام آماری مورد نظر دو سئوال زیر مطرح و مورد بررسی قرار گیرد .

الف ـ آیا جمع‌آوری اطلاعات در مورد هر یک از اقلام مورد نظر بایدازطریق سرشماری انجام شود ؟

جهت رسیدن به پاسخ سئوال فوق باید در مورد هریک از اقلام آماری مورد نیاز موارد زیر مطرح و مشخص شود .

الف ـ ۱ ـ سطح جغرافیائی قلم آماری مورد نیاز چیست ؟ بدیهی است چنانچه قلم آماری مورد نظر در سطوح کوچک جغرافیائی مورد نیاز باشد ، آن قلم آماری باید از طریق سرشماری جمع‌آوری شود و چنانچه قلم آماری مورد نظر در سطوح جغرافیائی بزرگ مورد نیاز باشد ، امکان جمع‌آوری آن از طریق نمونه‌گیری وجود دارد .

الف ـ ۲ ـ قلم آماری مورد نظر با چه دقتی مورد نیاز است ؟ بدیهی است چنانچه قلم آماری مورد نظر با دقت بالا مورد نیاز باشد بدین معنی که فاقد خطای نمونه‌گیری باشد ، در این صورت آن قلم آماری باید از طریق سرشماری حاصل شود . در اینجا ذکر این نکته لازم است که اگر چه در سرشماریها خطای نمونه‌گیری وجود ندارد ولی باید توجه داشت که معمولا " میزان

خطاهای غیرنمونه‌ای در سرشماریها بیشتر از نمونه‌گیریها است ، بنابراین در مورد میزان دقت قلم آماری مورد نیاز باید مسئله فوق را مورد توجه قرار داد .

ب ـ آیا جمع‌آوری اطلاعات در مورد هر یک از اقلام مورد نظر از طریق سرشماری مطلوب و عملی است ؟

بمنظور بررسی و پاسخ به سئوال فوق ، باید برای هریک از اقلام آماری مورد نیاز محاسبات و بررسیهای زیر انجام پذیرد :

ب ـ ۱ ـ هزینه‌های مربوط به جمع‌آوری اقلام آماری مورد نظر در سرشماری و نمونه‌گیری .

ب ـ ۲ ـ زمان مورد نیاز برای پرسش اقلام آماری مورد نظر در سرشماری و نمونه‌گیری .

ب ـ ۳ ـ هزینه‌های مربوط به چاپ فرمها و پردازش اطلاعات در سرشماری و نمونه‌گیری

ب ـ ۴ ـ فشار حاصل بر کادر اجرائی سرشماری و پاسخگویان و اثر فشار مزبور بر کیفیت داده‌ها .

به‌عبارت دیگر در این بخش مشخصات قلم آماری مورد نیاز که در بند الف تعیین شده است در برخورد با واقعیات و توانائیهای موجود و امکانات موجود مورد ارزیابی قرار می‌گیرد و احتمالاً با توجه به بررسیهای بند ب ، بعضی از مشخصات قلم آماری مورد نیاز شامل سطح جغرافیائی و دقت مورد نظر تغییر خواهد کرد . توجه به مسائل فوق در هنگام طراحی یک سرشماری به‌طور قطع و یقین می‌تواند علاوه بر کاهش هزینه‌های جمع‌آوری و پردازش اطلاعات ، در بالا بردن کیفیت اطلاعات نیز تاثیر بسزائی داشته باشد . تجربیات کشورهائی که از این تکنیک استفاده کرده‌اند گواه این واقعیت است که نمونه‌گیری به‌همراه سرشماری هنگامی میتواند سودآور باشد که در این ترکیب کلیه بررسیهای فوق به‌دقت انجام گرفته باشد .

۵ـ تقسیم اقلام آماری مورد نظر بین سرشماری و آمارگیری نمونه‌ای

براساس بررسی‌های بندهای الف و ب بخش ۴ باید کلیه اقلام آماری مورد نظر را به دو سیستم جمع‌آوری اطلاعات یعنی سرشماری و نمونه‌گیری منتسب کرد. چنانچه براساس نتایج بررسی‌های انجام‌شده، جواب سئوالهای بندهای الف و ب مثبت باشد باید آن قلم آماری را از طریق سرشماری جمع‌آوری کرد و در صورتی که حتی پاسخ یکی از سئوالها منفی باشد، اطلاع مزبور را باید از طریق نمونه‌گیری جمع‌آوری کرد.

پس از مشخص شدن روش جمع‌آوری اقلام مورد نظر باید مقایسه‌ای بین هزینه‌ها و دست‌آوردهای حاصل بعمل آید. درباره‌ای از موارد تعداد اقلامی که باید از طریق نمونه گیری جمع آوری شوند به قدری کم است که در مجموع از نظر اقتصادی صلاح در گردآوری آنها از طریق سرشماری است. در هر حال باید توجه داشت که ایجاد توازن بین هزینه‌ها و دست‌آوردهای حاصل از این روش به عوامل زیادی بستگی دارد که باید کلیه عوامل شناخته شوند و بر آن اساس در صورتی که دست‌آوردهای حاصل از این روش نسبت به هزینه‌های استفاده از این روش مقرون به صرفه باشد، از روش نمونه‌گیری همراه با سرشماری استفاده نمود.

۶ـ تعداد نمونه و روش نمونه‌گیری

بعد از تشخیص ضرورت انجام نمونه‌گیری همراه با سرشماری و تعیین اقلامی که باید در طرح یا طرحهای نمونه‌گیری جمع‌آوری شوند، باید نسبت به تعیین تعداد نمونه و روش نمونه‌گیری بررسی‌های لازم بعمل آید.

تعداد نمونه و روش نمونه‌گیری به عوامل زیر بستگی دارد.

الف ـ امکانات اعم از مالی، نیروی انسانی، امکانات پردازش و ...

ب ـ سطح جغرافیائی اطلاعات

ج ـ دقت مورد نظر برای برآوردهای مختلف جامعه

د ـ پراکندگی صفات مورد نظر در جامعه

هـ ـ روشهای اجرائی ترکیب نمونه‌گیری و سرشماری

ز ـ سهولت یا پیچیدگی پرسش اقلام مورد نظر

لازم به تذکر است که معمولا " روشهای اجرائی ترکیب نمونه‌گیری و سرشماری، روش نمونه‌گیری و اقتدار نمونه متاثر از یکدیگر می‌باشند و معمولا " بیشترین پیچیدگی ها در مرحله انتخاب روش نمونه‌گیری به وجود می‌آید .

انتخاب روش نمونه‌گیری به عوامل زیادی بستگی دارد که تجزیه و تحلیل عوامل موثر خود به بررسی های زیادی نیاز دارد . لیلا " چند روش متداول نمونه‌گیری که معمولا " در ترکیب های نمونه‌گیری همراه با سرشماری مورد استفاده قرار می‌گیرد ذکر و مناسب و نامناسب هر کدام از این روشها نیز شرح داده می‌شود .

۱ـ۶ـ نمونه‌گیری تصادفی ساده

چنانچه اقلامی که باید در برنامه نمونه‌گیری مورد پرسش قرار گیرند ساده باشد یعنی به گونه‌ای باشد که با سرشماری بتوان در یک مراجعه به همراه سرشماری از شاخص های نمونه آنرا تکمیل کرد ، در این صورت روش نمونه‌گیری تصادفی ساده روش مناسبی است . در غیر این صورت یعنی چنانچه سئوالات مورد نظر در طرح نمونه‌گیری به گونه‌ای باشد که نتوان در یک مراجعه به همراه سرشماری از خانوارهای نمونه آنرا تکمیل کرد و نیاز به مراجعه مجدد داشته باشد ، با توجه به پراکندگی واحدهای آماری نمونه ، هزینه مراجعه بعدی در این روش بسیار زیاد خواهد بود .

۲ـ۶ـ نمونه‌گیری خوشه‌ای

این روش هزینه یابی بیشتری نسبت به روشهای دیگر نمونه‌گیری دارد و اجرای آن نیز آسانتر است . خصوصا " زمانی که مجبور هستیم به علت پیچیدگی سئوالها این طرح نمونه‌گیری

مامورآن جداگانه‌ای داشته‌باشیم ؛ درعین حال باید توجه داشت که معمولا " به‌علت وجود همبستگی داخلی بین خوشه‌ها ، واریانس برآوردهای حاصل ازاین روش بیشتر از سایر روشهای نمونه‌گیری است .

۳ـ۶ـ نمونه‌گیری چندمرحله‌ای

این روش ازنظر هزینه وسهولت انجام آن نسبت به نمونه‌گیری تصادفی ساده ارجحیت دارد ولی درمقایسه بانمونه‌گیری خوشه‌ای از هزینه وسختی کار بیشتری برخوردار است . یکی ازمشکلات عمده این روش انتخاب واحدهای رده بالاتر از واحد آماری است . چنانچه بخواهیم واحدهای آماری باشانس مساوی انتخاب شوند ، دراین صورت انتخاب واحدهای رده بالاتر کاری بسیار دشوار خواهد بود . درغیراین صورت ، یعنی چنانچه واحدهای آماری با احتمال نامساوی انتخاب شوند ، دراستخراج نتایج دشواری خواهیم داشت .

نهایتا " ممکن است برای حصول به اهداف درنظرگرفته شده به بیش ازیک طرح نمونه‌گیری نیاز داشته باشیم . دراین گونه موارد برای انتخاب واحدهای نمونه برای طرحهای مختلف ، ترکیبهای مختلفی رامیتوانیم درنظربگیریم .

الف ـ انتخاب واحدهای آماری مشترک برای کلیه آمارگیریهای نمونه‌ای .

دراین روش هزینه آمارگیری کاهش می‌یابد وضمنا " مامورآمارگیرباپی بردن به روابط بین متغیرها در طرحهای مختلف میتواند مانع بروز اشتباهات ازطرف پاسخگو شود . ولی دراین روش هر مامورآمارگیر باید مرحله آموزشی مربوط به تمام طرحها را طی کند وضمنا " به پاسخگویان فشار زیادی جهت پاسخگوئی به کلیه سئوالهای طرحهای مختلف وارد می‌شود .

ب ـ انتخاب واحدهای نمونه مشترک رده بالاتر از واحد آماری و انتخاب واحدهای آماری متفاوت برای هر کدام از طرحها

دراین روش واحدهای آماری توزیع مناسبی درجامعه خواهند داشت ولی برنامه‌ریزی

واجرای طرح را با مشکلات زیادی روبرو می‌کند .

ج ـ انتخاب واحدهای نمونه جداگانه رده بالاتر از واحد آماری

این روش از سهولت زیادی برخوردار است . خصوصاً "اگر برای هر طرح آماری گیران مختلفی در نظر گرفته شده باشد . لکن با یک هزینه ثابت برای آمارگیری ، تعداد واحدهای نمونه رده بالاتر از واحد آماری تقلیل می‌یابد و در نتیجه واریانس برآوردها افزایش خواهد یافت .

محمدباقر حقیقی

محاسبهٔ شاخص نابرابری توزیع کالا (روغنهای نباتی) به تفکیک مناطق شهری و روستایی

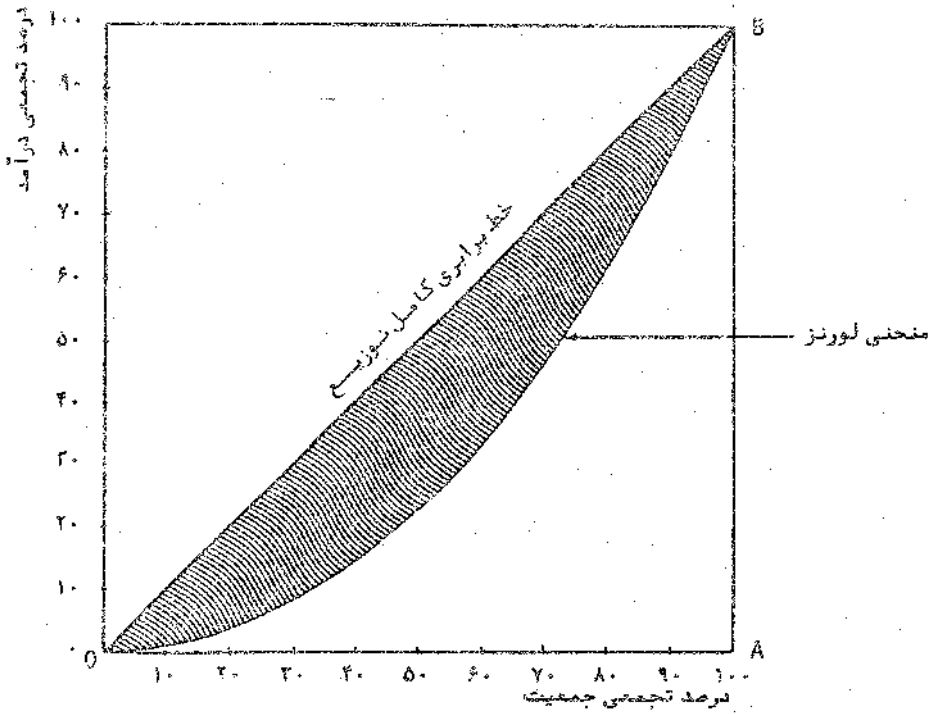
طاهره دفتری وطن‌پرست

مرکز آمار ایران

در این بررسی با اقتباس از منحنی لورنز "Lorenz Curve" وضریب‌جینی با ضریب‌تمرکز "Gini Coefficient" که‌از روشهای آماری جهت‌اندازه‌گیری نابرابری توزیع درآمدمی‌باشد، شاخص نابرابری توزیع‌مقدارمصرف‌روغنهای نباتی درواحدآماری موردنظر (خانوار) به‌تفکیک مناطق شهری وروستائی موردمحاسبه‌قرارگرفته‌است. دربررسی توزیع کیفی روغنهای نباتی، ارزش مقداری آن (ونه‌ارزش پولی) مبنای محاسبه‌وتحلیل واقع شده‌است. با توجه‌به‌تکنیک وابزار بکار گرفته‌شده در این بخش از مطالعه، ذیلاً منحنی لورنزوضریب جینی به‌اختصار توضیح داده‌می شود.

ـ منحنی لورنز "Lorenz Curve"ـ این منحنی رابطه‌بین‌نسبتهای تجمعی متغیرمالی (درآمد) ونسبتهای تجمعی واحدهای دریافت‌کننده (خانوارها) رانشان می دهد. با استفاده‌ازمنحنی لورنزمی‌توان وابستگی جمعی فرادانیها ومتغیرهائی که‌توزیع آنها موردنظراست راتوأما" نشان داد. برای مثال، با تجمعی کردن توزیع درآمدمی توان درصد درآمدی که‌بوسیله‌طبقات مختلف درآمدی تحصیل می شودرا نمایش داد، به‌بیان دیگرمی توان گفت که‌چنددرصدازخانوارهای موردبررسی، چند درصدازکل درآمدرا کسب می کنند. برای رسم منحنی پس‌ازمرتب‌نمودن درآمدهای خانوارها به‌ترتیب صعودی، درهرمرحله‌هم‌زمان با محاسبه‌درصد تجمعی تعدادخانوارها، درصدتجمعی درآمدمتناظرخانوارها نیزمحاسبه می شودوبه‌این ترتیب مختصات نقاط موردنظرتعیین می گردد. بدیهی است خط منتسب‌از (۰) که‌به‌خط برابری کامل توزیع درآمدموسوم‌است‌ازمجموعه نقاطی تشکیل شده‌که‌درصدتجمعی جمعیت ودرصدتجمعی درآمدمتناظرآنها

برابر است . این منحنی درحالت برابری کامل توزیع درآمددریک جامعه برخط برابری کامل توزیع درآمدانطباق می‌یابد (خط $\overline{OB}$) ودرجامعه‌ای با عدم تعادل کامل توزیع درآمدبرخط شکسته OAB قرارمی‌گیرد.



ـ ضریب جینی " Gini Coefficient " ـ ضریب جینی از شاخصهای مهم اندازه‌گیری نابرابری توزیع درآمدمی باشد . این ضریب درمنحنی لورنزدرصدتمرکزرا نشان داده‌وبنا برتعریف مقایسه‌ای‌است بین سطح شکاف بین منحنی لورنزوخط برابری کامل توزیع درآمد (سطح هاشورخورده) با سطح مثلثی که منحنی لورنزدرآن قرارداردد . ضریب جینی بین دوحدصفروبک قرارمی‌گیردوچنانچه این ضریب برای توزیعی که‌بوسیله منحنی لورنزنشان داده‌شده‌صفرباشد ، خط واقعی توزیع روی خط برابری کامل منطبق است

(نشان دهنده عادلانه ترین توزیع) وچنانچه این ضریب برابر یک باشد نابرابری کامل درتوزیع متغیر موردبررسی درجامعه موردنظروجودخواهد داشت .

درعمل هنگام محاسبه ضریب جینی ، با گروههای هزینه‌ای (= درآمدی) مواجه می شویم وبنابراین منحنی به حالت خط شکسته خواهدبود . این مقدار معمولا" کمتر از مقدار واقعی ضریب جینی است زیرا خط شکسته بجای منحنی واقعی فرض شده که حد پائین " Lower Bound " ضریب جینی است ورابطه آن براساس فرمول Miller به صورت زیراست[1]:

$$G_L = 1 - \sum_{i=1}^{k} (X_{i+1}-X_i)(Y_{i+1}+Y_i)$$

که درآن :

G_L ـ حد پائین ضریب جینی

X_i ـ درصدتجمعی جمعیت درگروه درآمدی

Y_i ـ درصدتجمعی درآمد (هزینه)

L.Gastwirth[2] یک راه حل آماری برای محاسبه حد بالای " Upper Bound" منحنی فوق الذکر بوا سا س رابطه زیر ارائه داده است :

$$G_U = G_L + \bar{D}$$

$$\bar{D} = \frac{1}{\mu} \sum_{i=1}^{k} (X_{i+1}-X_i)^2 \frac{(\mu_{i+1}-g_i)(g_{i+1}-\mu_{i+1})}{(g_{i+1}-g_i)}$$

۱ـ سهرابی ، حمیدرضا فاجانزاده ، مدیک ، (۱۳۶۱)، " مطالعه تطبیقی وضع درآمدها وهزینه‌ها درمناطق شهری ایران ، ۱۳۵۷-۱۳۵۱ ، سازمان برنامه وبودجه ـ دفتربرنامه ریزی اجتماعی ونیروی انسانی ، ص ۲۲

2. Gastwirth,J.L.(1972)," The Estimation of Lorenz Curve and Gini Index", Review of Economics and Statistics. pp.306-316

که در آن :

G_U ـ حد بالای ضریب جینی

G_L ـ حد پائین ضریب جینی

$\bar{D}$ ـ سطح بین منحنی لورنز و خط شکسته جایگزین منحنی لورنز

μ ـ میانگین کل درآمد

μ_i ـ میانگین درآمد در گروه درآمدی

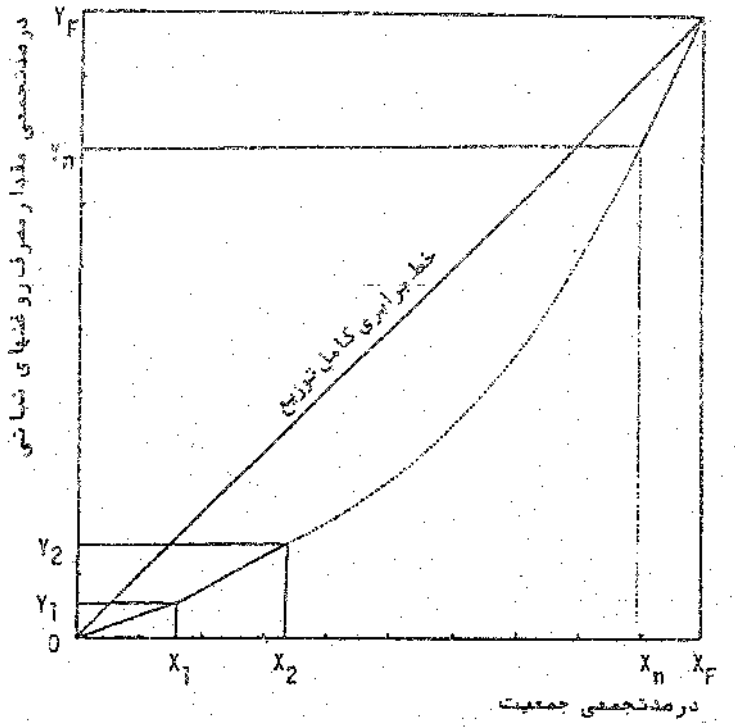
X_i ـ درصد تجمعی جمعیت

$g_{i+1}-g_i$ ـ حدود گروههای هزینه (= درآمدی) که با X فاصله معین شده است .

استفاده از تکنیک و روش فوق الذکر در اندازه‌گیری توزیع درآمد در رابطه با بسیاری از متغیرهای دیگر نیز می تواند کاربرد داشته باشد ، ولذا می توان از آن برای بررسی شاخص نابرابری مقدار مصرف روغنهای نباتی استفاده کرد .

چنانچه صرفاً از طریق اقتباسی از منحنی لورنز و ضریب جینی چگونگی مقدار مصرف روغنهای نباتی در بین جمعیت را مشخص نمائیم به وضعی ساده و ابتدائی برخورد می کنیم به این مفهوم که روی یک محور درصد تجمعی جمعیت و در محور دیگر بجای درصد تجمعی درآمد ، درصد تجمعی مقدار مصرف روغنهای نباتی متناظر جمعیت مذکور را نشان می دهیم آنگاه از مقایسه سطح بین منحنی لورنز و خط برابری کامل توزیع با سطح مثلثی که منحنی لورنز در آن قرار دارد ، شکاف طبقاتی در مورد توزیع مقدار مصرف روغنهای نباتی مشخص می شود .

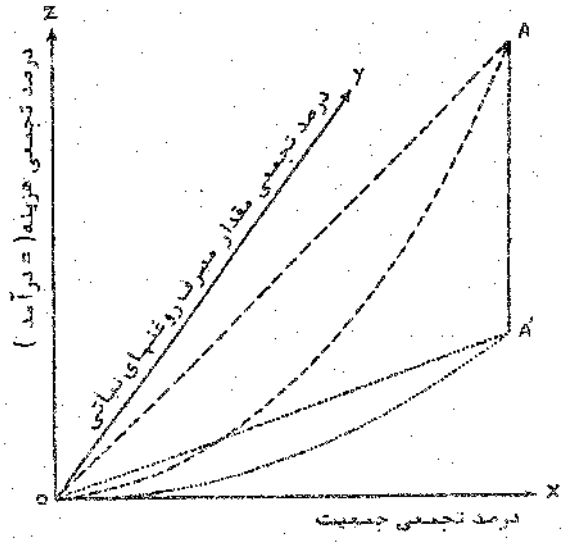
محاسبه سطح مورد نظر از نظر هندسی ، به صورت زیر است :



$$\text{سطح زیر منحنی} = S = \frac{0X_1 \cdot 0Y_1}{2} + \frac{X_1X_2(0Y_1+0Y_2)}{2} + \cdots\cdots + \frac{X_nX_F(0Y_n + 0Y_F)}{2}$$

همانگونه که ملاحظه می شود عبارت اول درصورت تمام کسرها یعنی $0X_1$ ، X_1X_2 و ..... X_nX_F درصدهای نسبت جمعیت به کل وجمله دوم درتمام کسرها یعنی $0Y_1$ ، $0Y_1+0Y_2$ و ..... $0Y_n + 0Y_F$ جمع دوتائی و پی درپی درصدهای تجمعی مقدار مصرف روغنهای نباتی جمعیت متناظرمی باشد . به دلیل آنکه اطلاعات مربوط به گروههای مقدار مصرف روغنهای نباتی بطور مستقل دردسترس نمی باشد ومقدار مصرف روغنهای درقالب گروههای هزینه‌ای (= درآمدی)

موجود است ، با منظور نمودن بعد درآمد و مشخص شدن این نکته که موقعیت مصرف کننده و طرز رفتار وی در مورد مقدار مصرف روغنهای نباتی تا چه حد تحت تاثیر درآمد می باشد ، با یک فضای سه بعدی مواجه می باشیم . بدین ترتیب که درصد تجمعی جمعیت روی یک محور و درصد تجمعی مقدار مصرف روغنهای نباتی روی محور دیگر و درصد تجمعی هزینه (= درآمد) در روی محور سوم منظور می شود و بر این اساس با یک منحنی فضائی سروکار خواهیم داشت که به صورت زیر به تصویر کشیده شده است :



در عمل و اجرا محاسبه شاخص نابرابری در منحنی فوق دشوار می باشد و لذا برای آنکه بتوانیم شکاف طبقاتی را در مورد توزیع مقدار مصرف روغنهای نباتی بررسی نمائیم از تصویر این منحنی فضائی در صفحه XOY استفاده می کنیم و تصویر را مورد مطالعه قرار میدهیم . لذا درصد تجمعی جمعیت در هر گروه هزینه ای (= درآمدی) روی محور OX و درصد تجمعی مقدار مصرف روغنهای نباتی در هر گروه روی محور OY منظور می شود و سپس از مقایسه منحنی -

نمودار با خط نیمساز زاویه $X\hat{O}Y$ به توجیه شکاف نابرابری توزیع می‌پردازیم، و بدیهی است که هرچه سطح بین خط برابری کامل توزیع و منحنی بیشتر باشد، شکاف طبقاتی بیشتر خواهد بود. منحنی نمودار با منحنی لورنز دو فرق عمده دارد:

ـ منحنی لورنز فقط با منظور نمودن درآمد، وضعیت جمعیت را توجیه می کند اما منحنی نمودار حالتی پیچیده‌تر از منحنی لورنز است به این مفهوم که مقدار مصرف روغنهای نباتی در هرگروه هزینه‌ای (= درآمدی) با جمعیت مربوط می گردد.

ـ منحنی لورنز همواره در زیر خط برابری کامل توزیع قرار می گیرد درحالیکه منحنی نمودار ممکن است در مواردی خط برابری کامل توزیع را در یک یا چند نقطه قطع کند و یا حتی کاملاً در بالای خط برابری کامل توزیع قرار گیرد که هرکدام می تواند رفتار مصرف کننده را با توجه به هزینه (=درآمد) توجیه نماید. حال علاوه بر حالت فوق که بطور مطلق شکاف مقدار مصرف روغنهای نباتی بین گروههای هزینه (= درآمد) را نشان می دهد، بایستی مشخص شود که بطور نسبی، با افزایش هزینه (= درآمد) مصرف کننده:

ـ آیا مقدار مصرف روغنهای نباتی کاهش می یابد؟

ـ آیا مقدار مصرف روغنهای نباتی افزایش می یابد؟

ـ آیا در سطوح هزینه‌ای (= درآمدی) با ثبات مقدار مصرف روغنهای نباتی کم می شود و به تدریج که درآمد افزایش یافت، به مقدار مصرف افزوده می گردد؟

محاسبه شکاف طبقاتی در مورد توزیع مقدار مصرف روغنهای نباتی بطور مطلق پاسخگوی نکات فوق نمی باشد لذا بررسی شکاف طبقاتی در رابطه با توزیع مقدار مصرف روغنهای نباتی بطور نسبی (نسبت به جمعیت طبقات) برای تحلیلهای اقتصادی ضرورت می یابد. محاسبه شکاف طبقاتی بطور نسبی

مشخص می کند که بطور متوسط توزیع به نفع یا زیان چه طبقاتی صورت گرفته است و برای این منظور به های درصدتجمعی مقدار مصرف روغنهای نباتی در هرگروه هزینه ای (= درآمدی) درصدتجمعی متوسط های سرانه مقدار مصرف روغنهای نباتی درنظرگرفته می شود یعنی در حالت اول مشخص می شدکه یک درصد جمعیتی معین چنددرصد از کل مقدار مصرف روشهای نباتی را دریافت کرده اند و اکنون این مسئله مطرح می شودکه همان درصدجمعیت ، چنددرصد از کل متوسط های مقدار مصرف سرانه روغنهای نباتی هرگروه هزینه ای (= درآمدی) را به خود اختصاص داده اند یعنی ستغییرجدیدی (متوسط مقدار مصرف سرانه) درگروههای درآمدی درنظرگرفته شده و با یکدیگر مورد مقایسه قرار گرفته اند .

با فرض ثابت بودن درصدهای مقدار مصرف سرانه روغنهای نباتی در هرگروه هزینه ای (= درآمدی) موارد زیر قابل طرح می باشد :

- درهر مقطع درصورتیکه درصدتجمعی مقدار مصرف سرانه روغنهای نباتی کمتر از درصدتجمعی جمعیت باشد منحنی زیرخط نیمساز زاویه $X\hat{O}Y$ واقع می شود ، در این حالت نتیجه می گیریم علاوه بر آنکه ممکن است توزیع بطور مطلق نابرابر باشد ، بطور نسبی نیز به نفع گروههای پردرآمد است . درعین حال سطح زیرمنحنی نشانگر موقعیتی است که هرچه بیشتر باشد ، درجه رفاه نسبی کمتر خواهد بود و حداکثر می تواند بر دو ضلع مثلثی که در آن قرار گرفته واقع شود ، در این وضعیت میزان شاخص ۱ خواهد بود .

- هرگاه منحنی بالای خط نیمساز زاویه $X\hat{O}Y$ قرار گیرد ، علیرغم آنکه ممکن است بطور مطلق توزیع نابرابر باشد ولی بطور نسبی به نفع طبقات کم درآمد خواهد بود ، در این حالت هرچه سطح بین منحنی و خط نیمساز زاویه $X\hat{O}Y$ بیشتر باشد ، نشانگر رفاه نسبی بیشتر است و حداکثر بر دو ضلع مثلث بالائی منطبق خواهد شد و در اینصورت شاخص برابر ۱- می باشد .

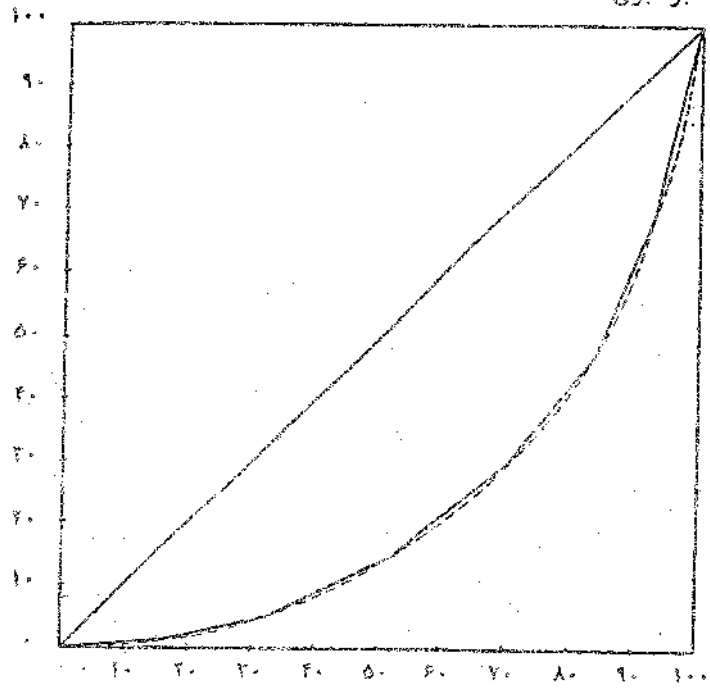
- چنانچه منحنی خط نیمساز زاویه $\hat{OY}$ را قطع نماید ، با توجه به مجموع سطوحی که بین منحنی وخط برابری کامل توزیع درو بالا و زیر خط نیمساز قرار می گیرند ، درجه عدم تساوی رفاه محاسبه و مشخص می شود .

همانطور که گفته شد شاخص نابرابری توزیع از مقایسه بین سطح شکاف بین منحنی وخط برابری کامل توزیع درآمد با سطح مثلثی که منحنی درآن واقع است حاصل می شود و برای این آسان شاخص نابرابری توزیع که با دو برابر سطح بین منحنی وخط برابری کامل توزیع درآمد برابر خواهد شد از رابطه زیر بدست آمده است :

$$I > I_L = 1 - T$$

که در آن : ۱۰۰۰۰/ مجموع ستون ۱۰ = T می باشد .

با توجه به شکل زیر ملاحظه می شود که شاخص مورد نظر (I) فقط در یک حالت خاص برابر طرف دوم رابطه فوق می شود و آن حالتی است که توزیع مقدار مصرف روغنهای نباتی بین جمعیت به گونه ای باشد که خط شکسته و منحنی برهم منطبق شوند و از این جهت مقدار شاخص نابرابری توزیع که بر مبنای اطلاعات جدول و فرمول مورد نظر بدست می آید (I_L) حد پائین شاخص نابرابری است .



با توجه به مطالب فوق الذکر ، برای محاسبه شاخص نابرابری توزیع مقدار مصرف روغنهای نباتی در چهار سال انتخابی ۱۳۵۶ ، ۱۳۵۸ ، ۱۳۶۱ و ۱۳۶۸

از نتایج طرح آمارگیری از هزینه و درآمد خانوارها در مناطق شهری و روستائی که توسط مرکز آمار ایران ارائه شده، استفاده گردیده است:

برخی از تعاریف و مفاهیم موردنظر در طرح مزبور به شرح زیر است[1]:

- خانوار: منظور از خانوار، خانوار معمولی ساکن است که بر حسب تعریف از یک یا چند نفر تشکیل می شود که با هم در یک اقامتگاه معمولی و ثابت زندگی می کنند و یا یکدیگر هم خرج هستند.

- مناطق شهری - در طرح مذکور، در سالهای ۱۳۵۶، ۱۳۵۸ و ۱۳۶۱ به کلیه نقاطی که در سرشماری آبان ماه سال ۱۳۵۵ مرکز شهرستان بوده (بدون در نظر گرفتن تعداد جمعیت) و یا در زمان سرشماری مذکور بیش از ۵۰۰۰ نفر جمعیت داشته اند مناطق شهری اطلاق شده است و در سال ۱۳۶۸ کلیه نقاطی که در سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ دارای شهرداری بوده اند جزو مناطق شهری به حساب آمده اند.

- مناطق روستائی - در طرح هزینه و درآمد خانوار در سالهای مورد نظر نقاطی که شهر تلقی نشده اند، جزء مناطق روستائی به حساب آمده اند.

- هزینه - ارزش پولی تهیه کالا یا خدمت توسط خانوار، هزینه محسوب می شود. تهیه شامل خرید، تولید خانگی، دریافت برخدمت، از محل کسب، پایاپای و مجانی می باشد که به صورت پولی برآورد و جزء هزینه منظور شده است.

---

[1] سرشوری، محمدباقر. (۱۳۶۹)، مروری بر شاخصهای نابرابری توزیع درآمد در کل کشور، مناطق شهری و مناطق روستائی، سالهای ۱۳۶۸-۱۳۶۱. نشریه داخلی شماره ۱۳۶. تهران: مرکز آمار ایران، ص ۲

ـ گروههای هزینه (= درآمد) ـ درطرح مذکورگروههای هزینه‌سالانـــه در سالهای موردنظرطبق جدول زیراست :

گروههای هزینه سالانه	سالهای ۱۳۵۶ و ۱۳۵۸	سالهای ۱۳۶۱ و ۱۳۶۸
گروه ۱	کمتر از ۳۰۰۰۰	کمتر از ۱۲۰۰۰۱
گروه ۲	۳۰۰۰۰ تا ۵۹۹۹۹	۱۲۰۰۰۱ تا ۲۴۰۰۰۰
گروه ۳	۶۰۰۰۰ تا ۸۹۹۹۹	۲۴۰۰۰۱ تا ۳۶۰۰۰۰
گروه ٤	۹۰۰۰۰ تا ۱۱۹۹۹۹	۳۶۰۰۰۱ تا ۴۸۰۰۰۰
گروه ۵	۱۲۰۰۰۰ تا ۱۴۹۹۹۹	۴۸۰۰۰۱ تا ۶۰۰۰۰۰
گروه ۶	۱۵۰۰۰۰ تا ۲۲۹۹۹۹	۶۰۰۰۰۱ تا ۹۰۰۰۰۰
گروه ۷	۲۳۰۰۰۰ تا ۳۵۹۹۹۹	۹۰۰۰۰۱ تا ۱۴۰۰۰۰۰
گروه ۸	۳۶۰۰۰۰ تا ۵۹۹۹۹۹	۱۴۰۰۰۰۱ تا ۲۸۰۰۰۰۰
گروه ۹	۶۰۰۰۰۰ تا ۱۱۹۹۹۹۹	۲۸۰۰۰۰۱ تا ۴۲۰۰۰۰۰
گروه ۱۰	۱۲۰۰۰۰۰ و بیشتر	۴۲۰۰۰۰۱ و بیشتر

ـ نتایج محاسبات

نتایج محاسبات مربوط به شاخصهای نابرابری توزیع مقطعی درآمد و مصرف و شاخصهای نسبی با استفاده به جداول شماره ۶۲ تا ۷۱ به شرح زیر تشخیص داده شده است :

سال	شاخص نابرابری توزیع به طور مطلق		شاخص نابرابری توزیع به طور نسبی	
	مناطق شهری	مناطق روستایی	مناطق شهری	مناطق روستایی
۱۳۵۶	۰/۴۲	۰/۰۴	۰/۲۹	۰/۰۲
۱۳۵۸	۰/۳۶	۰/۰۱	۰/۲۹	۰/۰۱
۱۳۶۱	۰/۰۸	۰/۰۴	۰/۰۷	۰/۲۸
۱۳۶۸	۰/۳۲	۰/۱۰	۰/۴۲	۰/۱۴

تذکار : این نتایج نشاندهنده حداکثین شاخص نابرابری می باشدو محاسبه حداقل به‌دلیل عدم امکان دسترسی به‌اطلاعات مربوط به‌گروههای مقدار مصرف روغنهای نباتی مستقل ازگروههای هزینه (= درآمد) میسرنمی باشد .

ازبررسی نتایج مربوط به‌شاخصها و منحنی های نابرابری مقدارمصرف روغنهای نباتی (بطورمطلق ونسبی) درسالهای ۱۳۵۶،۱۳۵۸،۱۳۶۱ و ۱۳۶۸ درمناطق شهری وروستائی موارددزیراستنتاج می شود :

ـ درسالهای موردبررسی درمناطق شهری بیشترین میزان شاخص نابرابری بطورمطلق درسال ۱۳۵۸ (اولین سال بعدازانقلاب و قبل از طرح سهمیه‌بندی) برابر ۰/۳۶ وکمترین آن درسال ۱۳۶۱ (بعدازاجرای طرح سهمیه‌بندی) برابر ۰/۰۸ می باشد . ارقام محاسبه‌شده‌برای شاخص نابرابری بطورمطلق نشاندهنده نابرابری بیشترتوزیع درسال ۱۳۵۸ و عادلانه‌ترشدن آن درسال ۱۳۶۱ است .

ـ میزان شاخص نابرابری بطورمطلق درمناطق روستائی درسالهای موردبررسی، روندی صعودی داشته‌ونشاندهنده نامتعادلترشدن توزیع در سال ۱۳۶۸ نسبت به‌سایرسالهای موردبررسی است .

ـ درمقایسه‌شاخص نابرابری مطلق بین مناطق شهری وروستائی‌ملاحظه می شودکه‌درسالهای مورد مطالعه‌همواره‌نابرابری توزیع بطورمطلق، در مناطق شهری بیش ازمناطق روستائی است .

ـ دامنه‌تغییرات‌شاخص نابرابری بطورمطلق درسالهای مورد بررسی درمناطق شهری بیش ازمناطق روستائی است به‌گونه‌ای که‌درمناطق شهری برابر ۰/۲۸ ودرمناطق روستائی ۰/۱۴ می باشد .

ـ بیشترین تفاوت وفاصله‌بین شاخص نابرابری بطورمطلق درمناطق شهری ومناطق روستائی مربوط به‌سال ۱۳۵۸، برابر ۰/۲۵ وکمترین آن در ـ

سال ۱۳۶۱ ، معادل ۱/۰ است .

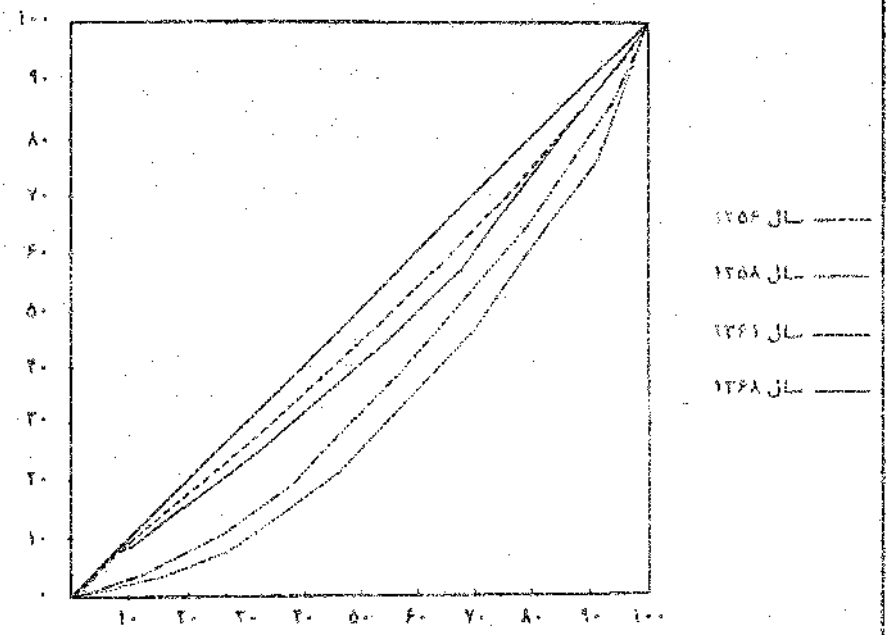
– بررسی شاخص نابرابری بطورنسبی درمناطق شهری نشان دهنده، نا متعادلترشدن توزیع درسال ۱۳۵۸ نسبت به سایرسالهای موردمقایسه و متعادلترشدن توزیع بطورنسبی درسال ۱۳۶۸ است ، مقدار این شاخص در سال ۱۳۵۸ برابر ۰/۳۹ ودرسال ۱۳۶۸ معادل ۰/۴۳ می باشد این ارقام نشانگر آن است که درسال ۱۳۵۸ مقدار مصرف روشنهای نباتی بطورنسبی به نفع اقشار مرفه وبه زیان خانوارهای کم درآمد بوده است .

– مقایسه شاخص نابرابری بطورنسبی درسالهای موردبررسی در مناطق روستائی نشان میدهدکه این شاخص درسال ۱۳۶۱ برابر ۰/۳۸ است که بیانگرنا متعادلترشدن توزیع بین گروههای هزینه (= درآمد) دراین سال نسبت به سایرسالهای موردنظرمی باشد ، ازسوی دیگراین شاخص درسال ۱۳۶۸ برابر ۰/۱۴ است که نشانگرتوزیع متعادلتربه نفع خانوارهای کم درآمد روستائی است .

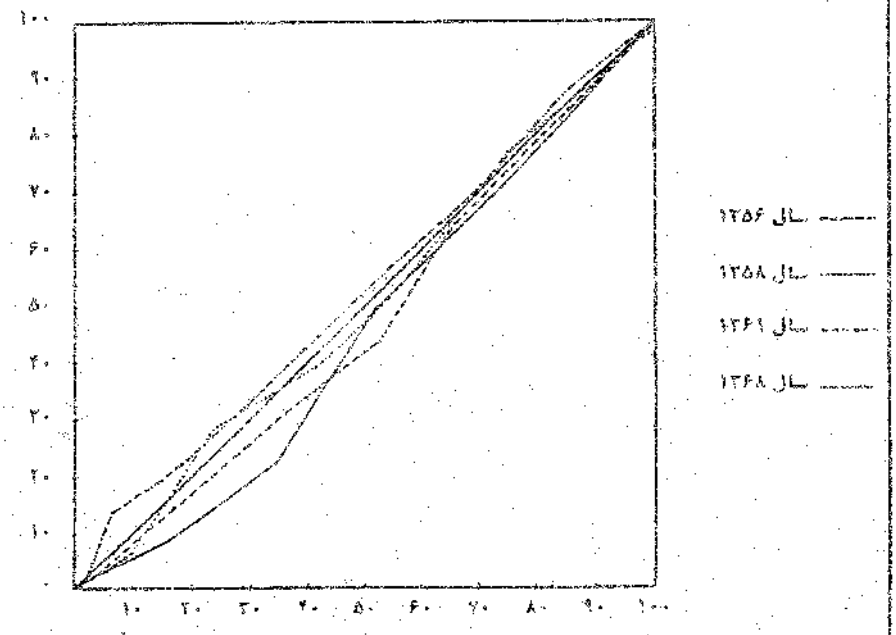
– مقایسه شاخصهای نابرابری درمناطق شهری درسال ۱۳۵۸ بطورمطلق ونسبی گویای نابرابرترین توزیع نسبت به سایرسالهای موردبررسی است ، به عبارت دیگردراین سال نه تنها توزیع درمناطق شهری بطوری مطلق نابرابرتر ازسایرسالهای موردمقایسه است بلکه بطورنسبی نیزبه نفع گروههای مرفه بوده است .

– درسال ۱۳۶۸ درمناطق روستائی علیرغم آنکه شاخص نابرابری بطور مطلق بیش ازسایرسالهای موردنظراست لکن شاخص نابرابری بطورنسبی نشاندهنده عادلانه ترشدن توزیع به نفع گروههای کم درآمدمناطق روستائی می باشد .

منحنی نابرابری توزیع مقدار مصرف روغن نباتی در مناطق شهری بطور مطلق



منحنی نابرابری توزیع مقدار مصرف روغن نباتی در مناطق روستایی بطور مطلق



آمار در بخش کشاورزی، مشکلات و راه‌حلها

علی‌رضا دهقان

مرکز آمار ایران

۵شنبه ۱۴:۳۰ تالار ۳

ترسیم وضع مطلوب اوّلین آمارگیری، به طرق علمی در خصوص کشاورزی ایران در سال ۱۳۳۹ انجام گرفته و از آن تاریخ تاکنون سرشماری‌های ۱۳۵۲ و ۱۳۶۰ و ۱۳۶۷ با اختلافات فاحشی در تعاریف و روش‌ها و اهداف انجام شده است. به‌طور قانونی مرکز آمار ایران مسئولیت اجرای آمارگیریهای سرشماری را درکشور به عهده دارد. جمع‌آوری آمار فعالیت‌ها و تولید آمار و...

سطح زیرکشت و تولید از طریق نمونه‌گیری به عهده وزارت کشاورزی و اخیراً در خصوص دام با جهاد سازندگی است. اصولاً همیشه بین آمار وزارت کشاورزی و مرکز آمار ایران در خصوص سطح زیرکشت، حجم تولید و عملکردها اختلافات فاحشی وجود دارد.

هر دو دستگاه اصلی تولیدکنندهٔ آمار در بخش کشاورزی، خواه سرشماری، خواه نمونه‌گیری از شیوه‌های سنّتی پرسش و پاسخ استفاده می‌کنند. درکشور ما به تجربه ثابت شده که این روش به علل گوناگون فرهنگی و اقتصادی با اشکال زیادی روبرو است. هم پرسشگر و هم پاسخگو تجربه و ایمان کامل به آمارگیری ندارند و خود موجب بروز انحرافات زیادی می‌شوند.

به نظر می‌رسد وقت آن رسیده که:

- هماهنگی کامل بین واحدهای تولیدکننده و جمع‌آوری‌کنندهٔ آمار در بخش کشاورزی به وجود آید
- روشهای آمارگیری، تجزیه و تحلیل و ادبیت و غیره با تکنیکهای آماری روز منطبق شوند
- از تکنیکهای نوین تولید آمار و اطّلاعات مثل استفاده از عکسهای ماهواره‌ای و هوایی و تلفیق آن با شیوه‌های موجود بهره‌گیری‌شده و به سمت رهایی هر چه بیشتر آمار از وابستگی به پرسش و پاسخ حرکت کنیم.

بررسی اجمالی وضعیّت آمارهای ثبتی کشور

عبدالکریم دی‌پیر

مرکز آمار ایران

اگر در هر زمینه مجموعه داده‌های کمی (عددی) متفرق را از اسناد اداری استخراج و با انجام عملیات آماری آنها را به صورت داده‌های جمعی در سطوح مختلف تنظیم نماییم، داده‌های آماری (آمار) حاصل می‌گردد که به خاطر منبع آن، " داده‌های آماری ثبتی " (آمار ثبتی) نامیده می‌شود.

بنابراین کلیه سازمانهای عمومی با استفاده از اطلاعات کمی ثبت شده در جریان فعالیت روزمره سازمان، می‌توانند مبادرت به تولید آمار (آمار ثبتی) نمایند. به عبارت دیگر آمارهای ثبتی به مفهوم آماری است که در جریان فعالیت نهادها و ارگانهای اجرایی بدست می‌آید.

داده‌ها و آمارهای ثبتی دارای خصوصیاتی بشرح زیر می‌باشند:

۱ـ ثبت داده‌ها در سازمانهای عمومی به ترتیب خاص صورت می‌گیرد، مثلا " وقایع تولد و مرگ بعد از وقوع و ازدواج و طلاق در حین وقوع ثبت می‌شود و یا در مورد مادرات ثبت خصوصیات قبل از خروج از کشور به عمل می‌آید.

۲ـ جمع آوری داده‌های ثبتی فعالیتی است که از وظایف اصلی سازمان، عموما " قابل تفکیک نیست. به عبارت دیگر جمع آوری داده‌های ثبتی ذاتا " و ماهیتا " از فعالیت اصلی سازمان نمی‌تواند منفک گردد. مثلا " سازمان ثبت احوال کشور بخاطر حفظ هویت و حقوق افراد قانونا " موظف به ثبت وقایع چهارگانه می‌باشد.

۳ـ جمع آوری داده‌های ثبتی غالبا " صورت استمرار دارد و بخاطر همین خصوصیت همیشه می‌توان داده‌های ثبتی را بصورت تراکمی (تجمعی) در مقطع زمانی ارائه نمود و تحولات را بصورتی جامع در طی زمان مورد بررسی قرار داد.

۴ـ ثبت داده‌ها در سازمانهای عمومی چون با حقوق افراد، خواه حقیقی یا حقوقی ارتباط دارد، لذا در یا یگانی سازمانها اصولا " باید اطلاعات فرعی در رابطه با فعالیت سازمان وجود داشته باشد.

۵ـ آمارهای ثبتی در مکان و زمان محدود تقریبا " همیشه قابل ارائه هستند و این از سخترین ویژگیهای این نوع از آمارها است که برای بسیاری از برنامه ریزیها و تصمیم گیریهای محلی و موضعی مفید و ارزنده می‌باشد.

مثلا " ثبت احوال اصولا " بایدازاین توانایی برخوردارباشدکه وقایع حیاتی رابرای هرده درهرزمان ارائه دهدویاوزارت امور اقتصادی ودارایی بایدقادرباشدوصولی مالیات رابرای هرده ویاهرشهردرهرزمان عرضه نماید .

۶ـ آمارهای ثبتی محصول فرعی فعالیتهای اصلی سازمانهای مختلف کشورمی باشد .به عبارت دیگر هدف ازایجادسازمان (ماموریت سازمان) تهیه وتولیدآمارنیست بلکه سازمان دراجرای وظایف اصلی خویش وبه ضرورت ثبت ونگهداری هرواقعه برای حفظ حقوق فردی واجتماعی اطلاعات مربوطه راثبت می کندوسپس ازبرخی ازاین اطلاعات برای تولیدآمار استفاده می نماید .

۷ـ تولیدآمارهای ثبتی عموما " باهزینه اندك می تواندهمراه باشد .زیراجمع آوری داده های منفردآن درجریان فعالیت اصلی سازمان فراهم می شودومستلزم هزینه جداگانه ای نمی باشد .درحالیکه بخش عمده ای ازهزینه های تولیدآمارازطریق آمارگیری صرف جمع آوری اطلاعات می شود .

۸ـ ازویژگیهای مهم آمارهای ثبتی این است که می تواندمهمترین ابزارمدیریت برای مسئولین سازمان باشدو کارآیی ونافعیت آن سازمان رانشان دهد .به عنوان مثال بااستفاده ازآمارهای مربوطه می توان نیازبه نیروی انسانی راحتی برحسب توزیع جغرافیایی تاکوچکترین سطح موردنظرنشان داد .

۹ـ تولیدآمارهای ثبتی چون درمحدوده یك سازمان عمل می کندعموما " می تواندباانعطاف همراه باشدوبنابر تشخیص مسئولان دستگاه می توان نوع اطلاعات درخواستی راتغییرداد (البته تعدادی ازسئوالات منشاء قانونی دارندکه غیرقابل تغییرمی باشند) ونیازهای جدیدبرنامه ریزی ومدیریت رادرآن گنجانید .

درضرورت توجه بیشتربه آمارهای ثبتی ازطرف ارگانهای دولتی وواحدهای بخش خصوصی ،اشاره می شودکه برخلاف آمارگیریهای نمونه ای ویاسرشماریهاکه معمولا " تاخیرزمانی زیادی بین زمان مصاحبه وزمان وقوع اطلاع وجوددارد، درآمارهای ثبتی معمولا " این دوزمان برهم منطبق هستند .بنابراین ازقلم افتادن اطلاع ویاجوابنادرست دادن بعلت فاصله زمانی زیادوفراموشی ازبین می رودولذا حدودا" اطمینان آمارها بالااست وباعنایت به اینکه اینگونه آمارهاهزینه های خاصی نداشته وضمن فعالیت های جاری درموسسه باین امراقدام می شود، ازنظرصحت اطلاع

وکم‌هزینه‌بودن دستیابی‌بآن درمدیریت اقتصادی جایگاهی خاص دارد .

درسال ۱۳۰۲ ازبیست وظیفه تهیه آماری که برای اداره احصائیه و سجل واحوال بنابه قانون تعیین شده بیش ازنیمی ازآنها مربوط به اطلاعات آماری می شدکه به روش ثبتی توسط اداره مربوطه وسایرو زارتخانه‌های وقت تهیه می شد . تاسال ۱۳۱۸ به استثنای چندمورد آمارگیری درسطح محدودکلیه آمارهای تولیدشده به روش ثبتی تهیه شده است .

این روندتاسال ۱۳۲۵ به همین منوال ادامه داشته است . بعدازآن ۲۷ سال اززمان تاسیس اداره احصائیه در فروردین ۱۳۳۲ قانون آمارو سرشماری به تصویب رسیدواز این زمان بودکه اساس نظام آماری ایران دستخوش تغییر گردیدوآمارگیری ازطرحهای مختلف آماری درزمینه‌های جمعیتی، کشاورزی، صنعتی وغیره جزو وظایف اصلی سازمان مرکزی آمار قرارگرفت . همچنین هماهنگی بین سازمانهای تولیدکننده آمارهای ثبتی نیز ازجمله مسئولیتهای فرعی اداره آمار عمومی که ازسال ۱۳۳۴ شکل گرفته بودقرارداده شد .

درسال ۱۳۴۴ باتاسیس مرکزآمارایران مقررشدکلیه آمارگیریها توسط این مرکزانجام شودوموسسات دیگر تحت نظرمرکزآمارایران صرفا" به تهیه وتولیدآمارهای ثبتی درمحدوده فعالیت خودبپردازند . که البته درعمل چنین نشد .

درسال ۱۳۵۳ قانون مرکزآمارایران به تصویب رسیدوبرطبق آن یازده وظیفه برای آن مشخص گردید . که تنها یک مورد مربوط به ارتباط مرکزبا واحدهای آماری وزارتخانه‌ها وموسسات عمومی وخصوصی مربوط می شدوآنهم درسطح راهنمایی ودرحدامکان .

درتهیه آمارهای ثبتی توسط ارگانهای اجرایی تاکنون نظارت وهدایت مسئولانه‌ای ازطرف سازمان مرکزی اعمال نشده وواحدهای آماری موسسات مختلف بدون ارتباط با آن اقدام به تولیدآمارکرده‌اند . بدین جهت درمراحل مختلف تهیه آمارهای ثبتی دربارهای ازاین واحدها همواره شیوه‌های کهنه وغیرعلمی متداول بوده ودرمواردی ازابتدای شروع ثبت اطلاع درآن سازمانها تاکنون ، هیچگونه تغییری درنحوه کارصورت نگرفته است . این امرباعث ثبت نشدن

وبااستخراج نشدن بسیاری از اطلاعات ضروری گردیده است .

از جانب دیگر عدم وجود یک خط مشی کلی و اصولی تثبیت شده از مشکلات عمده کنونی می باشد ، زیرا به علت نبودن این خط مشی ، تولید آمارهای ثبتی تابع سلیقه های شخصی و تصمیمات فردی مسئولان اجرایی وقت می شود و چون دوران تصدی آنها نیز غالبا " کوتاه می باشد ، نتیجتا " جابجایی مسئولان موجب می گردد که جریان تولید آمار دستخوش تشتت شود و در نتیجه برنامه مدون و مشخصی برای تولید اینگونه آمارها وجود نداشته باشد .

با توجه به نکات فوق چنین استنتاج می شود که توسعه آمارهای ثبتی در کشور منوط به تحول اساسی در نظام آماری کشور جهت نظارت مستمر و دادن خط مشی های لازم به سازمانهای تولیدکننده آمار می باشد .

تحقیق در مورد آمارهای تولید شده توسط سازمانهای مختلف در شرایطی که اطلاعات کاملی از آنها در دسترس نیست کاری بس مشکل می باشد . برای رسیدن به چنین مقصودی نیاز به جمع آوری اطلاعات جامعی از مختصات این سازمانها ضرورت پیدا می کند . بواسطه اهمیت پرداختن به موضوع مورد بحث سعی شده است با حداقل اطلاعات آماری که در این زمینه وجود دارد فتح البابی بر آن شده باشد .

در حال حاضر کلیه واحدهای آماری ارگانهای دولتی با استفاده از داده های منفرد جمع آوری شده در جریان فعالیتهای جاری ارگان متبوع خود به تولید آمارهای ثبتی می پردازند . با بررسی فهرست آمارهای ثبتی تولید شده[1] مشاهده می گردد که در اکثر زمینه ها آمارهای ثبتی تهیه می شود .

معدودی از سازمانهای اجرایی وجود دارند که علاوه بر تولید آمارهای ثبتی به جمع آوری داده های منفرد از طریق آمارگیری نیز می پردازند (از جمله بانک مرکزی ، وزارت کار و امور اجتماعی و وزارت کشاورزی) . سهم آمارهای تولید شده توسط اینگونه واحدها در کل آمارهای تولید شده توسط واحدهای آماری موسسات بخش عمومی و خصوصی بسیار ناچیز و قابل صرف نظر کردن است . بدین جهت در بررسی زیر که از نتایج " طرح آمارگیری از امکانات آماری کشور " که در سال ۱۳۶۷ توسط مرکز آمار ایران به اجرا درآمد و تنها اطلاعات موجود در بر نگاه می است که تا حدودی به موضوع مورد بحث مربوط می شود ، هر جا به " تهیه آمار " اشاره شده ، منظور " تهیه آمارهای ثبتی " می باشد .

---

(۱) ـ فهرست آمارهای ثبتی تولید شده به تفکیک ۱۵ بخش در نشریه " فهرست اقلام آماری مورد نیاز سازمانهای دولتی و نهادهای انقلاب اسلامی " در مرکز آمار ایران قابل دسترسی است .

بررسی نتایج طرح آمارگیری از امکانات آماری کشور در سال ۱۳۶۷ نشان می دهد که از مجموع ۲۱۴۰ سازمان عمده عمومی و خصوصی مورد بررسی، ۷۵۳ سازمان در استان تهران و ۱۳۸۷ سازمان در سایر استانها واقع شده اند. از سازمانهای مستقر در استان تهران، ۶۶ سازمان تحت عنوان سازمانهای مرکزی و ۶۸۷ سازمان به عنوان سایر سازمانهای این استان طبقه بندی شده است.

نتایج حاصل همچنین نشان می دهد، در ۱۶۳۰ سازمان (۷۶/۲ درصد) فعالیت تهیه آمار انجام می شود و در ۵۱۰ سازمان (۲۳/۸ درصد)، این فعالیت صورت نمی گیرد. همچنین نتایج بررسی حاکی از آن است در حالیکه در کلیه سازمانهای مرکزی فعالیت تهیه آمار انجام می گیرد، در ۶۵/۱ درصد سایر سازمانهای استان تهران و ۸۰/۵ درصد سازمانهای واقع در سایر استانها فعالیت تهیه آمار صورت می گیرد. از مجموع سازمانهای تهیه کننده آمار، ۹۲/۴ درصد دارای مدیریت عمومی و ۷/۶ درصد دارای مدیریت خصوصی هستند.

الف- تعداد و درصد سازمانهای مورد بررسی برحسب وضع فعالیت تهیه آمار

شرح	جمع		فعالیت تهیه آمار انجام می شود		فعالیت تهیه آمار انجام نمی شود	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
کل کشور ۰۰۰۰	۲۱۴۰	۱۰۰	۱۶۳۰	۷۶/۲۰	۵۱۰	۲۳/۸
استان تهران ۰۰۰۰	۷۵۳	۱۰۰	۵۱۳	۶۸/۱	۲۴۰	۳۱/۹
سازمانهای مرکزی ۰۰۰۰	۶۶	۱۰۰	۶۶	۱۰۰	۰	۰
سایر سازمانها ۰۰۰۰	۶۸۷	۱۰۰	۴۴۷	۶۵/۱	۲۴۰	۳۴/۹
سایر استانها ۰۰۰۰	۱۳۸۷	۱۰۰	۱۱۱۷	۸۰/۵	۲۷۰	۱۹/۵

از کل سازمانهای تهیه کننده آمار، ۸۲۳ سازمان (۵۱/۱ درصد) دارای واحد آمار هستند و ۲۹۷ سازمان

(۴۸/۹ درصد) بدون داشتن واحد آمار به تهیه و جمع آوری اطلاعات آماری اقدام می نمایند ، درصدهای مزبور برای سازمانهای مرکزی به ترتیب برابر ۹۰/۹ و ۹/۱ ، برای سایر سازمانهای استان تهران به ترتیب برابر ۸/۹ و ۳۱/۱ و برای سازمانهای واقع در سایر استانها به ترتیب برابر ۴۱/۶ و ۵۸/۴ است .

ب ـ تعداد و درصد سازمانهای تهیه کننده آمار برحسب سازمانهای دارای واحد آمار و بدون واحد آمار

شرح	جمع		سازمانهای دارای واحد آمار		سازمانهای بدون واحد آمار	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
کل کشور	۱۶۳۰	۱۰۰	۸۳۳	۵۱/۱	۷۹۷	۴۸/۹
استان تهران	۵۱۳	۱۰۰	۳۶۸	۷۱/۷	۱۴۵	۲۸/۳
سازمانهای مرکزی	۶۶	۱۰۰	۶۰	۹۰/۹	۶	۹/۱
سایر سازمانها	۴۴۷	۱۰۰	۳۰۸	۶۸/۹	۱۳۹	۳۱/۱
سایر استانها	۱۱۱۷	۱۰۰	۴۶۵	۴۱/۶	۶۵۲	۵۸/۴

براساس اطلاعات بدست آمده ، فعالیت اصلی ۶۶/۵ درصد از مجموع سازمانهای تهیه کننده آمار در زمینه فعالیت خدمات عمومی ، اجتماعی و شخصی ، ۲۳/۹ درصد در زمینه فعالیت صنعت و ۱۰ درصد در سایر زمینه های فعالیتهای اقتصادی گزارش شده است .

مقایسه سازمانها برحسب فعالیتهای اقتصادی از نظر وضعیت تهیه آمار نشان می دهد که فعالیت اصلی بیشترین درصد (۶۶/۲) سازمانهایی که فعالیت تهیه آمار در آنها صورت نمی گیرد مربوط به فعالیت عمده فروشی ، خرده فروشی ، مرمت وسائل و منازل دار است .

ج ـ تعداد و درصد سازمانهای مورد بررسی بر حسب فعالیت اصلی و وضعیت فعالیت تهیه آمار

فعالیت اصلی	جمع		فعالیت تهیه آمار انجام میشود		فعالیت تهیه آمار انجام نمیشود	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
کل	۱۲۴۰	۱۰۰	۱۶۳۰	۷۶/۲	۵۱۰	۲۳/۸
کشاورزی، دامپروری، شکار و ماهیگیری ...	۷	۱۰۰	۶	۸۵/۷	۱	۱۴/۳
استخراج معادن	۴	۱۰۰	۴	۱۰۰	۰	۰
صنعت	۶۱۲	۱۰۰	۳۸۹	۶۳/۶	۲۲۳	۳۶/۴
آب و برق و گاز	۳۷	۱۰۰	۲۹	۷۸/۴	۸	۲۱/۶
ساختمان	۴۰	۱۰۰	۳۰	۷۵/۰	۱۰	۲۵/۰
عمده فروشی، خرده فروشی، رستوران و هتلداری ...	۱۸	۱۰۰	۶	۳۳/۳	۱۲	۶۶/۷
حمل و نقل، ارتباطات و انبارداری	۶۲	۱۰۰	۴۹	۷۹/۰	۱۳	۲۱/۰
خدمات مالی، بیمه، ملکی و حقوقی	۳۱	۱۰۰	۲۶	۸۳/۹	۵	۱۶/۱
خدمات عمومی، اجتماعی، شخصی	۱۳۱۳	۱۰۰	۱۰۷۸	۸۲/۱	۲۳۵	۱۷/۹
فعالیتهای غیر طبقه بندی	۱۶	۱۰۰	۱۳	۸۱/۳	۰۳	۱۸/۷

" مقایسه سازمانها بر حسب فعالیتهای اقتصادی از نظر دارا بودن واحد آمار نشان می دهد که بیشترین درصد سازمانهای دارای واحد آمار (۱۰۰ درصد) مربوط به سازمانهایی است که فعالیت اصلی آنها " استخراج معادن " و " آب و برق و گاز " و کمترین آن با ۳۹/۸ درصد مربوط به سازمانهایی است که فعالیت اصلی آنها خدمات عمومی، اجتماعی و شخصی است .

د ـ تعداد و درصد سازمانهای تهیه‌کننده آمار بر حسب فعالیت

فعالیت اصلی	جمع		سازمانهای دارای واحد آمار		سازمانهای بدون واحد آمار	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
کل	۱۶۳۰	۱۰۰	۸۳۳	۵۱/۱	۷۹۷	۴۸/۹
کشاورزی، دامپروری، شکار و جنگلداری	۶	۱۰۰	۴	۶۶/۷	۲	۳۳/۳
استخراج معادن	۴	۱۰۰	۴	۱۰۰	۰	۰
صنعت	۳۸۹	۱۰۰	۲۶۶	۶۸/۴	۱۲۳	۳۱/۶
آب و برق و گاز	۲۹	۱۰۰	۲۹	۱۰۰	۰	
ساختمان	۳۰	۱۰۰	۲۶	۸۶/۷	۴	۱۳/۳
عمده‌فروشی، خرده‌فروشی، رستوران و هتلداری	۶	۱۰۰	۴	۶۶/۷	۲	۳۳/۳
حمل و نقل، ارتباطات و انبارداری	۴۹	۱۰۰	۴۰	۸۱/۶	۹	۱۸/۴
خدمات مالی، بیمه، مستغلی و حقوقی	۲۶	۱۰۰	۲۳	۸۸/۵	۳	۱۱/۵
خدمات عمومی، اجتماعی و شخصی	۱۰۷۸	۱۰۰	۴۲۹	۳۹/۸	۶۴۹	۶۰/۲
فعالیتهای غیر قابل طبقه‌بندی	۱۳	۱۰۰	۸	۶۱/۵	۵	۳۸/۵

از مجموع فعالیتهای آماری خاتمه یافته[1] سازمانها، ۵۹/۵ درصد مربوط به فعالیتهای آماری مستمر و ۴۰/۵ در صد مربوط به فعالیتهای آماری موردی برده است. بالاترین درصد در مورد فعالیتهای موردی متعلق به سازمانهای مرکزی بوده است.

---

(۱) ـ منظور از خاتمه یافته، فعالیتی است که نتایج آن تا پایان سال ۱۳۶۶ تهیه شده و در دسترس قرار گرفته است.

هـ ـ تعداد و درصد فعالیتهای آماری خاتمه یافته سازمانهای تهیه کننده آمار بر حسب نحوه انجام فعالیت :

۱۳۶۲ـ۶۶

شرح	جمع		مستمر		موردی	
	تعداد	درصد	تعداد	درصد	تعداد	درصد
کل کشور	۱۳۳۹	۱۰۰	۷۹۷	۵۹/۵	۵۴۲	۴۰/۵
استان تهران	۸۹۲	۱۰۰	۵۷۰	۶۳/۹	۳۲۲	۳۶/۱
سازمانهای مرکزی	۵۶۵	۱۰۰	۲۶۳	۴۶/۵	۳۰۲	۵۳/۵
سایر سازمانها	۳۲۷	۱۰۰	۳۰۷	۹۳/۹	۲۰	۶/۱
سایر استانها	۴۴۷	۱۰۰	۲۲۷	۵۰/۸	۲۲۰	۴۹/۲

بررسی فعالیتهای آماری نشان می دهد که ۲۳/۵ درصد از کل فعالیتهای آماری، در زمینه صنعت صورت گرفته است . جمعیت و نیروی انسانی با ۱۶/۸ درصد و آموزش و پرورش و فرهنگ با ۹/۷ درصد، بهداشت، درمان و خدمات رفاهی با ۸/۴ درصد، کشاورزی با ۷/۱ درصد، خرده فروشی، عمده فروشی، رستوران و هتلداری با ۵/۹ درصد در رتبه های بعدی قرار گرفته است .

و ـ تعداد و درصد فعالیتهای آماری برحسب بخش فعالیت آماری: ۶۶ـ ۱۳۶۴

بخش فعالیت آماری	تعداد	درصد
جمع	۱۸۵۶	۱۰۰
جمعیت و نیروی انسانی	۳۱۲	۱۶/۸
آموزش و پرورش و فرهنگ	۱۸۱	۹/۷
بهداشت، درمان و خدمات رفاهی	۱۵۶	۸/۴
کشاورزی	۱۳۳	۷/۱
استخراج معادن	۴۶	۲/۵
صنعت	۴۲۷	۲۳/۵
آب و برق و گاز	۸۴	۴/۵
ساختمان و مسکن	۹۵	۵/۱
خرده‌فروشی، عمده‌فروشی، رستوران، هتلداری و تجارت خارجی	۱۱۰	۵/۹
حمل و نقل، ارتباطات و انبارداری	۸۶	۴/۶
موسسات پولی، مالی، خدمات و ترازنامه‌ها	۱۰۱	۵/۴
حقوق و قضا	۲۵	۱/۳
فعالیتهای سیاسی و ارتباطات بین‌المللی	۳	۰/۲
مصرف خانوار	۱۵	۰/۸
حسابهای ملی و دولت	۶۲	۳/۹

تعدیل شاخص ازدحام با استفاده از ضریب نابرابری جینی برای استانهای کشور

فریده دیبائی

مرکز آمار ایران

خلاصه

شاخص " متوسط نفر در اتاق " یکی از معیارهای اندازه گیری کافی بودن مسکن است این شاخص که نسبت تعداد افراد به تعداد اتاقها را نشان میدهد در شکل فعلی خود دارای نارسائیهائی است از جمله اینکه عامل بعد و ساخت خانوار در آن نادیده گرفته میشود. در این مقاله با استفاده از ضریب جینی (شاخص نابرابری توزیع درآمد) ، شاخص مورد بحث تعدیل شده و با دو فرض متفاوت با استفاده از اطلاعات حاصل از سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ ، برای مناطق شهری استانهای کشور محاسبه شده است . در حالت اول ، فرض شده است که توزیع متعادل اتاقها بین افراد خانوار زمانی حاصل میشود که برای هر نفر یک اتاق در نظر گرفته شود. در حالت دوم تعداد اتاقهای مورد نیاز خانوار مورد توجه قرار گرفته و فرض شده است که با افزایش بعد خانوار تعداد اتاقها مورد نیاز خانوار متناسب با آن افزایش نمی یابد بلکه شدت افزایش آن کمتر از شدت افزایش بعد خانوار است . پس از انجام محاسبات برای دو حالت فوق نتایج بدست آمده نشان میدهد که در حالت اول شاخص متوسط نفر در اتاق برای مناطق شهری کل کشور ۱/۶ و شاخص تعدیل شده آن ۲/۴ است و در حالت دوم بترتیب ۱/۳۸ و ۱/۹ است .

* * *

مقدمه

بهبود کیفیت زندگی مردم یک کشور یکی از هدفهای اساسی است که در برنامه های توسعه اقتصادی و اجتماعی کشورها منظور میشود. تامین مسکن "کافی" همراه با بهبود شرایط تغذیه، پوشاک، آموزش و پرورش، بهداشت، ایجاد فرصت برای اشتغال، برخورداری از ساعات فراغت، امکان دسترسی به تسهیلات زیر بنائی و ... از عوامل تشکیل دهنده بهبود کیفیت زندگی است. مسکن "کافی" در اکثر موارد مسکن فاقد ازدحام بیش از حد تعریف می‌شود. بدین ترتیب شاخص ازدحام موسوم به "متوسط نفر در اتاق" یکی از معیارهای اندازه گیری کافی بودن مسکن قلمداد می‌شود که بعنوان یکی از ابزارهای کمی ارزیابی سیاست مسکن دولتها بکار میرود. شاخص ازدحام یا شاخص متوسط نفر در اتاق نسبت تعداد افراد به تعداد اتاقهای در اختیار آنان را نشان میدهد ونشانگر متوسط تعداد افراد ساکن در هر اتاق میباشد. محاسبه شاخص مزبور برای دوره های مختلف زمانی یک جامعه آثار سیاستهای مختلف بهبود شرایط کیفی مسکن را نشان میدهند. با این حال شاخص ازدحام در شکل فعلی خود دارای نارسائیهائی است. ازجمله اینکه در شاخص مورد بحث توزیع اتاقهای موجود بین خانوارها متناسب با بعد خانوارها در نظر گرفته شده است و درنتیجه عامل بعد و ساخت خانوار و به تبع آن تعداد اتاقهای مورد نیاز و واقعی خانوار نادیده گرفته میشود. یکی از روشهای تعدیل نارسائی مزبور استفاده از ضریب تبدیل جینی است که شاخص متوسط نفر در اتاق را با استفاده از آن اصلاح می کنند.

در این مقاله ابتدا نارسائیهای شاخص مورد نظر مورد بحث قرار گرفته سپس روش تعدیل و اصلاح آن ارائه خواهد شد. بمنظور نشان دادن نتیجه تعدیل و اصلاح شاخص فوق یک کار تجربی در مورد ایران انجام گرفته است. باین ترتیب که با استفاده از داده های آماری سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ ـ شاخص "متوسط نفر در اتاق" قبل و بعد از تعدیل در دو حالت مختلف و با دو فرض متفاوت ـ برای مناطق شهری کل کشور و استانها ی مختلف محاسبه شده است.

شاخص متوسط نفر در اتاق

شاخص "متوسط نفر در اتاق" بعنوان یکی ازمعیارهای اندازه‌گیری کافی بودن مسکن خانوارها بکار میرود و یکی از ابزارهای کمی سیاست گذاری دولتها در زمینه مسکن است . این شاخص که یک میانگین ساده است بصورت $I = \frac{P}{S}$ تعریف میشود که در آن P مجموع تعداد افراد ساکن در کل واحدهای مسکونی و S مجموع تعداد اتاقهای موجود در کل واحدهای مسکونی است .

این شاخص نشان دهنده تعداد افرادیست که بطور متوسط در هر اتاق زندگی میکنند . بعبارت دیگر این شاخص حاکی از آنست که کلیه اتاقهای موجود متناسب با بعد خانوار بین خانوارها توزیع شده است و کلیه خانوارهای جامعه اعم از خانوارهای یک نفره ، دو نفره ، سه نفره ، ... و ده نفره متناسب با تعداد افراد خود از اتاقهای موجود استفاده می کنند . لیکن واقعیت امر غیر از اینست . بعنوان مثال متوسط افرادی که در خانوارهای یک نفره در هر اتاق بسر می برند و متوسط افرادی که در خانوارهای ده نفره در هر اتاق بسر میبرند یکسان نیست و خانوارهای با ابعاد مختلف از امکانات یکسان برخوردار نیستند در این نابرابری از جنبه های دیگر نیز وجود دارد . توجه به این اختلاف ، نارسایی شاخص مورد بحث را برای ارزیابی کیفیت مسکن خانوارها نمایان میسازد . بایستن ترتیب که این شاخص عامل بعد خانوار را نادیده میگیرد . لذا کوششهائی برای تعدیل این شاخص و اصلاح آن انجام شده است از آن جمله رنزوریچی استاد دانشگاه فلورانس ایتالیا در مقاله‌ای[1] استفاده از ضریب نابرابری جینی را برای اصلاح این شاخص پیشنهاد می کند .

تعدیل شاخص

بمنظور مشاهدهٔ حاصل اصلاح وتعدیل شاخص مورد نظر ، با استفاده از داده های آماری سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ مرکز آمار ایران شاخص "متوسط نفر در اتاق" در دو حالت مختلف با دو فرض متفاوت ، با استفاده از ضریب نابرابری جینی "که معمولا" برای اندازه‌گیری میزان نابرابری توزیع درآمد بین افراد یک جامعه بکار میرود ، قبل از تعدیل و بعد از آن برای مناطق شهری کل کشور و استانهای مختلف محاسبه شده . هریک از حالت های مورد مطالعه ذیلا" مورد تجزیه وتحلیل قرار میگیرد .

فرض اول

در حالت اول فرض بر اینستکه توزیع عادلانه اتاقها بین افراد جامعه زمانی حاصل میشود که اتاق های موجود در جامعه بطور مساوی بین افراد جامعه تقسیم شده باشد .

در شکل شماره ۱ که برای مناطق شهری کل کشور رسم شده است فراوانی تجمعی افراد در روی محور افقی و فراوانی تجمعی اتاقهای در اختیار خانوار روی محور عمودی نشان داده شده است و منحنی لورنز

---

(۱) رجوع شود به مقاله زیر در ماخذ شماره ۱ Observations On The Significance And Use of the Indices of Crowding By Mr. Renzo Ricci, University of Florence, Italy.

توزیع اتاق واحدهای مسکونی بین افراد کشور (منحنی AC) رسم شده است .

خط AC توزیع متعادل اتاقهای موجود بین افراد است . بعبارت دیگر این خط نشانگر توزیع عادلانه اتاقهای موجود بین افراد کشور است و نشان دهنده این موضوع است که تعداد اتاقهای دراختیار هریك ازافراد درخانوارهای بابعد و ساخت متفاوت مساوی است و به بیان دیگر در این حالت تعداد افراد ساکن در هر اتاق درکلیه خانوارها (خانوارها بابعدهای مختلف) مساویست . خط AD خط توزیع متعادل ایده آل است با فرض اینکه هر نفر یك اتاق در اختیار داشته باشد . در اینصورت شاخص متوسط نفر در اتاق (I) عبارت خواهد بود از نسبت مساحت زیر خط AD به مساحت زیر خط AC زیرا :

$$I = \frac{S_{ABD}}{S_{ABC}} = \frac{\frac{1}{2} AB \cdot BD}{\frac{1}{2} AB \cdot BC} = \frac{BD}{BC} \qquad (1)$$

بنابراین خط AC نشان دهنده شاخص مورد بحث (قبل از تعدیل) است .

I =۱/۶۰۰ برای مناطق شهری کل کشور محاسبه شده است . شاخص تعدیل شده متوسط نفر در اتاق بر اساس ضریب نابرابری جینی نیز از رابطه زیر بدست خواهد آمد :

$$IR = \frac{S_{ABD}}{\text{ساحت زیرمنحنی}_{AC}} = \frac{S_{ABD}}{S_{ABC} - \varphi} = \frac{\frac{S_{ABD}}{S_{ABC}}}{1 - \frac{\varphi}{S_{ABC}}} \qquad (2)$$

φ مساحت بین منحنی لورنز توزیع اتاق بین افراد جامعه وخط توزیع متعادل اتاق در شرایط موجود جامعه است و نشان دهنده نابرابری توزیع اتاقها بین افراد جامعه است . در اینصورت نسبت میزان نابرابری توزیع اتاقها بین افراد جامعه (ضریب جینی) خواهد بود .

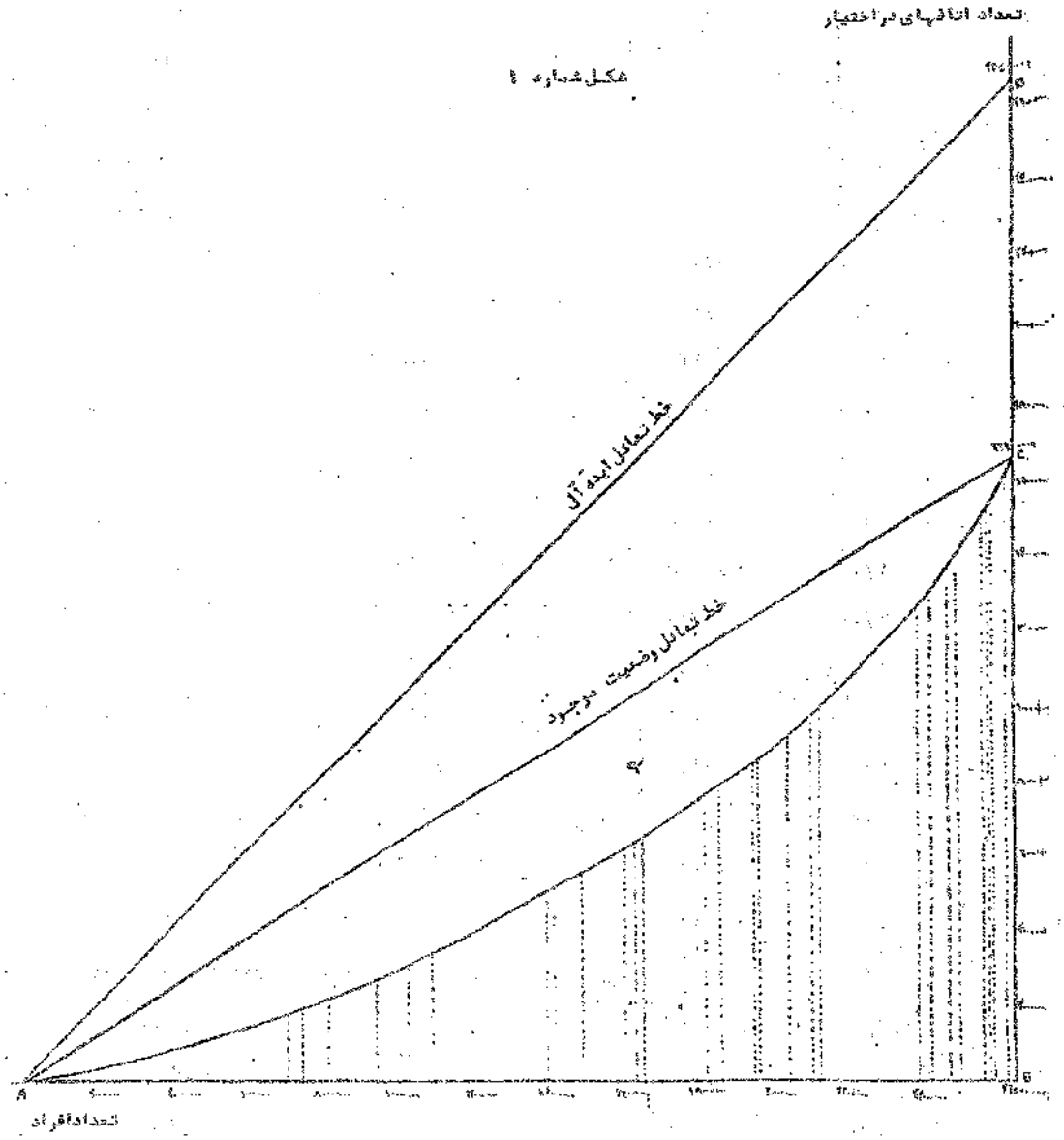
$$R = \frac{\varphi}{S_{ABC}} = \quad ۰/۳۳۲ \qquad (3)$$

بعبارت دیگر سطح φ نشان دهنده فاصله موجود بین واقعیت جامعه و شاخص مورد بحث قبل از تعدیل است . با توجه به رابطه (۱) و(۲) و (۳) شاخص تعدیل شده IR بصورت زیرخواهد بود :

$$IR = \frac{I}{1 - R} = ۱/۳۹۲$$

با استفاده ازاطلاعات مربوط به سرشماری سال ۱۳۶۵ هرکدام ازاقلام I ، R و IR برای کل کشور واستانهای مختلف محاسبه شده و درجدول شماره ۱ درج شده است .

شکل شماره ۱


تعداد اتاقهای در اختیار
خط تمایل ایده آل
خط تمایل وضعیت موجود
تعداد افراد

جدول شماره ۱ ـ شاخص‌های ازدحام درنقاط شهری استان‌های کشور در سال ۱۳۶۵
(فرض اول)

استان	متوسط نفر در اتاق	ضریب جینی	شاخص تعدیل شده
کل کشور	۱/۶۰۰	۰/۳۲۲	۲/۳۹۲
تهران	۱/۵۱۲	۰/۳۳۱	۲/۲۶۱
مرکزی	۱/۴۹۹	۰/۳۱۲	۲/۱۸۴
گیلان	۱/۷۲۴	۰/۳۱۳	۲/۵۱۱
مازندران	۱/۵۵۸	۰/۳۰۱	۲/۲۲۰
آذربایجان شرقی	۱/۹۲۸	۰/۳۴۵	۲/۹۴۵
آذربایجان غربی	۱/۸۵۰	۰/۳۲۰	۲/۷۲۰
باختران	۱/۹۷۶	۰/۳۳۵	۲/۹۷۱
خوزستان	۲/۰۱۸	۰/۳۲۵	۲/۹۹۱
فارس	۱/۴۷۱	۰/۳۱۴	۲/۱۴۴
کرمان	۱/۳۹۰	۰/۳۱۶	۲/۰۳
خراسان	۱/۴۶۵	۰/۳۱۶	۲/۱۷۵
اصفهان	۱/۴۰۴	۰/۳۱۹	۲/۰۶۱
سیستان و بلوچستان	۱/۹۴۵	۰/۳۵۳	۳/۰۰۸
کردستان	۲/۰۳۰	۰/۳۱۱	۲/۹۴۴
همدان	۱/۶۱۲	۰/۳۱۰	۲/۳۳۵
چهارمحال و بختیاری	۱/۶۴۰	۰/۳۱۴	۲/۳۵۹
لرستان	۲/۰۴۶	۰/۳۲۳	۳/۰۲۰
ایلام	۲/۲۱۸	۰/۳۱۰	۲/۲۱۲
کهگیلویه و بویراحمد	۱/۸۹۰	۰/۳۰۱	۲/۷۰۴
بوشهر	۱/۶۸۸	۰/۳۲۰	۲/۴۸۱
زنجان	۱/۶۸۴	۰/۳۲۱	۲/۴۷۹
سمنان	۱/۴۱۲	۰/۲۹۸	۱/۷۲۶
یزد	۱/۲۷۷	۰/۳۱۱	۱/۸۵۵
هرمزگان	۱/۷۹۲	۰/۳۲۱	۲/۶۲۱

در حالت اول فسوق بر این بود که در وضعیت توزیع متعادل ایده آل برای هر نفر یک اتاق در نظر گرفته شده است . لیکن در واقع با افزایش بعد خانوار نیاز سرانه به اتاق کاهش می یابد . بعنوان مثال تعداد اتاقهای مورد نیاز خانوار ده نفره لزوما " ده برابر تعداد اتاقهای مورد نیاز خانوار یک نفره نیست . بنابراین برای حالت توزیع متعادل ایده آل باید معیارهای واقعی تری را جستجو کرد که از آن جمله معیار متناسب بودن تعداد اتاقهای مورد نیاز خانوار بر اساس بعد و ساخت خانوار است در این معیار برحسب اینکه خانوار هسته ای باشد یا گسترده و با توجه جنس و سن فرزندان خانوار و در نظر گرفتن تعداد افراد خانوار تعداد اتاقهای مورد نیاز خانوار تغییر می یابد .

در این مرحله از بررسی شاخص متوسط نفر در اتاق با در نظر گرفتن تعداد اتاقهای مورد نیاز خانوارها بر مبنای معیار فوق الذکر اصلاح شده است باین ترتیب که در رابطه $I = \frac{P}{S}$ بجای تعداد افراد خانوار (P) ، تعداد اتاقهای مورد نیاز خانوار (P') بکار برده شده است . در اینصورت شاخص متوسط نفر در اتاق بصورت $I' = \frac{P'}{S}$ خواهد بود که در آن P' تعداد کل اتاقهای مورد نیاز خانوارها و S تعداد کل اتاقهای در اختیار خانوارها خواهد بود . علیرغم توضیح فسوق ، بعلت محدودیت اطلاعات در زمینه تعداد اتاقهای مورد نیاز برحسب ساخت خانوار فقط تعداد اتاقهای مورد نیاز خانوار براساس بعد خانوار مورد توجه قرار گرفته و شاخص IR اصلاح شده است . یکی از معیارهای پیشنهاد شده در این زمینه که مطابق آن تعداد اتاقهای مورد نیاز خانوار یک نفره ۲ و تعداد اتاقهای مورد نیاز خانوار ۲ نفره ۳ و ... به ترتیب جدول زیر است ، در این بررسی مورد استفاده قرار گرفته است . همانطور که در این جدول مشاهده میشود لزوما " تعداد اتاقهای مورد نیاز خانوار با تعداد افراد خانوار مساوی نیست و با افزایش تعداد افراد خانوار قدر مطلق افزایش تعداد اتاقها مورد نیاز خانوار به تدریج کمتر و کمتر میشود .

بعد خانوار	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
تعداد اتاقهای مورد نیاز	۲	۳	۳	۴	۴	۵	۵	۶	۶	۷	۷	۸	۸	۹	۹

منحنی های شکل شماره ۲ با استفاده از معیار فسوق و استفاده از اطلاعات سرشماری عمومی نفوس و مسکن سال ۱۳۴۵ برای مناطق شهری کل کشور رسم شده است . در این شکل تعداد اتاقهای مورد نیاز ، روی محور افقی و تعداد اتاقهای در اختیار خانوار روی محور عمودی نشان داده شده است . خط OA خط متعادل ایده آل است با فرض اینکه تعداد اتاقهای در اختیار هر خانوار پاسخگوی نیازهای خانوار به اتاق باشد . بعبارت دیگر در این حالت تعادل تعداد اتاقهای در اختیار خانوار برابر با

تعداد اتاقهای مورد نیاز آن است . خط $A'C'$ خط تعادل توزیع اتاقهـــا در وضعیت موجود است . بعبارت دیگر این خط حاکـــی از اینست که تعداد اتاقهای موجود در جامعه متناسب باتعداد اتاقها ی مورد نیاز هرخانوار توزیع شده است . لیکن درحال حاضر نحوه توزیع اتاقهای موجود جامعه بر حســـب اتاقهای مورد نیاز خانوارها بصورت منحنی $A'C'$ است . فاصله این منحنی تا خط تعادل در وضعیـــت موجود و خط تعادل اینده آن نشان دهنده شکاف موجود برای رسیدن به حالت های تعادل در زمینـــــه توزیـــع اتاقهـــا اســـت .

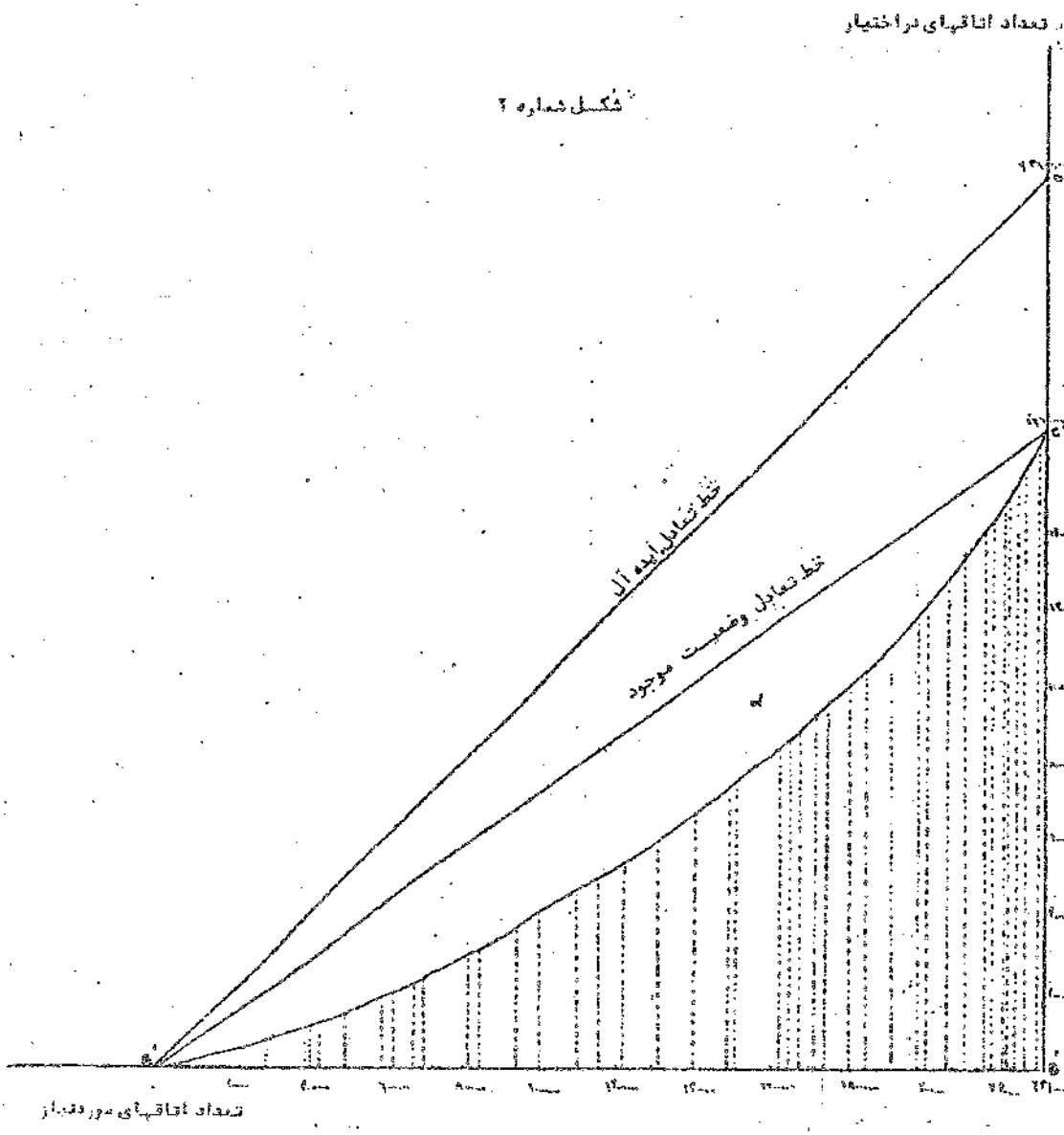
بر مبنای آنچـــه گفته شد در این مرحلـــه از بررسی شاخصهـــای I' ، R' و IR برای کـــــل کشـــور و در استانهای مختلف نقاط شهری محاسبه شده است که در جدول شماره ۲ مندرج است .

ازجمله کشورهائی‌که به محاسبه شاخصهای مورد بحث پرداخته اند کشور ایتالیا است ، شاخصهای محاسبه شده در کشور ایتالیا و چهار ایالـــت آن درجدول زیر درج شده‌است .

شاخصهای ازدحام در ایتالیا و چهار ایالت آن کشـــور درسال ۱۹۶۱

شاخصها / ایالت	I	I'	IR	$I'R$
تورینـــن	۰/۹۲	۰/۹۹	۱/۳۷	۱/۳۱
ژنـــوا	۰/۷۲	۰/۷۴	۱/۱۱	۱/۰۳
رم	۱/۰۱	۰/۹۷	۱/۵۵	۱/۳۶
باری	۱/۵۰	۱/۲۸	۲/۵۵	۲/۱۶
ایتالیا	۱/۰۴	۱/۰۰	۱/۶۶	۱/۴۵

شکل شماره ۲


تعداد اتاقهای در اختیار
خط تعادل ایده آل
خط تعادل وضعیت موجود
تعداد اتاقهای مورد نیاز

جدول شماره ۴ ـ شاخص های ازدحام در نقاط شهری استان های کشور در سال ۱۳۶۵
(فرض دوم)

استانها	متوسط نفر در اتاق	ضریب جینی	شاخص تعدیل شده
کل کشور	۱/۳۸۲	۰/۲۹۲	۱/۹۵۵
تهران	۱/۳۴۹	۰/۲۸۷	۱/۸۹۴
مرکزی	۱/۳۱۰	۰/۲۶۸	۱/۷۹۰
گیلان	۱/۵۲۸	۰/۲۷۹	۲/۱۱۸
مازندران	۱/۳۸۸	۰/۲۵۶	۱/۸۶۶
آذربایجان شرقی	۱/۶۴۶	۰/۳۰۲	۲/۲۵۸
آذربایجان غربی	۱/۵۷۱	۰/۲۸۲	۲/۱۸۹
باختران	۱/۶۸۱	۰/۳۰۵	۲/۴۱۸
خوزستان	۱/۶۵۴	۰/۲۹۴	۲/۳۲۹
فارس	۱/۲۵۶	۰/۲۷۹	۱/۷۴۲
کرمان	۱/۱۸۵	۰/۲۷۱	۱/۶۲۵
خراسان	۱/۲۷۱	۰/۲۸۶	۱/۷۸۰
اصفهان	۱/۲۳۲	۰/۲۸۳	۱/۷۱۸
سیستان و بلوچستان	۱/۶۰۲	۰/۳۲۳	۲/۳۴۹
کردستان	۱/۷۱۷	۰/۲۷۰	۲/۲۵۱
همدان	۱/۴۰۳	۰/۲۶۹	۱/۹۱۹
چهارمحال و بختیاری	۱/۳۴۲	۰/۲۷۹	۱/۸۸۸
لرستان	۱/۷۰۲	۰/۲۹۰	۲/۴۰۲
ایلام	۱/۸۱۵	۰/۲۸۶	۲/۵۲۲
کهگیلویه و بویراحمد	۱/۵۶۱	۰/۲۷۶	۲/۱۵۶
بوشهر	۱/۴۰۰	۰/۲۸۶	۱/۹۶۱
زنجان	۱/۴۲۸	۰/۲۸۰	۱/۹۸۸
سمنان	۱/۰۸۵	۰/۲۵۹	۱/۱۶۵
یزد	۱/۱۲۲	۱/۲۵۸	۱/۵۱۳
هرمزگان	۱/۵۱۲	۰/۲۹۱	۲/۱۲۵

بررسی ضرایب حاصل از محاسبات انجام شده بر مبنای فروض ۱ و ۲ نشانگر آنست که شاخص متوسط نفر در اتاق ، ضریب جینی و شاخص تعدیل شده ازدحام برای کل مناطق شهری ایران به ترتیب ۱/۴۰۰، ۰/۳۳۲ و ۲/۳۹۴ در فرض اول و ۱/۳۸۲ ، ۰/۲۹۲ و ۱/۱۵۵ در فرض دوم است .

به این ترتیب شاخص مورد بحث در حالت اول پس از تعدیل ، ۵۰ درصد و در حالت دوم ۴۳ درصد افزایش می یابد و شاید پس از دخالت دادن عامل ساخت وترکیب خانوار، بعضی تعداد اتاقهای مورد نیاز خانوار بر حسب ترکیب خانوار، که بمراتب فرض واقعی تری است ، درصد افزایش شاخص فوق از این مقدار هم بیشتر شود . بنابراین لازمست در مورد استفاده از شاخص مورد بحث در سیاست گذاری های مسکن با احتیاط عمل شود. زیرا چنین استنباط میشود که میانگین های ساده شاخصهای مقایسه ای برای نشان دادن واقعیت های جامعه نیستند .

منابع مورد استفاده

۱ـ HOUSING REQIRMENTS AND DEMAND : CURRENT METHODS OF ASSESSMENT AND PROBLEMS OF ESTIMATION. UN

و ترجمه مقاله : ملاحظاتی درباب اهمیت و کاربرد شاخصهای ازدحام . ترجمه فاطمه تقی زاده سازمان برنامه و بودجه استان فارس

۲ـ نشریات سرشماری عمومی نفوس ومسکن کل کشور و ۲۴ استان سال ۱۳۶۵ ـ مرکز آمار ایران .

۳ـ ابعاد اجتماعی مسکن ـ مترجم عباس منبر ـ سازمان برنامه وبودجه .

۴ ـ شاخصهای نابرابری و روش محاسبه آنها ـ ابوالفضل ابوالفتحی قمی ـ مدیریت آمارهای مالی و محاسبات ملی ـ مرکز آمار ایران .

ضریب همبستگی داخلی

توفیق رجائی راد
مرکز آمار ایران

مقدمه

درنظریه نمونه گیری میزان کارایی هرطرح نمونه گیری را با نمونه گیری تصادفی ساده ارزیابی می‌کنند . عامل ارزیابی ، نسبت واریانس برآوردهای حاصل از دوطرح نمونه گیری‌است . این نسبت را " اثر طرح[1] " می‌گویند و عکس اثر طرح شاخص ارزیابی میزان کارایی است . در نمونه گیری سیستماتیک وخوشه ای بخاطر گروه بندی واحدهای نمونه ای ، واریانس برآورد تحت تاثیر قرار می گیرد واین اثر از همبستگی بین واحدهای داخل گروه ناشی میشود . اثر طرح دراین حالت تابعی از میزان این همبستگی وتعداد واحدهای نمونه ای داخل گروه است . این نوع همبستگی را " همبستگی داخلی " گفته وعامل تعیین مقدار آن را " ضریب همبستگی داخلی[2] " می‌گویند . ضریب همبستگی داخلی علاوه بر تعیین میزان کارائی در محاسبه " حجم‌نمونه[3] " نیز نقش عمده دارد .

درکارهای عملی (نمونه‌گیری‌کاربردی) معمولا " محاسبه ضریب همبستگی بدلیل پیچیدگی ودردسترس نبودن اطلاعات موردنیازباشکلاتی‌همراه‌است . دراین مقاله ضمن‌ارائه روش محاسبه " ضریب همبستگی داخلی " ، مقادیری که درچند زمینه برای آن بدست آمده‌است نیز معرفی میشود .

ضریب همبستگی داخلی

همانطورکه درمقدمه اشاره شد عکس " اثرطرح " میزان کارائی را اندازه می‌گیرد . اگر برآورد

1- Design Effect

2-Intracluster (Intraclass) Correlation Coefficient

3- Sample Size

میانگین مورد نظر باشد در اینصورت در طرح های نمونه گیری سیستماتیک وخوشه ای اثر طرح عبارتست از :

$$Deff = V(\bar{x}_{Sy}) / V(\bar{x}_R) \quad \text{یا} \quad V(\bar{x}_{cl}) / V(\bar{x}_R)$$

که در آنها اندیس‌های Sy ، R و cl به ترتیب معرف " سیستماتیک " ، " تصادفی ساده " و " خوشه ای " است . در نمونه گیری سیستماتیک بخاطر فرض " توزیع تصادفی واحدهای جامعه " و به عبارت دیگر توزیع کاملا " تصادفی واحدها در " چارچوب [1] " مقدار $Deff$ برابر با یک فرض میشود . البته این فرض دوراز واقعیت نیست زیرا : اگر ترتیب قرار گرفتن واحدها در چارچوب از " الگوی خاصی " پیروی کند در اینصورت متخصص نمونه گیری از بکارگیری روش سیستماتیک اجتناب می‌کند واین امر یکی از دلایل متعددی است که در نظریه نمونه گیری " بررسی دقیق چارچوب " را توصیه می‌کند . اگر چنین کاری امکان پذیر نباشد در اینصورت :

$$V(\bar{x}_{Sy}) = \frac{N-1}{N} \cdot \frac{S^2}{n} \left[1 + (n-1)\rho\right]$$

که در آن ρ " ضریب همبستگی " بین هر دو واحد داخل حجم نمونه سیستماتیک است که به صورت ذیل تعریف میشود :

$$\rho = \frac{E(x_{ij} - \bar{X})(x_{ik} - \bar{X})}{E(x_{ij} - \bar{X})^2} \qquad k \neq j$$

همانطور که اشاره شد در نمونه گیری سیستماتیک عملا " نیازی به محاسبه این ضریب نیست ولیکن در نمونه گیری خوشه ای نقش تعیین کننده دارد . در نمونه گیری خوشه ای آن را " ضریب همبستگی داخلی خوشه " نامیده و با رابطه ذیل نشان میدهند :

$$\rho = \frac{E(x_{ij} - \bar{\bar{X}})(x_{ik} - \bar{\bar{X}})}{E(x_{ij} - \bar{\bar{X}})^2} = \frac{2 \sum_{j<k} \sum (x_{ij} - \bar{\bar{X}})(x_{ik} - \bar{\bar{X}})}{(M-1)(NM-1) S^2}$$

1- Frame

که درآن :

x_{ij} : مقدار "صفت[1]" واحد j ام درداخل خوشه i ام

$$\bar{\bar{X}} = \frac{1}{NM}\sum_i \alpha_i = \frac{\bar{X}}{M}$$

که درآن M تعداد واحدهای داخل خوشه و N تعدادکل خوشه درجامعه است ، و

$$S^2 = \frac{\sum_i \sum_j (x_{ij} - \bar{\bar{X}})^2}{NM-1}$$

واریانس برآورد میانگین یعنی $\bar{\bar{X}}$ عبارتست از :

$$V(\bar{\bar{X}}) = (1 - n/N) \cdot \frac{S^2}{Mn}\left[1+(M-1)\rho\right]$$

(خوشه های مساوی M عنصری)

واریانس برآورد درحالت نمونه گیری تصادفی ساده عبارتست از $(1-0^{n}/_{N}) \cdot \frac{S^2}{Mn}$ وبنابراین اثر طرح :

$$Deff = \frac{V(\bar{\bar{X}}_a)}{V(\bar{\bar{X}}_R)} = \left[1+(M-1)\rho\right]$$

تابعی از ρ و M (حجم خوشه) است .

از آنجائیکه درمحاسبه حجم نمونه نیاز به واریانس است ومحاسبه واریانس از طریق خوشه بندی مشکلات عملی بهمراه دارد لذا معمولا " از واریانس تصادفی ساده استفاده کرده وبه کمک مقادیر ρ و M ، حجم نمونه لازم را بدست می‌آورند . درحقیقت عامل $1+(M-1)\rho$ نشان میدهدکه مقدار واریانس درحالتیکه بجای نمونه گیری تصادفی ساده از نمونه گیری خوشه ای استفاده میشود چقدر تغییر می‌کند واین امر درمورد حجم نمونه نیز صادق است .

رابطه مربوط به ρ حالت دیگری نیز دارد که بیشتر درمحاسبه مقدار آن مورد استفاده قرار می‌گیرد وآن عبارتست از :

$$\rho = \frac{\bar{M}S_b^2 - S^2}{(\bar{M}-1)S^2}$$

1-Characteristic

که در آن $\bar{M}$ متوسط تعداد واحدها درهرخوشه و S_b^2 واریانس بین خوشه هاست .

ضریب همبستگی داخلی میزان " همگنی[1] " واحدهای داخل خوشه را اندازه می‌گیرد . از دیدگاه بررسی تغییرات آن موارد ذیل قابل ذکر است :

باتوجه به اینکه $S^2 = S_b^2 + S_w^2$ واریانس داخل خوشه) می‌باشد در اینصورت ضریب همبستگی داخلی به هردومقدار S_b^2 و S_w^2 ارتباط دارد . اگر $S_w^2 = 0$ باشد باین مفهوم که :

تمام واحدهای داخل خوشه از جهت صفت مورد بررسی شبیه هم هستند وهیچگونه تغییری بین واحدهای داخل خوشه وجود ندارد . دراین حالت $S_b^2 = S^2$ و $\rho = \frac{\bar{M}S^2 - S^2}{(\bar{M}-1)S^2} = 1$

بطریق مشابه اگر $S_b^2 = 0$ باشد به این معنی که میانگین تمام خوشه ها باهمدیگر برابر باشد . دراینصورت $S_w^2 = S^2$ و $\rho = \frac{-1}{\bar{M}-1}$

مشاهده میشود که وقتی واحدهای داخل خوشه باتوجه به صفت مورد بررسی زیاد شبیه هم باشند مقدار ρ برای آن صفت زیاد خواهد بود وبرعکس وقتی که واحدهای داخل خوشه باتوجه به صفت مورد بررسی بطور نسبی ناهمگن باشند مقدار ρ برای آن صفت مثبت وکم ویادرحالت های نادر منفی خواهدبود .

نتایج حاصل از چند بررسی

همانطور که اشاره شد مقدار ρ درمحاسبه حجم نمونه نقش تعیین کننده دارد درهمین مورد " هنسن وهاویتس[2] " طی بررسی های گسترده‌ای مقدار ρ را برای خوشه های باواحدهای متفاوت درچند زمینه برای خانوارهای آمریکا محاسبه وبصورت ذیل ارائه کرده‌اند :

1-Homogenity

2-Hansen & Hurwitz

میزان همگنی : ρ ، بین خانوارهای آمریکایی در ارتباط با چند صفت

صفت مورد بررسی	میزان ρ برای خوشه های متشکل از :			
	۲ خانوار	۹ خانوار	۲۷ خانوار	۶۲ خانوار (متوسط بلوک)
۱ـ نسبت خانوارهای مالک واحد مسکونی	۰/۱۷۰	۰/۱۷۱	۰/۱۶۶	۰/۰۹۶
۲ـ نسبت خانوارهای اجاره نشین با اجاره ۱۲ـ۱۰ دلار در ماه	۰/۲۳۵	۰/۱۶۹	۰/۱۰۷	۰/۰۶۲
۳ـ نسبت خانوارهای اجاره نشین با اجاره ۲۹ـ۴۰ دلار در ماه	۰/۲۳۰	۰/۲۴۹	۰/۲۲۳	۰/۱۱۲
۴ـ متوسط اندازه خانوار	۰/۱۲۰	۰/۱۸۶	۰/۱۴۲	۰/۰۶۶
۵ـ نسبت مردان بومی سفید پوست	۰/۱۰۰	۰/۸۸	۰/۷۷	۰/۰۵۸
۶ـ نسبت بیکاران مرد	۰/۰۶۰	۰/۰۷۰	۰/۰۴۵	۰/۰۳۴
۷ـ نسبت مردان بین ۳۴ـ۲۵ سال	۰/۰۲۵	۰/۰۲۶	۰/۰۱۸	۰/۰۰۸

دو مورد خانوارهای ایرانی نیز انجام چنین بررسی هایی ضروری است ولیکن محاسبه آن فقط زمانی مورد توجه قرار می‌گیرد که یک " طرح نمونه گیری خوشه ای " مورد نظر باشد . در همین راستا مقدار ρ در جامعه ایران برای چند صفت محاسبه شده است . این مقادیر عبارتند از :

ρ	تعداد خانوار خوشه	صفت مورد بررسی
		نسبت افراد از قلم افتاده در سرشماری :
۰/۰۰۶	۲۰	شهری
۰/۰۰۲	۴۰	روستایی
		نسبت مرگ و میر :
۰/۰۰۱	۱۰۰	در نقاط شهری
۰/۰۰۲	۱۰۰	در نقاط روستایی
		برای صفاتی که نسبت آنها در جامعه ۰/۰۵ است :
۰/۰۰۵	۲۶۰	در نقاط شهری
۰/۰۰۶	۲۶۰	در نقاط روستایی

محاسبه ρ به ویژه به تکنیک استان زمینه های مطالعاتی وتحقیقی ذیقیمتی را فراهم می‌کند که علاوه بر محاسبه حجم نمونه ، میتواند ایده های جدیدی را درپیمه تحقیقاتی مطرح کند.

پرسشنامه آمارگیری جاری جمعیت سال ۱۳۷۰ به گونه ای طراحی شده است که اطلاعات ماشینی آن زمینه بسیار مناسبی را برای مطالعات مربوط به ضریب همبستگی داخلی " صفات موجود درآن ایجاد کرده است . انجام فعالیت های جاری درمرکزآمارایران فرصت کمتری را برای محققان این مرکز بوجود می‌آورد تا در چهارچوب آن تحقیقات مورد نیاز را انجام دهند ولیکن این فرصت برای محققان سایر سازمانها و بویژه دانشجویان فراهم خواهد بود که انجام چنین مطالعاتی را درحیطه بررسی های خود قراردهند.

فهرست منابع ومآخذ

1-Cochran.G. William, Sampling Techniques, U.S.A, John Willey & Sons, 1977

2-Kish. Lesly, Survey Sampling, U.S.A, John Willey & Sons, 1965

3-Hansen,M.H,Hurwitz,W,N,and Madow,W,G, Sample Survey Methods and Theory, U.S.A, John Willey & Sons, 1955

4-Konijn. H,S, Statistical Theory Of Sample Survey Design And Analysis, North -Holland Publishing Company, London, 1973

پرسشنامه، سنگ زیربنای آمارگیریها و سرشماریها

پرویز رضایی تهرانی

مرکز آمار ایران

مقدمه

یکی از راههای ثبت اطلاع استفاده از پرسش است . خواه این اطلاع مربوط به گذشته، حال یا آینده باشد . عامل زمان در نحوه ساخته شدن یک پرسش اثر دارد . معمولا " برای جمع آوری یک یا چند اطلاع بسهولت فراوانی، گستردگی و چگونگی، وسیله ای بکار می رود که آنرا پرسشنامه می نامیم . صرفنظر از انواع پرسشنامه ها که هرکدام برای منظور خاصی تهیه می شود، بطور کلی پرسشنامه باید دارای خصوصیات عمومی و عمده ای باشد که خارج از آن هر پرسشنامه مورد نظر نمی تواند مفید افتد .

اصولا " پرسشنامه به عنوان ابزاری برای جمع آوری یک یا چند اطلاع بکار می رود تا هدف یک طرح را معنی دار نماید . بنابراین پرسشنامه سنگ بنای طرحهای آمارگیریها و سرشماریها و هرگونه وسیله اطلاع گیری می باشد . در این بحث همواره این منظور پی گرفته می شود و هرکجا لازم باشد بسه حواشی نیز پرداخته خواهد شد .

بر هیچ کس پوشیده نیست که نتیجه آمارگیریها و سرشماریها چون ابزاری برای برنامه ریزان و سیاست گزاران است ، و بدون این ابزار آنچه آنان می سازند و آنرا پروژه عمرانی می نامند ، شاید دقیقا " جرا بگوی نیاز جامعه نباشد و آنچه در این میان تلف می شود، غیر از زمان، ثروتی است که جایگزین کردن آن بسادگی ممکن نیست . این ابزار بوسیله پرسشنامه از واقعیاتی که بصورت اطلاع جمع آوری می شود بدست میاید و نتیجه میزانهائی را ارائه میدهد که خود پایه و اساس کار برنامه ریزی است . بدون شک پرسش غلط، تنظیم غلط، چگونگی پرسش، شکل پرسشنامه، طولانی بودن یا کوتاه بودن پرسشنامه، نامربوط بودن سوال، ناراحت کننده بودن سوال، پرسش هائی که کمیت را سوال می کند و بسیاری مسائل دیگر که منتقیها " نتایج حاصله از یک طرح را منحرف می کند، امر برنامه ریزی را بسمتی ناهنجار سوق می دهد .

نخست ببینیم پرسشنامه چیست ؟

بطور کلی پرسشنامه به دو قسم است . پرسشنامه ثبتی و پرسشنامه مبتنی بر پرسش مستقیم .

پرسشنامه‌های ثبتی عموما " اطلاع را با در موقع وقوع یک واقعه یا از سایر مدارک جمع آوری می‌کنند .
پرسشنامه مبتنی بر پرسش که معمولا " مفصل تر و دارای سوالات متفاوت است ، اطلاعات را مستقیما "
از پاسخگو گردآوری می‌نماید . بنابراین بطور کلی باید گفت که پرسشنامه سلسله پرسشهایی
استفسارهائی است که به عنوان راهنما و تعیین کننده جهت ، در یک بررسی برای جمع آوری اطلاعات
مورد استفاده قرار می‌گیرد .

در هر پرسشنامه باید شرایطی رعایت شود :

الف : پرسشنامه باید فرضیه‌ها و اهداف بررسی را در قالب پرسشهائی خاص بیان دارد .

ب : پرسشنامه باید چنان طرح شوند که طرف مورد مصاحبه را آزرده نکند . در چارچوب طبیعی ترین
گفتگوها بگذجد ، ارائه پاسخها را به طور آزادانه و عینی تسهیل کند .

ت : در پرسشنامه باید علامت گذاری پاسخها روشن و دقیق انجام گیرد تا موجبات تسهیل در کدگذاری
و استخراج فراهم آید .

ث : در یک پرسشنامه به قسمتهائی که مربوط به پرسش نیست جذل توجه خاص ضروری است : مانند
شماره ، آدرس و غیره .

خصوصیات پرسشنامه :

پرسشنامه‌ها خصوصیات متفاوتی دارند . این خصوصیات به دو قسمت اول خصوصیات ظاهری
و دوم خصوصیات مفهومی آن . یکی از مهمترین اثراتی که در تکمیل یک پرسشنامه دخالت دارد ، اثر
ظاهری آن است . اندازه ، شکل ، طرز قرار گرفتن سوالات ، جای کافی برای نوشتن ، قابلیت حمل و نقل ،
از جمله ظواهری است که اگر با هدف پرسش و طرح مطابقت داشته باشد می‌تواند نتیجه نیکوئی به بار
آورد . به عنوان مثال ، تحویل یک برگه پرسشنامه به قطع بزرگ به مامور آمارگیری ، در حالی که این
مامور باید در منطقه روستائی در فصل پائیز با زمستان کار کند یا اینکه سوالات نامرتب و با خط
بدخوانت نوشته شده باشد و ضمنا " جای کافی برای نوشتن پاسخ وجود نداشته باشد ، قابلیت حمل آن
مشکل بوده و پرسشنامه ناقص تکمیل می‌گردد ، کم کم مامور را خسته کرده ، و لاجرم گاه ممکن است نتیجه
عکس حاصل آید .

مفهوم پرسشنامه برای طراح آن ، اصولا " باید براساس آنچه در سیر انسانها و قشر مطلق های شهائی
مورد نظر است ، استفاده شود . بدین معنی که از بابت تکنیک کار هم از نظر اقتصادی ، پرسش و پرسشنامه
باید طوری تهیه شود تا با در صدهای آنچه را که مورد نظر طرح است کشف کند نه کم و نه زیاد .
اگر آنچه مورد نظر است کم دریافت شود مسلما " نمی‌توان تکنیک نیاز را پذیرش داد و اگر زیاد ، محققا " هم

اطّلاع رسانی ـ سالنامهٔ آماری کشور

سیدعلیرضا رفیع‌زاده

مرکز آمار ایران

مقدمه :

فرصت را مغتنم دانسته یکی از اشربات مهم مرجع آماری کشور را که پایگاه اطلاع رسانی مهمی است در نخستین کنفرانس آمار و ایران مطرح می‌کنم ، امیدآنکه با بذل توجه اساتید و فضلا و علاقمندان به توسعه و پیشرفت این مرز و بوم، قدمی در راه تامین نیازهای امت اسلامی و برآوردهای حق‌تعالی بردارید ...

اگرچه جای سخن در این زمینه بسیار است ، معذالک با زمان کوتاهی که برای تهیه این مقاله در اختیار داشته‌ام گفتار خود را در ۹ بند بشرح زیر برای تقدیم با حضار و تکمیل آن تقدیم میدارم .

الف) آمارهای رسمی در ایران و جهان

استفاده از آمار از زمانی شروع شد که انسان با مفهوم اندازه گیری را درک کرد و شمارش را یاد گرفت . با این ترتیب " اطلاعات عددی " یا بعبارت امروزی " آمار " در زندگی فردی و اجتماعی جوامع بشری اهمیت یافت .

شواهد تاریخی حاکی است که در زمان هخامنشیان واحدهائی برای اندازه گیری طول ، وزن و حجم ابداع شده است ، کشور به ۲۰ ایالت تقسیم شده و از روشهای خاصی برای جمع آوری اطلاعات آماری استفاده شده است . از هرودوت مورخ یونانی نقل می‌شود که " وقتی خشایارشا از تفلیس دیدن می‌کند ، محل مخصوصی را برای سرشماری افراد قشون خود مناسب دید و مقرر داشت تا ده هزار سپاهی شمرده شده در سطلی گرد آیند بنحوی که حتی المقدور بیکدیگر نزدیک باشند . سپس دور این محوطه را خطی کشیده و بعد از خروج آنان از داخل خط ، دیواری بسر روی خط رسم شده بنا نهند که بلندی آن تا کمر یک سپاهی برسد . چون حصار آماده شد ، سپاهیان گروه گروه در آن

داخل و خارج شوند تا تعداد آنها معلوم شود. مورخ یاد شده نتیجه این سرشماری را یک میلیون و هفتصد هزار نفر ذکر می‌کند که بعضی آن را نپذیرفته‌اند.

مورخین از دوره ساسانیان (خسرو انوشیروان) چنین یاد می‌کنند که فرمان داده شد کلیه نفوس کشور با شمار میزان ثروت هر کدام سرشماری شوند. مردانی که سن آنها بین ۲۰ و ۵۰ سال بود بر حسب میزان ثروت در گروه‌های مختلفی طبقه‌بندی می‌شدند و برای هر گروه مالیات خاصی تعلق می‌گرفت. همچنین دختران شوهردار به علت آنکه مالیات بر آنها تعلق می‌گرفت شمرده می‌شد.

در زمان خلفای اسلامی، عهد صفویه و دوره قاجاریه نیز شواهدی وجود دارد که تهیه آمار مورد توجه بوده است. از آنجایی که برای فعالیتهای دولت در زمنه قدیم هدفی بجز ایجاد نظم و امنیت اداره امور عمومی تصور نمی‌شد، اهداف سرشماریها و جمع‌آوری آمار نیز اغلب در شناسایی امکانات برای تعیین نیرو و آذوقه و پول خلاصه می‌شد. از این روست که هنوز در نقاط دور افتاده کشور در اذهان عمومی آمارگیری را از سر بازگیری مرتبط می‌دانند.

" از اوایل قرن نوزدهم در زمان تحول و تکامل جوامع صنعتی، حکومتها ملزم شدند که مسائل روبه‌افزایش پیچیده اقتصادی و اجتماعی را مورد توجه قرار دهند و بسیاری از کشورها به تاسیس اداره آمار مبادرت کردند تا اطلاعات دقیقی جمع‌آوری و سیاستهایشان را بر اساس اطلاعات حاصل پایه‌گذاری کنند. علاقمندان برای جمع‌آوری حقایقی که به غبیدگیهای اجتماعی، اقتصادی و مالی کمک می‌کرد مجامع آماری ملی تاسیس کردند. اگرچه کمابیش نکاتی در مورد موجودیت علم آمار وجود داشت لیکن شهد علاقمندان و نیازهای حکومتها، برای تدوین آینده این علم کافی بود. گامهای اولیه برای ظهور همکاری بین‌المللی رسمی در آمار توسط Adolph Quetelet در نمایشگاه بزرگداشت در سال ۱۸۵۱ برداشته شد.

به ابتکار او و با همکاری کمیسیون مرکزی آمار بلژیک اولین کنگره بین‌المللی آماری در تابستان ۱۸۵۳ (مطابق با شهریور ۱۲۳۲) در بروکسل تشکیل شد. مطالب این کنگره نه تنها بخاطر ارزش ظهیری آنها بلکه به علت ارائه موتر الگوی اموری که قرار بود در کنگره‌های بعدی، در حالیکه در فعالیتهای " موسسه بین‌المللی آمار " طی قرن نوزدهم انعکاس یابند، اهمیت داشتند. در حالی که آمارهای جمعیتی، اقتصادی و اجتماعی در دستور جلسات این کنگره گنجانده شده بود، موضوع سازمان آماری کار اصلی بود. در این کنگره " ایجاد کمیسیونهای ملی مرکزی آمار و ارتباط با یک کمیسیون بین‌المللی که تضمین ایجاد قابلیت مقایسه آمارهای تولید شده در کشورهای مختلف را به عهده داشته باشد " مطرح و پیشنهاد گردید.

کنگره‌های بعدی که بین سالهای ۱۸۵۵ (۱۲۳۴ هجری شمسی) و ۱۸۷۶ (۱۲۵۵ هـ . ش) تشکیل شد به‌طور عمده درگیر آمارهای اجرایی و اداری (Administrative and Official Statistics) بودند و در سال ۱۸۷۲ (۱۲۵۱ هـ . ش) یک کمیسیون دائمی برای تهیه برنامه‌های گردهمائیهای آینده و نظارت بر تهیه آمارهای بین المللی بوجود آمد .

در سال ۱۸۷۸ (۱۲۵۷ هـ . ش) مقرر شد که کمیسیون مزبور به یک ارگان بین المللی ... وسیله ارتباطی بین اعضاء تبدیل شود و تصمیمات کمیسیون ، با روشی که اکثریت دول در مورد چگونگی به اجرا در آمدن آنها تعیین خواهند کرد پذیرفته شود .

مذاکرات پاینده دول آلمان و سوئیس با این طرح عملاً " در زمینه‌هایی که فعالیتهای چنین کمیسیونی صرف نخالتهای بی مورد در امور داخلی آنها خواهد بود ، منجر به متلاشی شدن کمیسیون و پایان این کمیسیون‌ها شد . بدین ترتیب اولین جنبش وانگیزه بسوی همکاری بین المللی آماری در موج برخاسته ازناسیونالیسم که سی و شش سال آینده بر زندگی سیاسی اروپا حاکم شود ، فرو ریخت .

به هر صورت اعضاء متعهد به سعی خود برای سازمان دادن همکاری بین المللی موثر طی جلساتی که در پاریس و لندن در سال ۱۸۸۵ (۱۲۶۴ هـ . ش) تشکیل شد باقی ماندند . در نتیجه کار آنها اجلاس تاسیس " موسسه بین المللی آمار " در ۲۴ ژوئن ۱۸۸۵ (۱۲۶۴ هـ . ش) در لندن برگزار شد . اهداف موسسه جدید آن بود که در جایگاه یک انجمن بین المللی با هدف توسعه پیشرفت آمارهای علمی و اجرایی از طریق مشروحه زیر قرار گیرد :

۱ ـ معرفی و مطرح کردن همسانی در روشهای گردآوری و جمع بندی نتایج آماری و کاربرد آن در تدوین تشریحات آماری با هدف سنجش نتایج حاصل در کشورهای مختلف .

۲ ـ محول کردن نظر دولتها به مسائل متعدد قابل حل بوسیله مشاهدات آماری و طلب کردن اطلاعات در مورد موضوعهایی که سابق بر این به اندازه کافی در معرض بررسیها و راه حلهای آماری (Statistical Treatment) قرار نگرفته بودند .

۳ ـ تهیه تشریحات بین المللی به عنوان وسیله‌ای برای مرتبط کردن آمار شناسان کشورهای مختلف .

۴ ـ سعی و کوشش از طریق رسانه‌های مطبوعات و در صورت امکان به وسیله آموزش عمومی و سایر وسائل مناسب ، جهت

بالابردن سطح کلی دانش آماری و برانگیختن علاقه دولتها و افراد به مطالعه و بررسی پدیده های اجتماعی .

موسسه ای که به این ترتیب تاسیس شد یک انجمن آزاد علمی از هر صفت رسمی بود . با این وصف ، در این

نهاد وجود چند فراوانی برای بنیان گذاری پایه ای جهت متحدالشکل کردن آمارهای رسمی وجود داشت ...

تصمیمات متخذه در این موسسه بخاطر اعتبار شخصی فوق العاده اعضای منشکله آن و همچنین کارها و زحمات

با ارزش آنها از اهمیت قابل توجهی برخوردار بود .

بین سالهای ۱۸۸۵ و ۱۹۲۳ (۱۲۶۴ و ۱۳۰۲ هـ ش) عملا " کار موسسه بر تحقیق و استفسار رسمی ،

ارائه پیشنهاداتی برای ایجاد قابلیت مقایسه و ارائه طرحهای آماری متمرکز شد . علاقه اولیه ای که در آمارهای

نفوس تجلی یافته بود قرار شد در قرن بیستم ادامه یابد . گامهای پیشتازی که در این زمینه برداشته شد شامل کمکهای Korosy, Guillaume, Levasseur به روشهای سرشماری جمعیت

بود : کمیته ISI پیشنهادات و طرحها را به دولتها ارسال نمود که منجر به سرشماری ۳۱ دسامبر ۱۹۰۰

(۱۲۷۹ هـ ش) در ۱۰ کشور و چند کشور دیگر در تاریخهای نزدیک به آن شد . رویه کشورها برای انجام

سرشماری در اول هر دهه که امروزه هنوز بوسیله تعداد کثیری از کشورها دنبال می شود ، تا حدی یادآور طرحها

و پیشنهادات موسسه در قرن نوزدهم سرچشمه می گیرد . بعلاوه ، Nixon می نویسد " سرشماری ...

چشمه ها و نکاتی که در اجرای سرشماری نفوس پذیرفته و اقتباس شده است از قبیل نوع سئوالاتی که باید

پرسش شود ، روشهای جدول گیری ، فرمها و دستورالعملهائی که برای شمارش جمعیت در کشورها بدون انجام

یک سرشماری عادی به کار برده می شود ، و پیشرفتهای فراوانی که طی ۵۰ سال گذشته در این زمینه و در همه سطوح

بیشتر و قابلیت مقایسه سرشماریها بعمل آمده است ، همه به این اقدام پیشگام موسسه مدیون می باشند .

موسسه همچنین در زمینه آمارهای حیاتی فعال بوده و با دولتها در انتشار مقالات و کتابهایی در زمینه

حرکت جمعیت همکاری می کرد . این وظیفه سرانجام بوسیله دفتر دائمی جدیدالتاسیس ISI در سال

۱۹۱۳ (۱۲۹۲ هـ ش) به عهده گرفته شد . ضمنا " موسسه بتاحدی با موفقیت ، طرق پیشنهادی برای طبقه بندی

علل مرگ و میر در کشورها را برآمد . طی دوره مزبور ، عملا " ین اصلی ترویج این فعالیتها اغلب آمارشناسان رسمی

اروپائی بوده اند .

موجب اتلاف وقت شده است وهم از دقت آنچه حاصل مقصود است کاسته است . بنابراین بیهوده نیست که دست اندرکاران تهیه پرسشنامه ، آنرا نخست مورد آزمایش قرار می دهند تا سستی ها و افزونیهای آنرا سنجیده ، قیمت تمام شده هر سوال را برآورد کرده ، آنگاه به واقع آنرا بمورد اجرا می گذارند .

پرسشنامه ها و انواع آن بسیارند و هرکدام به منظوری بکار می آیند . در این مقال مجال تشریح مفصل این مهم نیست . آنچه بیشتر مورد توجه است و نظر عنوان این مقاله هم همین است ، تکیه بر سنگ بنا بودن پرسشنامه در طرحهای آماری به اخص و بطور کلی وسیله ای برای هرگونه اطلاع گیری است .

همانطور یکه قبلا " نیز اشاره شد ، پرسشنامه وسیله ای برای کسب اطلاع است . در اینجا ذکر یک موضوع و روشن کردن یک مطلب مهم ضروری است و آن اینکه تهیه کننده یک طرح آماری و در نتیجه طراح یک پرسشنامه با اشراف کامل به خواسته های برنامه ریز اما بدون دخالت نظر مستقیم یا غیر مستقیم نظر خود ، باید وضع موجود را نشان دهد و واقع مصالحی را برای برنامه ریز آماده سازد و این وظیفه برنامه ریز است که از این مصالح سود جسته آنها را کنار هم بگذارد ، مشکلات را باز شناسد و با ست گزاران را در تعیین خط مشی لازم یاری دهد . در غیر این صورت ، طرح موجود و در نتیجه پرسشها و پرسشنامه تهیه شده جهت دار تلقی خواهد شد که نتایجی دور از واقعیات را بدست می دهد . در اینجا تا حدودی مساله تهیه طرح و تهیه پرسشنامه ، که بی ارتباط به هم نیستند از موضوع مطروحه ما بدور افتاده اما بدلیل اینکه اغلب شایع است که تهیه طرح را همان تهیه پرسشنامه تلقی کنند ناچار توصیف این کار اشتباه پیش آمد . یک طرح غلط همراه یک پرسشنامه هم سنگ آن شاید به نوعی جوابگو باشد ، اما یک پرسشنامه غلط در یک طرح صحیح مانند بکار بردن قطعات ضعیف و ناخوانند در ساخت ماشینی است که قرار است دقیق کار کند . و آیا این ممکن است ؟

آنچه که در یک طرح آماری لازم است ، گذشته از موارد عینی و قابل اندازه گیری ، توجه به مراحل کار است . در یک طرح تنها توجه بارز به هدف کلی کافی نیست و باید همواره این اصل را در منظر گرفت که سنگ بنای طرحهای آمارگیری ها و سرشماریها همین پرسشنامه است که تمامی اصول علمی مربوط به آن باید در منظر گرفته شود . به ظاهر از بی اهمیت ترین مسائل مانند جنس کاغذ ، قطع و نوشتار ، تعداد سوالات ، طراحی مناسب ، گرفتن تاپی گیر بودن ، مرتبط بودن ، آسان بودن پرسش به شکلی که جواب را درست بدست آورد ، از مهمترین ارکان یک طرح است .

شاید مهمترین پرسشها و پرسشنامه ها ، آنانی باشند که به دستور العمل نیازمند نباشند ، ولی

این موضوع همیشه ممکن نیست ودر صورت تهیه دستورالعمل روشن بودن تعاریف ومفاهیم و مطابقت سوالات با دستورالعمل مربوط بخصوص یکسان بودن پرسش در پرسشنامه ودستورالعمل از اهمیت ویژه‌ای برخوردار است . سعی در بکارگیری زبانی حتی‌المقدور عامیانه برای تفهیم وتفاهم، بکارگیری جملات استفهامی مستقیم ومحدودکردن سوال به نحوی که با سخنگوکوتاه‌ترین و ساده‌ترین جواب را بدهد، از موزتهیه پرسشنامه موفق است . تهیه سوالات باز به همراه سوالاتی که جواب آنها بلی یا خیر است به خاطر دوگانگی روشن پرسش از دقت پرسشنامه می‌کاهد . بدین لحاظ است که پرسشنامه‌های تهیه شده برای ارزیابی عقاید هم از دقت کافی برخوردار نیست وهم قابل کدگذاری نمی‌باشد . اصولا " پرسشهائی دارای ارزش هستند که بشود آنان را با سادگی تبدیل به کد با کمیت کرد . پرسشنامه‌هائی که کدگذاری آنها بعدا " انجام می‌شود و با اصطلاح کدر خود Self - Code نیستند هم وقت زیادی را تلف می‌کنند و هم عدم دقت کدگذاری نیز اشتباهات را در مراحل محاسبات نیز آنها وقدر مطلق ها افزون می‌کند . در این مراحل از کارکه حد فاصل بین کار انسان وماشین است تمهیدات خاصی بکار رفته است که از جمله می‌توان پرسشنامه‌های Fosdic را که علامت گذاشته توسط با سخنگو مستقیما " بوسیله ماشین مخصوصی خوانده شده تحویل کامپیوتر می‌شود نام برد که این خود دلیل کافی برای تهیه پرسشنامه‌های با سوالات بسته ساده وقابل فهم همگان وبطور کلی دارای کمترین دستورالعمل می‌باشد .

نتیجه کلی اینکه، در کار کسب اطلاعات به خصوص در مورد دریافت وقایع به عنوان یک اطلاع پرسشنامه وسیله منحصر بفرد بوده وبهر طریقی که بتوان آنرا به عنوان یک وسیله ارتباط گیری مستقیم با پاسخگو (منبع کسب اطلاع) نزدیک وقابل فهم کرد، بهمان طریق نیز می‌توان امیدوار بود که نتیجه یک یا چند اطلاع در قالب جداول آماری ومیزانها وشاخصها دقیق‌تر باشد .

بطور خلاصه اینکه پرسشنامه اساسی‌ترین وسیله برای تهیه یک طرح بخصوص طرحهای آماری است .

مسلما " بدون پرسشنامه که صرفنظر از چند پرسشی آن مرجع آوری اطلاع ، وجود طرحی ممکن نیست . یک پرسشنامه غلط می‌تواند نتایج یک طرح را بصورتی که نه تنها نتایج عکس آنچه به واقع وجود دارد بدست دهد . بنابراین اگر سنگ بنا سست یا نامتعادل باشد، تزلزل وفروریختن آن متحمل‌تر خواهد بود .

بین سالهای ۱۹۱۶ و ۱۹۲۹ (۱۲۹۵ و ۱۳۱۸ هـ.ش) دفتر داشمی و موسسه بهمت مؤلف یک سلسله " سالنامه آماری و دیگر نشریات جمعیتی متعدد شدند . به مفهومی این فعالیت مرتبا " توسعه بیشتر و پیوستگی موسسه به آمارهای رسمی بود و گردآوری و تدوین منظم آمارها در سالنامه ها شاید بالاترین نقطه این ارتبـــاط را نشان می داد . این فعالیت تقریبا " ۲۵ سال ادامه داشت .

پس از جنگ بین الملل اول و تحت حمایت رئیس جمهور آمریکا Woodrow Wilson

از فکر تاسیس سازمانهای بین المللی جدیدی تاسیس شدند . جامعه ملل تازه به دوران رسیده ، یک کمیسیون بین المللی آماری را منصوب کرد که در سال ۱۹۲۰ تشکیل شد و توصیه نمود که " جامعه ملل باید یـــک کمیسیون بین المللی در آمار تاسیس کند که در کلیه سئوالات فنی آماری رایزنی نموده و جامعه ملل را حتی المقدور کمک نماید " لیکن به علت مقاومت گروههائی که در سطح ملی عهده دار تدوین آمارهای رسمی بودند ، مانند سال ۱۸۷۸ چنین کمیسیونی تشکیل نشد .

در سال ۱۹۲۲ (۱۳۰۱ هـ.ش) کمیته اقتصادی جامعه ملل ، با همکاری ISI ، کمیته مشترکی بـــرای تهیه برنامه مطالعه ای که شامل آمار تجارت بین المللی ، آمار تولید ، شاخصهای قیمت و شرایط اقتصـــادی بود تاسیس کرد .

گزارشهای تهیه شده به جامعه ملل فرستاده شد که جامعه نیز آنها را برای دولتهای عضو ارسال کـــرد . بخشهای نهائی در کنفرانس بین المللی آمارهای اقتصادی جامعه ملل که در تابستان ۱۹۲۸ (۱۳۰۷ هـ.ش) تشکیل شد انجام شد که از آن کنوانسیون بین المللی آمارهای اقتصادی سرچشمه گرفت و متعاقبا " اغلب کشورهای عمده صنعتی به آن پیوستند . شریک مساعی های مشابهی بین موسسه و دفتر بین المللی کار به عمل آمد .

به دنبال جنگ بین الملل دوم و تشکیل سازمان ملل متحد کمیسیون دائمی آماری سازمان ملل در سال ۱۹۴۷ تاسیس شد . کمیسیون مزبور اعلام داشت :

الف ـ سازمان ملل متحد و ادارات تخصصی (Specialised Agencies) مسئولیـــت جمع آوری ، تجزیه و تحلیل و انتشار آمارهائی را که در ادارهٔ وظائف محوله به آنها و یا پیشرفت کلی و سودیـــک سیستم بین المللی آماری مناسب و هماهنگ لازم است به عهده دارند .

ب - سازمانهای غیر دولتی بین المللی علاقمند به پیشرفت آمار باید خط علمی و حرفه ای خود را ادامه و توسعه دهند و فعالیتهایشان را بسوی پیشرفت متدولوژی آمار و استانداردهای علمی تبادل و نشر دانش، تربیت آمارشناسان و حفظ موجودیت و قابلیتهای اعلای حرفه ای متوجه سازند.

از آن پس موسسه بین المللی آمار نیز سعی کرد تعداد بیشتری از کشورها در آن شرکت کنند و همچنین پیوندهای بهتری با سازمانهای بین المللی مانند شورای اقتصادی و اجتماعی ملل متحد و یونسکو برقرار کند" (۱).

در ارتباط با وظایف جدیدی که در شئون مختلف اجتماعی و اقتصادی از دولتها انتظار می رفت، در ایران نیز از سال ۱۳۰۷ (۱۹۲۸ میلادی) تعدادی از فارغ التحصیلان دبیرستانها برای طی دوره های علمی بالاتر و مجمله آمار و ریاضی به کشورهای اروپائی اعزام شدند که شاید تعدادی از فارغ التحصیلان آن دوره ها در بین اساتید حاضر در این کنفرانس شرکت داشته باشند. اولین کتاب در زمینه آمار به اسم زیاد " کتاب روشهای آماری در آموزش و پرورش " است که توسط استادارجمند جناب آقای دکتر علی افضلی پور تالیف و در سال ۱۳۱۹ (۱۹۴۰ میلادی) در تهران منتشر شد.

در سال ۱۳۰۳ (۱۹۲۴ میلادی) اولین اداره آمار بنام " اداره آمار و ثبت احوال " در ایران تاسیس شد. در سال ۱۳۱۲ آئین نامه شورای عالی آمار تظیم شد و به موجب ماده اول آن کلیه وزارتخانه ها و ادارات مستقل دولتی به عضویت شورا درآمدند. از جناب آقای دکتر علی افضلی پور که در آن زمان عضو شورای مزبور بودند نقل می شود

" ... کسانی که به نمایندگی از طرف وزارتخانه ها در این شورا می آمدند، متاسفانه کم اطلاع ترین و بی کارترین افراد وزارتخانه ها نیز بودند ... " (۲)

در سال ۱۳۱۸ (۱۹۳۹ میلادی) اولین سرشماری جامع نفوس، مذهب، سواد، تابعیت، شغل و وضع فامیل بطور مرتب در ۲۴ شهر به مرحله اجرا درآمد که پس از بروز جنگ بین الملل دوم متوقف شد.

---

1) The International Statistical Institute (1885-1985) The Netherlands . June 1985

۲) بانک مرکزی جمهوری اسلامی ایران - آمار اقتصادی شماره ۱ سال اول سال ۱۳۶۸

همزمان با پایان جنگ بین‌المللی دوم و ایجاد نظم نوین جهانی، در نتیجه اقدامات ترومن رئیس جمهور امریکا، سازمان ملل متحد در سال ۱۹۴۵ (۱۳۲۴ هـ.ش) در شهر سانفرانسیسکو تشکیل شد. پیدایش دو ابرقدرت و تشکیل اردوگاه غرب و شوروی، ابعاد و جهات جدیدی در برنامه ریزیهای اجتماعی و اقتصادی در سطح مناطق مختلف جغرافیائی ایجاد کرد.

در نظام اطلاعات آماری جهان، کمیسیون دائمی سازمان ملل که در سال ۱۹۴۷ (۱۳۲۶ هـ.ش) تاسیس شده بود، هدفهای خاصی را تبیین کرد. بالاخره در کشورهای جهان سوم " اصل چهارم " ترومن که تحت عنوان " کمک به مناطق عقب افتاده جهان " در ۲۰ ژانویه ۱۹۴۹ (۱۳۲۸ هـ.ش) به کنگره امریکا پیشنهاد شده بود، نیازهای آماری و امکانات آماری جدیدی را به وجود آورد.

در شماری عمومی نفوس سال ۱۳۳۵ (۱۹۵۶ میلادی) اولین گام در زمینه تهیه اطلاعات آماری در سطح کشور جهت تامین نیازهای آماری برنامه ریزیهای اجتماعی و اقتصادی آن زمان بشمار می آید. پس از آن آمارگیری نمونه ای کشاورزی سال ۱۳۳۹ (براساس توصیه های سازمان ملل در زمینه سرشماری کشاورزی جهانی ۱۹۶۰) و آمارگیری صنعتی کشور در سال ۱۳۴۲ (۱۹۶۳ میلادی) به انجام رسید و سرانجام اطلاعات آماری جامعی در مورد متغیرهای اجتماعی و اقتصادی کشور فراهم گردید. در سالهای بعد با توسعه امکانات آماری در سطح سازمانهای دولتی و ایجاد سازمان " مرکز آمار ایران "، بدون آنکه نیازهای آماری کاملا " مشخص شود آمارگیریهائی متنوعی انجام شد که نتایج آنها کموبیش در برنامه ریزیها، تصمیم گیریها و ارزشیابی نتایج فعالیتهای اجرائی مورد استفاده قرار گرفته است.

پس از پیروزی انقلاب اسلامی، ضرورت وجود بسیاری از سازمانها منجمله " مرکز آمار ایران " مورد سئوال قرار گرفت. از اینرو سازمانهای دولتی برای توجیه فعالیتهای خود توجه بیشتری به استفاده از آمار نشان دادند. در جریان بازنگری برنامه فعالیتهای گذشته سازمانها و برنامه ریزی برای فعالیتهای آینده، ضرورت تشخیص نوع و میزان نیازهای آماری سازمانهای دولتی و لزوم برآورد امکانات آماری موجود، مرکز آمار ایران را بر آن داشت که در این موارد بررسی و اطلاعات وسیعی گردآوری و منتشر نماید.

در سال ۱۳۶۱ کمیته ای تحت عنوان " کمیته تشخیص نیازهای آماری سازمانهای دولتی و نهادهای انقلاب اسلامی " متشکل از ۶ نفر از کارشناسان منتخب معاونتها و دفاتر مرکز آمار ایران تشکیل شد. کمیته در برخورد با بهره گیری از تجارب

گذشته با تهیه پرسشنامه‌ای به جمع‌آوری نیازهای آماری سازمانها اقدام نمود. ۹۲ سازمان دولتی و ارگان انقلابی انتخاب و با همکاری ۶۹ نفر از کارشناسان زبده مرکز آمار ایران پس از طی یک کلاس آموزشی، در مورد تکمیل پرسشنامه‌ها اقدام شد. پاسخهای اولیه نشان می‌داد که هنوز برای بعضی از سازمانها "نیاز آماری" شناخته شده نیست و آن را با آمار پرسنلی یا موجودی انبار مترادف می‌دانند. تعداد ۸۰ سازمان پرسشنامه‌های مربوط را تکمیل کردند که اطلاعات مندرج با استفاده از طبقه‌بندی بین‌المللی فعالیتهای آماری (توصیه کمیسیون اقتصادی و اجتماعی برای آسیا و اقیانوسیه SCAP) و طبقه‌بندی سرشماری آماری استخراج و در سال ۱۳۶۳ منتشر شد.(۱) نتایج این طرح به دفتر مطالعات و تهیه طرحهای آماری نمونه‌گیری و تهیه پرسشنامه‌های سرشماری و بامور توجه و استفاده تهیه‌کنندگان طرحها قرار گرفت و در محدوده امکانات فنی و اجرائی طرحهای نیازهای آماری سازمانها در پرسشنامه‌های آماری منظور شد. کمیته تشخیص و جمع‌آوری نیازهای آماری توصیه کرد:

"برای آن که رسالت مرکز آمار ایران و دیگر سازمانهای تولیدکننده آمار در جمع‌آوری آمارهای مبنایی و اجرای طرحهای آمارگیری با عینیت بیشتری انجام گیرد، لازم است حداقل هر ۵ سال یکبار تصویری همگام از نیازهای آماری مملکت تهیه و نسبت به تأمین آنها اقدام گردد."

در تکریر مزبور اقلام آماری مورد نیاز، سازمان متقاضی، نحوه تولید آمار یا اقدام، اطلاع، سطح جغرافیائی اطلاع مورد نیاز، دوره زمانی مورد نظر و سازمان تولیدکننده آمار مشخص شده است و ضمناً "پاسخهای سئوال" کاربرد اقلام آماری مورد نیاز" به دلیل نارسا بودن درج نشده است.

در همین راستا طرح "شناسایی نیازها و امکانات آماری کشور" تهیه شد که در سطح ۲۵۶۸ سازمان عمومی و خصوصی در سراسر کشور در سه مرحله اجرا شد. در این طرح داشتن واحد آمار در سازمان، تهیه‌کنندگی و انسانی شاغل در تهیه آمار و استفاده از کامپیوتر در تهیه آمار و نیز فعالیتهای آماری سازمانها طی دوره ۵ ساله ۱۳۶۶ - ۱۳۶۲ به عنوان امکانات آماری سازمان تلقی شده است. طرح مزبور در طول سال ۱۳۶۷ اجرا شد و نتایج آن در جلد دوم سال ۱۳۶۸ منتشر شد.(۲) برای اطلاع از خلاصه یافته‌های این طرح به ضمیمه شماره ۱ پیوست مراجعه کنید.

۱- مرکز آمار ایران. فهرست اقلام آماری مورد نیاز سازمانهای دولتی و نهادهای انقلاب اسلامی. ۱۳۶۳.

۲- مرکز آمار ایران. طرح شناسایی نیازها و امکانات آماری کشور. نتایج آمارگیری از امکانات آماری کشور - سال ۱۳۶۷. تهران: ۱۳۶۸.

ج ـ نظام اطلاع رسانی ، منابع آماری و انتشار اطلاعات آماری

در قانون مرکز آمار ایران مصوب ۱۳۴۴ ، هدف از تاسیس مرکز آمار ایران " ایجاد یک نظام متمرکز آماری معتکی به منظور تهیه آمارهای جامع و صحیح در زمینه های مختلف اقتصادی و اجتماعی و بررسی علمی آنها در جهت برنامه ریزی های کشور " مقرر شده است . لیکن تاکنون نیل به این هدف بطور کامل تحقق نیافته است و علاوه بر آمارهای ثبتی که نتیجه فرعی فعالیتهای جاری دستگاههای دولتی است ، بعضی از سازمانها را " به آمارگیری در زمینه های موضوعی خاص اقدام می کنند مانند بانک مرکزی جمهوری اسلامی ایران .

جداول ضمیمه ۱ نشان می دهد که از ۲۱۴۰ سازمان عمده عمومی و خصوصی مورد بررسی در ۱۳۵۰ سازمان (۲۶/۲ درصد) فعالیت تهیه آمار انجام می شود . در حالی که در کلیه ۶۶ سازمان مرکزی ، فعالیت تهیه آمار انجام می شود فقط ۶ سازمان مرکزی بدون واحد آمار است .

استخراج ، تجزیه و تحلیل و انتشار نتایج فعالیتهای آماری سازمانهای تهیه کننده آمار اغلب به وسیله سازمانهای گردآوری کننده اطلاعات آماری انجام می شود و برخلاف بعضی از کشورهای پیشرفته ، در این زمینه فعالیتها به صورت غیر متمرکز و گاهی ناهماهنگ انجام می شود . جدول ۴ ضمیمه ۱ نشان می دهد که ۴۰/۷ درصد نتایج فعالیتهای آماری خاتمه یافته سالهای ۱۳۶۶ ـ ۱۳۶۲ سازمانهای مرکزی تهیه کننده آمار در داخل چاپخانه همان سازمانها و ۲/۹ درصد در سایر چاپخانه ها تکثیر شده است و وجود ۵۶/۲ درصد نتایج فعالیتهای آماری مربوط به روشی غیر از چاپ ، تکثیر گردیده است .

د ـ نشریات آماری و تشریفات اداری آماری

با وجود پیشرفتهای عظیمی که در تکنولوژی اطلاع رسانی ایجاد شده است ، کتاب و سایر انواع نشریات کماکان جایگاه مهمی در میان استفاده کنندگان از منابع اطلاعات آماری دارد . اغلب نشریات آماری در کشور به صورت موردی و بسیار غیر منظم منتشر می شود . با این وجود سازمانهائی که تشکیلات آماری منسجم و فعالیتهای آماری منظمی دارند سعی می کنند اطلاعات آماری خود را در چارچوبهای مدون و در فواصل زمانی تقریبا " سالانه در اختیار عموم قرار دهند .

مرکز آمار ایران نیز بخش عمده ای از نتایج سرشماری و آمارگیریهای نمونه ای را به صورت نشریات آماری چاپ و در دسترس متقاضیان اطلاعات آماری قرار می دهد که بعضی از آنها در دوره های زمانی معین منتشر می شود .

ـ تاریخچـــه :

سالنامه آماری ، معمولاً " نشانگر موجودیت یک سازمان آماری منسجم و پویاست . سازمانهای آماری کشورهـــای مختلف جهان عموماً از زمانی معین به تهیه و انتشار سالنامه آماری مبادرت ورزیده اند . حدود ۱۵۰ سال است کـــه سالنامه آماری بریتانیا که اکنون Annual Abstract of Statistics نامیده می شـــود مهمترین مرجع آماری این کشور محسوب می شود .

سالنامه آماری ایالات متحده آمریکا که از سال ۱۸۷۸ (۱۲۵۷ هـ ش) بطور منظم منتشر می شود ، اکنون بـــا نـــام Statistical Abstract of The United States کتاب آمار ملی و راهنمای منابع آماری کشور مزبور است .

سالنامه آماری ژاپن که در سال ۱۸۸۲ (۱۲۶۱ هـ ش) پایه گذاری شده است از سال ۱۹۴۹ (۱۳۲۸ هـ ش) بـا نام Japan Statistical Yearbook بطور منظم منتشر می شود .

سازمان ملل متحد و سازمانهای بین المللی تخصصی نیز حدود بیست سالنامه آماری منتشر می کنند که تاکنون حدود ۴۵ شماره Statistical Yearbook و ۵۰ شماره Yearbook of Labour Statistics و ۴۲ شماره از نشریه FAO Production Yearbook منتشر شده است.(۱)

در ایران نیز با تشکیل ادارات آمار در سازمانهای دولتی ، سالنامه های آماری موضوعی ، انتشار یافت ، نمونـــه سالنامه وزارت معارف و اوقاف و صنایع مستظرفه با نام " سالنامه احصائیه سال ۱۳۱۲ـ ۱۳۱۳ و ۱۳۱۳ـ ۱۳۱۴ " یـا سالنامه آمار بازرگانی خارجی ایران ـ اداره کل آمار ایران معرف اینگونه سالنامه های تخصصی است .(۲)

تهیه سالنامه آماری کشور به عنوان یکی از وظایف مشخص در قانون مرکز آمار ایران مورد تصویب قرار گرفت . بـــر این اساس اولین سالنامه آماری کشور ، در سال ۱۳۴۵ (۱۹۶۶ میلادی) پس از تاسیس مرکز آمار ایران در ۲ جلـــد منتشر شد که جلد اول آن ۲۷ فصل بود . در مقدمه این سالنامه تاکید شده است که " منطق ایجاب میکند که مقامات مسائـــل اجتماعی و اقتصادی علیرغم اینکه در هر زمان مستقیماً از مسائل متعدد جامعه آگاهی و بررسی میکنند در عین حال بکلیـــه

---

۱ ـ به ضمیمه شماره ۲ رجوع شود .

۲ ـ به ضمیمه شماره ۳ رجوع شود .

- آمارنامه استان

براساس الگوی سالنامه آماری کشور، مجموعه‌ای از اطلاعات پایه‌ای در سطح هر استان برحسب شهرستانهای ذی‌ربط توسط سازمان برنامه و بودجه استانها تهیه و منتشر می‌شود.

- نشریات " خلاصه آمارهای پایه‌ای کشور " ، " تازه‌های آمار " و " پرتلن آمارهای فصلی کشور " نیز از سالنامه آماری کشور تهیه می‌گیرند.

ز - تمرکز و عدم تمرکز در نظام اطلاع رسانی

بخش عظیمی از اطلاعات آماری بطور ثبتی یا تحقیقات دفتری توسط تعداد کثیری از سازمانهای دولتی تهیه می‌شود که انتشار و سپس آنها توسط سازمانهای ذی‌ربط مقرون به صرفه و امکان پذیر نمی‌باشد. بنابراین مرکز آمار ایران کلیه حدود ۳۰ درصد آمارهای جداول سالنامه را تامین می‌کند می‌تواند جمع‌آوری اطلاعات آماری از سایر سازمانها نظام اطلاع رسانی مطلوبی را در زمینه آمارهای تولید شده، از نظر تمرکز و عدم تمرکز اداری و جغرافیایی تحقق بخشد.

ح - تاثیر سالنامه آماری در هماهنگی اطلاعات آماری

کوشش در زمینه تهیه تعاریف و مفاهیم بدون تشخیص اقلام آماری و مطالعه تغییرات زمانی متغیرهای اقتصادی و اجتماعی و تعیین روش جغرافیایی اطلاعات به هماهنگی در روشهای تولید و انتشار آمارهای ثبتی منجر می‌شود.

ط - نتایج پژوهشی انجام شده در زمینه تهیه الگوی استاندارد برای سالنامه آماری کشور

بعضی از کشورهای پیشرفته جهان موفق به تهیه الگوی استاندارد برای سالنامه آماری کشور خود شده‌اند و تغییرات کم و مشخصی در سالنامه‌های آماری سالهای متوالی آنها مشاهده می‌شود. این موفقیتها که موجبات تسهیل تحقیقات آماری پژوهشگران را فراهم می‌سازد با عملکرد سیستم و سنخ نظام تولید آمار کشورهای مزبور ارتباط تنگاتنگی دارد.

برای نیل به یک نمونه استاندارد، به عنوان اولین گام، نتیجه بررسیهای انجام شده تحت عنوان " راهنمای تهیه فصلهای سالنامه آماری کشور " تدوین شده است که در میان بسیاری از سازمانها مشخصه تکامل فعالیت آماری در این زمینه تلقی می‌شود. در این راهنما سعی شده است در مورد ضوابط لازم برای تهیه جداول و فصلهای سالنامه آماری، تدابیری شود به امید آنکه راهنمائیهای صاحبنظران و مسئولان سازمانهای تولید کننده آمار در نمای مشخصی برای فعالیتهای آماری ترسیم شود و هماهنگی بیشتری در تامین نیازهای آماری کشور ایجاد شود.

متدولوژی تألیف و تحقیق

«فرهنگ آماری کلمات قرآن کریم» و دست‌آوردهای آن

A STATISTICAL DICTIONARY OF QUR'ANIC WORDS

کتاب سال جمهوری اسلامی ایران و کتاب برگزیده دانشگاههای کشور

دکتر محمود رجائی

دانشگاه علوم پزشکی مشهد

پژوهشهای آماری پیرامون قرآن کریم در فرهنگ اسلامی به عبارتی می‌تواند پیشینه‌ای به درازی تاریخ قرآن داشته باشد. از همان روزها که در نخستین دهه‌های سدهٔ اول اسلامی، قرآن مجید مدوّن گردید و از لوح سینه‌ها به سینهٔ دفترها منتقل شد، ابتدا شمار سوره‌ها و آیات و ترتیب و جایگاه هر یک مورد توجه قاریان و عامهٔ مسلمانان قرار گرفت و از آن پس هم، تعداد کلمات و حروف و مدّات و جزءها و وقوف و رکوعات، و سرانجام پس از پیدا شدن شکل و اعراب و تنقیط خط، شمار نقطه‌ها و شکلها و دیگر نشانه‌های سجاوندی و تجویدی. این بود که دلباختگان وحی، شوق و عمر و اعتقاد خود را به انگیزهٔ پاداش اخروی و یا به منظور سیراب کردن حس کنجکاوی به کتاب آسمانی خویش تسلیم داشتند، و به موازات کوششهایی که در زیبانویسی و تذهیب و تجلید هنرمندانهٔ قرآن صورت می‌گرفت، برخی نیز با شمارگریهای زمانگیر جنبه‌های کتبی قرآن را می‌آزمودند، همان که بعضاً نتایج کوشش آنان در کتابهای مستقل و یا صدر و ذیل قرآنهای خطّی به یادگار مانده است.

بعدها که کتابهای مستقل پیرامون قرآن و قرآن پژوهی سامان یافت، بسیاری از این آمارها هم بدان کتابها راه یافت یا پدید آورندگان آن آثار، خود به تن خویش دست به این گونه شمارگریها زدند و ارقام دیگری در هر مورد به دست‌آوردند که بالمآل دامنهٔ تنوّع رقمهای پیشین را وسعت می‌بخشید. آنچه در مورد این آمارها نباید از نظر دور بماند اینکه چون تهیهٔ آنها بیشتر از طریق شمارش مستقیم، آن هم با شیوه‌هایی ابتدایی ناملمتس صورت می‌گرفته است، هر گز نمی‌توانسته است از دقّت ریاضی لازم برخوردار باشد، بگذریم از اینکه به دست آوردن این اعداد و ارقام بیشتر چیزی از مقولهٔ تفنّن به شمار می‌رفته و کمتر مبنای علمی حساب شده‌ای بر آن مترتب بوده است.

مرحلهٔ نوین فهرست نگاری و مثلاً تهیهٔ کشف الآیات و کشف المطالب برای قرآن مجید را باید به همین اواخر و زمانی مربوط دانست که اروپاییان غیرمسلمان، برای رفع حوایج علمی خود، مراجعهٔ سریع و مطمئن را به کتاب آسمانی مسلمانان ضروری دیدند. از پاره‌ای کوششهای مقدماتی از ناحیهٔ خود مسلمانان که بگذریم، نخستین کشف الآیات علمی و مستند به موازین لغت نویسی عربی راه که برای یافتن هر کلمه و تعیین موضع آن، به مادهٔ واژه باید رجوع کرد گوستاو فلوگل دانشمند آلمانی پدیدآورد که به سال ۱۸۴۳ میلادی در اروپا به چاپ رسید. کار فلوگل که تا امروز سرمشق مسلمانان برای عموم معجمهای قرآنی بوده، بعدها با فهرست موضوعی زیر

نتایج و حوادث نیز احاطه داشته باشند تا بتوانند ارتباط آنها را نتیجه و قیقتری درباره مسئله مورد نظر بدست آورند و اینکار فقط وقتی میسر است که محقق آمار و اطلاعات گذشته و حال جامعه را بطور کامل در اختیار داشته باشد.(۱)

تهیه سالنامه‌های بعدی با دستور العمل خاصی با همین سلسله از سال ۱۳۴۲ ادامه یافت. در سال ۱۳۵۰ تغییرات خاصی در شکل و محتوای سالنامه ایجاد شد و بین سالهای ۱۳۶۶ (سالنامه آماری سال یکهزار و سیصد و شصت و پنج کشور) به آن تقریبا " به همان روال ادامه میافت.

سالنامه‌های مزبور علیرغم هماهنگی در شکل ظاهری الگوی فصول، از نظر محتوا نقاط ضعف و کمبش فراوانی داشت که از نگرش کارشناسان واحدهای تهیه کننده فصول مختلف سرچشمه می گرفت.

علیهذا از سال ۱۳۶۲ که " سالنامه آماری کشور ۱۳۶۶ " تهیه می شد با توجه به فروم بازنگری کلی وجامع در مورد این نشریه آماری مرجع، گروهی از کارشناسان با تجربه مرکز آمار ایران تعیین و سردبیری نشریه مزبور را بعهده دارند شدند تا در چارچوب و محتوای آن تغییرات هماهنگی بیشتری ایجاد کنند.

در این بررسی محتوای بسیاری از سالنامه‌های آماری صاحب نظر جهان نیز مورد توجه قرار گرفت و تغییرات بطور کلی و بطور اساسی در " سالنامه آماری کشور " آغاز شد که اکنون به برخی از مشخصات آن اشاره می شود.

- طبقه‌بندی اطلاعات آماری، پوشش موضوعی و تکرار اطلاعات.

پس از مطالعه طبقه‌بندی بین المللی فعالیتهای اقتصادی و طبقه‌بندی فصول سالنامه‌های آماری کشورهای مختلف و بررسی وضعیت موجود، ۲۵ فصل برای سالنامه آماری کشور تعیین شد و مقرر گردید از درج آمار مربوط به بیش از موضوع مشخص در فصول مختلف و یا تکرار اطلاعات در فصول مختلف اجتناب شود. بنابراین آمارهای مربوط به اعتبارات، بازرگانی خارجی، شاخصهای قیمت و حسابهای ملی و نیز آمارهای بین المللی که برای هر کدام از سالنامه آماری کشور فصلی اختصاص داده شد از سایر فصول سالنامه حذف شد. برای هر فصل مقدمه‌ای با چارچوب مشخص تهیه و تعاریف و مفاهیم آمارهای مندرج در هر فصل در پایان مقدمه ارائه شد.

- پوشش جغرافیایی.

مقرر شد آمارهای در سالنامه ارائه شود که کل کشور را پوشش دهد و اطلاعات آماری آخرین سال با منطقه آماری، حسب استان یا تقسیمات موضوعی تفصیلی ارائه شود.

---

۱ - به صفحه شماره ۲ رجوع شود.

- دوره زمانی اطلاعات :

برای ایجاد امکان مقایسه روند تغییرات اقلام آماری در طی زمان و به منظور هماهنگی اطلاعات آماری فصول مختلف مقرر شد اطلاعات آماری به صورت سریهای زمانی بر اساس ۵ مقطع زمانی آخر که آمار آن در دسترس است حداقل به صورت سالانه ارائه شود .

- نمودارها ، فهرست نمودارها و نمایه موضوعی

برای افزایش کارآئی سالنامه و تسهیل استفاده از مندرجات آن ، مقرر شد در هر فصل نمودارهائی بر اساس ضوابط معین شده ، ارائه شود و فهرست مندرجات ، فهرست جداول و فهرست نمودارها برای سالنامه تهیه شود . ضمنا " برای نخستین بار قرار شد از واژه های مهم سالنامه ، نمایه موضوعی لازم تهیه و بر حسب حروف الفبا تنظیم و در پایان سالنامه چاپ شود .

- منابع آماری :

آمار و اطلاعات ارائه شده در سالنامه آماری کشور نتیجه همکاری حدود ۲۰ سازمان با مرکز آمار ایران است . بنابراین این کتاب آماری مرجع می تواند راهنمای خوبی برای مصرفی منابع آماری در زمینه های مختلف باشد . از اینرو قرار شد علاوه بر درج سازمان ارائه دهنده اطلاعات آماری در ذیل جداول ذی ربط ، در صورتیکه اطلاعات آماری از نشریات رسمی سازمانها استخراج شده باشد ، فهرستی از منابع و مآخذ هر یک از فصول سالنامه تهیه و در پایان سالنامه ارائه شود .

- تدوین و انتشار :

تسریع در انتشار سالنامه آماری کشور به عنوان یکی از اصول در کلیه مراحل تهیه و تدوین این نشریه مد نظر قرار گرفته و سعی می شود با همکاری سازمانها و واحدهای ذی ربط نشریه مزبور با بیشترین آمارهای هنگام به زمان مطلوب انتشار منتهی میان اطلاعات آماری نزدیک شود . ضمنا " کوشش می شود این نشریه به عنوان یک کتاب مرجع از کیفیت تدوین و صفحه آرائی و چاپ خوب برخوردار شود .

نشریاتی که از سالنامه آماری کشور سرچشمه گرفته اند

در سیر تکاملی انتشارات ادواری آماری ، سالنامه آماری کشور منشاء ایجاد چندین نشریه مشروح و جزیی بوده است :

- آمار ایران در آئینه آمار (منتخبی از آمارهای پایه ای مندرج در سالنامه با نمودارهای رنگی در قطع جیبی) که به زبانهای عربی و انگلیسی این نشریه در سطح بین المللی برای استفاده خارجیان و ایرانیان مقیم خارج از کشور تهیه و منتشر می شود .

لابوم Jules Labeaume (۱۸۰۶ ـ ۱۸۷۶) خاورشناس فرانسوی، که برای نخستین بار به سال ۱۸۷۷ میلادی چاپ و منتشر شد، از جهتی تکمیل گردید، و کار پژوهشهای قرآنی را که پیش از آن در فرهنگ اسلامی ذوقی و استحسانی صورت می‌پذیرفت به سوی سامانها و شیوه‌های پژوهشی جدید علمی و دقیق هدایت کرد. امروز استفاده از کشف الآیات و طبقات الآیات، بویژه معجمهای پرآوازه و ظاهراً علمی و دقیق مثل المعجم المفهرس لالفاظ القرآن الکریم فؤاد عبدالباقی، و در نزد ایرانیان و فارسی زبانان علاقه‌مند به شیوه‌های القبایی، استفاده از کشف‌الآیات القبایی مرحوم دکتر محمددوامیار و اخیراً ترتیب المعجم المفهرس، به کوشش محسن بیدارفر و یا در نوع کشف المطالب، فرهنگ موضوعی قرآن مجید فراهم آوردهٔ کاصران فانی و بهاءالدین خرمشاهی، امری رایج و روزمره و در زمرهٔ بدیهیات و مسلّمات در زمینهٔ پژوهشهای متعلّق به قلمرو فرهنگ اسلامی به معنای وسیع و جامع کلمه است.»*

با طرح این مقدمهٔ مختصر ادای احترام به عظمت کار قرآن پژوهان بزرگی مانند زرکشی، سیوطی و زرقانی در فرهنگ اسلامی و فضل تقدم دانشمندانی چون فلوگل و ژول‌لابوم از حیث متدولوژی، در این مقال سخنی دربارهٔ انگیزهٔ تألیف و تحقیق کتاب «المعجم الإحصایی لألفاظ القرآن الکریم» می‌آوریم، آن‌گاه برخی مزایا و دست‌آوردهای آن را بطور اجمال بر می‌شماریم.

المعجم الإحصایی برپایهٔ اندیشهٔ بررسی آماری قرآن کریم و تجاربی در زمینهٔ معجم‌های قرآنی و کشف‌الآیاتها تدوین شده‌است، اندیشه‌ای که متجاوز از بیست سال ذهن نگارنده را به خود مشغول داشته است.

بررسی معجمهای قرآنی موجود نشان داد، مرجع اطمینان‌بخشی که کلیهٔ الفاظ قرآن و آمارهای مربوط به آن را در برگیرد و درعین حال برای گروههای بیشتری قابل استفاده باشد وجود ندارد.

در جریان تجربه نیز دیده می‌شد که آشنایی اکثریت افراد جامعه با معجمهای قرآنی و طرز استفاده از آنها ـ بویژه معجمهایی که براساس مادهٔ کلمات تنظیم شده ـ بسیار ناچیز است.

همین ضرورتها نظر مؤلّف را به تدوین معجم فوینی جلب کرد که در برگیرندهٔ کلیهٔ الفاظ قرآن و آمارهای مربوط به آن بوده و با روشی ساده تنظیم شده‌باشد.

بر این اساس هنگامی که شالوده‌ریزی کتاب «المعجم الإحصایی» آغاز گردید، تمام کوشش ما بر آن بود تا روشی را برگزینیم که ضمن رفع آن نیاز و حلّ این مشکل، اطّلاعات و نتایج حاصل از آن بتواند زیربنایی باشد برای مطالعات توصیفی آماری قرآن کریم. به همین دلیل طرح این کتاب با آنچه در غالب معجمهای قرآنی و کشف الآیاتها آمده، تفاوتی آشکار دارد، چه هدف عمدهٔ آنها معمولاً تسهیل امر دست‌یابی به آیات کلام الهی بوده است، درحالی که المعجم الإحصایی متضمّن مطالب دیگری هم هست که به اهمّ آنها اشاره می‌شود:

الف ـ بخش الفاظ. متن کتاب (ج ۲ و ۳)

ـ فهرست القبایی کلیهٔ کلمات قرآن با ضبط دقیق آنها
ـ فراوانی کلمات در قرآن و تفکیک آن برحسب مکّی و مدنی
ـ موضع کلمات در سوره‌ها و آیات برحسب نزول مکّی و مدنی
ـ ریشهٔ کلمات قرآن

* دکتر محمد جعفر یاحقّی، تحلیل تازه در پژوهشهای آماری قرآن کریم، کیهان فرهنگی، سال هشتم، شمارهٔ ۶، ص ۲۱، مهرماه ۱۳۷۰.

ـ الفاظ آیات مستثنیات (وجود آیات مکّی در سوره‌های مدنی و بالعکس) و مواضع آنها
ـ تعداد کلمات آغاز شده با هر یك از حروف الفبا با ذکر مجموع فراوانی آنها و نیز تفکیك این فراوانی به مکّی و مدنی

ب ـ المدخل (ج ۱)

۱ ـ مقدّمه
ـ تعداد آیات و کلمات هر یك از سوره‌ها
ـ تعداد آیات و کلمات قرآن در مآخذ دیگر
ـ تعداد و نام سوره‌ها به تفکیك مکّی و مدنی
ـ تعداد کلمات مشتق هر یك از سوره‌های قرآن برحسب حرف اوّل مادّهٔ آنها
ـ تعداد کلمات مشتق و غیرمشتقّ قرآن برحسب سوره‌ها
ـ تعداد آیات آغاز شده با هر یك از حروف الفبا
ـ تعداد و نام سوره‌هایی که در شمارهٔ آیات آنها اتّفاق‌نظر هست
ـ تعداد و نام سوره‌هایی که در شمارهٔ آیات آنها اتّفاق‌نظر نیست
ـ توزیع تعداد کلمات مدنی در سوره‌های مکّی به تفکیك بابهای الفبایی
ـ توزیع تعداد کلمات مکّی در سوره‌های مدنی به تفکیك بابهای الفبایی
ـ احصاء دقیق لفظ جلالهٔ (الله) در قرآن
ـ ذکر مزایا، نارسایی‌ها و اشکالات این معجمها:
الف ـ المعجم المفهرس لألفاظ القرآن الکریم
ب ـ فهرس الألفاظ «ملحق به قرآن مبین»
ج ـ نجوم الفرقان فی أطراف القرآن
د ـ الموسوعة القرآنیة المیسّرة (الجزء الخامس)
هـ ـ معجم الألفاظ و الأعلام القرآنیة
و ـ معجم ألفاظ القرآن الکریم
ز ـ المرشد إلی آیات القرآن الکریم و کلماته

۲ ـ ملاحظات
ـ الفاظ مشتق از هر یك از مادّه‌ها، تعداد و فراوانی مجموعی آنها، و نیز تفکیك این فراوانی به مکّی و مدنی
ـ تعداد مادّه‌های آغاز شده با هر یك از حروف الفبا
ـ فراوانی مجموعی الفاظ مشتق اومواد آغازشده باهریك ازحروف الفبا،ونیزتفکیك این فراوانی به مکّی و مدنی
ـ ارائهٔ آراء صاحبان معاجم قرآنی دربارهٔ کلماتی که درتعیین مادّهٔ آنها اختلاف نظر وجود دارد
ـ احصاء اعلام موجود در قرآن با ذکر مادّه و فراوانی آنها، و نیز تفکیك این فراوانی برحسب مکّی و مدنی
ـ معرفی اجمالی قدیمی‌ترین معجمهای قرآنی و کشف الآیاتهای زیر:
۱ ـ آیة الآیات فرقانی (اوّلین کشف الآیات)
۲ ـ ترتیب و بیا ـ دلیل الحیران فی الکشف عن آیات (آی) القرآن
۳ ـ هادیهٔ قطبشاهی در استخراج آیات کلام الهی (رسالهٔ هادیهٔ قطبشاهی)

۴ ـ کشف الآیات محمدرضا بن عبدالحسین النصیری الطوسی

۵ ـ نجوم القرآن (نجوم الفرقان)

۶ ـ کشف الآیات محمّد شاهی

۷ ـ نجوم الفرقان فی أطراف القرآن

۸ ـ مفتاح کنوز القرآن

۹ ـ مفتاح التفاسیر و مصباح الآیات (مصباح الآیات الجلیلة الفرقانیّة و مفتاح التفاسیر الجمیلة القرآنیة)

۱۰ ـ فتح الرّحمن لطالب آیات القرآن

۱۱ ـ ارشاد الرّاغبین فی الکشف عن آی القرآن المبین

۱۲ ـ المرشد إلی آیات القرآن الکریم و کلماته

۱۳ ـ المعجم المفهرس لألفاظ القرآن الکریم

۱۴ ـ کشف الآیات عن القرآن الکریم و التفاسیر (مفتاح التفاسیر و کشف الآیات)

۱۵ ـ معجم الفاظ القرآن الکریم

۱۶ ـ معجم الألفاظ و الأعلام القرآنیّة

۱۷ ـ فهرس الألفاظ (ملحق به قرآن مُبین)

۱۸ ـ مفتاح القرآن

۱۹ ـ الدّلیل الکامل لآیات القرآن الکریم

۲۰ ـ الموسوعة القرآنیّة المیسّرة (ج۳ و ۵)

۲۱ ـ کشف الآیات نوین (راهنمای کامل آیات با حروف ابجد)

۲۲ ـ معجم القرآن الکریم المفهرس (معجم أبجدی للآیات)

به این ترتیب **المعجم الإحصایی لألفاظ القرآن الکریم** به گونه‌ای تدوین شده‌است که هم خواص وصحّقان را به کار آید و هم عامّهٔ کسانی را که به نوعی با قرآن سروکار دارند سودمند افتد.

هر چند شناخت قلمرو و امتیازهای «المعجم الإحصایی» مطالعه مقدّمهٔ مفصّل کتاب را می‌طلبد با این حال اشاره‌ای مختصر به پاره‌ای از مزایای المعجم الإحصایی را خالی از فایده نمی‌دانیم:

ـ نخستین معجمی است که کلیّهٔ کلمات قرآن کریم را با ذکر مواضع آنها براساس قرائت کوفی، به دست می‌دهد، درحالی که دیگر معجمها مبتنی بر قرائت بصری مواضع آیات را مشخّص کرده‌اند و یا اگر طریقهٔ کوفی را هم به کار برده‌اند، تمام کلمات قرآن را احصا نکرده‌اند.

ـ در تعیین مواضع الفاظ برحسب مکّی و مدنی، به تفکیک ـ نه مثل دیگر معجمها توأماً ـ عمل شده‌است.

ـ آیات مستثنیات را با روش خاصّ خود نشان داده است.

ـ نحوهٔ ارائهٔ اعلام قرآن در این کتاب برآن دسته از معجمهای قرآنی که اعلام را مشخّص کرده‌اند مزیّتی آشکار دارد.

ـ ترتیب الفبایی این کتاب، دشواری استفاده از معجمهای تنظیم شده برحسب مادّه را برای کسانی که به زبان عربی آشنا نیستند، از پیش پای برداشته است.

ـ شماره‌گذاری مسلسل الفاظ در بابهای الفبایی ـ که تاکنون در هیچ یک از معجمهای قرآنی صورت نگرفته ـ امر ارجاع به کتاب را آسان کرده است.

ـ بخشی مربوط به فروش استثناء از معجم حاضر با مثالهای متعدّد به ضمیمهٔ «جداول راهنما» بد

خوانندگان در امر استفادهٔ بهتر و سهلتر از کتاب مدد می‌رساند.

ـ دو زبانه بودن مقدّمهٔ تحقیقی کتاب ـ فارسی و عربی ـ همراه با پیشگفتار انگلیسی، دامنهٔ استفاده از آن را وسیعتر کرده‌است.

قبل از آن که این گفتار را به پایان ببریم بی‌مناسبت نمی‌دانیم به برخی از نتایج علمی و فوایده‌ام حاصل از طرحی که در این کتاب دنبال شده است اشاره نماییم:

۱ ـ ارائهٔ جداگانهٔ مواضع الفاظ برحسب مکّی و مدنی، ضوابط جدیدی را برای شناخت این‌گونه آیات به دست داده است که می‌تواند ـ با استفاده از تفاسیر معتبر و کتب اسباب‌النّزول و... سرآغازی هم باشد برای مطالعهٔ تاریخ تنزیل قرآن که از سالها پیش توجّه محقّقان را به خود جلب کرده‌است.

۲ ـ ارائهٔ آیات مستثنیات باروشی که در این کتاب به کار رفته است امکان هرگونه جابجائی را برحسب مکّی و مدنی برای آن گروه از محقّقان علوم قرآنی که احیاناً از اقوال دیگری تبعیّت می‌کنند بسهولت فراهم می‌آورد.

۳ ـ با توجّه به این که بعضی از الفاظ قرآنی در واقع مَدخَلِ موضوعی هم به حساب می‌آیند، احصاء مواضع این‌گونه کلمات بویژه در ارتباط با آیات مکّی و مدنی، تا حدود زیادی حکم کشف المطالب هم پیدا کرده است. همچنین ذکر فراوانی تجمّعی کلماتِ مشتق از هر ماده در سوره‌ای می‌تواند سرآغازی برای مطالعهٔ موضوعی قرآن باشد.

۴ ـ ذکر مواضع الفاظ به تفکیک مکّی و مدنی، امکان مطالعهٔ دقیق وجوه معانی یك کلمه یا وجوه متضادّ آن را از این حیث فراهم می‌آورد.

۵ ـ اطّلاعات آماری دقیقی که ـ به گواهی فهرست جدولهای کتاب ـ برای اوّلین بار در این معجم فراهم آمده است می‌تواند به عنوان داده‌های اوّلیّه، فرضیّه‌ساز مطالعاتِ ریاضی ـ آماری قرآن، از جمله در ارتباط با نزول مکّی و مدنی، باشد.

۶ ـ با استفاده از این معجم، مطالعهٔ نظمهای آماری کلمات قرآن ـ از قبیل آنچه کتاب الإعجاز العَدَدی لِلقرآن الکریم به دست داده است ـ با توجّه به تفکیك آیات مکّی و مدنی قابل گسترش است.

۷ ـ و سرانجام اطّلاعات این کتاب می‌تواند بر روی دیسك یا دیسکت مغناطیسی کامپیوتر ضبط گردد و در دسترس مسلمانان و عموم علاقه‌مندان به علوم قرآنی قرار گیرد.

* * *

با همهٔ تلاش و ابتکاری که در شیوهٔ تنظیم این معجم به عمل آمده است، هرگز دعوی ما آن نیست که تنها با استفاده از این شیوه می‌توان به مطالعهٔ آماری قرآن پرداخت. آنچه اهمیت دارد نفس این اندیشه است که می‌تواند با استفاده از پیشرفتهای روزافزون علمی و فنّی به کمال مطلوب خود نایل آید. با این حال شایان ذکر است که راه صحیح مطالعه و تأمّل در قرآن را باید از خود قرآن جُست که «نُورٌ مُبِین» است و «تِبْیَاناً لِکُلِّ شَیْءٍ» و هیچ مرجعی همچون قرآن که «هُدًی لِلْعَالَمِینَ» است نمی‌تواند چشم آدمی را به روی معارف آن باز کند. استفاده از شیوه‌های علمی و فنّی تنها نگاه ما را برای استفادهٔ بیشتر از معارف قرآنی تجهیز می‌کند.

در این کتاب کوشش ما بر این بوده‌است که همه‌جا از منابع اصیل و متقن استفاده کنیم تا نتایج حاصله مطمئن و در خور اعتماد باشد، و هرجا که در موردی میان محقّقان اختلاف‌نظر وجود داشته است، آراء گوناگون را نقل کرده‌ایم تا هرکسی بتواند مطابق قول مورد قبول خود عمل کند.

تألیف و چاپ «المعجم الإحصائی» بر روی هم متجاوز از دوازده سال و بیش از بیست و شش هزار (۲۶۰۰۰) ساعت همّ و غمّ شبانه‌روزی مؤلّف را به خود مصروف داشته است.

گذشته از جستجوهای بسیار و مطالعهٔ کتابهای فراوان ـ که مقدمهٔ یکصد و سی و دو صفحه‌ای کتاب نمودار آن است ـ به منظور تعیین فراوانی و ثبت دقیق مواضع الفاظ قرآن کریم، کنترل و مقایسهٔ آنها با منابع دیگر و... در این تحقیق متجاوز از دو میلیون (۲/۰۰۰/۰۰۰) کلمه (کلمه ـ دفعه) به شمارش درآمده است. چون در مراحل عملی تدوین این معجم، بویژه در مرحلهٔ کنترل آماری تعداد و مواضع الفاظ، برای اطمینان از صحت آمارهای به دست آمده با قرآنی کار کردیم که در صفحهٔ ۲۱ جلد اول کتاب بدان اشاره شده‌است و نحوهٔ کنترل کلمات آن قرآن (نمونهٔ صفحه ۴۴ کتاب) تا حدی می‌تواند نمایانگر بخشی از تلاش خستگی ناپذیر مؤلف باشد، آن را به منظور ابقای نشانه‌ای از این کوشش مخلصانه به موزه قرآن آستانقدس رضوی اهداء می‌کنیم.

نظر به اینکه تنها این کتاب بر ارائهٔ آمار صحیح الفاظ قرآن از جهات گوناگون است و در این امر دقت و تفحص بسیار به عمل آمده است، آن را المعجم الإحصایی لألفاظ القرآن الکریم نامیدیم.

با همهٔ کوششها و موشکافیهایی که به هنگام تدوین و چاپ این کتاب برای برخیز از خطا و خلل صورت گرفته است، نمی‌توان ادعا کرد که از هرگونه سهو و اشتباهی کاملاً مصون مانده است که به فرمایش مولای متقیان علی علیه‌السلام:

إِذَا تَمَّ أَمْرٌ دَنَا نَقْصُهُ ... تَوَقَّعْ زَوَالاً إِذَا قِيلَ تَمّ*

امید است محققان و پژوهندگان علوم قرآنی ـ به سنّت دیرین عالمان ربانی ـ با اظهارنظر و تذکار نواقص و نقایص احتمالی کتاب، برمؤلف منت گذارند و پاداش و توفیق از پروردگار سبحان باز یابند.

... رَبَّنَا لَا تُؤَاخِذْنَا إِنْ نَسِينَا أَوْ أَخْطَأْنَا... (۲ : ۲۸۶)

---

* «چون تمام شود کار نزدیک شود نقصان او آشکار می‌گردد». متوقع زوال باش» چون گویند که تمام شد.

رک: دیوان حضرت امیر(ع)، ص ۱۲۲، جامع: ابوالحسن قطب الدین محمد بیهقی، نسخهٔ خطی، سال تحریر ۹۴۰ ق.

قسماً در بعضی نسخه‌ها چنین آمده است:

اذا تم امر بدا نقصه ... توقع زوالاً اذا قیل تم

وجوه تشابه و افتراق سرشماریهای عمومی نفوس و مسکن کشور

ابوالقاسم سادات دربندی

مرکز آمار ایران

تعریف سرشماری

نخست این توضیح عمده را متذکر می شوم که در این گفتار فقط در مورد سرشماریهای عمومی نفوس و مسکن صحبت می شود . و سایر انواع سرشماریها مانند سرشماری صنعتی یا سرشماری کشاورزی مورد نظر نیست

ساده ترین تعریفی که برای سرشماری نفوس و مسکن می توان عرضه کرد این عبارت است که :

ـ شمارش کلیه واحدهای مسکونی و جمعیت در یک کشور یا یک منطقه و محدوده مشخص را سرشماری می نامیم .

در یک تشریه سازمان اسکاپ که به نام : " توصیه هایی برای سرشماری های نفوس و مسکن سالهای دهه ۸۰ در کشورهای آسیایی و حوزه اقیانوس آرام " چاپ و منتشر شده است ، در مورد تعریف سرشماری آمده است که :

ـ " ... سرشماری جمعیت عبارت از عملیاتی شامل گردآوری ، تدوین ، استخراج و محاسبه ارزشیابی ، تجزیه و تحلیل و انتشار آمارهای اجتماعی ، اقتصادی و دموگرافی مربوط به تمام افراد یک کشور یا قسمت معینی از یک کشور در زمانی معین می باشد . "

در کتاب مذکور اشاره شده است که سرشماری جمعیت گاهی همراه سرشماری مسکن انجام می شود در مورد سرشماری مسکن هم اشاره شده است که :

ـ " ... سرشماری عبارت از عملیاتی شامل گردآوری ، تدوین ، استخراج و محاسبه ، ارزشیابی تجزیه و تحلیل و انتشار اطلاعات آماری درباره کلیه مکانهای مسکونی و خصوصیات آنها در تمام کشور یا قسمت مشخصی از یک کشور در زمانی معین می باشد . " (۱)

این تعاریف با سرشماریهایی که در قرون معاصر انجام شده است مطابقت دارد ولی سرشماریهای قدیمی و حتی باستانی با این تعاریف مطابقت ندارد .

سابقه سرشماریهای نفوس

قدیم‌ترین سند تاریخی که در آن به سرشماری نفوس اشاره شده است کتاب مقدس تورات است . در فصل مهاجرت آمده است که حضرت موسی برای مهاجرت قوم بنی اسرائیل از مصر کلیه افراد قوم یهود را سرشماری کرد . این سرشماری در سال ۱۴۹۱ قبل از میلاد (در حدود ۳۵۰۰ سال قبل) انجام گرفته است .

جالب است که به این نکته اشاره شود که در قدیم‌ترین کتابت تورات موجود " سِفْر یا فصل اعداد " شماره اقوام بنی اسرائیل و نفوس هر قوم را ارائه میدهد .

در همین فصل اشاره شده است که تعداد کسانی که قادر به جنگ بودند یعنی مردان ۲۰ ساله به بالا ، ۶۰۳۵۵۰ نفر بوده‌اند (فصل ۱ ـ آیه ۴۶) .

از همین جمله میتوان به خوبی هدف سرشماری نفوس را در آن دوران استنباط کرد که تعیین شمار مردان جنگی بوده است .

دومین گزارش ثبت شده در کتاب مقدس مربوط به زمان پیامبری حضرت داود است که اشاره به سرشماری نفوس در ۱۰۱۷ قبل از میلاد است . (۲)

در امپراطوری روم باستان نیز در فاصله سالهای ۲۳۵ قبل از میلاد تا ۱۴ بعد از میلاد نیز سرشماری نفوس بطور منظم هر پنج سال یکبار انجام می شده است . (۳)

---

(۱) ـ فهرست منابع و ماخذ در پایان گفتار می آید .

تصویر شماره ۱ ـ ۲ برگ از کتاب مقدس تورات . (تصویر صفحه بعد)

کتاب چهارم موسوم
معروف بسفر اعداد و مشتمل بر
سی و شش فصل

(فصل اول مشتمل بر نوزده و چهار آیه)

(۱) و در روز اول ماه دوم سال دوم خروج ایشان از زمین مصر خداوند موسی در بیابان سینا در خیمه
مجمع خطاب کرده فرمود (۲) که حساب تمامی جماعت بنی اسرائیل را موافق قبائل و خاندان آبای
ایشان با تعداد اسمای همگی ذکوران مطابق رؤس ایشان بگیرید (۳) از بیست ساله و زیاده تمامی
کسانی که در اسرائیل بجنگ بیرون میروند تو و هارون ایشان را مطابق افواج ایشان سان بینید (۴) و
از هر سبط یک نفر بهمراه شما باشد که هر یکی رئیس خاندان آبایش بوده باشد ۵ (۵) و اسمای مردمانی
که با شما خواهند ایستاد اینانند از روبین الیصور پسر شدیئور (۶) از شمعون شلومی ئیل پسر صوری شدای
(۷) از یهودا نحشون پسر عمیناداب (۸) از یساکار نتنئیل پسر صوعار (۹) از زبولون الیاب پسر حیلون
(۱۰) از پسران یوسف از افریم الیشامع پسر عمیهود از منسی گملی ئیل پسر پداهصور (۱۱) از بنیامین
ابیدان پسر جدعونی (۱۲) از دان اخیعزر پسر عمیشدای (۱۳) از اشیر فگیئیل پسر عکران (۱۴) از جاد
الیاساف پسر دعوئیل (۱۵) از نفتالی اخیرع پسر عینان (۱۶) اینانند ناموران جماعت و سروران اسباط
آبای ایشان و رؤسای هزارهای اسرائیل ۵ (۱۷) و موسی و هارون این مردمانی که اسماً مذکور شدند
گرفتند (۱۸) و تمامی جماعت را روز اول ماه دوم جمع کردند و دفتر تناسلها را موافق قبائل و خاندان
آبای ایشان از بیست سال و زیاده مطابق رؤس ایشان بتعداد اسمها بیان نمودند (۱۹) و موسی بنحوی
که خداوند باو امر فرموده بود ایشان را در بیابان سینا سان دید ۵ (۲۰) پس تناسلهای پسران روبین
اول زاده اسرائیل موافق قبائل و خاندان آبای ایشان هر ذکوری از بیست سال و زیاده تمامی کسانی
که بجنگ بیرون میرفتند به تعداد اسمها بر وفق رؤس ایشان این شد (۲۱) سانیدیده شدگان سبط
روبین چهل و شش هزار و پانصد نفر ۵ (۲۲) از پسران شمعون موافق دفتر تناسلهای قبائل و خاندان
آبای ایشان سان دیده شدگانش مطابق رؤس شان هر ذکوری از بیست سال و زیاده تمامی کسانی

نمونه‌ای از ساختمان دستنویس توراة
کرتهٔ سبتی تورات
مقدس

جنگ بیرون میرفتند به تعداد اسمها (۳۳) سال دیده شدگان سبط شمعون بنجاه هزار و سیصد * (۳۴) از پسران گاد موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بعداد اسمها (۳۵) سال دیده شدگان سبط گاد چهل و پنج هزار و ششصد و پنجاه * (۳۶) از پسران یهودا موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بعداد اسمها (۳۷) سال دیده شدگان سبط یهودا هفتاد و چهار هزار و ششصد * (۳۸) از پسران یساکار موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۳۹) سال دیده شدگان سبط یساکار بنجاه و چهار هزار و چهارصد * (۴۰) از پسران زبولون موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۴۱) سال دیده شدگان سبط زبولون بنجاه و هفت هزار و چهارصد * (۴۲) از پسران یوسف یعنی از پسران افرایم موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۴۳) سال دیده شدگان سبط افرایم چهل هزار و پانصد * (۴۴) از پسران منسی موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۴۵) سال دیده شدگان سبط منسی سی و دو هزار و دویست * (۴۶) از پسران بنیامین موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۴۷) سال دیده شدگان سبط بنیامین سی و پنج هزار و چهارصد * (۴۸) از پسران دان موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۴۹) سال دیده شدگان سبط دان شصت و دو هزار و هفتصد * (۵۰) از پسران اشیر موافق دفتر تناسل‌های قبائل و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۵۱) سال دیده شدگان سبط اشیر چهل و یک هزار و پانصد * (۵۲) از پسران نفتالی موافق دفتر تناسل‌های و خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که بجنگ بیرون میرفتند بتعداد اسمها (۵۳) سال دیده شدگان سبط نفتالی پنجاه و سه هزار و چهارصد بودند * (۵۴) اینانند سال دیدگانی که موسی و هارون با سروران اسرائیل دوازده نفر هر یک نفر از برای خاندان آبایش سان دیدند (۵۵) و تمامی سال دیدگان بنی اسرائیل موافق خاندان آبای ایشان از بیست سال و زیاده تمامی کسانی که در اسرائیل بجنگ بیرون میرفتند چنین بودند (۵۶) یعنی تمامی سال دیدگان ششصد و سه هزار و پانصد و پنجاه نفر بودند * (۵۷) اما لاویان از سبط آبای خودشان درمیان آنها سان دیده نشدند (۵۸) چونکه خداوند موسی را چنین فرموده بود (۵۹) که مذرستی که سان سبط لاوی را نمیبین و حساب ایشان را در میان بنی اسرائیل مگیر (۶۰) اما لاویان را بر مسکن شهادت با تمامی اسبابش و بر هر چه که بدست او دارد نصب نما که ایشان مسکن و تمامی اسبابش را بر

در امپراطوری ایران باستان نیز سرشماری های عمومی جمعیت برای دریافت مالیات سرانه و تعیین شمار مردان قادر به جنگ انجام می شده است .

هرودوت ، مورخ یونانی که ۴۸۴ تا ۴۲۵ قبل از میلاد می زیسته است در تاریخ خود پیرامون دوره تاریخی هخامنشی ها می نویسد که :

" ... وقتی خشایارشاه از تنگه داردانل می گذشت ، محل مخصوصی را برای سرشماری افراد قشون خود مناسب دید مقرر داشت تا ده هزار سپاهی شمارش شده در آن محل گرد آیند به نحوی که حتی المقدور به یکدیگر نزدیک باشند ، سپس دستور داد دور این عده را خط بکشند . بعد از خروج آنان از داخل خط دیواری که بلندی آن تا کمر یک سپاهی برسد بر روی آن بنا کنند . چون حصار آماده شد ، سپاهیان گروه گروه در آن داخل و خارج شدند تا مقدار آنها معلوم گردد ... ۳ .

هرودوت نتیجه این سرشماری را یک میلیون و هفتصد هزار نفر ذکر می کند . (۴)

سرشماری برای دریافت مالیات سرانه در امپراطوری ایرانیان رواج فراوان داشت . کریستین سن ایرانشناس دانمارکی ، که در تاریخ دوره ساسانیان تحقیقات فراوان دارد نوشته است که در دوره خسرو انوشیروان ساسانی که همزمان با تولد حضرت محمد (ص) و آستانه ظهور اسلام است ، احصائیه های دولتی برای تعیین تعداد افراد ۲۰ تا ۵۰ ساله که مشمول مالیات سرانه می شدند و تعداد دختران و حیوانات و ثروت ساکنین هر ایالت انجام می شد . این احصائیه ها که در دفاتر مخصوصی ثبت و نگهداری می شد به نام شاهنامار به معنای سرشماری سلطنتی خوانده می شد (۵)

در دوره اسلامی و بعد از آن نیز گمان بیش این رسم دیرینه ادامه یافته است و هموار ه ملاک و معیار مالیاتهای سرانه و جزیه و سایر امور سیاسی بوده است .

در سفرنامه ها و تاریخهای دوره های اسلامی به شمار دقیق دکانها و خانه ها و کاروانسراها و بعضی اماکن عمومی مثل مساجد و حمامها اشاره شده است که نشان می دهد شمارش و سرشماری رواج داشته است .

ولی لفظ سرشماری و مسکن امروزه معنای دقیق تری یافته است و همانطور که در مقدمه این گفتار اشاره شد به مجموعه فعالیتها و خدماتی اطلاق می شود که شامل طراحی ، برنامه ریزی ،

گردآوری ،گذاری وطبقه‌بندی ،اوزشیابی واصلاح وتجزیه وتحلیل ونمایندا" چاپ وانتشـــــار اطلاعات آماری ازجمعیت ومسکن می‌باشد .

این‌گونه‌سرشماری مطابق توصیه‌های سازمان‌ملل‌متحددراکثریت‌کشورهای جهـــــــان در فاصله های‌مساوی ده‌ساله (ودربعضی‌کشورهاپنج‌ساله) ازابتدای‌قرن‌بیستم‌رایج شده‌است .

درایران میعادازنهضت مشروطیت باقدامات قانونی‌برای‌سرشماری‌وتهیه‌آمارهای‌جمعیتی ازسال ۱۲۹۵ خورشیدی آغازشده‌است .

درآن‌سال‌برای‌تاسیس‌ثبت احوال‌وآمارتصویب‌نامه هیات وزیران‌صادرشدواقدامات‌قانونی برای‌ثبت وقایع حیاتی ومدورشناسنامه که سجل‌احوال‌خوانده‌می‌شدآغاز گردید . این اقدامـــات ازتهران شروع شدوبه‌تدریج به‌سایرنقاط‌کشورراه‌یافت . درسال ۱۳۰۲ تصویب نامه‌منکـــــور درمجلس‌شورای‌ملی‌مطرح‌شدورسمیت‌بیشتری‌یافت .

درسال ۱۳۱۴ طرحی‌برای‌تشکیل‌شورایعالی‌آمارارائه‌شدوبرای‌تربیت‌آمارگیرآموزشگـــــاه اختصاصی‌وزارت‌کشور درسازمان‌ثبت‌احوال تشکیل‌گردید .

درهفتم خردادماه ۱۳۱۸ قانون‌سرشماری در ۷ ماده و ۲ تبصره‌درمجلس‌شورایملی‌تصویب شد وبه‌دولت ابلاغ گردید .

سرشماری ۱۳۱۸ که به‌نام سرشماری یکروزمعروف شدتاسوم‌شهریور سال ۱۳۲۰ ادامـــه داشت . دراین‌سرشماری حدود ۲۰ هزارنفرازکارمندان‌ادارات دولتی‌شرکت داشتند . سرشماری از شهرهاشروع شد . درهرشهریکروزتعطیل‌عمومی‌اعلام‌شدومردم‌بایستی درخانه‌های‌خودبمانندتـــا سرشماری شوند . بااین‌روش ۳۵ شهرسرشماری‌گردید .

گزارش نتایج این‌سرشماری به‌صورت‌نشریه‌جداگانه‌منتشرنشده‌ولی‌سرجمع‌ارقام‌درمقدمـــه گزارش‌مشروح سرشماری ۱۳۲۵ ذکرشده‌است . به‌همین‌جهت درتاریخچه سرشماریهای‌عمومـــی ایران‌ومقایسه‌هایی‌که‌مطرح‌می‌شود،سرشماری‌یکروزه‌منظورنمی‌شود .

آنچه‌امامفهوم‌سرشماری‌بامفهوم‌جهانی‌وامروزی‌آن‌مطرح است سرشماری‌عمومی‌سالهـــــای

۱۳۳۵ ـ ۱۳۴۵ ، ۱۳۵۵ و ۱۳۶۵ کشور است که با فاصله های ده ساله انجام شده و نتایج آنها به صورت گزارشات مرتب و منظمی انتشار یافته است . مقایسه این سرشماریها ، و تعیین تفاوتها و شباهتها آنها راهنمای بسیار مفیدی برای بررسی وضع کشور و تعیین روند تغییرات طی این سالها می باشد .

روشهای سرشماریها

برای شمارش افراد در سرشماریها معمولا " یکی از دو روش زیر بکار می رود :

الف ـ روش سرشماری جمعیت حاضر یا واقعی (دوفاکتو De Facto) .

ب ـ روش سرشماری جمعیت قانونی (دوژور De Jure) .

در روش الف ، جمعیت حاضر در مکان خود فقط از اینکه دائمی یا موقت است ، حتی اگر مهمان است ، سرشماری می شود . یعنی غائبین شمارش نمی شوند . به همین جهت باید تعطیل عمومی اعلام شود و سفرها و جابجائی ها متوقف شود یا لااقل به حداقل ممکن برسد . در حال حاضر این روش طرفداران قلیلی دارد . نمونه ای از این روش سرشماری ، سرشماریهای کشور همسایه ، ترکیه است که با کمک نیروهای نظامی انجام می شود .

در روش ب ، افراد در محل قانونی خود یعنی محل زندگی یا محلی که اقامتگاه رسمی آنان است سرشماری می شوند . مثلا " مسافرین هتلها و مهمانخانه ها شمارش نمی شوند . بیشتر سرشماریهای عمومی جمعیت که در کشورهای مختلف جهان انجام شده است با این روش بوده است .

در ایران ترکیبی از این دو روش بکار می رود . زیرا سرشماری جمعیتهای غیرساکن به مراتب گرانتر و با روش دوژور نامقدور است با محتاج جستجوی فراوان و هزینه بسیار است . و به همین جهت جمعیتهای غیرساکن در سرشماریهای عمومی ایران با روش " دوژور " سرشماری می شوند ، یعنی هر کجا مشاهده شوند سرشماری می شوند . ولی بقیه جمعیت کشور با روش دوفاکتو سرشماری می شوند .

زمان مناسب برای مراجعه

در صورت تاریخ مراجعه و مصاحبه در سرشماریهای عمومی ایران معمولا " فصل پائیز به عنوان زمان

بهینه اعلام شده است. زیرا در این تاریخ در مناطق روستایی کار در مزرعه و صحرا تمام شده است، عشایر از ییلاق به قشلاق بازگشته اند و در شهرها به علت بازشدن مدارس مسافرت ها به حداقل رسیده است.

سرشماری های ۱۳۳۵ - ۱۳۴۵ - ۱۳۵۵ در مهرماه آبان بمدت ۲۰ روز انجام شد. ولی در سال ۱۳۶۵ چون آمارگیران بیشتر از دانشجومعلمان مراکز تربیت معلم بودند برای اجتناب از تعطیل کلاسها در میانه سال این زمان به ۱۵ روز نخست مهرماه تغییر داده شد. در سال ۱۳۷۰ نیز زمان بین ۱۸ شهریور تا ۵ مهرماه تعیین شده بود که عملا" تا ۱۰ مهرماه به طول انجامید.

پرسشهای سرشماریها

مقایسه ای از سوالات سرشماریها در جدول پیوست ارائه می شود. بررسی تطبیقی اقلام این جدول تغییرات سرشماریها را نمایان می سازد که در صورت فرصت داشتن در این جلسه مورد بحث قرار می گیرد.

فهرست اقلام مورد سئوال در سرشماریهای نفوس و مسکن ایران و توصیه های
سازمان ملل متحد (منطقه آسیا)

علائم : + توصیه شده یا پرسش شده
- توصیه نشده یا پرسش نشده

اقلام مورد سئوال	سرشماری ۱۳۳۵	سرشماری ۱۳۴۵	سرشماری ۱۳۵۵	توصیه آسیا	سرشماری ۱۳۶۵	سرشماری ۱۳۷۰
الف : اطلاعات جغرافیائی						
نام استان	+	+	+	+	+	+
نام شهرستان	+	+	+	+	+	+
نام شهر	+	+	+	+	+	+
نام بخش	+	+	+	+	-	-
نام دهستان	+	+	+	+	+	+
نام آبادی	+	+	+	+	+	+
نام خیابان	+	+	+	+	+	+
نام میدان	+	+	+	+	+	+
نام کوچه ـ بن بست و ...	+	+	+	+	+	+
شماره نامه	-	-	+	+	-	-
شماره شناسه	+	+	+	+	+	+
شماره قسمت	+	+	+	+	+	+
شماره حوزه	-	+	+	+	+	+
شماره بلوک شهری	-	+	+	+	+	+
شماره برزن شهرداری	+	-	-	+	-	-
شماره پلاک شهرداری	-	-	-	+	+	+
شماره اصلی و فرعی مکان	+	+	+	+	+	+
ب : اطلاعات مربوط به خانوار						
نوع خانوار	+	+	+	-	+	+
تعداد افراد خانوار	+	+	+	+	+	+
تعداد موتورهای خانوار	-	+	-	+	-	-
تعداد افراد مشغول در خانوار	-	-	+	-	-	+
وجود تأسیسات خانگی در خانوار	+	+	+	-	+	+
نوع صنعت خانگی	+	+	+	-	+	+

اقلام مورد سئوال	سرشماری ۱۳۲۵	سرشماری ۱۳۴۵	سرشماری ۱۳۵۵	تجربه اسکاپ	برنامه ریزی ۱۳۶۵	آمارگیری ۱۳۷۰
تعداد افراد شاغل در صنعت خانگی	+	-	+	-	-	-
ج : اطلاعات فردی						
نام و نام خانوادگی	+	+	+	+	+	+
بستگی و خویشاوندی	+	+	+	+	+	+
وضع اقامت در میان مورد سرشماری	+	+	+	+	+	+
جنسیت	+	+	+	+	+	+
سن - تاریخ تولد	+	+	+	+	+	+
محل تولد	+	+	+	+	+	+
وضع تابعیت	+	-	-	+	+	-
دین و مذهب	+	+	+	+	+	-
تاریخ اقامت در محل مورد سرشماری	-	-	-	+	+	S +
نام محل اقامت قبلی (پنجسال قبل)	-	-	+	+	+	S +
وضع زناشوئی	+	+	+	+	+	S +
تاریخ ازدواج	-	-	-	+	-	S +
تعداد فرزندان زنده متولد شده زنان	-	-	-	+	+	S +
وضع سواد	+	+	+	+	+	+
وضع تحصیل	+	+	+	+	+	S +
پایه و مدرک تحصیلی	+	+	+	+	+	S +
رشته تحصیلی	-	+	+	+	+	S +
وضع فعالیت اقتصادی	+	+	+	+	+	+
نوع شغل	+	+	+	+	+	S +
نوع فعالیت عمدهٔ محل کار	+	+	+	-	+	S +
وضع شغلی (سمت در شغل)	+	+	+	+	+	S +
ساعات کار هفتگی	-	+	+	+	-	-
زبان مادری	-	-	-	+	+	-
نوع معلولیت	-	-	+	-	+	-
غیرمعلولیت	-	-	+	-	+	-

اقلام مورد سئوال	سرشماری ۱۳۳۵	سرشماری ۱۳۴۵	سرشماری ۱۳۵۵	توصیه اسکاپ	سرشماری ۱۳۶۵	سرشماری ۱۳۷۵
د : اطلاعات مسکن و ساختمان						
نوع واحد مسکونی (خانه، آلونک، زاغه ...)	-	+	+	+	+	-
نحوه تصرف (ملکی، اجاری، ...)	+	+	+	+	+	-
تعداد خانوار در واحد مسکونی	-	+	+	+	+	-
تعداد اطاق در اختیار خانوار	-	+	+	+	+	-
تعداد اطاق در واحد مسکونی	-	+	+	+	+	-
مصالح عمدهٔ ساختمان واحد مسکونی	-	+	+	+	+	-
سن ساختمان واحد مسکونی	+	+	+	+	+	-
نحوهٔ آب رسانی واحد مسکونی	-	+	+	+	+	-
منابع تأمین آب واحد مسکونی	-	+	+	+	+	-
نوع روشنائی واحد مسکونی (برق)	-	+	+	+	+	-
هـ : سایر اطلاعات جنبی						
داشتن رادیو - تلفن - تلویزیون	-	-	+	-	-	-
نوع سوخت مصرفی خانوار	-	+	+	-	+	-
مالکیت زمین کشاورزی	+	-	-	-	-	-
داشتن دام و طیور خانوار	+	-	-	-	-	-

توضیح : توصیه های اسکاپ (سازمان ملل) از نشریهٔ : توصیه های سرشماریهای جمعیت ومسکن - سری M شماره ۶۷ - به زبان فرانسه که در ۱۹۸۰ در نیویورک منتشر شده است آمده شده است .

+ S : این سئوال در نمونه گیری منظور شده است .

فهرست منابع وماخذ :

۱ ـ توصیه‌هایی برای سرشماریهای نفوس ومسکن

سازمان ملل متحد ـ اسکاپ ـ بانکوک ۱۹۷۸

ترجمه فارسی توسط : مرکز آمار ایران ـ خرداد ۱۳۶۳

۲ ـ ترجمه کتاب مقدس (از اصل عبرانی به زبان فارسی)

سفر خروج وسفر اعداد

۳ و ۴ ـ اصول سرشماری ـ جزوه درسی دانشگاه تهران

دانشکده علوم اجتماعی وتعاون ـ نیمسال دوم تحصیلی ۵۲ / ۱۳۵۲

۵ ـ تاریخ تمدن ایران ساسانی ـ تالیف سعید نفیسی

چاپ دانشگاه تهران ـ سال ۱۳۳۱

۶ ـ اسناد ونشریات سرشماری عمومی نفوس سال ۱۳۳۵ ـ وزارت کشور ـ اداره کل آمار عمومی

۷ ـ اسناد ونشریات سرشماریهای عمومی نفوس ومسکن سالهای ۱۳۴۵ ـ ۱۳۵۵ ـ ۱۳۶۵ وآمارگیری

جاری جمعیت سال ۱۳۷۰ .

* پایان *

منابع خطا در امارگیریها

محمدباقر سخاوت

مرکز آمار ایران

۱ ـ مقدمه

همه افرادی که با بررسیهای آماری و به ویژه تهیه و اجرای طرحهای تولید آمار سروکار دارند به خوبی می‌دانند که در کار آمارگیریها موضوع پراهمیتی تحت عنوان میزان صحت و اعتبار داده های آماری وجود دارد که هیچگاه نباید از آن غافل باشیم و بلکه باید توجه داشته باشیم که اگر اهمیت نقش آمار در دنیای کنونی بدان پایه است که نمی‌توان به جهت کاربرد آن در اکثر زمینه ها از آن چشم پوشید ، ناگزیر باید به فکر صحت آمار و اطلاعاتی که جمع آوری می‌کنیم نیز باشیم و کاری نکنیم که صرفا" برپایه خوش بینی بیش از حد ، شالوده ای سست و نامطمئن را پایه بنایی عظیم قرار دهیم .

بدیهی است که میزان صحت داده های آماری بستگی به نوع و میزان خطاهائی دارد که در مراحل مختلف کار ممکن است رخ دهد و هر چقدر بتوانیم با استفاده از شیوه های مناسب ، منابع مختلف خطا را به گونه ای بهتر تحت کنترل درآوریم بر میزان صحت و اعتبار داده ها افزوده ایم .

در اینجا ذکر دو نکته ضروری به نظر می‌رسد . نکته اول اینکه هدف از کنترل خطا برطرف کردن کامل خطا نیست بلکه کاهش خطای کل تا حدی است که نتایج معقول و قابل قبولی حاصل شود ، نکته دیگر اینکه مطرح کردن موضوع خطاهای مختلف به هیچ وجه به معنای نفی استفاده از اطلاعات در برنامه ریزیها و بررسیهای اجتماعی ـ اقتصادی نیست بلکه بدین خاطر است که با توجه دادن دست اندرکاران بررسیهای آماری به لزوم شناسائی منابع خطا و کنترل آنها ، از داده‌های آماری و علم آمار که امروزه می‌رود تا فرا راه بشری در حل بسیاری از معضلات و مشکلات قرار گیرد استفاده معقولی به عمل آید .

۲ ـ هدف و اهمیت موضوع

چنانچه در آمارگیریها فرض کنیم مقداری که برای هر متغیر در صورت هرکدام از واحدهای آماری به دست می‌آید دارای هیچ گونه خطایی نیست در این صورت در زمینه سرشماریها می‌توان پذیرفت آنچه که تحت عنوان نتایج سرشماری مسرای کل جامعه منتشر می‌شود با مقادیر واقعی تفاوتی نخواهد داشت ، و خود

آمارگیریهای نمونه‌ای با قبول فرض فوق تنها می‌توان گفت که برآورد های حاصل دارای خطایی است که اصطلاحا " اشتباه یا خطای نمونه‌ای نامیده می‌شود وخطای مزبور ناشی از این امر است که به جای اینکه همه واحدهای جامعه ، مورد آمارگیری قرار گیرند، تنها واحد به عنوان نمونه انتخاب ونتایج حاصل به کل جامعه تعمیم داده می‌شود . درعمل فرض فوق تنها درآمارگیریهای بسیار ساده‌ای صادق است که کیفیت کار درحد بسیار بالائی باشد وعملا " احتمال بروز خطا درمراحل مختلف کار ، صفر یا نزدیک به صفر باشد . در آمارگیریهای پیچیده به ویژه وقتی که به نحوی با مردم که بالطبع دارای رفتار وکردارهای متفاوت هستند سروکار داشته باشیم فرض فوق باواقعیت یعنی آنچه درعمل رخ می‌دهد تفاوت بسیار زیادی خواهد داشت و درنتیجه علل متعددی برای وجود خطا درداده های آماری ظهور پیدا خواهد کرد که برخی از این علل عبارتنداز :

ـ نامناسب بودن چارچوب آماری ، فرمها ودستورالعمل های مورد استفاده

ـ عدم موفقیت درآمارگیری از پاره‌ای از واحدهای آماری

ـ وجود خطا درانداز‌ه‌گیری متغیرهای موردبررسی

ـ وجود خطا درمراحل ادیت ، کد گذاری واستخراج نتایج

خطاهای فوق که بروز آنها هم درسرشماریها وهم درنمونه گیریها محتمل است اصطلاحا " اشتباهات یا خطاهای غیرنمونه‌ای نامیده می‌شوند ودرواقع این خطاهامی‌توانند درکلیه مراحل یک بررسی آماری یعنی از مرحله برنامه‌ریزی وتهیه طرح تامرحله استخراج وانتشار نتایج روی دهند .

برپایه مطالب فوق ملاحظه می‌شود که درسرشماریها خطاهای غیرنمونه‌ای ودرآمارگیریهای نمونه ای علاوه برخطاهای غیرنمونه‌ای ، خطاهای نمونه‌ای نیز وجود دارد . هرچند بنابه نظر اکثر متخصصین آمار ، دراغلب آمارگیریهای نمونه‌ای ، خطاهای غیرنمونه‌ای بیشتر از خطاهای نمونه‌ای است وبخش عمده‌ای ازکل خطارا خطاهای غیرنمونه‌ای تشکیل می‌دهند ، لکن از آنجا که تکنیک کنترل خطاهای نمونه ای وبحاسبه آنها درمقایسه با خطاهای غیرنمونه‌ای ساده تراست لذا به خطاهای نمونه‌ای اهمیت بیشتری داده می‌شود وتالبا " برای کاهش آن ، تعداد نمونه‌ها را افزایش می‌دهند ، بدون توجه به آنکه این امر موجب افزایش خطاهای نوع دوم یعنی خطاهای غیر نمونه‌ای می‌شود . بدیهی است این موضوع درمورد

مقایسه سرشماریها با آمارگیریهای نمونه‌ای نیز صادق است زیرا نباید فراموش کرد که وقتی جمع آوری اطلاعات براساس نمونه‌گیری انجام می‌پذیرد امکان کنترل کار وانجام نظارتهای لازم به مراتب بیشتر از زمانی است که سرشماری انجام می‌شود وبدین ترتیب می‌توان انتظار داشت نتیجه ای که از نمونه گیری حاصل می‌شود به مقدار واقعی نزدیک تر باشد.

آنچه از مطالب فوق استنباط می‌شود این است که شناسایی منابع خطا درآمارگیریها وبه کارگیری شیوه های مناسب برای کنترل آنها از اساسی ترین کارهای هرآمارگیری است ودرواقع مقاله حاضر نیز درهمین راستا تهیه شده است تا علاقه مندان را هرچند به طور مختصر با منابع خطا درآمارگیریها و به ویژه خطاهای غیرنمونه‌ای آشنا سازد. ضمناً" ذکر این نکته ضروری است که هرچند دراین مقاله فقط منابع مختلف خطا به اجمال مورد بحث قرار گرفته است لکن درارتباط با بررسی موضوع خطا درآمارگیریها، روشهای کنترل وارزیابی خطاها نیز مقوله های دیگری هستند که باید همگام با شناخت منابع خطا مورد توجه وبررسی قرارگیرند.

۲ ـ اریبی‌ها وخطاهای تصادفی

خطاهای مختلف اعم از نمونه‌ای وغیرنمونه‌ای رامی‌توان به دوسته خطاهای تصادفی واریبی‌ها تقسیم بندی کرد. درعمل اریبی های غیرنمونه‌ای وخطاهای تصادفی نمونه‌ای مهمتراز دونوع دیگر یعنی اریبی‌های نمونه‌ای وخطاهای تصادفی غیرنمونه‌ای می‌باشند. بین اریبی‌ها وخطاهای تصادفی، تفاوتهای زیادی وجود دارد. اریبی‌ها ثابتهایی هستند که از شرایط اصلی آمارگیری تبعیت می‌کنند ومیزان آنها غالباً" مجهول باقی‌می‌ماند. کل اریبی، جمع جبری اریبی های منابع مختلف است که بعضی ازآنها ممکن است مثبت وبرخی دیگر منفی باشند وبنابر این اغلب بخشی از اریبی ها به وسیله یکدیگر خنثی می‌شود. ازطرف دیگر خطاهای تصادفی فرم مثبت $\sum \frac{S_h^2}{m_h}$ را بسته خودمی‌گیرند که S_h^2 و m_h به ترتیب واریانس وتعداد نمونه مربوط به مولفه h ام می‌باشد. اریبی هارامی‌توان به وسیله بهترکردن چیزی مثلاً" بهبود کیفیت بعضی‌از عملیات، آموزش بهتر ماموران آمارگیری، معاجه بهتر باپاسخگویان وبابین نگهداشتن اظهار نشده ها کاهش داد ولی کم کردن خطاهای تصادفی ازطریق افزایش بعضی‌از چیزها ازقبیل تعداد نمونه میسر است. گاهی اوقات به وسیله کم کردن واریانس

ازطریق طبقه بندی واحد ها نیز می‌توان خطاهای تصادفی را کاهش داد .

۴ـ منابع خطاهای غیرنمونه‌ای

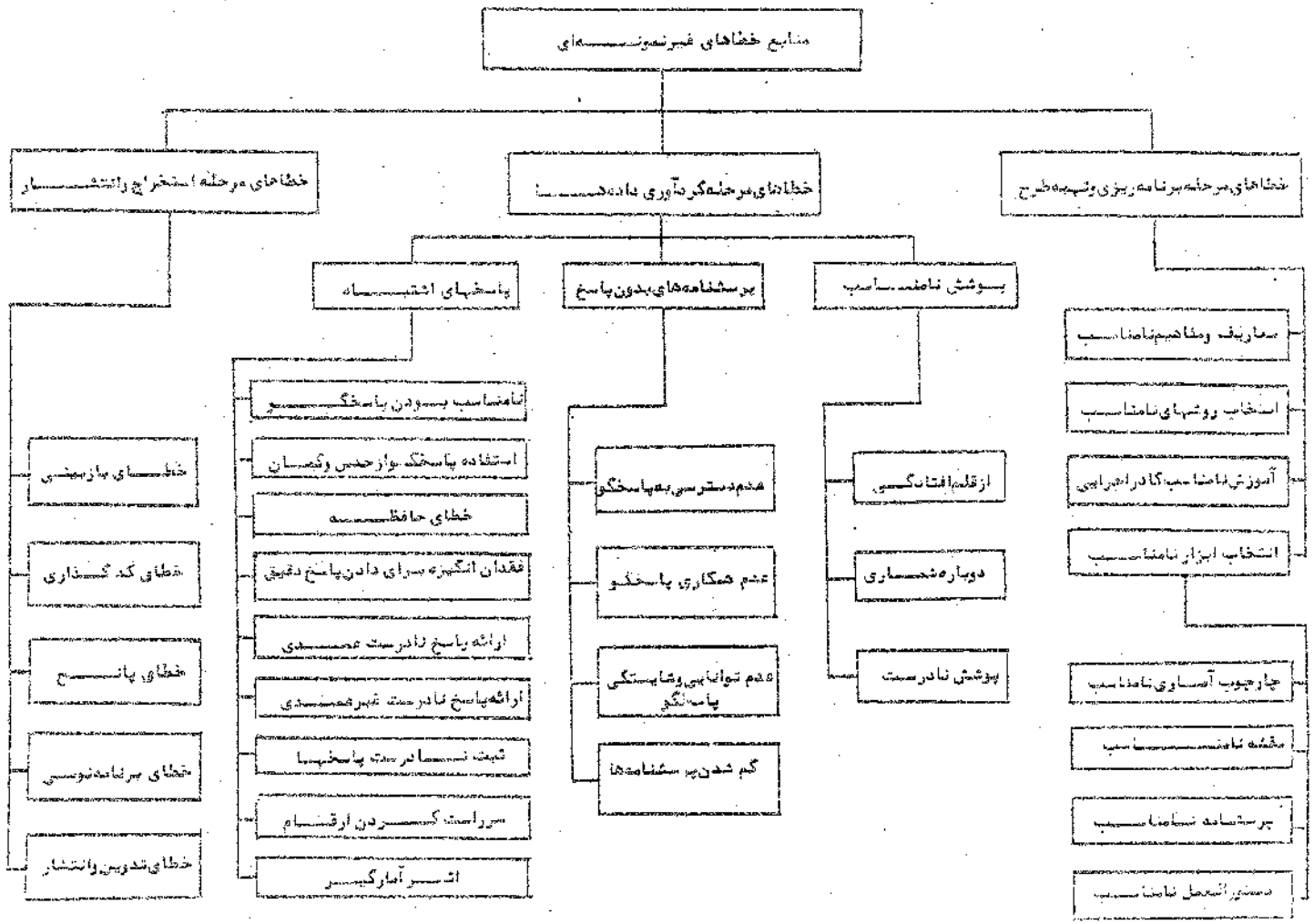
همان طورکه قبلا " گفته شد خطاهائی که از منبعی بجز نمونه‌گیری ناشی می‌شوند اصطلاحـــــا " به نام خطاهای غیرنمونه‌ای نامیده می‌شوند . دراصل ، هرکدام از عملیات مربوط به آمارگیریها مجمـــــوع بالقوه‌ای برای خطاهای غیرنمونه‌ای به حساب می‌آیندوبه لحاظ اهمیتی که کنترل این گونه خطاهـــــا درآمارگیریهای مختلف اعم از سرشماریها و آمارگیریهای نمونه‌ای دارد لازم است که منابـــع بروز آنهـــا به خوبی شناخته شود ، هرچند گاهی اوقات به دلیل وسعت وگستردگی خطاهای غیرنمونه‌ای شناخت همـــه آنها خود ، کار دشواری است .

منابع خطاهای غیرنمونه‌ای رامی‌توان به سه گروه عمده خطاهای مرحله برنامه‌ریزی وتهیه طـــــرح ، خطاهای مرحله گردآوری داده ها وبالاخره خطاهای مرحله استخراج و انتشار نتایج تقسیم بندی کرد . در ذیل ، نمودار مربوط به انواع خطاهای غیرنمونه‌ای ارائه شده ودرباره زمینه‌های بروز هرکدام، به طورمختصر توضیحاتی داده شده است .

۱ـ۴ـ خطاهای مرحله برنامه ریزی وتهیه طرح

دراین مرحله ، اهداف طرح ، جامعه آماری ، چارچوب آماری وروش جمع آوری آمار واطلاعات تعییـــن می‌شود ، همچنین فرمها ، نقشه‌ها ، پرسشنامه‌هاودستورالعملها تهیه وتنظیم می‌گردد . درمورد آمارگیریهـــای نمونه‌ای ، روش نمونه‌گیری ، روش انتخاب واحدهای نمونه ، فرمولهای برآورد مشخص می‌شود وبالاخره دراین مرحله مأموران آمارگیر وپرسنل سایر رده‌های اجرایی انتخاب وبه آنان آموزشهای لازم داده می‌شود .

بدیهی است که در هریک ازموارد فوق احتمال بروز خطاوجوددارد ودرمجموع می‌توان گفت خطاهایی که دراین مرحله به دلیل پائین بودن کیفیت طرح روی می‌دهد غالبا " ازنوع خطاهایی هستندکه درنتایج ، اریبی ایجـــاد می‌کنندو به همین دلیل درموارد ی ممکن است نتایج رابه میزان زیادی تحت تاثیر قرار دهند . به عنوان مثال اگـــر بعضی ازمفاهیم به گونه‌ای تعریف شوندکه برداشت نادرستی را القانمایدویا اصولا " تعاریف آنها مبهم باشد این امـــــر می‌توانـــد نتایـــج را دچـــــار انحراف سازد . حتـــــی انتخاب زمان نامناسب ـــــــــرای بررسی ممکن است روی نتایج تاثیـــــــر نامطلوب داشتـــــه باشد ، مثـــــلا "


منابع خطاهای غیرنمونه‌ای
خطاهای مرحله برنامه‌ریزی و تهیه طرح
خطاهای مرحله‌گردآوری داده‌ها
خطاهای مرحله استخراج و انتشار
سرپوشش نامناسب
پرسشنامه‌های بدون پاسخ
پاسخهای اشتباه
تعاریف و مفاهیم نامناسب
انتخاب روشهای نامناسب
آموزش نامناسب کادر اجرایی
انتخاب ابزار نامناسب
چارچوب آماری نامناسب
نمونه نامناسب
پرسشنامه نامناسب
دستورالعمل نامناسب
از قلم افتادگی
دوباره‌شماری
پوشش نادرست
عدم دسترسی به پاسخگو
عدم همکاری پاسخگو
عدم توانایی و بایستگی پاسخگو
گم شدن پرسشنامه‌ها
نامناسب بودن پاسخگو
استفاده پاسخگو از حدس و گمان
خطای حافظه
فقدان انگیزه برای دادن پاسخ دقیق
ارائه پاسخ نادرست عمدی
ارائه پاسخ نادرست غیرعمدی
ثبت نسنجیده پاسخها
سرراست کسر دن ارقام
اثر آمارگیر
خطای بازپرسی
خطای کدگذاری
خطای پانچ
خطای سرشماره‌نویسی
خطای تدوین و انتشار

اگر بررسی درزمینه درآمد افراد به فاصله کوتاهی بعداز تصویب پاره‌ای مقررات مالیاتی انجام پذیرد بابخشهاخواه‌اخواه تحت تاثیر مطالبی‌که افراد درمورد آن مقررات خوانده یا شنیده اند قرارمی گیرد . بنابراین حتی‌فضای اجتماعی حاکم برجامعه روی نتایج اثر می‌گذارد .

درتهیه طرحهای آماری ، به منظور کاهش هرچه بیشتر خطاهای غیرنمونه ای باید کوشش شود کلیه تجارب بررسیهای گذشته درکشور وحتی تجربیات سایر کشورها مورد استفاده قرارگیرد زیرا این کارموجب می‌شود که حداقل بعضی از خطاهای مهمی که درگذشته رخ داده است تکرارنشود . به هرصورت می‌توان گفت که تهیه واجرای یک طرح آماری کاری است که درآن ضمن رعایت اصول علمی ، تجارب موجودنیز بایه مورد استفاده قرارگیرد .

خطاهای مرحله برنامه‌ریزی وتهیه طرح یعنی مرحله قبل از گردآوری داده ها رامی‌توان به نحو زیر طبقه بندی نمود :

۱-۱-۲. خطای تعاریف ومفاهیم نامناسب

درهرآمارگیری ، تعاریف ومفاهیم ، پایه واساس کاررا تشکیل می‌دهد . دراین مرحله باید هدفهای آمارگیری به روشنی مشخص شود وهمزمان باآن درباره شکل ومحتوای جداول نهایی تصمیم گیری به عمل آید . علاوه براین لازم است جامعه آماری دقیقا" مشخص گردد تا اطمینان حاصل شود که باهداف بررسی هماهنگی کامل دارد . باتوجه به اینکه دربررسیهای آماری همواره یکنواخت کردن شیوه کار نیز مورد نظراست ، بنابراین دردستورالعمل آمارگیری باید تعاریف مربوط به هرکدام از اصطلاحات آورده شود تا از به‌کار بردن تعبیرات ذهنی مختلف توسط ماموران آمارگیر وعدم هماهنگی بین برداشتهای آنان ازاصطلاحات مورد استفاده جلوگیری به عمل آید . ضمنا" درآمارگیریها باید کوشش شود که حتی‌الامکان تعاریف و مفاهیم استانداردمورد استفاده قرار گیرد تا علاوه براطمینان از مناسب بودن آنها ، امکان مقایسه داده های جمع آوری شده با اطلاعات مشابه فراهم باشد .

۱-۱-۳. خطای انتخاب روشهای نامناسب

دربررسیهای آماری ، یکی از مسایل بسیارمهم ، تعیین روش جمع آوری اطلاعات است زیرا بیشترین خطای غیرنمونه‌ای درمرحله گردآوری داده ها روی می‌دهد وبه همین دلیل است که درمورد این مرحله از کار

علاوه برهمه کنترلهای معمول ، نظارت آماری مستقیم نیز ضرورت پیدا می‌کند . تعیین روش جمع‌آوری اطلاعات از جمله مسایل مهم آمارشناسان است زیرا تصمیماتی که دراین ارتباط اتخاذ می‌شود خطاهای مرحله گردآوری داده ها را تحت تاثیر قرار می‌دهد .

چگونگی اندازه گیری صفات مورد بررسی ومقیاسها و معیارهای انتخاب شده برای اندازه گیری نیز روی دقت نتایج اثر می‌گذارد . مثلا " درتعیین سن افراد ، روشن است که اگر سن آنان ازروی شناسنامه تعیین شود به مراتب بهتر از این است که به پاسخگو واگذار شود زیرا حداقل تاثیر گمراهی افراد به اظهار اعداد خاص ، مثلا " اعدادی که به ارقام صفر و پنج ختم می‌شوند تا حدود زیادی کاهش خواهد یافت .

روش انتخاب واحدهای نمونه وهمچنین شیوه‌ای که درمورد واحدهای بدون پاسخ اتخاذ می‌شود خود نیز از مسایلی است که میزان خطای غیرنمونه‌ای را تحت تاثیر قرارمی‌دهد . ساده ترین نحوه عمل در برخورد با واحدهای بدون پاسخ این است که به این امر اساسا " توجهی نشود وبرواقع نتایج براساس اطلاعات موجود ، استخراج ومحاسبه گردد . بدیهی است این روش زمانی صحیح است که وضع واحدهایی که آمارگیری از آنها میسر نگردیده است به طور متوسط با واحدهای آمارگیری شده تفاوت چندانی نداشته باشد . دربیشتر روش دیگر این قبیل واحدها به وسیله واحدهای دیگری جایگزین می‌شوندوبدیهی است که بااین کار فرض می‌شود که واحدهای جایگزین تا حدود زیادی دارای همان خصوصیات واحدهای بدون پاسخ هستند .

درپاره‌ای از بررسیهای آماری ، اطلاعات مربوط به اقلامی که به دلایل مختلف دربرگشته ثبت نشده است به روشهای مختلفی برآورد می‌گردند . علاوه براین گاهی اوقات پاسخهایی که غلط به نظر می‌رسد تغییر داده می‌شوندبه جای آن اطلاعاتی که معقول به نظر می‌رسد جایگزین می شود ، تغییرات های عمل درهردو مورد فوق می‌تواند مشمول خطا باشد زیرا احتمال اینکه از اطلاعات نادرستی برای جایگزینی اطلاعات استفاده به عمل آید وجود دارد وحتی ممکن است اطلاعی که به گمان نادرست بودن تغییر داده می‌شود واقعا " درست باشد .

۳-۱-۴-خطای آموزش نامناسب کارمراجرایی

درپرسشنامه‌هایی که از آمارگیر برای پرکردن پرسشنامه استفاده می‌شود آموزش این افراد ونیز کسانی که سرپرستی آنان را برعهده دارند از اهمیت زیادی برخوردار است زیرا دراین مرحله کارکنان اجرایی باروش آمارگیری،تعاریف ومفاهیم وچگونگی تکمیل پرسشنامه آشنامی‌شوند . بدیهی است هرگونه سهل انگاری وبی‌دقتی درآموزش کارکنان اجرایی موجب می‌شود که یکسانی وهماهنگی کار ازبین برود و درنهایت ، دقت وصحت نتایج بررسی درمعرض خطر جدی قرارگیرد .

۳-۱-۵-خطای انتخاب ابزار نامناسب

چارچوب آماری ، نقشه ، پرسشنامه ودستورالعمل ، عمده ترین ابزار یک آمارگیری به حساب می‌آیند واصولا " موفقیت بررسی به میزان زیادی بستگی به چگونگی تهیه واستفاده ازاین ابزاردارد .

چارچوب آماری نرمصورت کهنه ، ناقص یانادرست بودن ، مبنای بسیاری ازخطاهای غیرنمونه‌ای قرارمی‌گیرد ونامناسب ترین آن وقتی است که چارچوب آماری اصولا " مربوط به جامعه آماری مورد نظر نباشد .

نقشه های مورد استفاده درآمارگیری باید به نحو مناسب تهیه شده وعلاوه براین حتی‌المقدوربهنگام باشند تا ضمن سهولت دراستفاده ازآنها اشتباهات احتمالی به حداقل برسد .

هنگام تهیه وتنظیم پرسشنامه ، تمامی جوانب وجزئیات باید مطالعه شوند . پرسشنامه باید قبل از شروع آمارگیری عملا " مورد آزمایش قرارگیرد تا ایرادها ونقایص احتمالی موجود درآن مشخص وبرای آنها چاره‌جویی گردد . در تنظیم پرسشنامه ترتیب سئوالات ونیز نحوه تدوین آنها فوق العاده مهم است . همچنین سئوالها باید به گونه‌ای تنظیم شوند که گرایش به پاسخهای خاصی را ایجاد نکنند . درمورد سئوالهایی که وضعیت به گونه‌ای است که اطلاعات مورد نظر را نمی‌توان به طور مستقیم کسب نمود ، بهتراست سئوالها به نحوی تنظیم شوند که بتوان پاسخها را به طور غیر مستقیم بدست آورد . زیادبودن تعداد سئوالها وطولانی شدن مدت مصاحبه نیز از مواردی است که خطاهای غیرنمونه‌ای را افزایش می‌دهد . انتخاب اندازه مناسب برای پرسشنامه ودرنظر گرفتن فضای کافی برای درج پاسخها نیز از دیگر عواملی هستند که باید به هنگام طراحی پرسشنامه ،مورد توجه قرارگیرند .

دستورالعمل تکمیل پرسشنامه نیز حائز اهمیت زیادی است . درتهیه وتنظیم دستورالعمل آماری رعایت سادگی ، روشنی وقابل‌فهم بودن تعاریف ، عبارات وجملات از جمله اصولی است که همواره بایستی مد نظر قرارگیرد . این امر به ویژه وقتی که از آمارگیران غیر حرفه‌ای استفاده می‌شود از اهمیت بیشتری برخوردار است. زیرا این افراد معمولا " با مفاهیم خاص آماری آشنا نیستند . بدیهی است درمواردی که آمارگیری به گونه‌ای است که پاسخگو باید با مطالعه آن ، پرسشنامه را تکمیل کند این امر ضرورت بیشتری پیدامی‌کند.کامل بودن دستورالعمل آمارگیری از اصول دیگری است که همواره باید مورد توجه قرارگیرد . زیرا درغیراین صورت هرکس مواردی راکه درباره آنها توضیح داده نشده است بنابه سلیقه وهمچنین تعبیر وتفسیرشخصی‌خویش حل وفصل خواهد کرد وبدین ترتیب هماهنگی ویکنواختی جریان کار ازدست می‌رود .

۲-۴-خطاهای مرحله گردآوری داده‌ها

بخش اعظم خطاهای غیرنمونه‌ای درمرحله عملیات میدانی روی می‌دهد وبه همین جهت لازم است بیشترین کنترل ونظارت دراین بخش ازکار اعمال شود تا اولا " اطمینان حاصل شود که اطلاعات موردنظر مطابق با تعاریف ، مفاهیم و معیارهای مقرر جمع آوری می‌شود وثانیا " نقاط ضعف آمارگیران مشخص وبا آموزش وارشاد آنان‌از وقوع یا تکرار اشتباهات جلوگیری به عمل آید .

خطاهایی‌راکه‌درمرحله گردآوری داده ها روی می‌دهد می‌توان به‌سه‌گروه‌خطاهای‌پوشش‌نامناسب ، خطاهای پرسشنامه های بدون پاسخ وخطاهای مربوط به پاسخهای اشتباه طبقه بندی نمود . درذیل درباره هرکدام از خطاهای مزبور به طور مختصر توضیح داده‌شده است .

۱-۲-۴-خطاهای پوشش نامناسب

خطاهای مربوط به پوشش جامعه به آن دسته از خطاهای غیرنمونه‌ای اطلاق می‌شود که ناشی از عدم توانایی درپوشش جامع ومانع جامعه آماری است . پوشش نامناسب می‌تواند به سه صورت از قلم افتادگی، دوباره شماری ویا پوشش نادرست روی دهد .

از قلم افتادگی درسطوح مختلف وبه علل گوناگونی ممکن است رخ دهد . عدم آمارگیری از بعضی از نقاط به علت کهنه بودن نقشه های جغرافیایی ، عدم مراجعه آمارگیران به نقاط واقع درمناطق

صعب العبور ، عدم آمارگیری از افراد متحرک از جمله عواملی هستند که موجب ازقلم افتادگی افراد می‌شوند .

گاهی اوقات ، سن ورسوم ویا تصورات خاص موجب ازقلم افتادگی می‌شود . مثلا " بعضی از خانوارها ممکن است در گزارش کردن فرزندان دختر اکراه داشته باشند ویا اینکه در شرایط خاصی ازگزارش جوانان دوسن سربازی خودداری ورزند .

دوباره شماری که درواقع به معنای شمارش بیش از یک مرتبه یک واحد آماری است نیز می‌تواند در سطوح مختلف وبه علل گوناگونی روی دهد . دوباره شماری خانوارهایی که بیش از یک اقامتگاه دارند ونیز دوباره شماری خانوارهای متحرک از جمله مواردی است که خطای مزبور روی می‌دهد .

دراینجا ممکن است عده‌ای تصور نمایند که از قلم افتادگی ودوباره شماری به این دلیل که در دوجهت مختلف روی می‌دهند وبایکدیگر جمع جبری می‌شوند از اهمیت چندانی برخوردار نیستند . این تصور نمی‌تواند صحیح باشد زیرا اولا " میزان از قلم افتادگی ودوباره شماری لزوما " یکسان نیست وثانیا " خصوصیات واحدهای ازقلم افتاده بااحتمال زیاد همان خصوصیات واحدهای دوباره شماری شده نمی‌باشد .

نوع سوم خطای مربوط به پوشش یعنی پوشش نادرست عبارت از پوشش واحدهایی خارج از جامعه آماری است . به حساب آوردن مسافران دربرزمینهایی که افراد باید دراقامتگاه معمولی خود آمارگیری شوند از جمله موارد مربوط به پوشش نادرست می‌باشد .

۲-۲-۴-خطای پرسشنامه های بدون پاسخ

بخش بسیار مهمی از خطاهای غیرنمونه‌ای که درمرحله گردآوری داده ها روی‌می‌دهد خطاهای ناشی‌از پرسشنامه هایی است که هردلیلی تکمیل نشده است . این خطا می‌تواند ناشی از عدم دسترسی به پاسخگو عدم همکاری پاسخگو ، عدم توانایی وشایستگی پاسخگو ویا دیر گم شدن پرسشنامه ها باشد .

عدم دسترسی به پاسخگو می‌تواند ناشی از مراجعه دروقت نامناسب ویا جابه جایی بعضی ازواحدهای آماری باشد .

عدم همکاری پاسخگو درپاسخ به سئوالها معمولا " به این دلیل است که افراد نمی‌خواهند به سئوالاتی پاسخ دهند که نه تنها هیچ نفعی درآن نمی‌بینند بلکه درمواردی احساس خطر هم می‌کنند . دوهر حال عواملی نظیر نحوه برخورد آمارگیر ، زمان مصاحبه ، اطمینان از محرمانه ماندن اطلاعات ونحوه

تبلیغات از جمله عواملی هستند که می‌تواند درکاهش عدم همکاری پاسخگویان موثر باشد .

درباره‌ای موارد پاسخگو مایل است که درپاسخ به سئوالها همکاری کند ولی بنا به دلایلی از قبیل مشکل زبان ، بیماری وغیره قادر به این کار نیست ویا اگر هست پاسخهای وی نمی‌تواند از اعتبار کافی برخوردار باشد . این قبیل موارد عدم توانایی وشایستگی پاسخگو درپاسخ به سئوالات را تشکیل می‌دهد .

یکی دیگر از مواردی که به منزله پرسشنامه بدون پاسخ است ، گم شدن پرسشنامه های تکمیل شده توسط آمارگیران است . مسئله گم شدن پرسشنامه ها مرفنظر از اینکه یک خطای غیرنمونه ای به حساب می‌آید، درمواردی به دلیل دسترسی افراد به اطلاعات خصوصی ومحرمانه اشخاص دیگر می‌تواند بعضا" مشکلات زیادی را برای سازمان آمارگیری ایجاد کند . درمرحال چون مراجعه مجدد واخذ دوباره اطلاعات عملا" کاردشواری است بنابراین باید ترتیبی اتخاذ شود که از گم شدن پرسشنامه ها جلوگیری به عمل آید .

۳-۲-۴. خطای پاسخهای اشتباه

خطای پاسخهای اشتباه که اصطلاحا" " خطای محتوی " نیز نامیده می‌شود به خطاهایی اطلاق می‌گردد که ناشی از اعلام ، اندازه گیری یا ثبت نادرست پاره‌ای از اقلام آماری است . برخی از علل خطای پاسخهای اشتباه می‌تواند ناشی از خطاهای مرحله برنامه ریزی وتهیه طرح باشد . درمرحله جمع آوری داده ها نیز عواملی وجود دارد که دربروز خطای محتوی نقش عمده‌ای را ایفا می‌کنند . این عوامل درذیل شرح داده شده است :

نامناسب بودن پاسخگو - پاسخگو علاوه بر داشتن آگاهیها واطلاعات کافی باید تمایل به ارائه پاسخهای صحیح نیز داشته باشد . بنابراین مثلا" درآمارگیری از خانوارها کسب اطلاعات از افراد غیر مطلع خانوار وهمچنین همسایگان وکودکان به دلیل اینکه آگاهیهای لازم را ندارند ویا آگاهیهای آنان از اعتبار چندانی برخوردار نیست از جمله موارد پرسش از پاسخگوی نامناسب تلقی می‌شود .

استفاده پاسخگواز حدس وگمان - بعضی اوقات ، پاسخگو درمورد برخی از سئوالات که پاسخ آنها را نمی داند به حدس وگمان متوسل می‌شود واین امر بیشتر زمانی اتفاق می‌افتد که پاسخگو از عدم آگاهی خود شرم دارد وترجیح می‌دهد که جوابی صحیح ویا ناصحیح بدهد . روشن است که چنین اظهاراتی

گاهی چندان دقیق نیست ومی‌تواند دارای خطا باشد.

خطای‌حافظه ـ اصولا" انسان به مرور زمان ، وقایع وحوادث را فراموش می‌کند ولذا میزان دقت جوابهایی که پاسخگو می‌دهد تا حدود زیادی بستگی به این دارد که این وقایع تا چه حد در حافظه وی باقی مانده است . دوره زمانی سئوال از جمله عواملی است که میزان خطای حافظه را تحت تاثیر قرار می‌دهد . منظور از دوره زمانی ، دوره ای است که اطلاعات مورد نظر برای آن فاصله زمانی جمع آوری می‌شود . دوره پرسش را از دوسیدگاه طول دوره و همچنین نزدیکی یادوری آن به زمان پرسش می‌توان مورد بررسی قرارداد . درمورد دوره زمانی ذکر این نکته ضروری است که گاهی پاسخگو قادرنیست به خاطر بیاورد که وقوع حادثه موردنظر درداخل دوره وياخارج ازآن صورت گرفته است . به عنوان مثال اگر فراسفند ماه از فردی سئوال شود که اعضای خانوار وی دردی ماه چند بار به پزشك مراجعه کرده‌اند بیشترین خطا مربوط به مراجعاتی خواهد بود که اعضای خانوار احتمالا" در اواخر آذر یا اوایل دی ونیز اواخر دی یا اوایل بهمن داشته اند.

فقدان انگیزه‌برای‌دادن پاسخ دقیق ـ درآمارگیریهایی که از طرف سازمانهای دولتی صورت می‌گیرد پاسخگو به موجب قانون ، ملزم به همکاری وارائه اطلاعات صحیح می‌باشد . درسایر آمارگیریها که پاسخگو در مورد جواب دادن به سئوالها الزام قانونی ندارد شرکت وهمکاری وی دردادن پاسخ به سئوالها وبه ویژه ارائه اطلاعات صحیح کاملا" اختیاری است وبه نظر خود او بستگی دارد . لذا هرگاه پاسخگو هیچ انگیزه‌ای برای همکاری نداشته باشد معمولا" از پاسخ به سئوالها طفره می‌رود ویا درصورت قبول پاسخ دادن به سئوالها هیچگونه تلاشی درمورد ارائه اطلاعات صحیح نخواهد کرد . به هر صورت ایجاد انگیزه برای همکاری مطلوب پاسخگو از مسایل مهم درآمارگیری است .

ارائه پاسخ‌نادرست عمدی ـ یکی دیگر از مهمترین موارد مربوط به پاسخهای اشتباه ، پاسخ نادرستی است که پاسخگو آگاهانه وارزوی عمد در اختیار آمارگیرمی‌گذارد . این موضوع اغلب زمانی اتفاق می‌افتد که پاسخگو احساس کند که موضوع به گونه‌ای با حیثیت وپرستیژ وی ارتباط پیدا می‌کند . برای مثال برخی از افراد بی‌سواد به دلایل حیثیتی اظهار می‌دارند که خواندن ونوشتن می‌دانند . برخی ازافراد نیز به همین دلیل دربیان شغل خود مرتکب خطا می‌گردند وشغلی غیراز شغل واقعی خود اظهارمی‌دارند .

ارائه پاسخ نادرست غیرعمدی - درپاسخگوی موارد پاسخگو بدون اینکه عمدی درکار باشد به دلیل برداشت ناصحیح از سئوال ، ناخودآگاه پاسخ نادرست می‌دهد . به عنوان مثال ممکن است وقتی درآمد فردی که دارای چندین منبع درآمد است سئوال شود سهوا "به‌نظر درآمد حاصل از فعالیت اصلی خود اکتفا کند ودرآمد حاصل از سایر منابع را گزارش نکند .

ثبت نادرست پاسخها - این نوع خطا به اشتباهات غیر اختیاری آمارگیر هنگام تکمیل پرسشنامه مربوط می‌شود . اشتباه درتعیین کد مناسب درمورد سئوالاتی که پاسخ آنها باکد مشخص می‌شود ونیز درج پاسخ مربوط به یک سئوال درمقابل سئوال دیگر از جمله این گونه خطاها می‌باشد .

سرراست کردن ارقام- سرراست کردن ارقام توسط پاسخگو درکشورهای مختلف جهان بافرهنگهای گوناگون متداول است . موضوع سرراست کردن ارقام مربوط به سن به اعدادی که به مضرباپنج ختم می‌شوند از جمله این موارد می‌باشد . این امر به وضوح از روی هرمهای سنی جمعیت قابل تشخیص است . مثال دیگر اینکه اگر از خانوارها درمورد مدت تماشای تلویزیون در روز سئوال شود پاسخ غالبا " مضربی از ساعت است .

اثر آمارگیر - این عامل ، منبع دیگری برای بروز خطا درمرحله گردآوری اطلاعات است . جمع آوری آمار از طریق مصاحبه به دلایل مختلف منجمله سرعت درانجام کار ، کاهش تعداد پرسشنامه های ناقص ونیز بدون جواب به روش تکمیل پرسشنامه توسط پاسخگو برتری دارد . لیکن برفرض از مواردی که به کارگیری روش مصاحبه الزامی است درسایرموارد ، مزایای این روش هنگامی است که آمارگیران صالح ، باتجربه وآموزش دیده به خدمت گرفته شوند . با توجه به اینکه احساس مسئولیت ودارابودن تعهد ووجدان کار از شرایط اولیه کار ماموران آمارگیر است لذا واضح است که این افراد باید درکار خود صدیق وصمیمی باشند . اما نباید فراموش کرد که عدم احساس مسئولیت وفقدان صداقت غیرمحتمل هم نیست . گاهی اوقات که برای کسب برخی از اطلاعات ، مراجعه مجدد به واحد آماری مورد لزوم است ، مامور آمارگیر ممکن است به منظور اجتناب از مراجعه مجدد ، پاسخ سئوالهای باقیمانده را خود حدس بزند . برخی مواقع وضع بدتر ازاین است وآن وقتی است که آمارگیر پاره‌ای از سئوالات را اصلا " نمی‌پرسد وسرعت عمل درکار را درهمان زمینه آمارگیری ویادروقت دیگر حدس وگمان خودرا درپرسشنامه درج

می‌کند . دوپاره‌ای موارد حتی ممکن است عدم مراجعه آمارگیر به واحد آماری وتکمیل کلیه سئوالات پرسشنامه توسط خود وی نیز مشاهده شود . دوپاره‌ای از بررسیهای اجتماعی ، گرایشهای ذهنی یا عقیدتی آمارگیر ممکن است اطلاعات را دچار خطا سازد واین امر هنگامی افزایش می‌یابد که دستورالعمل روشن وصریحی درمورد نحوه پرسش وتکمیل سئوالها دراختیار آمارگیران قرارنگرفته باشد .

۲-۲-خطاهای مرحله استخراج وانتشار نتایج

از زمانی که کار آمارگیری به پایان می‌رسد تاوقتی‌که نتایج منتشر می‌گردد یک سلسله عملیات کدگذاری ، کنترل ونقل وانتقال اطلاعات صورت می‌گیرد . هرکدام از عملیات مزبور درمعرض پاره‌ای از خطاهای غیرنمونه‌ای است . این خطاها را می‌توان در۵ گروه ذیل طبقه بندی نمود .

۱-۲-۲-خطای بازبینی

آمارگیری ، کار مشکل وپیچیده ای است که درهرحال باید پاره‌ای خطاهای ناشی از بروز اشتباه درمراحل مختلف کاررونیز وجود نقص وناهماهنگی دراطلاعات را محتمل دانست وهمین امر ایجاب می‌کند که پرسشنامه های تکمیل شده مورد بازبینی قرارگیرد . بازبینی می‌تواند درزمینه پاسخهای ثبت نشده (ازقلم افتاده) و اطلاعات غلط ، غیر منطقی ویا ناهماهنگ صورت پذیرد . دستورالعملی که برای بازبینی درمورد هرکدام ازاین زمینه ها درنظر گرفته می‌شود فوق العاده مهم است زیرا دراین مرحله ، تصمیمات انسان ، عامل تعیین کننده است . اگر او بتواند موارد ناصحیح وغیر معقول را تشخیص دهد واگر قادر باشد اطلاعات ناهماهنگ را بازیابد وخارج کند تازه این سئوال مطرح می‌شود که چه اطلاعی جایگزین شود تا به واقعیت نامعلوم نزدیک شود .

۲-۲-۲-خطای کد گذاری

کد گذاری به معنای تبدیل کلمات به اعداد است . کد گذاری بعضی از اقلام مثل جنس ، وضع سواد ووضع زناشویی درصورتی که ضرورت داشته باشد آسان است . لکن درمعضی موارد مثل شغل و فعالیت محل کار ، کد گذاری کار ساده‌ای نیست ومی‌تواند به دلایل مختلف همراه با اشتباه باشد . این اشتباه ممکن است ناشی از درج اطلاعات مبهم یا ناقص توسط مامور آمارگیر ویا بی‌دقتی وکم تجربگی کد گذار باشد . به هرصورت از آنجا که کنترل کار کدگذاری ، ساده نمی‌باشد بنابراین لازم است که ضمن استفاده

از افراد باتجربه ، روشهایی انتخاب گردد که در هر مورد خطای کدگذاری به حداقل کاهش یابد .

۳-۲-۴-خطای پانچ

درمرحله پانچ نیز احتمال بروز خطا وجود دارد . ارقام وکدهای ناخوانای پرسشنامه ، هنگام پانچ غالبا " توسط پانچیست حل وفصل می‌شوند . پانچ پاره‌ای از اطلاعات ممکن است فراموش شود و درنتیجه این موارد جزو اطلاعات ازقلم افتاده منظور گردد . به هرصورت از آنجا که از طرفی درکار پانچ سرعت درانجام کار مهم است واز طرف دیگر همین عامل مهمترین علت بروز خطا درایین مرحله از کار است لذا همواره باید بانجام اقدامات مناسبی کار پانچ مورد کنترل قرارگیرد . دراینجا باید توجه داشت که چون داده های حاصل از آمارگیری به صورت اطلاعات حاصل از کارپانچ، درجریان استخراج وجدول‌گیری قرار می‌گیرد لذا ضرورت وجودکنترل کافی دراین مرحله ازکار ، بیشتر احساس می‌شود .

۴-۲-۴-خطای برنامه نویسی

استفاده از کامپیوتر برای استخراج نتایج ، خطای ویژه خودرا به همراه دارد . علاوه برخطای پانچ ، دستورالعملی که برای استخراج اطلاعات تهیه می‌شود نیز می‌تواند دارای اشتباه باشد . خطاهای برنامه نویسی عمدتا " ناشی از عوامل است . بخصوص کم تجربه بودن برنامه نویس درانجام وظایف محوله ودوم درک نادرست وی ازآنچه که باید استخراج شود. درسالهای اخیر به علت استفاده روزافزون از کامپیوتر ، آشنایی آمارشناسان باکار برنامه نویسی کاملا " ضروری به نظر می‌رسد زیرا هم می‌توانند برنامه نویس را درانجام هرچه بهتر وسریعتر کار یاری دهند وهم اینکه باطلاع وآگاهی از برنامه تهیه شده ، خطاهای احتمالی موجود دربرنامه را کشف نمایند .

۵-۲-۴-خطای تدوین وانتشار

آخرین کار یک بررسی آماری ، انتشار جداول حاوی نتایج است . دراین مرحله نیز می‌تواند به علت اشتباه درتنظیم جداول ، اشتباه درتایپ اطلاعات وغیره خطاهای غیرنمونه‌ای بروز کند وبه همین دلیل لازم است که جداول قبل از ارسال برای تکثیر ، مورد بررسی نهایی قرارگیرند .

منابع ومآخذ مورد مراجعه

۱-کلریز ، حسن ، خطاهای غیرنمونه‌ای درآمارهای اقتصادی ، تهران : دانشکده اقتصاد دانشگاه تهران ،

موسسه توسعه وتحقیقات اقتصادی ، ۱۳۵۸ .

۲- مرکز آمار ایران ، دفتر بررسی طرحها وحاسبات آماری ، اشتباهات غیرنمونه‌ای .

3. Cochran, William G. , Sampling Techniques, New York : John Wiley and Sons, Third Edition , 1977.

4. FAO , Sampling Methods For Agricultural Surveys , Rome : FAO Statistical Development Series No 3 , 1989.

5. United Nations , Economic and Social Commision for Asia and the Pacific, Manual on Training of Statisticians, New York : United Nations, 1988.

مدیریت داده‌ها در مقیاس بزرگ

جمشید میراوندی‌راد

مرکز آمار ایران

در اغلب کشورها ارتباط و همکاری نزدیک بین آمار و پردازش اطلاعات وجود دارد و در هر یک از کارهای روزانه سازمانهای آماری ، جائیکه برای حجم زیاد اطلاعات ضرورت اجرای عملیات و یا تکرار بمنظور دریافت نتایج وجود داشته باشد ، کاربرد کامپیوتر اجتناب ناپذیر می‌شود. هدف از مقاله ارائه شناختی از جمع آوری اطلاعات و روش های لازم در ایجاد محیط های مناسب و امکانات در پردازش و محاسبات است .

زمانیکه برنامه‌ریزی برای استفاده از کامپیوتر در امور پرسنلی یک واحد ، یک اداره و یا یک سازمان است فعالیتهای قابل کنترل و محدودی وجود دارد و کافیست براساس نیاز ، اطلاعات ذکر شده در فرمهای پرسنلی جمع آوری شود و با استفاده از برنامه‌هائی که تهیه و نوشته میشود بازیابی مورد استفاده قرار گیرد . در این مورد موضوعهائی مانند گستره جغرافیائی ، نیروئی که برای جمع آوری اطلاعات باید روانه کردند ، استفاده از امکانات سازمانهای دیگر بصورت گسترده و خیلی از موضوعات دیگر کاربردی نخواهد داشت ، همینطور در استفاده از کامپیوتر برای ثبت نام و انتخاب واحدها در دانشگاهها که حول محور یک برنامه کامپیوتری و تغذیه منابع اطلاعاتی و پردازش آن قرار داشته و نیازی به دخیلشدن و جلب همکاری جامعه و کمک گرفتن از ارگانها و سازمانهای دیگر و امثال آن نیست . اما درصورتیکه هدف پوشش و تامین خواسته‌های سیاستگذاران و برنامه ریزان یک کشور در زمینه آمار و اطلاعات باشد ، در آنصورت انجام فعالیت در گستره وسیع (LARGE AREA) و در مقیاس بزرگ (Large Scale) شکل گرفته و اجتناب ناپذیر می‌شود .

در زمینه تهیه و تولید آمار ، اگر داده‌ها را ، اطلاعات و ستانده‌ها را ، آمار بنامیم مطابق تعریف کلی تکنولوژی که عبارتست از عامل تبدیل کننده داده‌ها به ستانده‌ها می‌توان گفت در هر فعالیت آمار مجموع اجزای چهارگانه تکنولوژی یعنی فن افزار (Technoware) ، انسان افزار (Humanware) ، اطلاعات افزار (Inforware) و سازمان افزار (Orgaware) یک فرآیند تبدیل را بوجود می‌آورند* . در این فرآیند ، فن افزار شامل ابزار ، تجهیزات ، تاسیسات و تسهیلات فیزیکی و انسان افزار مجموعه مراتب ...

---

* توضیح بیشتر در این مورد در نشریه شماره ۶ پروژه طراحی سیستم اطلاعات آماری شیلات توسط مرکز آمار ایران ، می‌باشد .

تجربه ها ، دانشها ، ارزشها ، انگیزه‌های افراد و اطلاعات افزار مجموع اطلاعاتی به شکلهای گوناگون مانند اسناد، مدارک ، نقشه ، جزوه و کتاب درباره همه مراحل شامل طراحی ، ساخت ، راه اندازی ، نگهداری است و بدین ترتیب همه فرآیندها و مجموعه‌های نرم افزاری را در بر می‌گیرد و بالاخره سازمان افزار مجموعه عوامل سازمانی را که به نوعی در توسعه ، انتقال ، جذب ، بهره برداری تاثیر دارند شامل می‌شود . فعالیتها در این فرآیند تولید ، به زیر مجموعه‌هائی تقسیم می‌شوند که اهم آنها عبارت است از :

۱ـ جمع آوری نیازها در زمینه‌های مختلف اجتماعی ، اقتصادی ، ...

۲ـ تشخیص و تعیین و دریافت اولویت ها برای هریک از نیازها .

۳ـ تعیین و پوشش جغرافیائی (ناحیه‌ای ، منطقه‌ای ، استانی ، کل کشور و امثال آن) و تهیه نقشه .

۴ـ طبقه بندی و تقسیم نیازها به موضوعات مختلف بمنظور طراحی خاص برای هریک از آنها .

۵ـ روش جمع آوری اطلاعات .

۶ـ تعیین زمان جمع آوری اطلاعات و دوره‌های زمانی .

۷ـ نیروی انسانی لازم و تخصص‌های آنها .

۸ـ تهیه فرمهای مناسب برای پرسش و دریافت پاسخ .

۹ـ آموزش نیروها مرکلیه سطوح و با توجه به نوع وظیفه .

۱۰ـ تعیین روش نظارت بر اجرای فعالیتها .

۱۱ـ تعیین روش پردازش اطلاعات .

۱۲ـ آماده‌سازی فرمها و اطلاعات جمع آوری شده برای انتقال اطلاعات .

۱۳ـ انتقال اطلاعات جمع آوری شده بر روی وسائل جانبی برای پردازش کامپیوتر .

۱۴ـ کنترل کیفیت در انتقال اطلاعات .

۱۵ـ کنترل کیفیت مجموعه اطلاعات و کنترل هماهنگی در اطلاعات جمع آوری شده .

۱۶ـ ذخیره سازی و ایجاد بانکهای اطلاعاتی با امکان دسترسی مطلوب .

۱۷ـ ارائه گزارشات و انتشار نتایج و یافته‌ها .

۱ـ جمع آوری نیازها در زمینه‌های مختلف اجتماعی ،اقتصادی ، ...

سازمانها برای هماهنگی و تلفیق سیاستگزاری و برنامه ریزی خود با نیازهای جامعه ،محتاج به داشتن آمـار و اطلاعات هستند هرچند آمار درهرحال اطلاعات (Information) است ولی این عبارت باین مفهوم نیست که اطلاعات فقط باین صورت است ، آمار مقادیر عددی است ولی اطلاعات می‌تواندغیر عددی وقیاسی (Analogue) هم باشد ، در خیلی از موارد بجای داشتن اطلاعات صرفا" رقمی و عددی ، نیاز بـــه دانستن اطلاعات قیاسی است .

۲ـ تشخیص و تعیین و دریافت اولویت ها در نیازها :

برای نیازها و میزان فوریت آن از نظر حاصل شدن و دستیابی وبالامکان پذیری آن با توجه به امکانـــات و محدودیتـها وپارامترهاومتغیرهای موجود ،باید اولویت جمع آوری وفراهم‌شدن تعیین شود .

۳ـ تعیین پوشش جغرافیائی و تهیه نقشه :

اطلاعاتی که باید جمع آوری شوند بسته به کاربرد وبافراوانی درسطوح مختلف جغرافیائی ، به اطلاعـــات منطقه‌ای ، ناحیه‌ای ، استانی و مانند آن تقسیم می‌شوند و براین اساس سطح جغرافیائی که باید پوشش داده ـــ شود تعیین می‌گردد و نقشه‌های آماری جهت شناخت محدوده‌های فعالیت تهیه می‌شود .

۴ـ طبقه بندی و تقسیم نیازها به موضوعات مختلف :

بمنظور استفاده بهتر از انجام فعالیت جمع آوری اطلاعات مورد نیاز ، نیازهای تعیین شده برحســـب موضوعی طبقه بندی و تلفیق می‌گردد . این طبقه بندی ،مجموعه نیازهای اعلام شده درزمینه‌های اجتماعـی ، اقتصادی ، کشاورزی ، صنعتی ، بازرگانی ، فرهنگی وآموزشی ، ... را شامل خواهدشد ودرنتیجه برای هریک روش جمع آوری خاص پیش بینی می‌شود .

۵ـ روش جمع آوری اطلاعات :

برای جمع آوری اطلاعات روش های مختلفی ازسرشماری یا نمونه‌گیری بکار برده میشودبرای مراجعه متداول‌ترین نوع ، مراجعه مستقیم و تکمیل پرسشنامه می‌باشد . ممکن است فرمها جداگانه بوسیله ماموران و یا پست ارسال شده و بعدا" جمع آوری شود (مانند فرمهای ثبت نام کنکورهای سراسری دانشگاهها) درهریک از روشهـا با برنامه‌های تبلیغاتی وآشناسازی ،نوعی آموزش عمومی جهت همکاری جامعه انجام میشود .

۶- تعیین زمان جمع آوری اطلاعات و دوره‌های زمانی :

زمان مراجعه برای جمع آوری اطلاعات با توجه به شرایط اقلیمی و محیط و با توجه به موضوع مورد پرسش اهمیت زیادی داشته و باید مورد نظر قرار گیرد. همینطور ضروری است که جمع‌آوری اطلاعات در یک فاصله زمانی معین تکرار شود.

۷- نیروی انسانی لازم و تخصص های آنها :

پیش بینی و تعیین نیروی انسانی لازم در سطوح مختلف از فعالیتهائی است که در زمان طراحی پیمش بینی میشود این افراد بسته به وظایفی که بعهده خواهند داشت از نیروهای ساده تا افراد با تخصصهای مورد نیاز و مدیران اجرائی و ستادی را شامل می‌شود.

۸- تهیه فرمهای مناسب برای موضوعات مختلف جهت پرسش و دریافت پاسخ :

گردآوری سوالات در یک زمینه و هماهنگی آنها در یک فرم سبب تمرکز فکر در پرسشگر و پاسخگو می‌شود. از طرف دیگر فرم طراحی شده علاوه بر آنکه باید درج پاسخ را آسان نماید (مانند طبقه بندی پاسخها ، مشخص کردن آنها با علامت × و امثال آن) باید بصورتی باشد که استفاده‌های بعدی از آن در مراحل دیگر نیز آسان باشد.

۹- آموزش نیروها در کلیه سطوح با توجه به نوع وظیفه :

افرادی که باید در جمع آوری اطلاعات همکاری نمایند نیاز به آموزش دارند این آموزش شامل شناختن روش کار ، حوزه عمل هر فرد ، آشنائی با نقشه ، آشنائی با فرمها و پرسشنامه‌ها ، سوالاتی که باید پرسش شود و نحوه ثبت پاسخها ، تنظیم فعالیت در مدت زمان تعیین شده و امثال آن می‌باشد . افرادی که سرپرستی آمارگیران را بعهده خواهند داشت نیز باین آموزش نیاز دارند و در سایر فعالیتهای مربوط به اجراء و یا انتقال اطلاعات که نیروهای بیشتری بصورت موقت به خدمت گرفته می‌شوند آموزش های مربوط به خود را خواهند داشت .

۱۰- تعیین روش نظارت بر اجرای فعالیتها :

منظور از نظارت ، کنترل کیفیت و بهبود آن و یا شناخت میزان دقت و صحت بکار رفته است . این نظارت می‌تواند بصورت اعظام نیروهای خاص و بررسی عملیات انجام شونده از آنکه ما تهیه و آموزش دستورالعملهائی این کار را بصورت غیرمتمرکز و با استفاده از سرپرستان انجام داد .

۱۱ـ تعیین روش پردازش اطلاعات :

برای پردازش اطلاعات می‌توان تصمیم گیریهای لازم را از قبل بعمل آورد ، باید نسبت به شکل و الگـــوی پرسشنامه‌ها (Fram Design) و محتویات آن اعلام نظر شود و برای چگونگی تکمیل آن و روش های بعدی برای انتقال اطلاعات پرسشنامه‌ها به وسایل جانبی و قابل استفاده در کامپیوتر پیش بینی بعمل آید. معمولا " پردازش بصورت متمرکز انجام شده است ولی درصورت وجود امکانات ، تمام یا بخشی از آن می‌تواند بصورت غیرمتمرکز انجام شود.

۱۲ـ آماده سازی فرمها و اطلاعات جمع آوری شده برای انتقال اطلاعات :

بعد از جمع آوری ، اغلب فعالیتهائی وجود دارد که باید بصورت دستی و یا استفاده از نیروهای آموزش دیده برای این منظور انجام شوند . معمولا " پرسشها بصورتی هستند که دارای پاسخ عددی و یا اعدادی که از قبـــل تعیین شده می‌باشند (Self Code) ولی فرضا " شرح نوشته شده برای صفاتی چون شغل و مدرک رشتـــه تحصیلی افراد بگونه‌ای است که باید به کد تبدیل شوند و در اینصورت افرادی این فعالیت را انجام می‌دهند .

۱۳ـ انتقال اطلاعات بر روی وسایل جانبی برای پردازش :

یکی از فعالیتهائی که هم بزمان و هم به نیروی انسانی زیاد نیاز دارد انتقال اطلاعات بر روی وسایـــل جانبی است تا امکان پردازش های بعدی فراهم شود درصورتیکـــه فرمهـــــا بـــــرای استفـــاده از Mark Reader و یا Charater Reader طراحی شده باشند این فعالیت حذف شـــده و فرمها عینا " قابل خوانده شدن توسط دستگاههای جانبی کامپیوتر می‌باشند ولی بعلت وجود محدودیت های موجود در این روش ، از چنین فرمهائی کمتر استفاده می‌شود . زمان لازم برای انتقال اطلاعات با محاسبـــه تعداد ضربه (Stroke) در واحد زمان و یا تعداد وکورد برآورد می‌شود . با فرض انتقال اطلاعـــات مربوط به ۱۰۰ فرد در یک ساعت توسط یک نفر ، برای جمعیتی معادل ۶۰ میلیون نفر ، زمانی حدود ۶۰۰ هزار نفر ساعت پیش بینی می‌شود .

۱۱ ـ کنترل کیفیت درانتقال اطلاعات :

با توجه به اهمیت انتقال اطلاعات و نیروی انسانی شاغل در آن علاوه برآموزشهای لازم ،توجه به کیفیت کار ضروری می‌باشد. این کنترل ممکن است با استفاده از ماشینهای برنامه پذیر وهمزمان با انتقال اطلاعات بعمل آید و خطاهای وارد شده تشخیص و اعلام شود و با این روش دقت افزایش یابد و درصورت بکار گرفتن ماشینهای بدون برنامه کنترل ، باید از روش انتقال دو باره (Verification) استفاده کرد و دراین روش الزام به انجام آن بصورت ۱۰۰٪ نبوده و ممکن است بصورت موردی و یا قسمتی از اطلاعات مورد بررسی و کنترل قرار گیرند.

۱۲ـ کنترل کیفیت درمجموعه اطلاعات و اصلاح اشتباهات :

با تمام کوششی که برای بالا بردن دقت درمراحل مختلف می‌شود معذا خطاهای احتمالی وجود دارد . بهترین روش برای یافتن این خطاها و حذف یا اصلاح آن استفاده از کامپیوتر است .بااین منظور با تهیه روابط منطقی که بین صفات مختلف درمجموعه اطلاعات باید وجود داشته باشد وتهیه برنامه براساس آن ، خطاهای وارد شده قابل شناسائی می‌شوند . این روابط منطقی شامل بررسی هماهنگی بین اطلاعات (Consistency) وحدود تغییرات مجاز (Range check) است . پس از شناسائی خطاهای وارد شده یک راه اصلاح ، ارائه گزارش آن است تا با مقایسه با اصل مدارک اصلاح شود و راه دیگر که در حجم زیاد کاربرد مناسبتری دارد اصلاح آن براساس روابط تعیین شده است (Imputation) . دراین اصلاح ماشینی از روشهای مختلفی که عمدتا" با تکیه براطلاعات صحیح مشابه یا شبیه سازی است* استفاده می‌شود.

---

* از جمله این روش ها از Cold Deck که جایگزین کردن اطلاعات از قبل تعیین شده و ثابت به جای موارد اشتباه می‌باشد و از Hot Deck که جایگزین کردن اطلاعات درست و غیر شابت بجای موارد اشتباه است که درآن ، اطلاعات درست ، با استفاده از مجموعه اطلاعات جامعه مورد پردازش قبلا" تعیین داده می‌شود.همچنین از روش اصلاح با حداقل تغییر و جابجائی Minimal Set Of Fields For Imputation می‌توان نام برد.

۱۶ ـ ذخیره سازی و ایجاد بانک های اطلاعاتی با امکان دستوسی مطلوب :

برای تعیین شده ذخیره سازی و نگهداری بعدی آن چند عامل موثر وجود دارد که باید به آن توجه شود از آنجمله :

شناخت اطلاعات از نظر حجم ، فراوانی کاربرد ، میزان فوریت لازم برای دستوسی ، ...

طبقه بندی اطلاعات برحسب نوع ، حجم ، میزان دسترسی ، واحدهای مجاز استفاده کننده : ...

تعیین مشخصات لازم برای مراکز نگهدارنده اطلاعات از نظر امکانات محاسبه افزاری ، نیروی انسانی ، خطوط ارتباطی ، فضا و مکان ، محدوده و پوشش جغرافیائی ، ...

تعیین ضوابط نگهداری و دسترسی بصورت تهیه مستندات وجزوه‌های راهنما ، مقررات و آئین نامه های لازم در ارتباط با استفاده کنندگان ، ...

تعیین شیوه ذخیره سازی اطلاعات بصورت متمرکز ، غیرمتمرکز ، ماشینی ، غیرماشینی ، ...

دوره زمانی نگهداری اطلاعات

دوره زمانی بهنگام سازی و به روز رسانی

امنیت و حفاظت اطلاعات

۱۷ـ ارائه گزارشات و انتشار نتایج و یافته‌ها :

مجموعه اطلاعات پردازش شده ونتایج حاصل از بررسی وتجزیه و تحلیلها بصورت گزارش‌ها و نشریه‌های چاپ و منتشر می‌شود تا در دسترس عموم قرار گیرد وعلاوه برآن تهیه ونتایج مطابق خواسته متقاضیان ، براساس ضوابطی که تعیین می‌شود ، امکان پذیر خواهد بود .

با توجه به موارد گفته شده ، درپایان فکر چند نکته ضروری است . در سرشماریهائی که اجرا ، شده است معمولا " حدود ۶۰ درصد از کل هزینه برای فعالیت جمع آوری اطلاعات (Field data Collection) بوده است و از طرف دیگر سهم عمده‌ای از فاصله زمانی بین جمع آوری اطلاعات تا ارائه نتایج برای فعالیت انتقال اطلاعات (Data Entry) می‌باشد . از این رو توجه به این دو فعالیت با عنوان جمع آوری و انتقال اطلاعات (Field data Collection/Entry) مورد نظر بسیاری از سازمانهای آماری است . معرفی امکانات جدید بصورت کامپیوترهای شخصی کوچک و سبک وزن که قابل انتقال و نگهداری با یکدست و کار با آن با دست دیگر است و با نیروی چند باطری قلمی روشن شده و انتقال اطلاعات با استفاده از صفحه کلید انجام می‌شود، تولید بخش تغییری اساسی می‌باشد هرچند که درحال حاضر استفاده از آن در مقیاسهای کوچک و یا طرحهای نمونه‌ای جای عملی است .

درزمینه طراحی و تهیه و تنظیم برنامه های کامپیوتری برای کنترل و تصحیح و استخراج نتایج نیز گرایش به برنامه های از پیش نوشته شده یا بسته‌های نرم افزاری (Software Packages) که با بهبود و توسعه‌ای که یافته‌اند اغلب کارآئی لازم را برای سازمانهای مختلف دارند بسیار شدیدتر از گذشته شده است در ضمن آنکه استفاده از بسته‌های نرم افزاری ، هم درکشورهای پیشرفته وهم درکشورهای درحال توسعه راهی برای جبران کمبود نیروی انسانی بوده است .

این حرکت درزمینه چاپ و انتشار نیز مشاهده می‌شود واژه نگارها یا کلمه پردازها (Word Processors) عمل تنظیم صفحات و تعیین وتغییر در اندازه متن ها و جداول ، تلفیق شکلها و نمودارها و نوشته‌ها را براحتی و با صرف زمانی بسیار اندک فراهم ساخته‌اند . مجموعه این حرکتها وامکان عملهای گسترده‌ای که بوجود آمده است تمام با تمرکز زدائی درعین حال که سرعت اجرای فعالیتهای مختلف را افزایش داده است ضرورت بازنگری درتفکر و فلسفه‌های مبتنی بر تجربه های گذشته را نیز ایجاب می‌نماید .

ضمیمه ۱ ـ بررسی خطا درانتقال اطلاعات :

میزان خطا در مجموعه انتقال اطلاعات را می‌توان از رابطه

$$E = N.M.P$$

محاسبه نمود که درآن :

N ـ تعداد رکورد منتقل شده

M ـ تعداد صفت اطلاعاتی درهر رکورد

P ـ میزان درصد خطای فرضی است .

بعنوان مثال ، درانتقال اطلاعات مربوط به ۶۰ میلیون نفر که هر یک حدود ۱۰ صفت اطلاعاتی داشته باشند و با خطای معمول و متداول در حد 10^{-2} درصدکه اغلب درچنین فعالیتی فرض می‌شود مقدار E برابر خواهد بود با

$$E = 12 \times 10^{6}$$

با این میزان خطا ، با فرض وجود بیشترین خطا در کمترین تعداد افراد ، عدد 12×10^{5} نفر و با فرض وجودکمترین خطا در بیشترین تعداد ، عدد 12×10^{6} نفر حاصل می‌شود . با توجه به چنین اعداد بزرگی است که کنتـــرل و هماهنگی اطلاعات آماری ضروری تشخیص داده می‌شود . در ضمن آنکه روش تنظیمی برای این کنترل باید بگونه‌ای باشد که سبب تاخیر در ارائه نتایج نشود .

ضمیمه ۲ - روش جایگزینی اطلاعات برای پیش برنده خطا :

در روش استفاده از Cold Deck پیش فرضهائی اولیه بوجود می‌آید و در طی اجرای مرحله کنترل و اصلاح از آن استفاده می‌شود .

بعنوان مثال یک خانوار فرضی در یک سرشماری در زمینه جمعیت را در نظر می‌گیریم که از ۴ عضو تشکیل شده نفر اول سرپرست خانوار و نفر دوم همسر و نفر سوم فرزند سرپرست خانوار گزارش شده است . در یک فرض اولیه پیش بینی می‌شود که اختلاف سن بین پدر و فرزند ۲۵ سال و برای مادر و فرزند ۲۰ سال باشد . چنانچه در خانوار فرضی بالا ناهماهنگی و یا عدم پاسخگوئی در سن یکی از افراد مشاهده شود با استفاده از این پیش فرض جایگزینی عددی درسن آن فرد انجام می‌شود .

در روش Hot Deck همین اطلاعات اولیه بعنوان پیش فرض پذیرفته می‌شود ولی در طی اجرای مرحله کنترل و اصلاح هر بار که در یک خانوار ترکیب یکی از والدین و فرزند مشاهده شود اختلاف سن مشاهده‌شده بجای پیش فرض اولیه قرار می‌گیرد و این عدد با هر بار مشاهده تغییر می‌کند درنتیجه همیشه یکسان و به یک اندازه باقی نمی‌ماند و در هر بار استفاده ، اعداد گوناگونی جایگزین می‌شود و حاصل به متوسط جامعه نزدیکتر است .

حال در خانوار فرضی بالا که یک فرد با همسر و فرزند او گزارش شده است حالتی را در نظر می‌گیریم که سن فردی که بعنوان " فرزند " مشخص شده است از سن دو نفر دیگر بیشتر است که خطا می‌باشد . یک راه اصلاح استفاده از یکی دو روش فوق است (یعنی گزارش " فرزند " برای فرد سوم پذیرفته شود) ولی پی آمد بعدی این اصلاح ممکن است تغییرات دیگری را نیز بوجود آورد . مثلا " در اطلاع مربوط به " آیا دارای فرزند است " نیز باید برای فرد دوم که مناسب جایگزین کرد و از طرف دیگر چون پس از محاسبه سن ، عدد حاصل کمتر از ۱۰ سال شده است ، اطلاع مربوط به اشتغال فرد سوم نیز که گزارش شده است باید حذف گردد و این مجموعه اصلاحات پی در پی منظر غیر منطقی است . در اینصورت راه حل به روش فیلگی برزسی می‌شود که به " مجموعه تغییر " (Nimmal Set Of Tiems For Imputation) شناخته شده است و برای یافتن آن از روابط برنامه‌ریزی خاصی استفاده می‌شود . مثلا " اگر در این مثال برای فرد سوم ، گزارش " فرزند " بودن بستگی دیگری نیز از " فرزند " تبدیل شود فقط با این تغییر سایر صفات بدون تغییر باقی می‌مانند .

مورث فوق جنبه مثال داشته و برای محوری وشبا ارائه شده است و ممکن است واقعی نباشد. ولی درهمین چارچوب صاحبنظران موضوعی آمار و تهیه کنندگان برنامه‌های کامپیوتری قادر خواهند بود با ایجاد رابطه‌های منطقی میزان دقت در اسلاح را تا حد زیادی افزایش دهند.

استنباط آماری از واریانس در حالت بزرگ نمونه‌ای

بهرام سلیمانی

دانشگاه علوم پزشکی اصفهان

روش شناخته‌شده برای محاسبه فاصله اطمینان واریانس یک متغیر تصادفی مانند x و انجام آزمون در مورد آن مبتنی است بر فرض نرمال بودن توزیع این متغیر. اگر این فرض صحیح باشد متغیرهای $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ و $\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{\sigma^2}$ دارای توزیع χ^2 بترتیب با درجه آزادی n و $n-1$ خواهند بود. شرط نرمال بودن توزیع x برای رسیدن به این نتیجه ضرورتی اساسی است. اگر توزیع x نرمال نباشد حدود اطمینان σ^2 را از این طریق نمیتوان بدست آورد. برای نیل به یک راه حل کلی، متغیر x را از جامعه نامحدودی با توزیع غیر مشخص در نظر گرفته و متغیر y را بصورت $y=\frac{(x-\mu)^2}{\sigma^2}$ تعریف میکنیم. در اینصورت میانگین و واریانس y عبارتست از:

$$\mu_y = E(y) = E\left[\frac{(x-\mu)^2}{\sigma^2}\right] = 1$$

$$\sigma_y^2 = E(y-\mu_y)^2 = E\left[\frac{(x-\mu)^2}{\sigma^2}-1\right]^2 = \frac{1}{\sigma^4}E\left[(x-\mu)^2-\sigma^2\right]^2$$

$$= \frac{1}{\sigma^4}\left[E(x-\mu)^4 - 2\sigma^2 E(x-\mu)^2 + \sigma^4\right] = \frac{\mu_4}{\sigma^4} - 1 = \gamma - 1$$

که در آن $\gamma = \frac{\mu_4}{\sigma^4}$ در نظر گرفته شده است. اکنون با توجه به استقلال x_i ها و در نتیجه y_i ها خواهیم داشت:

$$E\left(\sum_{i=1}^{n} y_i\right) = \sum_{i=1}^{n} E(y_i) = \sum_{i=1}^{n} E\frac{(x_i-\mu)^2}{\sigma^2} = n \qquad (1)$$

$$Var\left(\sum_{i=1}^{n} y_i\right) = \sum_{i=1}^{n} Var(y_i) = n(\gamma-1)$$

ازطرفی میدانیم براساس قضیه حد مرکزی برای هرمتغیری مثل y_i هنگامیکه تعداد دفعات زیاداست توزیع $\bar{y}=\frac{\sum_{i=1}^{n} y_i}{n}$ بسمت نرمال میل میکند ضمنا" $\mu_{\bar{y}}=\mu$ ، $\sigma_{\bar{y}}^2=\frac{\sigma_y^2}{n}=\frac{\gamma-1}{n}$.

ازآنجاکه $\sum_{i=1}^{n} y_i = n\bar{y}$ بنابراین $\sum_{i=1}^{n} y_i$ نیز دارای توزیع نرمال بامیانگین n وواریانس $n^2\times\frac{\gamma-1}{n}=n(\gamma-1)$ خواهدبود:

بدین ترتیب چنانچه n بزرگ باشد، معمولا" ۲۰ یابیشتر، متغیر $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ تقریبا" دارای توزیع نرمال بامیانگین n وواریانس $n(\gamma-1)$ میباشد. γ معیاری برای تعیین کشیدگی توزیع x بوده درحالت نرمال مقدارآن برابر ۳است، ازاینرواگر x نرمال باشدواریانس $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ برابرخواهدبودبا $n(\gamma-1)=n(3-1)=2n$ ، وچون هنگامیکه x نرمال‌است متغیرفوق دارای توزیع کای اسکوربامیانگین n وواریانس $2n$ میباشد وضمنا" میدانیم باافزایش n توزیع کای‌اسکوربسمت نرمال میل میکنداین نتیجه با نتایجی که قبلا" دراین زمینه وجود داشته مطابقت دارد.

بنابراین برای هرتوزیع نامشخص وقتی n بزرگ باشدمیتوان حدودزیررابرای سطح اطمینان $1-\alpha$ نوشت.

$$P\left[-z_{\alpha/2} < \frac{\sum_{i=1}^{n} y_i - n}{\sqrt{n(\gamma-1)}} < z_{\alpha/2}\right] \approx 1-\alpha$$

که درآن $z_{\alpha/2}$ نقطه‌ایست که سطح سمت راست‌آن درتوزیع $N(0,1)$ برابر $\alpha/2$ است.

درنتیجه:

$$P\left[-z_{\alpha/2} < \frac{\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} - n}{\sqrt{n(\gamma-1)}} < z_{\alpha/2}\right] \approx 1-\alpha$$

$$P\left[n - z_{\alpha/2}\sqrt{n(\gamma-1)} < \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} < n + z_{\alpha/2}\sqrt{n(\gamma-1)}\right] \approx 1-\alpha$$

$$P\left[\frac{1}{n+z_{\alpha/2}\sqrt{n(\gamma-1)}} < \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\mu)^2} < \frac{1}{n-z_{\alpha/2}\sqrt{n(\gamma-1)}}\right] \approx 1-\alpha$$

بنابراین حدود σ^2 برای سطح اطمینان $1-\alpha$ عبارتست از:

$$P\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n+z_{\alpha/2}\sqrt{n(\gamma-1)}} < \sigma^2 < \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-z_{\alpha/2}\sqrt{n(\gamma-1)}}\right] \approx 1-\alpha \qquad (۲)$$

در توزیع یکنواخت $f(x)=\frac{1}{2\alpha}$ که دامنه تغییرات x از $-\alpha$ تا α است،

اگر مقدار α نامعلوم باشد فاصله اطمینان واریانس بشرح زیر قابل محاسبه است.

چون:

$$\mu = E(x) = \int_{-\alpha}^{+\alpha} x \cdot \frac{dx}{2\alpha} = 0$$

$$\mu_2 = E(x-\mu)^2 = \int_{-\alpha}^{+\alpha} (x-0)^2 \frac{dx}{2\alpha} = \frac{\alpha^2}{3}$$

$$\mu_4 = E(x-\mu)^4 = \int_{-\alpha}^{+\alpha} (x-0)^4 \frac{dx}{2\alpha} = \frac{\alpha^4}{5} \quad \Rightarrow \quad \gamma = \frac{\mu_4}{\mu_2^2} = \frac{\frac{\alpha^4}{5}}{(\frac{\alpha^2}{3})^2} = \frac{9}{5} = 1.8$$

و باتوجه به رابطه (۲) خواهیم داشت:

$$P\left[\frac{\sum_{i=1}^{n} x_i^2}{n+z_{\alpha_2}\sqrt{n(1.8-1)}} < \sigma^2 < \frac{\sum_{i=1}^{n} x_i^2}{n-z_{\alpha_2}\sqrt{n(1.8-1)}}\right] \approx 1-\alpha$$

و سپس:

$$P\left[\frac{\sum_{i=1}^{n} x_i^2}{n+z_{\alpha_2}\sqrt{0.8n}} < \sigma^2 < \frac{\sum_{i=1}^{n} x_i^2}{n-z_{\alpha_2}\sqrt{0.8n}}\right] \approx 1-\alpha$$

بطور مثال با استفاده از جدول اعداد تصادفی چهل نمونه دو رقمی بشرح زیر استخراج گردید و سپس متغیر x از تفاضل این ارقام و میانگین کل اعداد صفر تا ۹۹ یعنی $\mu = 49/5$ بدست آمد.

۸	۱۸	۳۴	۶۴	۲۱	۸۵	۷۳	۴۰	۶۸	۷۲
۴۴	۱۹	۲۷	۵۸	۸۰	۳۸	۸۸	۶۲	۱۷	۱
۹۸	۶۸	۶۵	۲۲	۹۰	۹۲	۵۲	۲۵	۱۷	۱۳
۶۸	۳	۲۱	۸۴	۶۳	۹۲	۱۱	۲۲	۲۶	۲۸

نتایج حاصله آنگاه بشرح زیراست : $\sum_{i=1}^{n} x_i^2 = ۲۳۱۲۶,۷۷۵$

$$P\left[\frac{۲۳۱۲۶,۷۷۵}{۴۰+۱,۹۶\sqrt{\frac{۲\times۵}{۵}}} < \sigma^2 < \frac{۲۳۱۲۶,۷۷۵}{۴۰-۱,۹۶\sqrt{\frac{۲\times۵}{۵}}}\right] \approx ۰,۹۵$$

$$P\left[۶۴۸,۴۳۲۹۳ < \sigma^2 < ۱۱۴۵,۷۵۷۰۱\right] \approx ۰,۹۵$$

$$P\left[۲۵,۴۶ < \sigma < ۳۳,۸۵\right] \approx ۰,۹۵$$

(ازنظربیرناچیز پیوستگی صرفنظرشده‌است)

ومیدانیم دراین توزیع‌مشخص‌که $f(x)=\frac{۱}{۹۹}$ وتغییرات x از ۴۹/۵ تا ۱۴۹/۵ است $\sigma^2 = ۸۳۳/۲۵$ و $\sigma = ۲۸/۸۷$ میباشد.

اگرنمونه‌های‌فوق مربوط به توزیعی‌یکنواخت بامقدارنامعلوم α ودامنه تغییرات R میبود،

چون درآن حالت $\sigma^2=\frac{\alpha^2}{۱۲}$ و $R=۲\alpha$ ، فواصل‌اطمینان‌برای α و R نیزقابل محاسبه میگردید.

درفرم عمومی‌توزیع‌یکنواخت بصورت $f(x)=\frac{۱}{d-c}$ ابتدابایدابتدا $\mu=\frac{c+d}{۲}$ محاسبه گردد، ازآنجاکه مقادیر c و d نامعلوم‌اند μ رامیتوان بوسیله $\bar{x}$ برآوردوجایگزین نمود، نظربه‌اینکه تعداد نمونه زیاداست انتظار میرود خطای ناشی‌ازاین‌جایگزینی‌اندک باشد.

متغیر $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ همواره مثبت بوده وتوزیع‌آن درسمت راست محور افقی قرار دارد، حدپائین‌آن‌صفروحدبالاباتوجه به نوع توزیع x متفاوت‌است، درتوزیع‌یکنواخت حد بالا ۲۴ میباشد، دراین‌حالت اگر n بزرگ‌باشد سطح زیرمنحنی‌نرمال‌برای‌مقادیربیش از ۲۴ بسیارناچیز خواهدبودوازآنجاکه برای‌محاسبه فواصل اطمینان فقط به مقادیرمیانی توزیع که درطرفین میانگین قرارداردنیازاست این تفاوت جزئی باتوزیع‌نرمال مشکل‌قابل بحثی‌راایجادنمیکند.

آزمون فرضیه $\sigma_1^2 = \sigma_2^2$

برای مقایسه واریانسهای دوجامعه مستقل، درصورتیکه فرضیه صفر $\sigma_1^2 = \sigma_2^2$ صحیح باشدملاک زیردارای توزیع نرمال $N(0,1)$ است.

$$\frac{\left[\sum_{i=1}^{n_1}\frac{(x_{1i}-\mu_1)^2}{\sigma_1^2} - \sum_{i=1}^{n_2}\frac{(x_{2i}-\mu_2)^2}{\sigma_2^2}\right] - (n_1-n_2)}{\sqrt{n_1(\gamma_1-1)+n_2(\gamma_2-1)}}$$

چون مقادیر σ_1^2 و σ_2^2 نامعلوم اندمیتوان از واریانس ادغام‌شده S_p^2 استفاده نمود. در اینصورت چنانچه مقدار:

$$\frac{\frac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2 - \sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{S_p^2} - (n_1-n_2)}{\sqrt{n_1(\gamma_1-1)+n_2(\gamma_2-1)}} \qquad (۴)$$

درخارج ازفاصله $Z_{\alpha/2}$ و $-Z_{\alpha/2}$ قرارگیرد فرضیه صفر رد میشود. این آزمون را با استفاده از حدود اطمینان زیر نیز میتوان انجام داد:

$$\frac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2 - \sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{(n_1-n_2)+Z_{\alpha/2}\sqrt{n_1(\gamma_1-1)+n_2(\gamma_2-1)}} < \sigma^2 < \frac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2 - \sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{(n_1-n_2)-Z_{\alpha/2}\sqrt{n_1(\gamma_1-1)+n_2(\gamma_2-1)}} \qquad (۵)$$

که در آن $\sigma^2 = \sigma_1^2 = \sigma_2^2$ است. اگر S_p^2 خارج از این فاصله قرار داشته باشد فرضیه صفر رد میشود.

برای آزمونهای یک دامنه با فرض صفر $\sigma_1^2 \geq \sigma_2^2$ نیز از همین روش میتوان استفاده نمود جز آنکه نقاط بحرانی Z_α و $-Z_\alpha$ خواهند بود.

ازآنجاکه واریانس همواره مثبت‌است از رابطه (۱) نتیجه میگردد $n(\gamma-1) \geq 0$ ،

چون n نیز مثبت است پس $\gamma \geq 1$ ، این نتیجه مستقل ازنوع توزیع و n است بنابراین :

$$\mu_4 \geq \sigma^4 \quad \text{یا} \quad \mu_4 \geq \mu_2^2$$

بنیانترتیب برای کلیه توزیعهای احتمال، روی‌میدان تغییرات متغیرنامساوی‌زیربرقراراست :

$$\int_S (x-\mu)^4 f(x)\,dx \geq \left[\int_S (x-\mu)^2 f(x)\,dx\right]^2$$

ودرتوزیعهای‌ناپیوسته :

$$\sum_{i=1}^{n} (x_i-\mu)^4 p(x_i) \geq \left[\sum_{i=1}^{n} (x_i-\mu)^2 p(x_i)\right]^2$$

نامساویهای‌فوق راهمچنین میتوان نشان داد .

$$\int_S x^4 f(x-\mu)\,dx \geq \left[\int_S x^2 f(x-\mu)\,dx\right]^2$$

و :

$$\sum_{i=1}^{n} x_i^4 p(x_i-\mu) \geq \left[\sum_{i=1}^{n} x_i^2 p(x_i-\mu)\right]^2$$

درتوزیعهائی‌که‌دارای‌میانگین‌صفرهستند مانندتوزیع Z روابط فوق بصورت زیرخواهدبود .

$$\int_S x^4 f(x) \geq \left[\int_S x^2 f(x)\,dx\right]^2$$

و :

$$\sum_{i=1}^{n} x_i^4 p(x_i) \geq \left[\sum_{i=1}^{n} x_i^2 p(x_i)\right]^2$$

پس‌دراینگونه توزیعهاهمچنین $\mu_4 \geq \mu_2^2$ میباشد .

از روابطفوق نامساویهای‌جالب توجهی‌میتوان بدست‌آورد .مثلا " یکمک توزیع پواسن

با $\lambda = 1$ نامساوی زیربدست میآید .

$$\frac{\left[1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots\right]}{\left[1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots\right]^2} \leq e$$

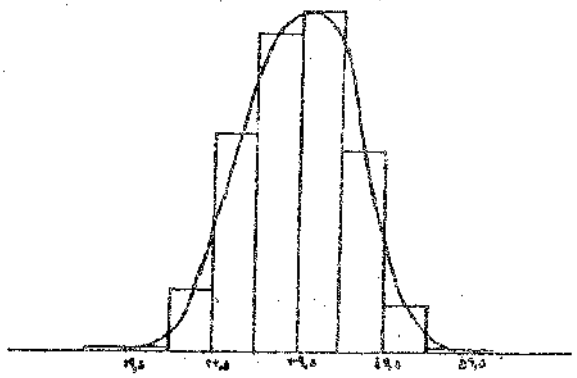
درتوزیع یکنواخت اگر $n=40$ باشد حدود $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ بین صفر تا ۱۲۰ است

اما چون $\sqrt{n(n-1)}\cong 5/66$ ، تقریباً ۹۹/۷ درصد از مقادیر در فاصله $n \pm 3\sqrt{\frac{n}{5}}$

یعنی از ۲۳/۰۲ تا ۵۶/۹۸ قرار دارند و فقط ۰/۳ درصد از مقادیر در خارج از این فاصله واقع میباشند.

بمنظور تحقیق درمورد رابطه (۲)، ۲۶۷ نمونه چهل تائی از اعداد دورقمی جدول اعداد تصادفی استخراج گردید و با توجه به اینکه $\mu=49/5$ و $\sigma^2=822/25$ است، $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ محاسبه شد، آنگاه نتایج بشرح زیر گروهبندی و توزیع فراوانی آن رسم گردید.

$\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$	فراوانی
۱۵ـ۱۹	۱
۲۰ـ۲۴	۱
۲۵ـ۲۹	۱۴
۳۰ـ۳۴	۴۹
۳۵ـ۳۹	۷۱
۴۰ـ۴۴	۷۴
۴۵ـ۴۹	۴۵
۵۰ـ۵۴	۱۱
۵۵ـ۵۹	۱



این نمودار که برابر با داده‌های مشاهده رسم شده است بخوبی یافته‌های نظری را تائید مینماید. بدین ترتیب انتظار میرود گرایش $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ بسمت نرمال برای توزیعهای کوهی شکلی که عدم تقارن آنها زیاد نیست بیشتر بوده و حتی در نمونه‌های با حجم کمتر نیز تقریب نرمال از نظر عملی بهاندازه کافی قابل استفاده باشد.

برای یافتن از توزیعهای χ^2 مقدار مشخص است، همانطور که ذکر شد در توزیع نرمال مقدار آن ۳ و در توزیع یکنواخت ۱/۸ است همچنین در توزیع پواسن $\frac{1}{\mu}=\chi^2$ است.

چنانچه μ و σ معلوم باشند از رابطه (۲) میتوان فاصله اطمینان را برای واریانس بدست آورد. اگر این پارامترها معلوم نباشند با توجه به بزرگبودن n برآوردهای این دو پارامتر که از داده‌های موجود بدست میآیند از نظر عملی به اندازه کافی قابل استفاده خواهند بود. بویژه انتظار میرود بعلت تقسیم $\hat{\mu}$ بر $\hat{\mu}'$، مقدار خطا در برآورد کاهش یابد. در اینگونه موارد میتوان توزیع $\sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2}$ را بدست آورد. با توجه به اینکه نرمال بودن x مفروض نیست، $\bar{x}$ و s^2 از هم مستقل نبوده و در مورد توزیع $\frac{(n-1)s^2}{\sigma^2}$ ابهام وجود دارد ولی بعلت بزرگ بودن n تفاوت $\bar{x}$ با μ ناچیز است و از اینرو نتایج حاصل از روش ذکر شده تقریب نسبتاً خوبی برای حدود σ^2 میباشد.

آزمون فرضیه $\sigma^2 = \sigma_0^2$

بمنظور آزمون برابری واریانس جامعه با عدد ثابت σ_0^2 میتوان از رابطه (۲) استفاده نمود. چنانچه مقدار σ^2 بین حدبالا و پائین قرار داشته باشد فرضیه $\sigma^2 = \sigma_0^2$ رد نمیشود ولی اگر خارج از این دو حد قرار گیرد فرضیه صفر با احتمال خطای α رد میشود. همین نتیجه را میتوان بامحاسبه:

$$\frac{\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma_0^2} - n}{\sqrt{n(2-1)}}$$

ومقایسه آن با مقادیر بحرانی مربوط به سطح معنی‌دار α در جدول سطح زیر منحنی نرمال بدست آورد. اگر مقدار آن بین $z_{\alpha/2}$ و $z_{1-\alpha/2}$ باشد فرضیه صفر رد نمیشود در غیر اینصورت این فرضیه با احتمال خطای α رد میشود.

این آزمون را بصورت یک دامنه نیز میتوان انجام داد. در چنین حالتی مقدار حاصل از شاخص فوق متناسباً با مقادیر z_α یا $z_{1-\alpha}$ مقایسه میگردد.

منابع :

1 - Snedecor, G.W., and Cochran, W.G. , Statistical Methods, 6 th edn. ames, Iowa : University press , 1973 .

2 - Mood , A.M. , Graybill, F.A. , and Boes, D .C. Introduction to the theory of Statistics, Mc Graw - Hill Book company , NewYork , 1974 .

3 - Dixon, Wilfrid, J., and Massey , Jr. Frank J., Introduction to Statistical Analysis, 4 th edn: Mc Graw - Hill , 1983 .

4 - Armitage, P. and Berry , G., statistical methods in medical research, 2 nd edn : Blackwell Scientific publication , 1987 .

۵- دکتر خواجه نوری ، عباسقلی ، آمار ریاضی ، جلد اول ودوم ، موسسه آموزش عالی آمار تهران ، ۱۳۴۱ .

۶- هاگ ، روبرت، وکریگ ، آلن ، مقدمه‌ای برآمار ریاضی ، ترجمه دکتر آبزد دوستدار، نوروز ، دانشگاه تهران ، چاپ سوم ، ۱۳۶۶ .

۷ـ دکتر بهبود بان ، جواد ، آمار ریاضی ، انتشارات امیرکبیر ، تهران ، چاپ اول ، ۱۳۷۰ .

۸ـ فروند، جان ، و والپول ، رانلد ، آمار ریاضی، ترجمه دکتر عمیدی ، علی ، ودکتر وحیدی اصل ، محمد قاسم ، مرکز نشر دانشگاهی ، تهران ، چاپ اول ، ۱۳۷۰ .

۹ـ وواکات ، توماس ، اچ . و وواکات ، رانلدجی . ، آمار مقدماتی ،ترجمه دکترمشکانی ، محمد رضا ، جلد اول ودوم ، مرکزنشردانشگاهی ، چاپ دوم ، ۱۳۷۰ .

۱۰ـ باتاچاریا ، گوری ، ك . ، وجانسون ، ریچارد ا . ، مفاهیم و روشهای آماری ، ترجمه دکتر ابن شهرآشوب ، مرتضی، ومیکائیلی ، فتاح ، جلد اول ودوم ، مرکزنشردانشگاهی، تهران ، چاپ اول ، ۱۳۶۶ .

/ك/۲/۲۸

تعیین مدل توابع تقاضای گاز طبیعی در بخشهای صنعتی-تجاری و خانگی در ایران

محمد رضا سینائی

مرکز آمار ایران

چکیده

کاربرد روشهای آماری در ارتباط با داده‌های اقتصادی و تجزیه و تحلیل این داده‌ها اساس بحث اقتصاد سنجی می‌باشد، این تجزیه و تحلیل می‌تواند به صورت آزمون نظریه‌های مختلف اقتصادی ، استفاده از نتایج در تصمیم گیری و امکان پیش بینی متغیرهای اقتصادی تجلی نماید . از آنجائیکه با رشد جمعیت و گسترش فعالیتهای اقتصادی تقاضای انرژی در بخش های مختلف اقتصادی افزایش یافته است نیاز به برنامه‌ریزی در بخش انرژی بیش از پیش احساس می‌شود و برای اینکار پیش بینی میزان مصرف هریك از انواع انرژی در سالهای آینده لازم است .

دراین مقاله ضمن اشاره به روشهای مختلف پیش بینی مصرف انرژی با استفاده از یکی از روشهای اقتصاد سنجی و با توجه به آمار موجود در زمینه مصرف گاز طبیعی در سالهای گذشته توابع تقاضای گاز طبیعی در بخش های صنعتی، خانگی و تجاری برآورد گردیده و کشش پذیری تقاضا نسبت به قیمت و تولید ناخالص داخلی محاسبه شده است .

توضیح : این مجموعه با راهنمائی جناب آقای دکتر فخرائی استاد موسسه عالی پژوهش در برنامه‌ریزی و توسعه تهیه شده است .

بسمه تعالی

مقدمه :

رشد جمعیت وافزایش فعالیتهای اقتصادی باعث افزایش مصرف انرژی دربخشهای اقتصادی وبخش خانگی خواهدشد تامین انرژی مورد نیاز این بخش ها یکی از اقدامات اساسی و زیر بنائی دولت محسوب می‌شود . برنامه‌ریزی و سرمایه‌گذاری در زمینه‌های مختلف زیر بنائی و تولیدی و بهره برداری ازاین سرمایه گذاریها نیاز به مدت زمان طولانی (حدود ۸ تا ۱۰ سال) دارد دراین فاصله اگر تقاضای هریک از انواع انرژی مشخص نباشد سرمایه گذاری متناسب با مصرف با مشکل مواجه خواهد شد، از طرف دیگر کشور ما از نظر منابع انرژی محدودیت قابل توجهی ندارد لیکن برای تبدیل انرژی اولیه به انرژی نهائی نیاز به ایجاد تاسیسات و پالایشگاههائی داردبنابراین در تخصیص انرژی متناسب با وضعیت انرژی مملکت دقت بیشتری لازم است انجام شود. بطورمثال درحالیکه ذخائر تثبیت شده گازطبیعی ایران درسال ۱۳۶۹ حدود ۱۷ ترلیون (17×10^{12}) مترمکعب برآوردمی‌شودکه درحدود ۱۴/۳ درصد ذخائر گاز طبیعی جهان است و مقام دوم را دارامی‌باشد اختصاص ارز قابل توجهی برای واردکردن مقدار ۲۵/۳ میلیون بشکه معادل نفت خام فرآورده‌های نفتی درسال ۱۳۶۹ بهیچ وجه اقتصادی بنظر نمی‌رسد، به این منظور دولت به سرعت درحال ایجاد تاسیسات گاز رسانی برای مصارف خانگی ، تجاری و صنعتی است تابتواند مصرف گازطبیعی را به جای مصارف سوخت دیگرمانندنفت سفیدوگازوئیل جایگزین نماید .

سهم مصرف فرآورده‌های نفتی دربخشهای خانگی، تجاری وصنعت درکل مصرف انرژی هریک از این بخشها به ترتیب از ۲۳/۲ درصد و ۵۲/۵ درصد درسال ۱۳۶۸ به ۲۱/۱ درصد و ۴۶/۷ درصددرسال ۱۳۶۹ تقلیل یافت امابا افزایش سهم گازطبیعی درکل مصرف انرژی بخشهای مزبور از ۱۲/۷ درصد و ۲۲/۱ درصد درسال ۱۳۶۸ به ۱۲/۷ درصد و ۲۶/۲ درصد درسال ۱۳۶۹ ، برنقش این سوخت در تامین نیازهای انرژی کشور افزوده شد . سهم گازطبیعی درکل تولیدانرژی اولیه کشوراز ۲۰/۸ درصد درسال ۱۳۶۸ به ۲۲/۲ درصد درسال ۱۳۶۹ ارتقاء یافت از این مقدار به ترتیب ۱۰/۲ درصد و ۱۰/۵ درصد مربوط به گاز سوزانده و تزریق شده بوده‌است .

بنابراین سهم گاز طبیعی تحویلی به شبکه انتقال و توزیع درکل تولیدانرژیهای اولیه ۱۰/۶ درصد و ۱۱/۷ درصد در سالهای ۱۳۶۸ و ۱۳۶۹ بوده‌است .

روند نزولی واردات انرژی که از سال ۱۳۶۲ شروع شده بود درسال ۱۳۶۹ نیز ادامه یافته‌ومقدار آن از ۲۶/۸ میلیون بشکه معادل نفت در سال ۱۳۶۸ به ۴۵/۲ میلیون بشکه معادل نفت تنـــــزل پیدا کرد. اما صادرات انرژی به حداکثر مقدار خود در ده سال گذشته افزایش یافت ودرسال ۱۳۶۹ مقدار ۹۲۶/۵ میلیون بشکه‌معادل نفت خام انرژی صادر گردید.

پیش بینی تقاضای گاز طبیعی دربخش های مختلف خانگی ، تجاری وصنعت در سالهای آینـــده بخوبی می‌تواند افق روشنی‌برای تخصیص منابع سرمایه‌ای به منظور ایجاد تاسیسات گاز رسانـــی باشد به این منظور دراین مقاله ضمن اشاره به‌روشهای مختلف برآورد تقاضای انرژی، توابـــع تقاضای گاز طبیعی در بخشهای خانگی، تجاری وصنعت محاسبه‌خواهد شد. که بادردست داشتن این توابع به راحتی می‌توان مصرف گاز طبیعی رادرسالهای موردنظرمحاسبه کرد.

مقایسه مصرف فرآورده‌های نفتی‌وگاز طبیعی‌درسالهای ۱۳۵۶ تا ۱۳۶۹

بررسی ارقام مربوط به مصرف فرآورده‌های نفتی و مصرف گاز طبیعی‌بخوبی‌نشان‌دهنـــده سرعت جایگزینی این دونوع انرژی می‌باشد، درحالیکه مصرف فرآورده‌های نفتی دربخش خانگـــی و تجاری درسال ۱۳۵۶ معادل ۵۱/۲ میلیون بشکه نفت خام بوده و درسال ۱۳۶۹ به ۱۰۵/۲ میلیون بشکه نفت خام رسیده یعنی چیزی حدود ۱۰۵/۲ درصد در طی این چهارده‌سال رشد داشتـــه، مصرف گازطبیعی دراین بخش درسال ۱۳۵۶ ۱/۲ میلیون بشکه معادل نفت خام بوده‌ودرســـال ۱۳۶۹ به ۲۱/۹ میلیون بشکه معادل نفت خام رسیده است یعنی‌حدود ۱۵۸۴/۶ درصد رشـــد داشته است.

این‌ارقام در بخش صنعت نیز نشان دهنده رشد بیشتر مصرف گاز طبیعی نسبت به مصـــرف فرآورده‌های نفتی است. (جدول ۱)

جدول ۱ ـ مصرف فرآورده‌های نفتی و گاز طبیعی در بخش‌های خانگی تجاری و صنعت در سالهای مختلف

(واحد: یک میلیون بشکه معادل مقدار خام)

نوع انرژی / نام بخش	۱۳۵۶	۱۳۶۱	۱۳۶۶	۱۳۶۷	۱۳۶۸	۱۳۶۹
فرآورده‌های نفتی						
خانگی و تجاری	۵۱/۲	۶۱/۵	۸۲/۵	۸۲/۱	۱۰۲/۶	۱۰۵/۴
صنعت	۳۵/۱	۴۲/۳	۵۴/۹	۵۴/۱	۴۲/۷	۴۶/۴
گاز طبیعی						
خانگی و تجاری	۱/۲	۲۲/۲	۱۶/۷	۱۶/۷	۲۲/۸	۲۱/۹
صنعت	۶/۱	۹/۶	۱۵/۵	۱۵/۵	۲۲/۸	۲۳/۶

حالت کلی تابع تقاضا

برای محاسبه تقاضای یک کالا، معمولاً فرض می‌شود که مقدار تقاضای کالا، تابعی از سطح درآمد، قیمت کالای مربوطه و قیمت سایر کالاهای مکمل یا جایگزین است. این موضوع را به شکل ریاضی به صورت زیر میتوان نشان داد.

$$Q = F(Y, P_x, P_1, P_2)$$

که در معادله فوق، Q مقدار مصرف کالای مورد نظر، Y میزان درآمد، P_x قیمت کالای مورد نظر، و P_1 و P_2 قیمت کالاهای مکمل و جایگزین کالای مورد نظر می‌باشند.

روش فوق براساس تئوری کلاسیک رفتار مصرف کننده قرار دارد. براساس این تئوری، هر مصرف‌کننده که دارای درآمد خاصی می‌باشد از هر کالا به مقداری تقاضا می‌کند که تابع مطلوبیت خود را به حداکثر برساند.

روشهای پیش بینی مصرف انرژی

بطور کلی، روش های مختلف پیش بینی و برآورد مصرف انرژی براساس مدل سازی از روند تقاضای مشتقه مصرف انرژی که در قسمت فوق بدان اشاره شد، قرار دارند. روش های متداول پیش بینی مصرف انرژی عبارتند از:

۱ - روش برون‌یابی روند Trend Extrapolation و روش سری زمانی Time Series Forecasting

دراین روش ها براساس روند های گذشته مصرف انرژی و معمولا " بدون استفاده از متغییرهای مستقل مصرف آینده انرژی پیش بینی میشود . این روش ها شامل روش های باکس - جنکینز (Box - Junkins) و تلفیق فرآیندهای اتورگرسیو - میانگین متحرک (Autoregressive - Moving Average) و دیگر روش های مشابه می‌باشند . اکنون با رواج مدلهای پیشرفته‌تر، از این روش‌ها فقط برای پیش بینی‌های کوتاه مدت (یک تا دوساله) استفاده می‌شود .

۲ - روش کلی اقتصاد سنجی (Aggregate Econometric Method)

در این روش مجموع مصرف انرژی بخش های مختلف بطور کل مورد بررسی قرارمی‌گیرد و مصرف کل هر نوع انرژی بصورت خطی یا لگاریتمی تابعی از متغییرهای مستقل مثل تولید ناخالص ملی، قیمت انواع انرژی و دیگر متغییرهای اقتصادی - اجتماعی فرض می‌شود .

بعلت رشد ناهمگون بخش های مختلف مصرف کننده و شدت مصرف متفاوت انرژی دراین بخش‌ها، تلفیق بخش های مختلف مصرف کننده با یکدیگر ، سبب ضعف این روش و غیر قابل اعتماد بودن پیش بینی بدست آمده‌می‌شود .

۳ - روش خرد اقتصاد سنجی (Disaggregate Econometric Method)

دراین روش مصارف انرژی بخش های مختلف بطور جداگانه مورد بررسی قرارمی‌گیرند و مصرف کنندگان به بخش های همگون و متجانس تقسیم می‌گردند تا عوامل رفتاری ، اقتصادی وتکنولوژیکی موثر بر مصرف انرژی هر بخش جداگانه تعیین گردد . دراین روش ضعف های ذکر شده روش کلی اقتصاد سنجی ، که ازتلفیق بخش های مصرف کننده ناشی می‌شود ، وجود ندارد . اضافه بر آن بعلت مشخص بودن تعداد مصرف کنندگان هر بخش ، می‌توان از این اطلاعات اضافی برای مدل سازی بهره گرفت . مثلا " می‌توان مدل سازی برای هرنوع انرژی را براساس متوسط مصرف هر

مصرف کننده قرارداد و مصرف کل هر نوع انرژی را، مطابق فرمول زیر به دوقسمت تجزیه کرد :

مصرف کل هر نوع

(تعداد مصرف کنندگان) ، (متوسط مصرف هرمصرف کننده) = انرژی هر بخش

مصرف کننده

که تغییرات مصرف کل درنوع انرژی هر بخش مصرف کننده را به تغییرات متوسط مصرف هر مصرف کننده و تغییرات تعداد مصرف کنندگان تجزیه کرده و بدین طریق میتوان تغییرات مصرف کل را بنحو بهتری توضیح داده و دنبال کرد ، حال آنکه درمدل کلی اقتصاد سنجی چنین امری امکان پذیر نیست .

لازم به توضیح است که دراین مدل ، متوسط مصرف هر مصرف کننده هر نوع انرژی ، در بخش های مختلف ، تابع متغییرهای مختلف می باشد ، مثلا " متوسط مصرف برق هر خانوار تابعی از متوسط تعداد وسائل برقی خانگی موجود، متوسط تعداد افراد خانوار، متوسط درآمد خانوار وقیمت انواع انرژی می باشد .

۲ - مدل وسایل مصرف کننده نهائی End Use Equipment Model

مدل های " وسایل مصرف کننده نهائی " تلفیقی از روش های تکنیکی، اقتصادی وحسابداری می باشد این روش ها، تاثیر تغییرات جمعیتی، اقتصادی ، و تکنولوژیکی را مستقیما " بر روی مصرف انرژی در بخش های مختلف مصرف کننده بوسیله ابزار و وسایل مصرف کننده انرژی اندازه می گیرند . بخش های مصرف کننده و ابزار مصرف کننده انرژی معمولا " بصورت واحدهای فیزیکی ، مثل تعداد و یا درصد خانواده هایی که دارای یخچال می باشند و میزان مصرف برق یک یخچال، اندازه گیری و ارزیابی می شوند . سپس نتیجه این ارزیابی ها برای سالهای آینده پیش بینی وشبیه سازی می شود . در هربخش مصرف کننده انرژی ، تعداد وسایل و دستگاههای عمده مصرف کننده انرژی مشخص می گردند و میزان مصرف انرژی هر وسیله جداگانه تخمین زده می شود . بنابراین ، برای مجموع

تمام وسایل مصرف کننده انرژی هر بخش :

(کارائی وسیله) (میزان مصرف انرژی هروسیله) (تعدادوسیله مصرف کننده) تعیین کننده تمام انرژی مصرفی درآن بخش می‌باشد . روش های مدل سازی برای قسمت های مختلف این معادله بقرار زیرمی‌باشند :

" تعداد (یادرصد) هروسیله مصرف کننده " با استفاده از روش رگرسیون لوجیت ساده یامرکب (Simple or Multinomial Logit) بعنوان تابعی از هزینه‌های خرید، نصب و استفاده از وسیله مصرف کننده (که بستگی به قیمت واحدانرژی مصرفی وسیله دارد) برآورد می‌گردد و مسئله جایگزینی وسائل مصرف کننده انواع مختلف انرژی، که کاریکسانی را انجام می‌دهند ، نیز مدل سازی می‌شود . " میزان مصرف انرژی هر وسیله " تابعی از متغییرهای رفتاری (میزان استفاده از وسیله)، اقتصادی (قیمت واحد انرژی) و تکنولوژیکی (حداقل میزان انرژی لازم) می‌باشد و از روش های مختلف مانندروش معادلات تقاضای شرطی (Conditional Demand Equation) برآورد و پیش بینی می‌شود . " کارآئی وسیله مصرف کننده " تابعی از عمر وسایل موجودو کارائی نسبی وسایل جدید بوده و با جایگزینی وسایل نوبجای وسایل کهنه ، متوسط کارائی وسائل مصرف کننده انرژی افزایش می‌یابد .

این روش مدل سازی اجازه میدهد که تعداد ومیزان مصرف وکارائی وسایل مختلف مصرف کننده انرژی بطور جداگانه درهر بخش شبیه سازی بشوندو روندهای تعیین کننده مصرف انرژی مثل افزایش تعداد وسایل مصرف کننده انرژی ، که تابعی از جمعیت ، درآمد سرانه و رشد اقتصادی مملکت است و افزایش متوسط کارائی وسایل مصرف کننده انرژی، که تابعی از رشد تکنولوژی در سطح جهانی و در سطح کشور می‌باشد، در نظر گرفته شوند .

نکته دیگر این است که درکشورهای درحال توسعه ، که روابط اقتصادی و تولیدی درحال تحول و دگرگونی هستند، در بعضی موارد لازم است میزان انرژی موردنیاز روندهای جدید تولیدی و

میزان مصرف وسائل جدید را با استفاده‌از اطلاعات خارج ازمدل محاسبه نمود. مزیتی‌که‌مدل‌های "وسایل مصرف‌کننده نهائی" به مدل‌های صرفا" اقتصاد سنجی دارندایین است که انجام چنین کاری را امکان پذیرمی‌سازند. بدین ترتیب که می‌توان اطلاعات خارجی‌را درمدل جایگزین ساخته و تاثیر آنها را بر مصرف انرژی محاسبه نمود. یکی‌دیگر از خواص این نوع مدل‌ها اینستکه با استفاده از آنها میتوان تاثیرات جایگزینی‌انواع انرژی را، بصورت جایگزینی‌وسائل مصرف کننده نوع‌های‌مختلف انرژی بجای یکدیگر، نشان داد و از آنها برای شناسائی‌امکانات وطرق مدیریت مصرف انرژی استفاده نمود.

منابع آمار :

آمار مصرف گاز طبیعی بخش های مختلف اقتصادی برای سال های ۱۳۴۶ تا ۱۳۶۸ از "ترازنامه انرژی کشور" بدست آمد. آمار جمعیت کل کشورنیز در ترازنامه انرژی موجود می‌باشد. منبع آمار درآمد و تولید ناخالص ملی، تولید بخش های مختلف اقتصادی و شاخص قیمت ها، "مجموعه اطلاعاتی (قوانین و آمار)" سازمان برنامه‌وبودجه می‌باشد. آمار قیمت‌گاز طبیعی از منابع مختلف جمع آوری شده و برای بخش خانگی و تجاری با استفاده از شاخص قیمت کالاها و خدمات مصرفی و برای بخش صنعتی با استفاده از شاخص بهای عمده فروشی کالاها و خدمات به قیمت ثابت سال ۱۳۶۱ تبدیل شده اند.

مدل های گاز طبیعی

جدول شماره ۲ برآورد مصرف گاز طبیعی در بخش های خانگی، تجاری و صنعتی ، برای دوره تاریخی ودوره برنامه ریزی با استفاده از مدل های انتخاب شده را نشان می‌دهد.

جدول شماره ۲

DEMAND FOR NATURAL GAS
MILLION CUBIC METER

YEAR	RES&COM	INDUSTRIAL	TOTAL
1358	611.7	306.8	918.5
1359	996.1	357.0	1353.1
1360	1443.1	539.6	1982.7
1361	1805.2	939.9	2745.1
1362	2239.5	1403.4	3642.9
1363	2410.8	1541.9	3952.7
1364	2618.2	1888.2	4506.4
1365	2625.7	2158.2	4783.9
1366	2706.6	2983.6	5690.2
1367	2667.1	3473.9	6141.0
1368	2882.6	4740.5	7623.1
1369	3216.9	5780.7	8997.7
1370	3510.1	6779.6	10289.7
1371	3798.7	7949.3	11748.0
1372	4093.3	9319.7	13413.0
1373	4345.4	10460.3	14805.8
1374	4578.7	11740.1	16318.9
1375	4798.2	13176.1	17974.3
1376	5006.9	14787.6	19794.5
1377	5207.3	16596.0	21803.2
1378	5359.0	17854.9	23213.9
1379	5505.1	19209.2	24714.3
1380	5646.4	20666.3	26312.7
1381	5783.7	22233.9	28017.5
1382	5917.4	23920.3	29837.7
1383	6048.1	25734.7	31782.8
1384	6176.1	27686.7	33862.8
1385	6301.7	29786.7	36088.5

الف - بخش خانگی و تجاری

برای مصرف گاز طبیعی بخش خانگی وتجاری، مدل زیر مورد استفاده قرارگرفت.

$$CONRC = -23797.3 + 697.6 \cdot Ln(TNCRC) + 2193.7 \cdot Ln(GDP)$$
$$(-2.37) \quad (4.83) \quad (1.77)$$

First order Autocorrelation Coefficient = 0.454

Adjusted R- Squared = 0.94

Durbin - Watson stat= 1.74

درتابع تقاضای فوق، CONRC میزان مصرف گاز طبیعی بخش خانگی وتجاری، TNCRC تعداد مشترکین بخش خانگی وتجاری و GDP تولید ناخالص داخلی (بدون نفت) به قیمت های ثابت ۱۳۵۲ میباشند. اعداد داخل پرانتز، آماره t قرایب محاسبه شده میباشند.

ب - بخش صنعتی

برای مصرف گازطبیعی بخش صنعتی، مدل زیرمورداستفاده قرار گرفت.

$$Ln(CONIND) = 11.462 - 1.396 \cdot Ln(INDPR) + 2.218 \cdot Ln(TGDP)$$
$$(-1.50) \quad (-5.61) \qquad (2.31)$$

First order Autocorrelation Coefficient = 0.587
Adjusted R - Squared = 0.944
Durbin - Watson Stat = 1.40
No. of observations = 22

درتابع تقاضای فوق، CONIND مصرف گازطبیعی بخش صنعتی، INDPR قیمت فروش گازطبیعی به بخش صنعتی وTGDP تولیدکل ناخالص داخلی به قیمت ثابت ۱۳۵۳ می باشند. اعدادداخل پرانتز، آماره t ضرایب محاسبه شده می باشند.

خط‌مشی حدّ نصاب صف M/G/1 با فراغتهای سرویس‌دهنده

غلامحسین شاهکار

دانشگاه فردوسی مشهد

چکیده:

در این مقاله حدّ M/G/1 با فراغتهای سرویس‌دهنده در حالتی که بازگشت سرویس‌دهنده از فراغت به مشتری سرویس‌دهی وابسته به تعداد مشتریان در سیستم است، مورد بحث قرار می‌گیرد. هدف اصلی بدست آوردن تابعی است که تحت ملاک متوسط هزینه در فاصله‌های تک فراغتی و چند فراغتی انجام می‌شود. تابعی که تعداد بهین مشتریان را برای شروع دوره اشتغال جدید تعیین می‌کند، ارائه شده است. معلوم می‌شود که تحت ملاک متوسط هزینه، تابع به دست آمده می‌تواند صعودیت داشته باشد، بدین معنی که برای رده‌ای کلی از مسائل نظیر آنچه در طرح نوشته و نظریه موجودی (برای ساختار فساد خریدهای که بحث شده است) بیشتر می‌آیند می‌توانند برقرار باشند.

خط مشی حد نصاب در صف M/G/1 با اولویت فراغت سرویس دهنده

یک مدل صف بندی M/G/1 را در نظر بگیرید به‌طوری که هر بار که سیستم خالی می‌شود ، سرویس دهنده فراغتی داشته باشد ، یعنی برای مدتی تصادفی سیستم را ترک کند . سرویس‌دهنده در بازگشت از فراغت سیستم را سرویس می‌کند ، اگر تعداد متقاضیان حاضر در سیستم کمتر از مقدار از پیش تعیین شده‌ای باشد به فراغت دیگر میرود و اگر تعداد متقاضیان حاضر به حد نصاب لازم برسد یا از آن تجاوز کند ، شروع به سرویس متقاضیان می‌نماید ، تا اینکه سیستم دوباره خالی شود (سرویس گامل) . متذکر میشویم که مدل با فراغت چندگانه توسط بچیالی [1] و کوپر [7] حالت خاص مدل فوق‌اند ، وقتی که حد نصاب مقدوش مساوی یک باشد.

به عنوان مثالی کاربردی که در خور مدل باشد ، مسئله کنترل کیفیت زیر را در نظر میگیریم . کارخانه‌ای کالای خاصی را تولید می‌کند ، که بعضی از آنها معیوب‌اند .کالاهای بدون عیب وارد بازار می شوند، در حالی که کالاهای معیوب پس از طرفکدن اشکالشان و رسیدن به وضع مطلوب در انبار نگهداری میشوند. فرض کنید که یکی از ماشینهای ساختار را بشود دوری تعمیر داده (با صرف هزینه‌ای) که بتواند از کالاهای معیوب رفع عیب کند . سئوال این است که حد نصاب تعداد کالاهای معیوب چقدر باشد تا این که ماشینی را از خط نتوانیم به خط تعمیر برگردانیم . بعلت هزینه لازم در تعمیر ماشین از خط نتوانیم به خط تعمیر ، ماشین بعد از ارجاع به خط تعمیر تا رفع عیب کلیه کالاهای معیوب بعلاوه کالاهای معیوب جدید کار میکند.

سپس بعد از تغییر لازم به خط تولید عودت داده می‌شود.

در تطبیق این مسئله کنترل کیفیت با مدل صف‌بندی مطرح شد، کالاهای معیوب را متقاضیان و ماشین خاص را سرویس دهنده تعبیر می‌کنیم، وقتی که سرویس دهنده تنها در صورتی قابل حمل باشد که ماشین در خط تعمیر قرار بگیرد. زمان سرویس همان زمان لازم برای رفع عیب تا حصول کیفیت مطلوب است. وقتی که ماشین در خط تولید است سرویس‌دهنده در شبانه‌ای از فاصله‌ها به سر می برد، که طول هر فاصله زمان لازم برای تولید یک کالاست. اگر فرض کنیم که تولید کالاهای معیوب در ماشین طبق توزیع پواسونی صورت می‌گیرد (ولذا تشکیل یک فرایند تجدید را می دهد)، آنگاه می‌توانیم فرایند خروجی کالاهای معیوب از ماشینهای مختلف را با یک فرایند پواسون تقریب کنیم. بنابراین ما با یک سیستم صف‌بندی M/G/1 با اوقات فراغت چند گانه مواجهیم وقتی که هر پایان هر فراغت تصمیم می‌گیریم که آیا :

(۱) فراغت دیگری داشته‌باشیم یا (۲) سرویس دهی کامل و بر این اساس که آیا تعداد متقاضیان حاضر در سیستم کمتر یا مساوی n است، شروع کنیم. به عبارت کلمه‌تر یکبار برهای بالقوه مدل تحت بررسی شامل هر سیستم صف‌بندی M/G/1 است که در آن سرویس دهنده با خالی شدن سیستم، بمنظور اشتغال در فعالی دیگر و یا داشتن فراغت از انجام وظیفه خود اری نموده شبانه در پایان یک فاصله زمانی تصادفی (که متغیرهای متعاقبی مستقل و همتوزیع‌اند) ممکن است جهت انجام وظیفه فراغت عزیمت نمود.

در اینجا ، مدل عضویتی معرفی شده تحت‌عنوان متوسط هزینه تشخیص می‌شود یکسه به نظر میرسد در خور مسئله کنترل کیفیت بحث شده اخیرا تحت عنوان "کل هزینه در یک دوره اشتغال" باشد. برای این مسائل بر اساس ساده‌ای که نقطه بیطرفانه‌ای در انتخاب را مشخص می‌نماید از آنکه شده است، ملاک متوسط هزینه تنها تحت قواعد ذیل نسبی ، یعنی قواعدی که ضامنه فراغت را تعیین می‌کنند تحلیل شده است. سپس نشان میدهیم که خط مشی عدم نمونه‌گیری در بین دوره فقط هنگامی که تصمیم گیری تنها در هنگام بازگشت از فراغت و یا در شروع هر فراغت دیگری اعمال می‌شود بهترین می‌باشد .

در بعضی موارد ، فراغتها به منظور حفاظت و نوار مسیر بیشتر سرویس دهنده انجام می‌شوند نه به جهت انجام کارهای تولیدی دیگر . در این حالت سرویس دهنده در هر زمان تنها یک فراغت مشاهده داشت که این همان مدل تک فراغت بررسی شده توسط لوکوویچیانی است . در حالت کلی در چنین محیطی محدودیتی برای شروع به کار سرویس دهنده هنگام بازگشت از فراغت وجود ندارد . اما در حالت خاص ، سرویس دهنده ممکن است تا وقتی که تعداد متقاضیان منتظر در سیستم به حد کافی برسد منتظر بماند به نحوی که شروع سرویس ارزش آن را داشته باشد . بنابراین مشابه مدل با فراغت چندگانه ای که در ابتدا تحت معیار متوسط هزینه در فراز مورد بحث قرار میگیرد ، مدل تک فراغتی را نیز تحلیل می‌نماییم . برای این مدل مشتاکر می‌نویسد که اگر سرویس دهنده هنگام بازگشت از

فراغت شده اند متقاضیان لازم را در سیستم نمیبیند، تا وقتی که تعداد متقاضیان به حد کافی برسد، در سیستم منتظر می ماند و در اولین لحظه ای که تعداد مراجعین به حد نصاب مورد نظر رسید، شروع به سرویس میکند تا اینکه سیستم خالی شود.

تعاریف و نمادهای مدل فراغت چندگانه :

یک مدل صف‌بندی M/G/1 با نرخ مراجعه λ و زمان سرویس S را در نظر بگیرید. فرض کنید زمانهای سرویس مستقل و همتوزیع با توزیع H(t) باشند. مطابق معمول فرایند ورود و زمانهای سرویس از هم مستقلند. هر بار که سیستم خالی میشود سرویس‌دهنده به مدت زمانی از فراغتها میرود. طول این فراغتها مستقل، همتوزیع و چون متغیری تصادفی با توزیع کلی هستند که آن را با V و تابع توزیعش را با V(t) نشان میدهیم. اولین لحظه‌ای که سرویس دهنده از فراغتی بر میگردد و حداقل n متقاضی را در سیستم می‌یابد (وقتی که n عدد صحیح و مثبت از پیش تعیین شده‌ای است) فوراً شروع به سرویس دهی کرده تا این که سیستم دوباره خالی شود. عدد n را حد نصاب مینامیم. طول فراغتها، مستقل از فرایند ورود و زمانهای سرویس فرض شده‌اند.

فرض کنید p_i و q_i به ترتیب احتمال مراجعه i متقاضی طی زمان سرویس و یک فراغت باشد. بنابراین

$$p_i=\int_0^\infty \frac{(\lambda t)^i}{i!}e^{-\lambda t}dH(t)\ ,\quad q_i=\int_0^\infty \frac{(\lambda t)^i}{i!}e^{-\lambda t}dV(t) \qquad (1)$$

است . بنابراین فرض کنید که

$$\rho=\lambda E(S) \quad (1$$

چرخه اشتغال و متتابع مربوطه :

فرض کنید θ فاصله زمانی متحدالتوزیعی باشد که از لحظه‌ای که سیستم خالی می‌شود ، شروع و در اولین لحظه پس از آن با مراجعه (ورود) یک متقاضی و خالی شدن دوره سیستم خاتمه پیدا کند . θ را چرخه اشتغال می‌نامیم، که θ^n حد متصاعد تعداد متقاضیان برای نیاز فرض میشود. متمرکز می‌شویم که اگر سیستم خالی باشد و متقاضیی هم فی فراغت ورود نشود ، در این صورت این فراغت یک چرخه اشتغال θ کاملی نیست.

θ و : طول دوره اشتغال صف معمولی M/G/1 در نظر بگیرید که دارای تابع توزیع $B(t)$ است. داریم :

$$\tilde{B}(s)=\tilde{H}\,[s+\lambda-\lambda\tilde{B}(s)] \qquad (2)$$

و

$$E(\theta)=\frac{E(S)}{1-\rho} \qquad (3)$$

فرض کنید K تعداد متقاضیان حاضر در سیستم ، هنگام بازگشت سرویس‌دهنده از اولین فراغت میباشد . اگر $n \leq K$ باشد ، آنگاه بلافاصله چرخه اشتغال θ^n مرکب از k دوره اشتغال $\theta(1)$ ، $\theta(2)$ و ... و $\theta(k)$ استکه توزیع هر یک از

آنها مانند توزیع θ است. زیرا که هر دوره اشتغال در اولین نخلهای که تعداد مشتاقانیان در سیستم یکی کمتر از تعداد آنان در شروع آن دوره اشتغال میشود ، خاتمه پیدا میکند اگر $n \geq 1$ باشد ، آنگاه بانتیجه‌ده فرقه اشتغال n تا ای توزیعی میجورت توزیع مجموع متغیرهای تصادفی مستقل θ و θ دوره اشتغال مستقل $\theta(1)$ و $\theta(2)$ و ... و $\theta(k)$ است که توزیع هر یک مانند توزیع θ میباشد . بنا بر فرمول کرمن این مطالب خواهیم داشت :

$$\tilde{\theta}_n(s) = \sum_{j=0}^{n-1} \int_0^\infty E \exp\left[-s\left(t + \theta_{n-j} + \sum_{i=1}^{j}\theta(i)\right)\right] p_j(t)dV(t)$$

$$+ \sum_{j=n}^{\infty} \int_0^\infty E \exp\left[-s\left(t + \sum_{i=1}^{j}\theta(i)\right)\right] p_j(t)\, dV(t)$$

$$= \sum_{j=0}^{n-1} \int_0^\infty e^{-(s+\lambda)t} \frac{(\lambda\tilde{\theta}(s)t)^j}{j!} (\tilde{\theta}_{n-j}(s)-1)\, dV(t) + \tilde{V}[s+\lambda-\lambda\tilde{\theta}(s)]$$

با استفاده از رابطه (4) و تکنیک مرتبه اول $\theta(\tilde{\theta})$ را به

$$\bar{\theta}_n = -\left.\frac{d\tilde{\theta}_n(s)}{ds}\right|_{s=0} = -\left[\sum_{k=0}^{n-1}\int_0^\infty e^{-\lambda t}\frac{(\lambda t)^k}{k!}(-\bar{\theta}_{n-k})dV(t) - (1+\lambda E\theta)E\bar{V}\right]$$

$$= \sum_{k=0}^{n-1} a_k\, \bar{\theta}_{n-k} + E\bar{V} + \lambda E\theta E\bar{V}$$

بنا به قرارداد $\theta_0 = 0$ میگیریم ، لذا با توجه به رابطه (3) میتوان نوشت :

$$\bar{\theta}_n = EV/(1-\rho) + q_0\bar{\theta}_n + \sum_{k=0}^{n-1} q_{n-k}\bar{\theta}_k$$

پس

$$\bar{\theta}_n = \left[\sum_{k=0}^{n-1} q_{n-k}\bar{\theta}_k + \frac{E(V)}{1-\rho} \right] / (1-q_0) \quad n\geq 1 \ , \ \bar{\theta}_0 = 0 \qquad (6)$$

فرض کنید متشکل‌پیرهای تصادفی V_n و N_n به ترتیب کل مدت زمان فراغت و تعداد فراغت‌های سرویس دهنده طی یک چرخه اشتغال n باشند . پس با توجه به رابطه (6) داریم :

$$E(V_n) = \left[\sum_{k=0}^{n-1} q_{n-k} E(V_k) + E(V) \right] / (1-q_0) \ , \ n\geq 1 \ , \ E(V_0) = 0 \qquad (7)$$

و چون با استفاده از معادله والد ($V_n = \sum_{i=1}^{N_n} V^{(i)}$ که در آن متغیر تصادفی $V^{(i)}$ طول تک فراغت است)

لذا ، از تقسیم طرفین رابطه (7) بر $E(V)$ به دست می‌آید :

$$E(N_n) = \left[\sum_{k=0}^{n-1} q_{n-k} E(N_k) + 1 \right] / (1-q_0), \ n\geq 1 \ , \ E(N_0) \geq 0 \qquad (8)$$

با توجه به روابط (7) و (8) و معادله والد داریم :

$$\frac{E(V_n)}{E(N_n)} = E(V) \ , \ n\geq 1$$

و با استفاده از روابط (6) و (7) نتیجه میشود

$$\frac{E(V_n)}{\bar{\theta}_n} = \frac{\sum_{k=0}^{n-1} q_{n-k} E(V_k) + E(V)}{\sum_{k=0}^{n-1} q_{n-k} \bar{\theta}_k (1-\rho) + E(V)} (1-\rho) \qquad (9)$$

اما میدانیم که در صف معمولی M/G/1 نسبتی از زمان که سرویس دهنده مشغول سرویس دهی نیست $(1-\rho)$ است بنابراین $E(V_1)/\bar{\theta}_1 = 1-\rho$. به استقراء نتیجه میشود

$$\frac{E(V_n)}{\bar{\theta}_n} = (1-\rho) \quad , \quad n \geq 1$$

یعنی جز نسبت $EV_n/\bar{\theta}_n = 1-\rho$ و $EV_n/EN_n = EV$ بستگی به مقدار n ندارد .

از دومین تساوی نتیجه میشود که نسبتی از زمان در دراز مدت که سرویس دهنده مشغول سرویس دهی نیست ، ρ است.

(همیشه دوره‌های اشتغال و بیکاری تشکیل یک فرایند تجدید مستقل میدهند). چنین نتیجه‌ای شاید عجیب باشد ، زیرا که برای هیچ سیستم پیچیده‌ای این موضوع درست است. با این حال روابط (6) تا (8) ، همانطور که بعداً خواهیم دید ، تنها به این منظور به دست نیامده‌اند .

فارمن و گسوین [3] شرایطی میدهند که تحت آن شرایط توزیع وضعیت پایای متقاضیان حاضر در یک سیستم با فرآیند ورود پیچش سه توزیع ساده‌تر به شرح زیر است:

۱) توزیع وضعیت پایای تعداد متقاضیان حاضر در سیستم‌های متناظر و معمولی M/G/1 .

۲) توزیع تعداد متقاضیانی که طی یک فاصله زمانی که به صورت پالایشیمانده عمر فرآیند توزیع شده‌است وارد سیستم میشوند.

۳) توزیع وضعیت پایای تعداد متقاضیان حاضر در سیستم در شروع فرآیند .

به سادگی میتوان تحقیق کرد که محل مشروط ما در همسه آن شرایط صدق میکند. ماتریس تغییر وضعیت $P=(p_{ij})$ برای تعداد متقاضیان حاضر در شروع فرآیندها به صورت زیر است

$$P = \begin{pmatrix} Q_0 & q_1 & q_2 & q_3 & \cdots & q_{n-1} \\ Q_1 & q_0 & q_1 & q_2 & \cdots & q_{n-2} \\ Q_2 & 0 & q_0 & q_1 & \cdots & q_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ Q_{n-1} & 0 & 0 & 0 & & q_0 \end{pmatrix} \qquad (10)$$

که در آن $q_i = 1 - \sum_{j=1}^{n-1} p_{ij}$ و $q_0 = 1$ با رابطه (1) تعریف شده‌اند، بین احتمالات وضعیت پایای زنجیر مارکوف فوق که آن را با Π_i نشان می‌دهیم در دستگاه معادلات

$$\begin{cases} \Pi_i = \sum_{j=1}^{i} q_{i-j} \Pi_j & 1 \le i \le n-1 \\ 1 = \sum_{j=0}^{n-1} \Pi_j \end{cases} \qquad (11)$$

صدق می‌کنند.

قضیه ۱: احتمالات وضعیت پایای تعداد دستگاههای حاضر در سیستم در شروع فرایند

$$\Pi_i = \frac{\bar{\theta}_{i+1} - \bar{\theta}_i}{\bar{\theta}_n} \quad , \quad i = 0,1,\ldots,n-1 \qquad (12)$$

است.

اثبات: با توجه به رابطه (5) می‌توان نوشت:

$$\frac{\bar{\theta}_{i+1} - \bar{\theta}_i}{\bar{\theta}_n} = \frac{\sum_{k=0}^{i+1} q_k \bar{\theta}_{i+1-k} - \sum_{k=0}^{i} q_k \bar{\theta}_{i-k}}{\bar{\theta}_n} = \frac{\sum_{k=0}^{i} q_k (\bar{\theta}_{i-k+1} - \bar{\theta}_{i-k})}{\bar{\theta}_n}$$

با تغییر زیر نویس k به $j=i-k$ داریم:

$$\frac{\bar{\theta}_{i+1} - \bar{\theta}_i}{\bar{\theta}_n} = \sum_{j=0}^{i} q_{i-j} \frac{\bar{\theta}_{j+1} - \bar{\theta}_j}{\bar{\theta}_n} \qquad (13)$$

اما چون ماتریس تشخیص وضعیت P (در رابطه (10)) مربوط به یک زنجیر مارکوف تحویل ناپذیر تناوری‌ای با فضای وضعیت متناهی‌است، جواب دستگاه معادلات (11) مثبت ، منحصر به فرد و برابر هر دو توزیع حدی و مانایست . پس ، از مقایسه دو رابطه (11) و (13) ، برقراری رابطه (12) نتیجه می‌گردد .

حال فرض کنید $g_n(z)$ تابع مولد تعداد متقاضیان حاضر در سیستم (تحت خط مشی حد نصابی) باشد. پس

$$g_n(z)=\left[\frac{(1-\rho)(1-z)}{\tilde{H}(\lambda-\lambda z)-z}\,\tilde{H}(\lambda-\lambda z)\right]\left[\frac{1-\tilde{V}(\lambda-\lambda z)}{\lambda E(V)(1-z)}\right]\left[\frac{\sum_{i=0}^{n-1}(\theta_{i+1}-\theta_i)z^i}{\theta_n}\right] \quad (14)$$

که در آن سه عامل سمت راست به ترتیب (از چپ به راست) عبارتند از :

۱) تابع مولد احتمال توزیع وضعیت پایای تعداد متقاضیان در نظام استاندارد و معمولی M/G/1 .

۲) تابع مولد احتمال توزیع تعداد متقاضیانی که در یک تناوب‌زمانی که بصورت باقیمانده عمر فرآیند توزیع شده‌است، وارد سیستم می‌شوند .

۳) تابع مولد احتمال توزیع وضعیت پایای تعداد متقاضیان حاضر در سیستم در لحظه خروج فرآیند

میانگین تعداد متقاضیان حاضر در سیستم ، با مشتقگیری،

$$g_n'(1)=\left[\frac{\lambda^2E(s^2)}{2(1-\rho)}+\rho\right]+\left[\frac{\lambda^2E(V^2)}{2E(V)}\right]+\left[n-\frac{\sum_{i=1}^{n}\theta_i}{\theta_n}\right] \quad (15)$$

فرض کنید بهداشی سیستم در اوقاتی که سرویس‌دهنده در فراغت است، در هر واحد زمان r باشد (این پاداش ممکن است بسته بشکل انجام کار سرویس‌دهنده در جایی دیگر باشد). فرض کنید که هزینه شایستی را برای ارجاع سرویس دهنده از فراغت به سرویس و بالعکس متحمل شویم . مجموع این دو نوع هزینه بنا $k(\theta)$ داده شده است همچنین فرض کنید $c(\theta)$ هزینه اشغال سیستم به وسیله هر مشتاقی در واحد زمان باشد . هدف مینیمم کردن امید ریاضی متوسط هزینه در واحد زمان و در دراز مدت است. ما تنها خط مشهای خد بتصادفی را در نظر میگیریم . خط مشهایی که بنا در نظر گرفتن یکدسته بتصادف η همچنان که در ابتدا شرح آن گذشت ، عمل میکنیم .

بعد آور میشویم گمری از زمان در دراز مدت که سرویس دهنده در فراغت بسر میبرد $(\rho-1)$ است و با استفاده از نظریه سکتمانتی تحدید ، ملاحظه میشود که متوسط تعداد نفرهای اشتغال (امید ریاضی) در دراز مدت درهر واحد زمان $1/\bar{\theta}_n$ است، پس امید ریاضی متوسط هزینه (وقتی که بتصادف η باشد) در دراز مدت برابر است با :

$$c_\eta = r(1-\rho) + \frac{k}{\bar{\theta}_\eta} + c g'_\eta(1) \qquad (16)$$

از معادلات (15) و (16) بسه سادگی شایستی میشود که مینیمم‌کردن c_η معادل با مینیمم کردن

$$\bar{c}_n = n+\left(\frac{k}{c} - \sum_{i=1}^{n} \bar{\theta}_i \right) / \bar{\theta}_n \quad (17)$$

است. حال چون

$$\bar{c}_n - \bar{c}_{n+1} = \left(\frac{k}{c} - \sum_{i=1}^{n} \bar{\theta}_i \right) \left(\frac{1}{\bar{\theta}_n} - \frac{1}{\bar{\theta}_{n+1}} \right) \quad , \; n \geq 1 \quad (18)$$

و چون برای $n \geq 0$ و $\bar{\theta}_{n+1} < \bar{\theta}_n$ (که می‌توان آن را از طریق معادله (6) به استقراء شابیت کرد) است ، بنابراین برای $n \geq 0$ و $0 < (1/\bar{\theta}_n - 1/\bar{\theta}_{n+1})$ است. اما برای یافتن مقدار حد n متناسب n ، در صورتی که حد اخیر هزینه را مینیمم نمایم ، باید $\bar{c}_{n+1} \leq \bar{c}_n$ باشد یعنی $\bar{c}_n - \bar{c}_{n+1} \geq 0$ باشد. لذا به قضیه زیر می‌رسیم.

قضیه ۲ : با توجه به معیار امید ریاضی متوسط هزینه هر هزار ساعت ، مقدار بهین حد نصاب (مشیت) برای سیستم با توزیع چندگانه با خط مشی حد نصابی به صورت زیر است :

$$\hat{n} = \min\left\{ n \geq 1 \mid \frac{k}{c} \leq \sum_{i=1}^{n} \bar{\theta}_i \right\} \quad (19)$$

متذکر میشویم که برای آن که این مدل با مدل مفروض مطابقت داشته باشد ، بهینه‌سازی را تنها نسبت به π_1 انجام میدهیم. برای $n = 0$ به مدل تک فرآیندی میرسیم که به وسیله لوی ووبچیالی تحلیل شده است و آنسان با در نظر گرفتن ساختار هزینه فوق نشان داده‌اند که اگر امید ریاضی متوسط هزینه در واحد مدت برای مدل بسته فرآیند چندگانه ($n=1$) بیشتر از صف متناظر و معمولی M/G/1 باشد، آنگاه همان رابطه هم بین مدل تک فرآیندو صف متناظر و معمولی M/G/1 برقرار خواهد بود . از طرف دیگر آنها همچنین نشان داده‌اند که اگر این هزینه ها برای مدل با فرآیند چند گانه کمتر از مقدار متناظرش برای صف معمولی M/G/1 باشد ، آنگاه از مقدار متناظر برای مدل تک فرآیند هم کمتر خواهد بود . بنابراین اگر ردخط مشی ها را آنقدر گسترش دهیم که شامل مدلهای تکفرآیند ($n=0$) و مدلهای بدون فرآیند و علاوه بر این خط مشیهای حد متعادلی هم بشود ، آنگاه میتوانیم حالت $n=0$ را نادیده بگیریم . چون خط مشی تک فرآیند هیچگاه بهین نیست.

اما برای ملاحظه این که چه وقت مدل بدون فرآیند بهین است، مرتبا " نیاز به مقایسه c (معادله (19) را بهینه کنید) با $Cg(1)$ است که هزینه برای صف معمولی است (که $g(1)$ امید ریاضی تعداد متقاضیان در سیستم در صف متناظر و معمولی M/G/1 در وضعیت پایا است) .

[1] Levy,Y.,and U.Yechiali .1975. Utilizatein of idle time in an M/G/1 Queueing System. Mgmt.Sci 22,202-211

[2] Cooper,R.B.1981.Introduction to Queueing Theory, Ed.2 North Holland , New York.

[3] Fuhrmann,S.W.,and R.B. Cooper.1985. Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations. Ops.Res.33,1117-1129.

آیندهٔ آمار در ایران

امیر شاپور شاهین
مرکز آمار ایران

۱ـ مقدمه

هرچند فعالیت های آماری" از دیر زمانی درکشور ما وجود داشته وسابقه ای طولانی دارد، مع الوصف اساسآمار علمی و روشهای جمع آوری ارقام خام آماری بصورت علمی و پیشرفتهٔ امروزی ، بین سالهای ۱۳۲۷ تا ۱۳۳۵ پی ریزی شده است .

در طول چهل سال اخیر روشهای جمع آوری ، پردازش متجزیه و تحلیل وانتشارآمار پیشرفت های چشمگیری داشته ، بطوریکه درآستانه سال ۱۳۷۱ نه تنها در دانشگاهها و دبیرستانها تدریس آمار رواج دارد، بلکه در بیش از دوهزار و یکصد موسسه ، تحقیق و بررسی آماری جریان دارد و آمار کاربردی بموازات آمار نظری و ریاضی بصورت ابزار لایفک تحقیق وتوسعه ، درجامعه ما درحال تجلی است .

در ده سال گذشته حجم فعالیتهای آماری در کشور ما بسیار گسترده بوده است . نتایج آمارگیری از " امکانات آماری کشور : ۱۳۶۶ " نشان داد که در تهران و شهرستانها ۲۱۴۰ موسسه به فعالیتهای مختلف آماری مشغول بوده اند . در ۱۶۳۰ موسسه عملا " جمع آوری ارقام خام آماری وتجزیه و تحلیل داده ها انجام میگردد حتی درمجموعهٔ نمودار سازمانی ۸۲۲ موسسه ازاین موسسات ، یک "واحدآمار" پیش بینی شده است .

(۲۲۹واحد از همین موسسات ، دولتی بوده انـد) ، از ۱۶۳۰ موسسه فوق دوسوم درشهرستانهـا و یک سوم درتهران فعالیت داشته‌اند . تعداد کارکنان این ۱۶۳۰ موسسه حـدود ۱۰ هـزار نفر بوده که $\frac{1}{2}$ آنها تحصیلات بالاتر ازلیسانس داشته وسن نیمی ازآنها از ۳۵ سال کمتر بوده است .

درمیان ۱۶۳۰ موسسه‌فوق بعنوان مثال نام "مرکزآمار ایران " ، " اداره کل آمارونیروی‌انسانی وزارت کارواموراجتماعی " ، " بانک‌مرکزی جمهوری اسلامی ایـران " ، " گمرک‌ایران " ، "سازمان هوا شناسی کشور "، " شرکت‌ایران خودرو" و ... وجود دارد . فعالیت آماری در این ۱۶۳۰موسسه سالانه بصورت مستمر و مداوم انجام‌میگیرد . زمینه فعالیتها و تحقیقات آماری بترتیب مربوط است به صنعت ، جمعیت و نیروی‌انسانی ،بهداشت و درمان ،آموزش و پرورش ، کشـــاورزی بازرگانی ، امور پولی ومالی‌این موسسات هم " آمار عمومی وکشوری " وهم "آمار اختصاصی" راجمع آوری وتجزیـــه وتحلیل می‌کنند .

اگر بخواهیم فعالیتهای عظیم این‌مؤسسات را (که قسمت عمده‌ای ازآمار ثبتی کشور رانیز فراهم می کنند) ازنزدیک مورد مداقه قرار دهیـم ، بطور نمونه میتوان از مرکزآمارایـران نام برد کـه تنها در پنج سال‌گذشته کارهای عظیمی‌ازجمله سرشماری عمومی نفوس‌ومسکن سال۱۳۶۵ وسرشماری کشاورزی سال ۱۳۶۷ وسرشماری های عشایـروصادر و ... رانجام رسانیده وحدود ۱۲۰ طـــرح آمارگیری نمونه‌ای را بثمر رسانیـده وحدود ۲۰۰ نوع مختلف نشریه آماری را (صرفنظرازعناوین فرعـــی) بمرحله انتشاررسانیده و حدود ۸۰ مجموعـه آماری را مدوّن و در دسترس قرارداده‌است تعداد نشریات نتایج حاصل‌ازسرشماری عمومی نفوس‌ومسکن بالغ بر ۴۴۵ ، جلد است و انتشـار چنین حجمی ازاطلاعات‌بانضمام تدارک مقدمات و مساعی‌مربوط به مصاحبـه وجمـــع‌آوری و اصلاح پرسشنامه‌ها ،سنگنه زنی وجدول بندی وتهیـه‌نمودارها وانشاء متن گزارشات آماری خود بیانگر وسعت فعالیتهای آماری اســت .

ازمیان طرحهای یادشده فقط بعنوان مثال میتوان به طرحهای سالانه آمارگیری نمونـــه‌ای هزینه و درآمد خانوارهای شهری و روستایـی کشور ، طرح آمارگیری نمونه ای از کارگاههـــای :

محل صرف غذا و نوشیدنی ـ محل خرید و فروش کالا ، طرح نمونه‌ای زراعت و باغداری ـ طرح آمارگیری از قیمت خرده فروشی کالاها و خدمات در روستاها و اعتال آن اشاره کرد .

صرفنظر از فعالیتهای ۲۱۴۰ واحد آماری فوق بایستی از فعالیت مظیم آموزش آمار در دانشگاهها و دبیرستانهای کشور و نیز فعالیتهای تحقیقات آمار نظری و ریاضی در پنجساله گذشته نام برد .

بطور خلاصه باید گفت که بین سالهای ۱۳۶۰ تا ۱۳۴۰ حجم فعالیتهای آماری کشور (با توجه به مقیاس نشر کتابهای درسی و نشریات آماری ـ اجرای آمارگیر بهداشتی ـ شتاب گرفتن ثبت آمار ثبتی ، استفاده از کامپیوترها و ...) آنچنان چشمگیر بوده است که در مجموع خود برابر همه فعالیتهای ۳۰ سال قبل از آن ، یعنی فعالیتهای بین سالهای ۱۳۳۰ تا ۱۳۶۰ است . بعبارت دیگر میتوان ادعا کرد حجم فعالیتهای ۱۰ سال گذشته معادل حجم فعالیتهای تدریجی یک دوره ۳۰ ساله قبل از آن بوده است . (نسبت تراکم فعالیت و ژرفاگرائی در ده سال گذشته $۳=\frac{۳۰}{۱۰}$) اگر این مدعا را قبول کنیم و اگر با همین قیاس محاسبه را پیش ببریم تا ۴۰ سال دیگر یعنی در آغاز سال ۱۴۱۰ ، در کشور ما حجم فعالیتهای آماری چیزی حدود ۸ برابر حجم فعالیتهای ۱۰ سال اخیر توسعه خواهد یافت . (یعنی $\frac{۳۰}{۱۰}$ بار ، دو برابر خواهد شد)

$$\text{برابر} \quad ۸=۲^{\left(\frac{۳۰}{۱۰}\right)}$$

البته تصور ۸ برابر شدن فعالیتهای ده سال گذشته قدری مشکل است ، ولی ما باید برای کارهای عظیم آینده خود را آماده کنیم .

سنگ راه ما گرچه سخت و راه طلب　　کوه و صحرا پیش سیل بیقرار ما یکی است

احتمالا " با امکانات مالی کافی هم برای روبرو شدن با این حجم کار را نخواهیم داشت (در واقع هم اکنون اعتبارات مراکز آماری بقیمتهای ثابت سال ۱۳۴۱ بسیار کمتر شده است .) اما باید بهر حال آمادگی کافی داشته باشیم و از هم اکنون نیروی انسانی آزموده برای همکاری فراهم کنیم .

اکنون که به عرضه فعالیتهای آماری اشاره کردیم باید نگاهی به تقاضا نیز داشته باشیم. همه درک میکنیم که تقاضای بالقوه برای فعالیتهای آماری بیش از تقاضای بالفعل است ، با آنکه منابع آماری بصورت وسیعی دردسترس است ، و باآنکه هرسال منظما " نشریاتی مانند " سالنامه آماری کشور " ، " گزارش اقتصاد کشور " و دهها نشریه آماری دیگری حاوی دامنه و سیعی از اطلاعات آماری منتشر می‌شود معهذا الوصف میزان استفاده از آمار ، و سطح کلی تجزیه وتحلیل آمار آنگونه که شایسته است پیشرفتی نداشته است .

درکشورما ، درهرزمینه باندازه‌کافی داده‌های آماری در دسترس است و حتی برای بهنگام کردن داده ها ، روشهای آماری مطمئن و متکاملی وجود دارد ولی بعضی کارشناسان هنوز بر اساس " حدس وگمان " تصمیم می گیرند .

هرآنکس که عیبش نگویند پیش " هنر " داند ازجهل آن " عیب " خویش

اگر پنج درجه دقت در داده‌های آماری (یعنی داده‌های حاصل از : سرشماریها وآمارهای ثبتی ، نمونه گیریها ، برآوردها ، حدس وگمان ، انتخاب سرسری و بی پایه) را درنظر بگیریم ، متاسفانه هنوز مشاهده می‌شود که بعضی از مقامات مجری یا تصمیم گیرنده ، سرنوشت جامعه را جدی نمی‌گیرند و بجای استفاده از ارقام صحیح موجود از ارقام نادرست استفاده می‌کنند .

تعاریف ومفاهیم استاندارد بین المللی یا طبقه‌بندیهای جهانی* را (مانند ISIC ، CPC ، HP و ...) بکار نمیبرند واین امر گاهی نارسائیهای اجتماعی ، اقتصادی بدنبال دارد .

عالم نادان پریشان روزگار به ز دانشمند نا پرهیز گار

حتی از داده‌های کم میتوان به نحو احسن استفاده کرد . ونیز مثلا " اگر سطح زیرکشت دردست نباشد میتوان از آمار مقدار تولید مربوط به آن سطح بعنوان " نماینده تخمین " استفاده کرد . تقاضای

---

* طبقه بندی کالای گمرکی ، طبقه‌بندی مرکزی محصولات ، طبقه‌بندی فعالیتها ، طبقه‌بندی مشاغل و ...

بالقوه برای کار رابایستی بااستفاده ازآمار مندرج در تشریحات آماری ، به تقاضای بالفعل تبدیل کرد .

نه تنها به داده‌های آماری حاصل از سرشماریها ونمونه‌گیریها وآمارهای ثبتی باید بهای لازم داده شود

بلکه اصولا " باید برای " تفکر آماری " ارزش شایسته ای قائل شد .

تفکرآماری چیست ؟ فرایند فکری است که انسان تشخیص میدهد در پدیده‌های محیط او ودرکارهای

روزمره او، غالبا " فراز ونشیب ، کم و زیاد ، زیر و زبر ، افت وخیز وخلاصه تحولات وجود دارد وآگاهی

ازحدمتوسطی یاحدپیش‌بینه‌ای ازآنها و شناخت دامنه تغییرات آنها ضروری است .

کنترل کردن این تغییرات وتجزیه وتحلیل وکنترل آنها بما فرصت میدهد که امور مربوط را تدریجا " توسعه

دهیم ودرجهت نهائی گام برداریم .

تفکرآماری مراحلی دارد . این مراحل عبارتند از :

- اندیشیدن درباره مسئله : چه امری است که رخ میدهد ونباید رخ دهد و یاباید رخ دهد و رخ

نمی دهد ؟

- شناخت تغییرات نسبت به حالت مطلوب (یامتوسط)

- تجزیه و تحلیل دامنه تغییرات (احتمالا " بوسیله استفاده ازابزار آماری)

- بسط دانش فنی درباره مسئله

- اقدام اجرائی درجهت رفع مشکل وشروع آزمونها

- توسعه و بهبود وضع

و البته مراحل فوق ممکن است چند دوره تکرار شود .

امروز تفکر آماری برای جامعه آماری برای جامعه محصول ما ضروری است و چه در کارخانه‌ها وچه

در خدمات دولتی به تفکر آماری نیازمندیم . تنها نباید گفت که " واریانس میانگین زیاد است "

گاهی در کنار آن باید گفت : " چه کنیم که واریانس را کمتر کنیم و کیفیت را بالا ببریم " دولت

معنی و علت پراکندگی ، را زیرشرفت واعتلاء کیفیت کار است . تفکرآماری ، همراه با استفاده از

داده‌های آماری منتشر شده ، تقاضا برای آمار و فعالیتهای آماری وارز قوه به فعل درمیاورد و بدین

ترتیب عرضه وتقاضا تدریجا" هماهنگ می‌شود.

۲ـ آینده آمار و آمارشناسان

سئوالی که غالبا" مطرح می‌شود اینست که بالاخره سرنوشت آینده آمار و آمارشناسان ایران چه خواهد بود؟

این سؤال اصولا" غلط است.

بجای سوال فوق باید موضوع را بشکل زیر مطرح کنیم:

من و شما برای آینده آمار و آمارشناسان چه برنامه‌ای داریم؟ از حاصل اعمال و رفتارمان چه خواهیم ساخت؟

" زندگی " در صدف خویش گهر ساختن است " عشق " ازاین گنبد سربسته بیرون تاختن است

اگر توجه خود را باین سؤال صحیح اخیر معطوف کنیم طبعا" بادرنظرگرفتن سه عامل: اوضاع واحوال جامعه خودمان - برون‌گیری روندهای گذشته - و بسط تفکر منطقی میتوانیم به سوال فوق پاسخ موثری بدهیم. بااحتمال زیاد پیش‌بینی می‌شود که خودمان آینده آمار را به شکل نوید بخشی بسازیم وهمه باید سعی کنیم برای نیل به مقاصد زیر کوششهای خود را همسو کنیم. کارشناسان آمار باخود اهل تحقیق هستند و بانوشه رسانان روشهای علمی تحقیق وتوسعه هستند و در زیر برای ۱۰ ساله آینده، آنچه را که احساس می‌شود بدهمت والای کارشناسان آمار و دست اندرکاران واستفادان آمار انشاء الله اتفاق خواهد افتاد بیان می‌کنیم.

الف ـ در زمینه آمار نظری و ریاضی

- نقش پرقدرت احتمال بعنوان " وزن " و" ضریب " در نظریه احتمال مشخص‌تر خواهد شد و موارد استفاده بیشتری خواهد یافت.

- استدلال آماری و قضاوت آماری وسیله مسئولان وتصمیم گیرندگان فزونی خواهد یافت و قوت خواهد گرفت.

- درمباحث نمونه گیری " روشهای کنترل کمی " و نیز بررسی سریهای زمانی " جای خود را در میان روشها بیشتر باز خواهد کرد .

- مدلهای پارامتری با تکنیک های بیش بینی ظریف تر و پیچیده تر در زمینه های جدید بکار گرفته خواهد شد .

- کاربرد روشهای آمار در پزشکی و بیوفیزیک و امور قضایی میدانهای نظری جدید باز خواهد کرد .

- پافشاری روی معادلات ریاضی و روشهای نظری صرف (وبدون مثالهای کاربردی) محدودتر خواهد شد .

- آزمون فرض ، در بررسیهای آماری قوت و قدرت بیشتری خواهد یافت .

سعدی هوچیزمیشکند قدر مرد را تصدیق بی وقوف وسکوت وقوف دار

- ازسایر ابزار وروشهای موجود علم آمار (نظریه و توزیعات احتمال - توزیعات نمونه ای - روشهای برآورد - تجزیه وتحلیل رگرسیونی - آنالیز واریانس - روشهای نمونه گیری - طرح آزمایشها - نظریه فرآیند های تصادفی - استنباط آماری، تحقیق عملیاتی و ...) در آینده موثرتر وکارآمدتر استفاده خواهد شد .

ب - آمارکاربردی (بیومتری - اقتصاد وکسب وکار - اجتماعی ونفوس و ...)

- " تحقیقات عملیاتی " مبتنی برآمار کاربردی بصورت بسیط مورد استفاده واقع خواهد شد .

- درچارچوب کاربرد ، ارج وحرمت و حفظ اصالت " روشهای مقدماتی آمار " مجدد ارزیابی مجدد قرار خواهد گرفت .

- کارشناسان آمار کاربردی ، روشهای " پیش بینی " و آزمایش صحت این روشها را ترسیم خواهند داد .

- درآمارهای دولتی عملکرد فعالیتهای آماری وانتظارات مقامات مهم تر و عمیقتر خواهد شد .

- ارزش واقعی " واحدهای آماری " در داخل موسسات اقتصادی روزبروز روشن تر خواهد شد .

- نماگرهای اساسی اقتصادی و اجتماعی ، برنامه‌های کوتاه مدت (ماهانه و فصلی) و شاخص‌ها منظما " و بطور مداوم انتشار خواهد یافت . (از جمله شاخص طریقه تامین سرمایه در گردش ، شاخص تولید سرمایه ، شاخص بهره وری)

- حسابهای ملی ، طیف وسیعی از آمار اقتصادی را بصورت یکدست و یکنواخت منسجم خواهد کرد .

- بتدریج که اهمیت آمار در جامعه روشن‌تر می‌شود ، فشار بر کارشناسان آمار کاربردی بیشتر خواهد شد .

- تلفیق و آمیختگی بیشتری بین آمارهای نقوس و آمارهای اقتصادی رخ خواهد داد . (مثلا " اثر تغییر بعد خانوار روی پس‌انداز ملی و تشکیل سرمایه ، بطور مداوم پیگیری خواهد شد .)

ج - پردازش آمار و استفاده از رایانه (کامپیوتر)

- اهمیت نگاه کردن ، بررسی و کنترل داده ها - نقطه یابی - رسم هیستوگرام - پیدا کردن انحرافات - حذف داده‌های مهمل و بی‌معنی - و ... بوسیله کامپیوتر بصورت بارزی جلوه‌گر خواهد شد .

- استفاده بهینه از کامپیوترهای موجود (جدید و قدیمی) گسترش خواهد یافت .

- قدرت پردازش و استفاده از نرم افزارهای مخصوص پردازش داده‌ها ، بیشتر خواهد شد .

- استفاده از زبان نرم‌افزاری BASIC روز افزون خواهد بود . بسته‌های نرم افزار استاندارد متداول خواهد شد .

- اخذ نتیجه تصمیم گیری با کامپیوتر مبتنی بر روشهای " شبیه سازی " و " مونت کارلو " متداول خواهد شد .

- روشهای " انیمیشن ANIMATION " و " هولوگرافی " با کامپیوتر توسعه پیدا خواهد کرد .

- برنامه‌های کوتاه مدت آماری و درس آمار بودن معلم وسیله کامپیوتر به بازار خواهد آمد .

- کامپیوترهای کوچک ویژه ، بمنظور استفاده در اجرای آمارگیریها (مانند قرائت کنتورها - ضبط و پردازش تردد شماری جاده‌ها و ...) بمیزان عمده خواهد شد .

- بانک‌های اطلاعاتی در شهرها و در سازمانها کاملتر و مرتبط‌تر خواهند شد .

۵- آموزش علم آمار

- دبیران شخصاً " به آزمایش راههای آنالیز داده‌ها، مسئله‌های پارامتریک و روشهای ساده پیش‌بینی و ... خواهند پرداخت .

- دبیران مثالهای واقعی ساده از جامعه را برای تعلیم روشهای آماری بکار خواهند برد و بعضی مسائل روزمره را با کمک متعلّمان حل خواهند کرد .

- در دبیرستانها بر ازنده کردن مدلهای ها به داده‌ها، استدلال و استنباط آماری، روشهای برآورد و نمونه گیری بصورت ساده عرضه خواهند شد .

- در دبیرستانها مفهوم "قریب بیس" (۱) به دانش‌آموزان تعلیم داده خواهد شد .

- در دانشگاهها روشهای جدید که برای جامعه فعلی ضروری است (از جمله راه تشخیص روندهای کوتاه مدت از بلند مدت - راه تلفیق دوگرایش غیرهمسو و ...) کشف خواهد شد .

ای ایستاده در پس این پرده شمار نزدیکتر بیا که هویدا بینیمت .

- پایا بودن و کافی بودن بسیاری از روشهای جدید و الگوهای ریاضی ، در دانشگاهها مورد ارزیابی قرار خواهند گرفت .

۶- تحقق خواستهای آمارگیریها

- طبقه‌بندیهای استاندارد در اغلب دوایر دولتی بالاخره یکدست خواهند شد و به پوشش کلی آمار

۱- در احتمالات "قریب بیس BAYS " عبارت است از : لگاریتم نسبت احتمال شرطی پسین به احتمال شرطی پیشین :

$$B = LOG \frac{P(E)}{P(E)}$$

هنگام اجرای هر گونه عملیات روی چند عدد (مثلاً " جمع کردن چند عدد با استفاده در الگوهای ریاضی) هر یک از آن اعداد را با " قریب بیس " وزن میدهند تا نتیجه صحیح تری بدست آید .

ثبتی جامعه‌ها توجه‌بیشتری معطوف خواهد گشت .

- تعاریف ومفاهیم ، در دوایر دولتی یکسان ترخواهدشد وآمار ثبتی یکنواخت تر بدست خواهد آمد .

- از نقشه‌ها و اطلس ها موثرتر استفاده بعمل خواهد آمد .

- تهیه و تنظیم طرحهای آمارگیری باتوجه به اهداف چند منظوره ، تخصصی تر خواهد شد .

- شخصیت واثر کارماموران آمارگیر واهمیت آموزش آنان بارزتر جلوه خواهد کرد .

- به سازمان اجرائی وسازمان نظارت باهمیت والاتر داده خواهد شد .

- انگیزه های همکاری پاسخ دهنده بمنظور هماهنگی اوبا آمارگیران مورد بررسی بیشتر قرار خواهد گرفت .

- جمع آوری داده ها باحجم کمتر - ولی دقیق تر-بهنگام تر-مرتبط-و با استفاده آسان تر انجام خواهد گرفت . تجزیه و تحلیل و بازبینی ومیزی داده‌ها توسعه خواهد یافت .

- صورتهای خلاصه ، علاوه بر داده‌های تفصیلی در دسترس قرار خواهد گرفت .

- روشهای اجرائی وسیله کارشناسان ، منظما " دقیق ترخواهد شد .

و - کارشناسان آمار و دست اندرکاران آمار

- ارزیابان گمرک به صحت نمرنه‌ها توجه بیشتری خواهند کرد (ومثلا " نیترات سدیم را اگر در باربا شد در طبقه " دارو " و اگر کودشیمیایی باشد درطبقه خود ، طبقه بندی خواهند کرد . باید گفت که هم اکنون برای اقلامی مانند هواپیما ، کشتی ، موتور هواپیما ماهواره و ... " برگ سبز گمرکی " صادر نمی شود و این اقلام جزء ارزش وارزات محسوب نمی شوند . صادرات نفت نیز از ارزش مندرج در آمار گمرکی اخیرا " حذف شده است .)

- دست اندرکاران آمار به‌مسئله " قیمت " توجه موثرتری خواهند کرد . وبعضی از سازمانها اختلاط شگفت آوری ازنوع قیمت‌کالاها (قیمت تولید کننده ، قیمت سرخرمن ، قیمت پشت‌گمرک

قیمت عمده فروشی ـ قیمت خرده فروشی ـ قیمت خریدار با مالیات ـ قیمتهای بدون مالیات) ارقام به چشم میخورد که باعث گمراهی مصرفکنندگان آمار می‌شود . برازنده کردن هرگونه الگوی ریاضی به ارقامی چنین متشتت هرگز نتیجه ای رضایتبخش بدست نخواهد داد . ارقام آماری باید مطمئن و توام با تعریف عرضه شود .

در باب این قوله سخن بیشمارهست اما نگنجد این سخنان در مقاله ها

ـ در سال آینده استفاده‌کنندگان و تولیدکنندگان آمار بیش از پیش دیگر خواهند شد .

ـ ریاضی دانان متوجه شدکه تجزیه و تحلیل داده‌ها نه تنها علوم ریاضی را جالب میکند بلکه بعنوان یک مهارت ضروری برای تحصیلکرده‌های ریاضی بشمار میرود . ریاضی دیگر منحصر به جبر و آنالیز و هندسه و مثلثات نخواهد ماند . " نظام دادن به داده‌های حاصل از اندازه‌گیریها و مشاهدات " نیز جزء دروس ریاضی خواهد آمد و ریاضیات را متوجه تر و جالب تر خواهد ساخت .

ـ کارشناسان آمار در محیط فعالیت خود (تولید صنعتی ـ تولید کشاورزی ـ بهداشت و درمان و ...) تخصص موضوعی پیدا خواهند کرد و روشهای آماری را با احاطه کامل در همان رشته‌ها بکار خواهند گرفت و مرزهای جدیدی از علم نمایان خواهد شد و نظام حقیقت جویی در متن حرفه پرورش خواهد یافت . کارشناسان آمار خود عرضه کننده راه حلها و آفریننده روشهای نو خواهند بود .

ـ استادان و کارشناسان آمار یافته‌های علمی خود را از طریق نشریات ویژه مبادله خواهند کرد .

ـ فرزانگان علم آمار با همکار طریق مکاتبه آشنایی بیشتری پیدا خواهند کرد . اسناد گرانبهای بررسیهای علمی در بایگانیها مدفون نخواهد ماند و بصورت مدون مورد استفاده آیندگان قرار خواهد گرفت .

بسیار کاروان که از این ره گذشت و رفت ز آنها نشان بجای بجز گرد ره نماند

* * * * * * * * * * * * *

زنام نیکان جاودانه‌ای بگذار تا نامه زندگی جاودان میسر نیست

- " تواضع بیش‌ازحد" دربین استادان وکارشناسان آمار تبدیل به " شکاکی شجاعانه " خواهد شد و در شهر های مختلف کارشناسان ، سمینارها و دوره‌های دیدو بازدید وتبادل نظر تشکیل خواهند داد.

قطره‌های پراکنده بصورت دریای کُهربار جلوه گر خواهند شدوکارشناسان خود را به مجامع آماری و اساتید آمار وابسته خواهند کرد.

قطره دریااست اگربادریا است ورنه او قطره و دریا دریا است

منابع وماخذ

۱ـ گزارش فعالیتهای مرکزآمار ایران : دوره ۱۳۶۵ تا ۱۳۶۹

۲ـ نتایج آمارگیری ازامکانات آماری کشور-مرکز آمار ایران : ۱۳۶۷

۳ـ سالنامه آماری کشور-مرکز آمارایران : ۱۳۶۹

۴ـ گزارش اقتصادی کشور- سازمان برنامه و بودجه : ۱۳۶۹

۵ـ نتایج تفصیلی سرشماری عمومی نفوس ومسکن مهرماه ۱۳۶۵ - مرکزآمار ایران : ۱۳۶۷

۶ـ آمار کارگاههای بزرگ صنعتی -مرکزآمار ایران : ۱۳۶۷

۷ـ آمار بازرگانی خارجی ایران -گمرک ایران : ۱۳۶۸

بررسی اثر عوامل اقتصادی-اجتماعی بر سطح باروری

سیدحسن شمس

مرکز آمار ایران

مطالعهٔ اثرات عوامل اقتصادی-اجتماعی بر سطح باروری که به منظور انجام تحقیقات علمی و تهیهٔ ابزاری مفید برای اتخاذ سیاست جمعیتی مناسب در راستای برنامه‌ریزیهای توسعهٔ اقتصادی و اجتماعی کشور انجام می‌شود، از دیرباز مورد توجه مسئولان سیاستگذاری جوامع بوده است و بررسیهای زیادی نیز هم در کشورهای توسعه‌یافته و هم در کشورهای درحال‌توسعه، توسط کارشناسان علم جمعیت‌شناسی، اقتصاد و ... با بکارگیری روشها و مجموعه‌های متفاوتی از عوامل مستقل انجام شده و نتایج متفاوتی به دست آمده است. در این بررسی که براساس نتایج سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ و با بهره‌گیری از روش همبستگی چند متغیره خطی انجام‌شده، هدف شناخت رابطهٔ بین شاخص سطح باروری با مجموعه‌ای از هفت متغیر مستقل اقتصادی و اجتماعی، تعیین درجه اهمیت هر یک از این عوامل در افزایش یا کاهش میزان باروری و نیز فراهم آوردن امکان مقایسه با نتایج مطالعه‌ای است که قبلاً براساس اطلاعات سرشماری عمومی نفوس و مسکن سال ۱۳۵۵ انجام شده است.

خطا و کنترل خطا در استخراج

حمیدرضا صاحب

مرکز آمار ایران

پیشگفتار

اشتباهاتی که در یک طرح آماری بروز می نماید بر دو دسته تقسیم میشود. دسته‌اول اشتباهات نمونه‌ای است که در اثر تعمیم نتایج بدست آمده از نمونه‌گیریها برای جامعه حاصل میگردد و دسته دوم اشتباهاتی که در مراحل مختلف تهیه طرح، اجراء و جمع‌آوری اطلاعات، استخراج و انتشار نتایج آماری بوجود می‌آید. این اشتباهات که هم در آمارگیریهای نمونه‌ای و هم در سرشماریها وجود دارد به "اشتباهات ویا خطاهای غیر نمونه‌ای" معروف هستند.

خطاهای غیر نمونه‌ای بخش عمده‌ای از کل خطا را تشکیل میدهند که اغلب بعلت عدم شناخت کافی کنترل روی اینگونه خطاها کمتر است و همانطور که گفته شد این خطاها از مرحله برنامه‌ریزی جهت تهیه طرح گرفته تا انتشار نتایج نهائی بروز می نماید.

مقدمه

صرفنظر از بودجه و دیگر امکانات که در یک طرح آماری بکار میرود، میزان دقت قابل حصول نیز دارای محدودیتهائی میباشد. اکنون که محدودیتهائی چون بودجه و امکانات وجود دارد بجا است که از طرق مختلف خطاهای گوناگون را به حداقل رسانیده و آنها را کنترل نمائیم.

تکنیکهای کنترل خطاهای نمونه‌گیری بخوبی شناخته شده و مورد استفاده قرارمیگیرد. ولی نظریه واحد و دستورالعملی برای کنترل خطاهای غیر نمونه‌ای موجود نیست. در صورتیکه خطاهای غیرنمونه‌ای که اغلب قسمت عمده‌ای از خطاهای کل را تشکیل می‌دهند، میتوانند درهریک از مراحل طرح از برنامه‌ریزی گرفته تا مرحله انتشار بروز نمایند.

هدف از کنترل خطا بر طرف کردن کل خطا نیست و نمی تواند باشد بلکه کاهش خطای کل تا میزانی است که نتیجه قابل استفاده‌ای حاصل شود. هر نوع خطا باید در رابطه با سایر انواع خطاها و به نسبت هزینه‌هائیکه جهت کاهش آن پرداخت میگردد، کنترل شده و کاهش یابد.

از وقتی که آمارگیری به پایان میرسد تا زمانیکه نتایج بررسی و منتشر گردد یک سلسله عملیات تحویل و تحول کنترل و تفکیک، آماده سازی، بازبینی، کدگذاری، ادیت و انتقال اطلاعات از پرسشنامه‌ها و فرمها به کامپیوتر صورت میگیرد. مجموع این عملیات که معمولا" در طرحهای بزرگ زمان قابل ملاحظه‌ای را به خود اختصاص میدهند در اصلاح "استخراج" نامیده میشود.

عملیات استخراج به جهت وسعت و گستردگی خاص خود بخصوص در سرشماریها نیز مشمول پاره‌ای "خطاهای غیر نمونه‌ای" است که در این مقاله مورد شناسائی و بررسی قرار میگیرد.

کلیه خطاهای ناشی از طراحی پرسشنامه، تحویل و تحول، کنترل و آماده‌سازی، بازبینی، کدگذاری، انتقال اطلاعات و برنامه‌نویسی صورت میگیرد، " خطاهای استخراج " شمرده میشوند و در ۶ گروه عمده در این مقاله مورد بررسی قرار میگیرند.

۱ ـ خطاهای ناشی از طراحی پرسشنامه‌ها و دستورالعملهای مربوطه

موفقیت یک طرح آماری بستگی زیادی به کیفیت پرسشنامه‌های آن دارد . لذا قبل از اجرای طرح می‌بایست کلیه جوانب پرسشنامه مورد بررسی و بحث قرارگیرد و عملاً بطور آزمایشی تکمیل گردد تا اگر چنانچه اشکالاتی وجود دارد قبل از اجرای اصلی برطرف و آماده گردد .

طراحی پرسشنامه رابطه مستقیم با استخراج نتایج دارد بطوریکه طراحی مناسب پرسشنامه در کاهش حجم خطاهای استخراج تاثیر بسیار مثبت داشته و عوامل موثری که موجب کاهش خطاهای استخراج میگردند عبارتند از :

الف ـ تعیین اندازه مناسب ستونها و سطرهائیکه می‌بایست اطلاعات در آنها ثبت گردد . بطوریکه در مواقع بازبینی پرسشنامه‌ها جای کافی جهت ثبت مجدد اطلاعات اصلاح شده توسط بازبین وجود داشته باشد که در زمان انتقال اطلاعات پانچیست کمتر دچار خطا گردد وازطرفی بازبین بتواند اطلاعات اصلاح شده را بطور مشخص در جای مربوطه ثبت نماید .

ب ـ طرح سوالات فرعی که بتوان بکمک آن سوالات اصلی وکلیدی را بمنظور کنترل کار مامورآمارگیر چک کرد . بدیهی است که با طرح اینگونه سوالات میتوان خطای منگنه‌زنی را کاهش داد .

ج ـ پرسشنامه بصورتی طراحی گرددکه در زمان انتقال اطلاعات ، پانچیست در یک مسیر سطری یا ستونی پانچ نماید و یا اینکه اگر چنین امکانی در طراحی پرسشنامه وجود ندارد پرسشنامه را به بخشهای مختلف تقسیم نمایند و هر بخش در یک طرف پرسشنامه طراحی گردد تا در هنگام پانچ دچار خطای کمتری گردد . و بخشهای مختلف بدنبال هم و مرتب منگنه گردد . بدین معنی که هر چه طرح پرسشنامه ساده و واضح باشد استخراج آسانتر خواهد بود .

د ـ واحد شمارش طوری انتخاب و در سرستونها مشخص گردد که در پرسشنامه از ثبت اعداد اعشار جلوگیری شود چون پانچ ارقام اعشاری پانچیست را مرتکب خطای زیادی خواهد نمود .

هـ ـ برنامه‌های ادیت و طراحی نوار منگنه‌زنی با توجه به دستورالعملهای آمارگیری تهیه میگردد پس ضروری است که در تهیه دستورالعملهای اجرائی آمارگیری نهایت سعی وتلاش ودقت کافی

بعمل آید.

۲ - خطای تحویل و تحول، تفکیک و آماده‌سازی

از زمانیکه عملیات میدانی یک طرح آماری خاتمه یافت مراحل مختلف تحویل و تحول مدارک انجام می‌گیرد. بدین معنی که رده‌های مختلف اجرائی از آمارگیر گرفته تا تحویل مدارک به بایگانی مرکز در هر مرحله از تحویل مدارک به رده‌های دیگر امکان گم شدن مدارک وجوددارد. با توجه به اینکه در هر یک از رده‌های مختلف که به رده بعدی تحویل میگردد حجم مدارک زیاد و زیادتر میگردد بطوریکه تحویل‌مدارک در مرکز شامل کلیه مدارک یک طرح بوده که بسیار زیاد و انبوه میباشد، می‌بایست برای تحول هرمرحله از کار نظام و سیستم یکدلختی وجود داشته باشد تا از بروز هر گونه خطای احتمالی جلوگیری بعمل آمده وکار قابل کنترل باشد. مرتب کردن وکنترل مدارک اگر تابع یک سیستم ونظام منسجم نباشد غیر ممکن بوده و امکان بروز خطای ناشی از گم شدن، پاره شدن ومخلوط شدن مدارک وجود دارد. پس می‌بایست از قبل جهت تحویل مدارک در رده‌های مختلف کاری اقدام به تهیه دستورالعمل و فرمهای مخصوص تحویل و تحول مدارک نمود تا پرسشنامه‌ای جابجا نگردد.

پس از تحویل مدارک بر اساس برنامه‌های از پیش تعیین شده به بایگانی مرکز می‌بایست براساس دستورالعملهای کنترل و تفکیک اقدام به تفکیک مدارک نمود. دراین‌مرحله‌ازکار می‌بایست دقت کافی در کنترل مدارک با قابل اصلی نمود تا پوشش جغرافیایی کامل باشد. در این قسمت از کار نیز امکان بروز خطا وجود دارد که جهت کنترل خطا می‌بایست دستورالعملهای کنترل و تفکیک را قبل از شروع بکاراصلی آزمایش نمود و پس از حصول از صحت آن اقدام به عملیات تفکیک مدارک نمود. به جهت اینکه دراین مرحله باید مدارک و پرسشنامه‌های دریافت شده با پوشه‌های نقشه که در اجرا، از آنها استفاده شده کنترل گردد و این کار بوسیله افراد و به روش دستی انجام میگیرد دقت در انجام کار وتوسط افراد با تجربه نیز از حجم خطا می‌کاهد. پس از کنترل، تفکیک وآماده‌سازی می‌بایست اطلاعات منگنه شده وارد کامپیوتر گردد.

۳ - خطای بازپرسی

اگرکلیه پرسشنامه‌ها در مرحله اجرا، بصورت کامل و دقیق تکمیل می گردید نیازی به بازبینی اطلاعات مندرج در آنها نمی‌بود ولی آمارگیری کاری مشکل و پیچیده است که در هر حال باید پاره‌ای خطاهای ناشی از ثبت ارقام، ناقص بودن اطلاعات، عدم احساس مسئولیت آمارگیر و ناهماهنگی میان ارقام را محتمل دانست. اگر چه دستورالعملهای مربوط به مکانیسم بازبینی غالباً" وجود دارد اما در مورد فلسفه بازبینی داده‌ها، متون بسیار ناچیزی در دست است. دلیل این امر احتیاط سازمانهای جمع‌آوری کننده آمار در بیان چگونگی ساختن اطلاعات ناموجود و دستکاری ارقام ناقص و ناهماهنگ است بیم سازمانهای آماری از آن جهت است که ممکن است متهم به کاربرد روشهای غیر علمی و یا سرهم کردن ارقام شوند. در هر حال بازبینی پرسشنامه‌های آماری باید دو هدف اساسی زیر را دنبال نماید.

ـ حتی‌الامکان در اطلاعات جمع‌آوری شده، تغییری داده نشود.

ـ ارقام و آمارهائیکه به وضوح غلط و ناهماهنگ هستند، حذف شوند.

ـ برای اطلاعات ثبت نشده، حتی‌الامکان از روی اطلاعات موجود برآوردهای دقیق انجام شود.

بدین ترتیب عمل بازبینی عمدتاً" شامل کنترل پاسخهای از قلم افتاده، ارقام غلط، اقلام ناهماهنگ ارقام و اطلاعات غیر منطقی میباشد. لذا خطاهای غیر نمونه‌ای در مرحله بازبینی میتوانند در هر یک از زمینه‌های مذکور بروز نمایند.

الف ـ پاسخهای از قلم افتاده

بسیاری از سوالاتی که به دلیل فراموشی آمارگیر بدون پاسخ مانده است در مرحله بازبینی پرسشنامه مشخص میشوند. با وجود این، احتمال اینکه بازبین نیز متوجه از قلم افتادگی نشود وجود دارد. وقتی انتخابهای پاسخ دهنده محدود است یافتن پاسخ سوال از قلم افتاده آسانتر است، مثلاً" وقتی پاسخ به سوال جنسیت پاسخ دهنده، که فقط میتواند مرد یا زن باشد از قلم افتاده باشد، یافتن پاسخ از روی دیگر پاسخها غالباً" میسر است ولی وقتی انتخابهای پاسخ دهنده در پاسخ به سوال محدود نیست یافتن پاسخ کاری دشوار است مانند پاسخ مربوط به سوال تاریخ تولد.

ب ـ اقلام غلط

یافتن اقلام و ارقام غلط در پاره‌ای موارد آسان است ، برای مثال وقتی به آمارگیر گفته شده است در مورد سوال مربوط به جنسیت اگر پاسخ مربوط به مردان میباشد در ستون مربوط حرف M و اگر پاسخ مربوط به زنان میباشد در آن ستون حرف F بنویس هر کد دیگری غیر از M و F در ایـن ستون ثبت نموده باشد، غلط است و باید حذف شده و بجای آن کد صحیح درج گردد .

خال اگر بازبین نیز متوجه اشتباهات اینچنین نشود خطای مرحله بازبینی روی داده است .

در مواردی یافتن ارقام غلط چندان ساده نیست ، برای مثال وقتی آمارگیر سطح زیـر کشت را سوال نموده و در سوال دیگری مقدار محصول تولید شده پاسخ داده شده، تصمیم گرفتن در مـورد اینکـه اصولا" این دو پاسخ درست هستند یا خیر و اگر درست نیستند کدامیك غلط است ، کاری بسیـار دشوار می‌باشد ولی غیر ممکن نخواهد بود .

ج ـ ناهماهنگی میان اقلام

ناهماهنگی میان اقلام پرسشنامه در صورتی روی میدهد که پاسخ یکی از سوالهـا ، پاسـخ (پاسخهـای) دیگری را نقض نماید . برای نمونه اگر آمارگیر وضع تحصیلـی فـردی را در حـال تحصیل گزارش کرده باشدولی در ستون دیگر پرسشنامه وضع سواد همان فرد بی‌سواد ثبت شـده باشد بین این دو پاسخ ناهماهنگی وجود دارد و یا اگر در پـرسشنامه‌ای سطح کـاشت یا نوع بذر و تولید محصول دارای ناهماهنگی باشد این خود تناقض آشکار در پرسشنامه میباشدکـه در چنیـن مواردی جهت کنترل خطا می‌بایست با استفادهاز دستورالعملهای قبلی بازبینـی حالت غالـب در ستونهای مختلف را اصل قرار داده و ستون مورد نظر را اصلاح نمود .

د ـ ارقام غیر منطقی

در این مورد اقلام مورد تـردید لزومـا " غلط نیستند بلکـه با معیارهای متداول ، غیرعقلائی بنظر میرسند . برای مثال وقتی آمارگیر شغل پاسخ دهنده را استاد دانشگاه ولـی حقوق او را رقمـی در حدود دستمزد کارگران ساده گـزارش نموده است اگر چه ممکن است اطلاعات گزارش شـده صحیـح باشد ولـی بهر حال مستلزم بررسی و تحقیق بیشتری است .

در کلیه موارد ذکر شده فوق که امکان بروز خطا در مرحله بازبینی وجود دارد نیز میتوان خطا را کنترل نموده و به حداقل ممکن کاهش داد. بدین ترتیب که در هریک از اشتباهات ناشی از قلم‌افتادگی، اقلام غلط، ناهماهنگی میان اقلام و اعداد و ارقام غیر منطقی میبایست با استفاده از برنامه‌ها ودستورالعملهای ادیت به دوصورت زیر عمل نمود:

ـ در تهیه دستورالعملهای ادیت ماشینی می‌بایست از قبل وقوع چنین مواردی را پیش‌بینی نمود و یک یا چند سوال پرسشنامه را کلید قرار داده وبقیه ستونها را با استفاده از آن ادیت وکنترل کرده و اصلاح نمود.

ـ با استفاده از برنامه‌های ادیت میتوان اطلاعات ناهماهنگ و غیر منطقی را حذف و اطلاعاتی که با استفاده از میانگین ارقام صحیح ساخته شده‌اند، جایگزین نمود. بکارگیری این روش که آنرا "تخمین اطلاعات ثبت نشده" می‌نامیم بدون اشکال نبوده و دارای خطا می‌باشد.

چنانچه در مواردی که پاسخ به سوال یا سوالاتی که اساسا" وجود ندارد یعنی یا آمارگیر پرسش آنها را از قلم انداخته و یا پاسخ دهنده از دادن پاسخ امتناع نموده است بهر حال در هر دو صورت فوق یعنی تخمین اقلام ثبت نشده یا حذف اقلام غلط و تخمین اقلام عقلائی‌تر برای آنها در ارتباط با روش محاسبات میتواند مشمول خطا باشد. تخمین پاره‌ای از اقلام و ارقام با استفاده از سایر اطلاعات موجود و همبستگی میان اقلام صورت میگیرد. درجه دقت تخمین بدست‌آمده به روش محاسبه بستگی دارد و در واقع در ارتباط مستقیم با میزان دقت و زحمتی است که مصروف این امر میشود.

برای پاره‌ای اقلام ثبت نشده بهر حال نمی‌توان رقمی برآورد نمود در این صورت شیوه‌های عملی متفاوتی بخدمت گرفته میشود. نخست اینکه ممکن است اقلام ثبت نشده را در جداول نتایج زیر عنوان "شناخته نشده" یا "نامعلوم" گروه‌بندی نمود. بدیهی است که هر قدر میزان ارقام ستون "نامعلوم" به نسبت زیادتر باشد نتیجه‌گیری و تحلیل ارقام بدست‌آمده، مشکلتر خواهد بود.

روش دیگری که ممکن است مورد استفاده قرار داد، توزیع مقادیر "نامعلوم" به نسبت مقادیر معلوم است.

روش دیگری که برای تخمین اطلاعات ثبت نشده ممکن است رعایت شود این است که رقم ثبت نشده در یک پرسشنامه از پرسشنامه دیگری که سایر خصوصیات مشابهی با پرسشنامه ناقص مورد بررسی دارد استخراج شود. در هر دو حالت فوق احتمال بروز خطا بخوبی قابل قبول و قابل پیش‌بینی است.

۲ ـ خطای کدگذاری

کدگذاری در واقع عمل تبدیل کلمات به اعداد است. بطوریکه اطلاعات پرسشنامه قابل انتقال بر روی کارت، نوار و دیسک باشد. بعضی اقلام پرسشنامه توسط کدهای ازپیش‌تعیین شده (SELF CODE) بر روی پرسشنامه‌ها ثبت و براساس دستورالعملهای مربوطه مشخص وقابل استخراج میباشند.

تجربه نشان داده است اگر چنانچه پرسشنامه بگونه‌ای طراحی شود که حتی‌الامکان بیشتر از کدهای از پیش تعیین شده استفاده گردد میزان بروز خطا در کدگذاری کمتر خواهد شد. مانند کدگذاری در ستون وضع فعالیت که بر اساس دستورالعمل مربوطه کد شاغل ۱، بیکار ۲، محصل ۳ و ... میباشد که کدگذاری آن بسیار ساده و آسان بوده و در هنگام اجرای عملیات میدانی و توسط مامور آمارگیر انجام میگیرد. البته در این شیوه کدگذاری نیز مامور آمارگیر مرتکب خطا خواهد شد که بحث مربوط به آن در بخش خطای کدهای از پیش تعیین شده توضیح داده میشود.

الف ـ کدگذاری با استفاده از کدهای طبقه‌بندی شده

در بعضی موارد این امکان وجود ندارد که از کدهای از پیش تعیین شده استفاده گردد و کدگذاری مستلزم بررسی دقیق پرسشنامه و سوالهای مختلف است تا بتوان کد صحیح را با استفاده از دستورالعملهای مربوطه کشف وثبت نمود. مثلا " مشاغلی مانند کارگر، مهندس، نگهبان و یا فعالیتهائی مثل فروشگاه، مغازه، اداره، شرکت و امثال آن را میتوان به انحاء مختلف تعبیر نمود. در این صورت می‌بایست طبقه‌بندی دقیقی از مشاغل، فعالیتها، رشته و پایه تحصیلی و غیره جهت کدگذاری دقیق وجود داشته باشد تا کدگذار با مراجعه به رشته فعالیت دریابد که برای

مثال کارگر در رشته "صنعت قند سازی" کار میکرده است و یا مهندس در رشته "ساختمان" فعالیت داشته است . آنگاه کد مناسب را به آنها اختصاص دهد . در بعضی موارد به اطلاعاتی که در پرسشنامه وجود دارد ولی برای آنها سوالی وجود نداشته است کد داده میشود . این اطلاعات که بطور تشریحی بدست آمده و پاسخ دهنده بطور اختیاری جواب داده است نمیتوانند در محاسبات و نتیجه‌گیریها وارد شوند .

عملیات کدگذاری به علت ماهیت یکنواخت بودن آن گاهی موجب کم توجهی و کم دقتی کدگذار میشود . نکته‌ای که باید توجه داشت این است که بنا به تجربیات بدست آمده کنترل پرسشنامه‌های کدگذاری شده نتیجه چندان موثری ندارد . زیرا کنترل کننده یا بازبین که غالبا " اشتباهات کدگذار را تکرار میکند . بهترین روش کنترل کدگذاری، کدگذاری دوباره و مجزای پرسشنامه‌ها و سپس مطابقت کدهای ثبت شده توسط کدگذار و بازبین میباشد که در این روش بعلت کدگذاری مجدد عملیات کدگذاری هزینه‌های زیادی را در بر میگیرد که با توجه به امکانات مالی طرح می‌بایست انجام گیرد .

نکته مهمی که در عملیات کدگذاری می‌بایست بدان توجه کافی نمود بازبینی کد در روزهای نخستین کدگذاری میباشد که باید فاصله زمانی بین کدگذاری و بازبینی بسیار کوتاه بوده تا اشکالات و خطاهای کدگذار بلافاصله به او گزارش شود تا اگر چنانچه آموزش کافی نبوده که در نتیجه کدگذار مرتکب خطای سیستماتیک در کدگذاری میگردد، اصلاح گردد تا خطا تکرار نشود و اگر چنانچه در ذهن کدگذار، کد به اشتباه نقش بسته است این خطا پیوسته تکرار نگردد . بدیهی است اگر چنانچه کیفیت کار کدگذاری پائین‌تر از سطح و حد رضایت‌بخش باشد باید پس از تعیین علت در صورت لزوم آموزش جدیدی به کدگذار داده شود و اگر او شخص مناسبی برای اینکار نیست از ادامه کار وی جلوگیری بعمل آید .

ب - کدگذاری توسط کدهای از پیش تعیین شده (SELF CODE)

اگر چنانچه پرسشنامه‌های آماری بگونه‌ای طراحی شوند که حتی‌الامکان پاسخ سوالات هنگام

آمارگیری بصورتکد ثبت شوند تا کدگذاری مجدد ضرورتی نداشته باشد عملیات استخراج نیز تحت تاثیر قرارگرفته و در استخراج نتایج تسریع و تسهیل حاصل میگردد. استفاده از کدهای از پیش تعیین شده در کم کردن خطاهای ناشی از کدگذاری موثر میباشد از طرفی انتخاب کدهای از پیش تعیین شده نیز از عوامل موثر در کاهش حجم خطا خواهد بود زیرا اگر چنانچه حروف و یا اعدادی که جهت کدهای از پیش تعیین شده انتخاب میشوند مطلوب نباشند هنگام ثبت کد توسط مامور آمارگیر و هم در هنگام پانچ اطلاعات توسط پانچیست مرتکب خطا خواهیم شد. بطور مثال اگر در دو ستون مجاور هم در پرسشنامه از یک کد استفاده شود بروز اشتباه و جابجا ثبت نمودن ارقام توسط آمارگیر و پنهان ماندن از دیده بازبین وجود دارد. مانند استفاده از کدهای ۱ و ۲ و ۳ و ۴ و ... برای دو ستون مجاور هم و یا عنوان‌های کاملا" متفاوت که امکان جابجائی ثبت کدها وجود ندارد بخصوص در زمان انتقال اطلاعات ممکن است پانچیست مرتکب خطا شده و اطلاعات در ستون مربوط به خود پانچ نشده و اصطلاحا" پانچیست مرتکب خطای شیفت میگردد. برای جلوگیری از اینگونه خطاها می‌بایست در ستونهای مجاور و نزدیک به هم حتی‌الامکان از کدهای یکسان استفاده نشود و سعی گردد در یک پرسشنامه از کدهای مختلف استفاده شده تا میزان بروز خطا کمتر گردد و یا اگر چنین امکانی وجود ندارد از حروف لاتین جهت بعضی از ستونها استفاده گردد مانند طرح آمارگیری جاری جمعیت که با توجه به تجربیات طرحهای گذشته از حروف لاتین بعنوان کدهای از پیش تعیین شده استفاده گردید.

در بعضی از ستونهای پرسشنامه نوع سوال خود بصورت عددی پاسخ داده میشود مانند نام ماه، سال تولد که در چنین موارد محدودیت جهت انتخاب کد مناسب وجود دارد که در این صورت میبایست کد ستونهای قبل یا بعد ستون نام ماه و یا سال تولد را تغییر داد.

۵ـ خطای انتقال اطلاعات (پانچ)

مرحله انتقال اطلاعات یکی از حساسترین مراحل کار استخراج میباشد. زیرا هر زمان که اطلاعات و ارقام از شکلی به شکل دیگر در می‌آیند احتمال بروز خطا بشدت افزایش می‌یابد.

بروز خطا در مرحله انتقال اطلاعات به عواملی نظیر سرعت در کار، خستگی ناشی از یکنواختی، ناخوانا بودن اعداد و ارقام و انتخاب افراد پانچیست بستگی دارد.

الف ـ تعرفه جهت سرعت در کار

در انتقال اطلاعات از روی پرسشنامه به ماشین منگنه‌زنی (DATA ENTRY) که بوسیله افرادی بنام پانچیست انجام میگیرد لازم است سرعت انتقال اطلاعات بجهت حجم فراوان آنها زیاد باشد تا اطلاعات در اسرع وقت پانچ، پردازش و منتشر شده و جهت برنامه‌ریزیهای مختلف بهنگام و قابل استفاده باشد. به این جهت لازم است جهت سرعت بخشیدن به انتقال اطلاعات انگیزه‌های مالی برای افرادیکه در این مرحله از کار انجام وظیفه مینمایند، ایجاد گردد. دراینصورت با قراردادن تعرفه‌های کاری بر حسب رکورد در ساعت و پرداخت دستمزد بیشتر برای پانچ رکوردهای بیشتر و سازاد بر تعرفه اقدام به سرعت بخشیدن بکار منگنه‌زنی می‌نمایند. در نتیجه بدلیل سرعت زیاد و بالارفتن کمیت ناچار کیفیت پائین آمده و خطاهای بیشتری حاصل میگردد که با بکارگیری شیوه‌های مختلف کنترل خطا اقدام به کاهش میزان اشتباه می‌نمایند. این شیوه‌ها عبارتند از پانچ مجدد و ویرایش (EDIT) که در بخش مربوط به کنترل خطا در انتقال اطلاعات توضیح داده میشود.

ب ـ خستگی و یکنواختی کار

کار پانچ بخاطر یکنواختی آن بسیار خسته کننده بوده و فرد پانچیست می‌بایست دارای صبر وحوصله‌زیادی جهت کار در قسمت منگنه‌زنی باشد تا بتواند ساعتها بطور یکنواخت وثابت نشسته و کار کند. اغلب مشاهده شده که فرد بدلیل ثابت بودن کار دچار خستگی بعضی از اعضاء بدن بخصوص چشم، کمر و گردن میگردد که این عوامل باعث بروز خطا در کار میگردد. خستگی ناشی از ناحیه چشم اثر مستقیم در دید فرد دارد. بطوریکه احتمال خطای دید افزایش یافته و در نتیجه انتقال اطلاعات دچار خطا میگردد. بطور مثال فرد ممکن است عدد ۲ را با ۳ و ۶ را با ۹ و ۷ را با ۸ اشتباه ببیند.

در مواردی نیز برای پانچیست رقم، عدد حاصل میگردد یعنی عدد ۲۳ را ۲۲ و ۵۶ را ۶۵

می‌بیند و یا اینکه سرعت گرفتن دست روی دکمه‌ها هماهنگی دید چشم با دست را بهم میزند و اعدادی غیر از اعداد و ارقام واقعی پانچ میگردد که برای کاهش خطای ناشی از خستگی فرد می‌بایست مدت زمان کار مناسب انتخاب گردد و از طرفی شرایط محیطی مناسبی نظیر نور کافی و مناسب، صندلی مناسب و فضای مساعد کار فراهم نمود تا پانچیست کمتر خسته و فرسوده گردد و در نتیجه مرتکب خطای کمتری شود.

ج - مغشوش و ناخوانا بودن اعداد و ارقام

زمانیکه ارقام پرسشنامه ناخوانا بوده و یا در اثر بازبینی اعداد خط‌خورده و دوباره‌نویسی شده باشند بخصوص وقتیکه جای کافی در پرسشنامه جهت ثبت دوباره وجود نداشته باشد و مامور آمارگیر و بازبین‌جهت نوشتن در مضیقه باشند، اینگونه موارد موجب میگردد که پرسشنامه‌ها کثیف، ناخوانا، مغشوش خط‌خورده و بدخط تکمیل گردد و در نتیجه تمام عوامل موثر جهت ایجاد خطا مهیا گردد. در اینصورت‌پانچیست نمی‌تواند اعداد و ارقام را بطور واضح خوانده و تشخیص دهد ودرنتیجه مرتکب خطا میگردد. برای احتراز از بروز خطا میبایست در زمان آمارگیری وطراحی پرسشنامه تدابیر لازم اتخاذ شده و در زمان بازبینی پرسشنامه‌ها با دقت کافی بازبینی شده و اصلاح گردد.

د - انتخاب افراد

انتخاب افراد سالم وکارآمد جهت انتقال اطلاعات نیز بنوبه خود جهت کاهش حجم خطا موثر میباشدبدین‌معنی که افراد پانچیست میبایست دارای چشمهای سالم، دست و انگشتان سالم‌وکشیده بوده و از بکارگیری افرادیکه دارای انگشتان پهن میباشند، خودداری نمود چون احتمال اینکه در موقع استفاده از صفحه کلیدی (KEYBOARD) انگشتان روی دوکلید ضربه وارد نمایند زیاد می‌باشد، که در نتیجه احتمال بروز خطا افزایش می‌یابد. افراد مبتلا به امراض استخوانی نظیر دیسک نباشند تا نشستن روی صندلی منگنه‌زنی برای مدت زمان طولانی آنان را خسته و فرسوده نکند. حتی‌الامکان افراد پانچیست از میان افرادی انتخاب شوند که اعتیاد به دخانیات نداشته باشندزیرا استعمال دخانیات هنگام کار غیر ممکن می‌باشد.

۶ ـ کنترل خطا در انتقال اطلاعات

بیشترین خطا در مراحل مختلف کار استخراج نتایج مربوط به خطای آمارگیر و خطای پانچیست میباشد که برای کنترل خطای ناشی از مرحله اجراء شیوه‌های بازبینی، کنترل، آماده‌سازی و نظارت بر اجراء اعمال میگردد ولی برای کنترل خطا در مرحله انتقال اطلاعات علاوه بر شیوه‌هائیکه قبلا " ذکر شد به سه روش کلی زیر میتوان عمل نموده و خطا را کاهش داد.

الف ـ پانچ مجدد (VERIFY)

زمانیکه اطلاعات منگنه‌شده ولی هنوز نوار و یا دیسک نشده است میتوان با پانچ مجدد صد در صد اطلاعات، اشتباهات منگنه‌زنی را کشف و اصلاح نمود. اعمال این شیوه در طرحهائی مثل سرشماریها بدلیل حجم زیاد اطلاعات غیر ممکن میباشد و علاوه بر صرف مدت زمان طولانی که در استخراج نتایج اثر نامطلوب میگذارد نیاز به صرف هزینه‌های سنگین جهت پانچ مجدد میباشد و از طرفی ماشینهای منگنه‌زنی که جهت پانچ مجدد مورد استفاده قرار میگیرد باید قدرت و ریف صددرصد را داشته باشند.

در طرحهای بزرگی چون سرشماریها میتوان قسمتی از کار را پانچ مجدد کرده و از طریق بررسی اشکالات موجود در پانچ از بروز خطاهای بعدی جلوگیری بعمل آورد. در طرحهای کوچک و بخصوص آمارگیریهای نمونه‌ای از این شیوه جهت کم کردن خطای انتقال اطلاعات میتوان بهره گرفت و در پانچ مجدد اطلاعات از اشتباهات پانچ اولیه که هنوز به مرحله نوار و یا دیسک اطلاعاتی نرسیده، جلوگیری کرد.

ب ـ کاهش خطا در حین انتقال اطلاعات

همزمان با انتقال اطلاعات میتوان از بروز پاره‌ای از خطاها جلوگیری بعمل آورده و یا آنرا کاهش داد بدین ترتیب که در هنگام پانچ و در زمانیکه هنوز اطلاعات وارد ماشینهای منگنه‌زنی نشده است با استفاده از سیستمهای نوین و پیشرفته DATA ENTRY که از نقطه‌نظر سخت افزار و قابلیت نرم‌افزاری برای ایجاد برنامه‌های کنترل و اصلاح آمادگی داشته و امکان پذیر باشد و قبل از اینکه اطلاعات وارد ماشین شوند آنها را از نظر درست بودن کنترل کرده و در صورت وجود ناهماهنگی اجازه

وارد شدن به ماشین را ندهند.

بکارگیری و استفاده از سیستمهای پیشرفته علاوه بر اینکه در کاهش حجم خطا موثر بوده از طرفی نیز ادیتهای استخراج را آسانتر و مراحل مختلف کنترل و ادیت را کم می‌نماید که در نتیجه هزینه‌های استخراج کاهش یافته و در انتشار نتایج تسریع حاصل میگردد.

ج -ویرایش (EDIT)

هنگامیکه پرسشنامه‌ها پانچ گردید و اطلاعات وارد کامپیوتر شد می‌بایست از صحت ثبت اطلاعات در کامپیوتر مطمئن شویم برای این منظور بوسیله اجرای برنامه‌های ادیت جهت اصلاح فایل اطلاعات اقدام به لیست‌گیری و یا آوردن اطلاعات روی صفحه نمایش (MONITOR) می‌نماید. در این حالت اپراتور مربوطه با استفاده از اطلاعات روی لیستها که در فایل اصلی قرار دارد وانطباق آنها با مدارک موجود اقدام به کنترل و اصلاح فایل اصلی می‌نماید و مجددا" اطلاعات اصلاح شده را جهت پانچ دوباره در اختیار پانچیست قرار میدهد.

در هر مرحله از کار پس از بهنگام نمودن (UPDATE) اطلاعات فایل اقدام به اصلاح فایل اصلی شده تا جائیکه اطلاعات پرسشنامه‌ها با فایل اصلی در کامپیوتر هم‌خوانی داشته‌باشد. این عملیات و رفت و برگشت کار ممکن است بدفعات تکرار شده تا فایل اصلی اصلاح گردد. حال برای احتراز از این آمد و رفت مدارک که خود موجب بروز خطاهای دیگری میگردد از روش ادیت بصورت ONLINE استفاده میگردد که این سیستم از تکرار مجدد پانچ، کنترل و اصلاح، UPDATE و لیست گیری به دفعات مکرر جلوگیری می‌نماید. با استفاده از سیستم فوق‌الذکر میتوان فقط یکبار عمل ادیت را مستقیما" روی صفحه نمایش انجام داده و فایل را اصلاح نمود.

۷-خطای برنامه‌های ادیت و استخراج جداول

استفاده از کامپیوتر برای استخراج نتایج، خطاهای ویژه خود را نیز دارد. بدین ترتیب که برنامه یا دستورکاری که برای کامپیوتر نوشته میشود میتواند مشمول خطاباشد. برنامه، دستورالعملی است که بزبان مخصوص ماشین نوشته میشود و راه حل مساله را در اختیار ماشین میگذارد. این

برنامه میتواند جهت انتقال اطلاعات ، ادیت و جدول‌گیری باشد . این بدان معنی است که برنامه‌نویس باید از طریق حل مساله آگاه باشد . در عمل آنچه که میخواهد از ماشین گرفته شود به برنامه‌نویس گفته میشود و او آنرا بزبان ماشین مینویسد . گاهی لازم است ابتدا ، مساله توسط فرد متخصص حل شود و سپس راه حل که بزبان معمولی ریاضی است برای تبدیل به زبان ماشین به برنامه‌نویس داده شود .

مهمترین عوامل خطای برنامه‌نویس آن است که برنامه‌نویس بدرستی آنچه را که از ماشین خواسته شده است ، درک نکرده باشد . البته کم‌تجربه بودن برنامه‌نویس نیز مزید بر علت خواهد بود که در اینصورت میتوان برنامه نوشته شده را بوسیله اطلاعات آزمایشی (TEST DATA) برای کنترل صحت برنامه‌های استخراج مورد بررسی و آزمایش قرار داد . اطلاعات آزمایشی که بمنظور کنترل و تست برنامه‌ها تهیه میشود باید دارای پوشش کامل بوده و کلیه اشکالات و حالات مختلفی را که در اطلاعات اصلی میتواند موجود باشد ، در برگرفته و دارا باشد تا کلیه جوانب برنامه را کنترل نماید و در صورت صحت برنامه‌نوشته شده از آن برای استخراج اطلاعات و داده‌های اصلی استفاده گردد . البته ناگفته نماند اگر چنانچه برنامه‌های تهیه شده از نظر منطقی غلط باشد ماشین خودبخود آنها را کشف و برنامه نیز غیر قابل اجراء میباشد و همچنین وقتی برنامه حاوی اشتباهات فاحشی باشد نیز از روی نتایج و جداول بدست آمده میتوان به آن پی‌برد ولی اگر چنانچه اشتباهات کوچک و یا ناپیدا باشند همچنان باقی خواهند ماند . یکی دیگر از عواملی که در تهیه برنامه‌های استخراج موجب کاهش خطا میگردد هماهنگی کامل سیستم آنالیست (SYSTEM ANALYST) با برنامه‌نویس (PROGRAMMER) میباشد ، بدینصورت که سیستم آنالیست میبایست کلیه دستورات لازم را بصورت مکتوب و مستند بمنظور تهیه برنامه‌ها در اختیار برنامه‌نویس قرار داده و برنامه‌نویس نیز میبایست اطلاعات کافی در مورد سیستم استخراج داشته باشد تا با کمک یکدیگر یک تیم کاری مناسب را تشکیل دهند .

عوامل مؤثّر در ارتقای کیفیّت و ارائهٔ بهنگام آمار

جعفر صادقی نامور
مرکز آمار ایران

مقدمه

امروزه اداره جوامع مدرن برپایه انواع اطلاعات استواراست . اطلاعات اساس برنامه ریزی ووسیله ای لازم برای تصمیم گیری درفعالیتهای جاری دولت ومدیران بخش خصوصی است این اطلاعات باید موثق وبهنگام باشد تا از حداکثر ارزش برخوردار شود . بهنگام بودن اطلاعات یعنی نزدیکی زمان انتشار آن به دوره آماری مربوط.دلالت برکارایی دارد که کار تولید آمار را برعهده دارند . مصرف کنندگان آمار همیشه آمار بهنگام را ترجیح میدهند و این حق آنهاست که ارائه اطلاعات کهنه را زیر سئوال ببرند . درجامعه ای که پدیده های اقتصادی واجتماعی درزمانهای کوتاه درحال تحول است اطلاعات به عنوان ابزار تصمیم گیربها خیلی سریع اعتبار خودرا از دست میدهد . ویژگی دیگر اطلاعات قابل استفاده ، دقت آن است قابل اعتماد بودن اطلاعات به روشهای جمع آوری وشیوه های استخراج وانتشار آنها بستگی دارد درآنجا تجربه ، دانش وهنرمندی آمارشناسان وکارشناسان آمار قانون تعیین کننده است تعیین نقطه تعادل بین دقت وبهنگام بودن اطلاعات به توانایی وتکنیک نیروی انسانی وامکانات دوگیر درتولید آمار ارتباط مستقیم دارد . تولید کنندگان آمار باشناخت از نوع نیازهای آماری از

یاد طرف وبادرنظـــر گرفتن امکانات موجود از طـــرف دیگر وظیفه سنگینی درحفظ موازنـــه ای معقول بین صحت آمار وسرعت انتشار آن دارند.

تکنولوژی درحال پیشرفت کامپیوتر امکان تجزیه وتحلیل واستخراج حجم عظیمـــــی از اطلاعات را درزمانهای کوتاه فراهم کرده ومی‌توانند درارائه بهنگام آمار تاثیـــر مطلوب داشته باشد مهمترین مسئله این است که چگونه ازایـن تکنولوژی استفاده شـود وبیشترین استفـــاده درصورتی حاصل می‌شود که ترکیب مناسبی از نیروی انسانی ماهر وتکنولوژی دراختیـــارباشـــد بهمین منظور دستیابی تولید کنندگان آمار به نیروی انسانی ورزیده برای بهره گیـــــــری‌از کامپیوتر دارای اهمیت خاص است . امروزه اگر چه حتی درکشورهای توسعه یافته پیشرفتـــــای تکنولوژی برآموزش نیروی انسانی در استفاده ازآن پیشی گرفته است ولـی کمبود نیروی‌انسانـی ماهر درکشورها بطور اعم ودرسازمانهای دولتی بطور اخص کامـــلا " چشم گیر است . نگاهی بـــه ظرفیت وتوانایی کامپیوترهای موجود درسازمانهاومجامع علمی کشور ومیزان استفاده از آنهـــــا تاییدی برمطالب فوق است از اینرو آموزش های مناسب پرسنل آماری وتبادل تجربیات ودانـــش آنها از طریق چنین سمینارهایی برای بالابردن میزان استفاده از تکنولوژی مدرن یک ضـــرورت است .

بحث درزمینه دقت وبهنگام بودن آمار را درسه مقوله جنبه های مدیریتی ، فنی واستفاده از تکنولوژی می‌توان طبقه بندی کرد.

جنبه های مدیریتی

باتوجه به اهمیت برنامه ریزی ومدیریت درکنترل عملیات تولید آمار لازم است کـــــه آموزش کافی دراین زمینه ارائه شود . آموزش به عنوان یک عنصر حساس دربهبود کیفیت وبهنگام

سازی ارائه آمار مطرح است . آموزش ضمن خدمت برای آشنایی با تکنولوژی مدرن ، روشهای نوین آماری وتجربیات دیگر کشورها با سازمانهای داخلی با استفاده از وسایل کمکی آموزشی ، نیروی انسانی ، آزمایشهای واقعی وفنون سمعی وبصری از ضرورتهای اجتناب ناپذیر است . امکانات این نوع آموزشها میتواند ازطریق سازمانهای عمده آماری وموسسات آموزشی فراهم آید . کنفرانس آمار وبرگزاری سمینارهای مناسب از طریق انجمن آمار میتواند مکمل فعالیتهای سازمانهای فوق باشد .

پیشرفت وافزایش دسترسی به کامپیوتر دوارتقاء کیفیت وسرعت انتشار آمار کاملا " موثر است ولی بطوریکه اشاره شد کمبود نیروی انسانی ماهر دربکارگیری این تکنولوژی مشکل اساسی است . آموزش آمارشناسان وکارشناسان برای انجام برخی کارهای کامپیوتری به عنوان کمک برنامه نویسان میتواند مشکل را کاهش دهد ضمن اینکه آگاهی این رده از مسایل موجود ارتباط با ماشین ، اثرات مثبتی درطراحی طرحهای آماری میگذارد همانگونه که آشنایی کادر کامپیوتری با مفاهیم موضوعی طرح میتواند مفید باشد . ولی درهرحال آموزش آمارشناسان وکارشناسان غیر کامپیوتری درزمینه استفاده از بسته های نرم افزار که بطور وسیعی درحال گسترش است درغلبه برمشکلات بکارگیری ماشین اثر تعیین کننده را خواهد داشت .

برای اطمینان ازکارایی عملیات درروند جمع آوری آمار لازم است درمورد کلیه فعالیتها ومسایلی که درمراحل مختلف آمارگیری وبعدازآن بروز میکند فکر وبرنامه ریزی شود . تهیه برنامه زمانبندی شده فعالیتها ، بطوری که به کلیه واقعیتهای موجود وموثر درتهیه طرح ، اجرا واستخراج نتایج توجه شود از مهمترین ضرورتهای برنامه ریزی تولید آمار است . متعهد بودن به اجرای فعالیتها طبق برنامه پیش بینی شده ضرورتی مهمتر میباشد . استفاده از تکنیک هایی نظیر CPM (روش مسیر بحرانی) این امکان را فراهم میکند که ضمن شناخت

کلیه فعالیتها وزمان لازم برای هرفعالیت ، پیش نیازهای فعالیتها نیز شناخته شود و فعالیتهایی که باید همزمان انجام گیرد وهمچنین فعالیتهایی که درمسیر بحرانی قرار می‌گیرند تحت کنترل مدیریت طرح درآید .

ارتباط بین تولید کنندگان ومصرف کنندگان آمار می‌تواند به ایجاد زبان مشترک بین آنها کمک کند . شناخت نیازهای مصرف کنندگان ضمن اینکه درتعیین اهداف طرح آمارگیری کمک کننده است دانش قبلی از نیازهای استفاده کنندگان دربرنامه ریزی استخراج نتایج طرح والویت بندی استخراج نیز سودمند می‌باشد تهیه واجرای طرح خاصی در زمینه شناخت نیازهای آماری سازمانهای مختلف وطبقه بندی آنها وهمچنین کوشش درشناخت نیازهای آماری برنامه های توسعه اقتصادی ، اجتماعی وفرهنگی جمهوری اسلامی ایران توسط مرکزآمار ایران تجربه مفیدی دراین رابطه است . همخوانی بین تولید کنندگان آمار واستفاده کنندگان ذینفع ازاین نظر مفید است که اطمینان دهد آمار مرتبط ، صحیح ومناسب مورد استفاده قرارگرفته است .

ایجاد هماهنگی بین سازمان مرکزی آمار وسازمانهای تولید کننده آمار درسطح جغرافیایی کوچکتر ازیک طرف وسازمان مرکزی آمار بادیگر سازمانهای تولید کننده آمار از طرف دیگر، یکی از مسایلی است که مخصوصا" درزمان فعلی کمک زیادی می‌تواند به بهبود کیفیت واراسته بهنگام آمار بکند . ایجاد ارتباط نزدیک بین سازمان مرکزی آمار وادارانی که تولید کنندگان آمارهای ثبتی هستند نیز توجه خاصی را طلب می‌کند این ارتباط به بهبود طبقه بندیها ومفاهیم استاندارد وارتقاء کیفیت آمارهای ثبتی کمک می‌کند .

انجام برخی از فعالیتهای آماری بعداز مرحله جمع آوری بصورت غیر متمرکز به سرعت کدگذاری ، ادیت وتجزیه وتحلیل اطلاعات بهبود می‌بخشد اگر چه ممکن است باعث غیراستاندارد

شدن برخی از تعاریف ومفاهیم وطبقه بندیها شود. سهولت دسترسی به میکرو کامپیوترها در برخی از استانها این امکان را بوجود می‌آورد که اطلاعات جمع آوری شده درمراحل بعدی بسته به صورت الکترونیکی به مرکز انتقال داده شود تا مراحل بعدی پردازش وتولید آمار در سطح ملی روی آن صورت گیرد.

باتوجه به اینکه برای انجام کارهای آماری نیروی انسانی ومنابع مالی وتجهیزات وسیعی مورد نیاز است مدیران پروژه های آماری باید برای کارهای خود اولویت بندی وتعیین برنامه کرده وقبل ازهر اقدامی مطمئن شوند که منابع لازم برای به پایان رسانیدن موفقیت آمیز پروژه پیش بینی شده است. درسالهای اخیر تعداد طرحهای آماری مختلفی که درسازمانهای متعدد شده ودرمراحل اجرا یا استخراج ناتمام باقیمانده است کم نیست.

به وجود آوردن انگیزه یکی از ابزاری است که می‌تواند کارایی پرسنل آماری را افزایش دهد انگیزه ها می‌تواند براساس حاصل کار انجام گرفته ویا کیفیت آن درمراحل مختلف جمع‌آوری آمار مطرح شود برخی از سازمانها از انگیزه های پولی درمراحل مختلف تولید آمار استفاده کرده ونتیجه مطلوبی گرفته اند.

اتخاذ خط مشی‌های درست برای آگاهی دادن به عموم درارتباط با جمع آوری آمار، با این هدف که ضمن به حداکثر رسانیدن میزان همکاری، پاسخهای باکیفیت بالا حاصل شود می‌تواند اثرات مثبت معنی‌داری داشته باشد خط مشی های تعیین شده می‌تواند تهیه بروشور وفلرچه‌های توضیحی برای پاسخگویان که به طور ساده اهداف جمع آوری اطلاعات را شرح دهد، باشد. استفاده از روزنامه، رادیو وتلویزیون ودیگر وسائط ارتباط جمعی شامل مصاحبه های مسئولان موسسات آماری می‌تواند کمک کننده خوبی باشد. اطمینان دادن به پاسخگویان که اطلاعات اخذ شده از آنها محرمانه خواهد بود درارتقاء کیفیت اطلاعات دارای اهمیت خاصی است و

می‌بایست به پاسخگویان منتقل شود . جایی که ماموران آمارگیر توانسته اند با توضیحات مناسب درمورد اهداف آمارگیری واصل محرمیت آمار ، اعتماد پاسخگو را جلب کنند نرخ پاسخگویی و کیفیت اطلاعات اخذ شده درسطح کاملا " مناسبی بوده است . درحالیکه استفاده از اختیارات قانونی درواداشتن پاسخگو به ارائه اطلاعات باعث ایجاد جو نامساعد واثر نامطلوب بر کیفیت آمار میشود .

استفاده از تعاریف ومفاهیم وطبقه بندیهای یکسان درآمارگیریهای مختلف توسط سازمانهای جمع آوری کننده آمار باعث میشود تا استفاده کنندگان بتوانند اطلاعات مختلف را به راحتی درکنار یکدیگر قرار داده ونیازهای آماری خودرا برطرف کنند . اگر درآمار واطلاعاتی که توسط سازمانهای آماری جمع آوری وارائه میشود از تعاریف یکسان استفاده نشود نتایجی متفاوت را نشان خواهد داد نوسان این نتایج گاه به حدی است که استفاده کننده را به تردید درصحت آنها وامی‌دارد . برای ایجاد امکان بالقوه مقایسه اطلاعات استفاده از طبقه بندیهای استاندارد ضرورتی اجتناب ناپذیر است .

<u>جنبه های فنی</u>

سرشماریها ، نمونه گیری ها وجمع آوری آمارهای ثبتی هرکدام مسایل خاص خودش دارد و باید بعنوان مقولاتی جداگانه مطالعه وبررسی شود تا مشکلات گردآوری آن به حداقل برسد . همچنین باید چارچوبهای آماری بهنگام شود واین خود باتوجه به تغییرات سریع اقتصادی واجتماعی جامعه به کوشش زیادی نیاز دارد یک طرح نمونه گیری خوب نیاز مند چارچوبی باکیفیت خوب ونقشه های بهنگام است وتهیه نقشه های بهنگام نیز مستلزم اختصاص منابع وسیع مالی ونیروی انسانی به آن می‌باشد .

طراحی خوب پرسشنامه ها دارای اهمیت خاصی است ودربهبود کیفیت وارائه بهنگام

آمار مستقیما" تاثیر دارد به تجربه دیده شده است چنانچه درطراحی پرسشنامه ها هماهنگی وهمفکری لازم بین آمار شناسان ،کارشناسان موضوعی ، طراحان ومتخصصین استخراج آمار موجود باشد مشکلات اجرایی واستخراج به حداقل میرسد وزمان زیادی صرفه جویی میگردد . در طراحی فرمها علاوه برملاحظاتی مانند ترتیب منطقی سئوالها ، شکل پرسشنامه ، مسایل کدگذاری ، میبایست نیازهای اداری وعملیات آمارگیری نیز درنظر گرفته شود . آزمایش قبلی پرسشنامه ها وفرمها میتواند دربهبود آنها کمک کننده باشد .

وجود دستورالعمل های مدون برای رده های مختلف شغلی در آمارگیریها ، علاوه بر کاهش خطاهای غیر نمونه‌ای درتسهیل عملیات اجرایی نیز کمک کننده است چگونگی استفاده از نقشه ، نحوه دستیابی به واحد آماری ، روش تکمیل فرمها وپرسشنامه ها وتعاریف ومفاهیم اصطلاحات بکار رفته درپرسشنامه ها مواردی هستند که دردستورالعمل باید عنوان شود همچنین میتوان چگونگی دریافت وتحویل فرمها ومدارک ونوع ارتباط وروش هماهنگی باردههای شغلی دیگر را در دستورالعمل یادآورد شد .

تصمیم برای استفاده از تکنولوژی بایستی برحسب نوع آمارگیری یا سرشماری وساختار سازمانی پیش بینی شده برای طرح اتخاذ شود ولازم است این تصمیم گیری با توجه به برنامه ریزی چگونگی جمع آوری آمار وبادرنظر گرفتن کارایی وهزینه استفاده از تکنولوژی صورت گیرد ودرهرحال برخورد باتکنولوژی جدید بدون درنظرگرفتن امکان جذب کارکنان متخصص نتیجه لازم را نخواهد داد از تکنولوژی سدرن جدا از عملیات ثبت واديت اطلاعات میتوان دربدست آوردن داده های آماری استفاده کرد واین خود برای پیشگام نگاهداشتن انتشار نتایج کمک موثری است .

خطاهای نمونه ای آمارگیری درمرحله طراحی طرح تعیین میشود ولی تعیین خطاهای

غیر نمونه‌ای کار دشواری است بهمین دلیل تاثیر آنها برروی آمار ناشناخته باقی می‌ماندانجام بازشماری بعداز آمارگیریهای بزرگ وسعت خطاهای غیرنمونه‌ای را مشخص می‌کند. برای ارزیابی کیفیت آمارهای جمع آوری شده از روشهای دیگر نیز مثل تجزیه وتحلیل سریهای زمانی ومراجعه به سایر منابع می‌توان استفاده کرد.

مراحل تدوین وچاپ اطلاعات به زمان نسبتا" زیادی نیاز دارد استفاده از جدول چاپی کامپیوتری علاوه برسرعت بخشیدن به انتشار نتایج باعث جلوگیری از نوشتن مجدد، بازبینی وسیع درمرحله انتشارات ووجود خطاها دربازنویسی وتایپ می‌شود.

بکارگیری روشهای نمونه‌گیری درانتشار نتایج زودرس سرشماریها می‌تواندبرد دستیابی به اطلاعات مورد نیاز مصرف کنندگان سرعت ببخشد. این نوع اطلاعات از دقت کمتری برخوردار است ولی برای رفع نیازهای فوری قابل قبول خواهد بود.

استفاده ازتکنولوژی

دراینجا قصد این نیست که جایگاه کامپیوتر درتولید آمار اشاره شود چراکه اهمیت، استخراج وتجزیه وتحلیل نتایج آمارگیریها مخصوصا" درسطح وسیع آن بدون بهره گیری ازاین تکنولوژی اصولا" غیرعملی می‌نماید. امروزه پیشرفتهای اصلی که درجمع آوری آمار حاصل شده است براساس استفاده از وسایل الکترونیکی به‌صورت مستقیم متمرکز است دراین رابطه می‌توان به استفاده از CAC[1] کدگذاری باکمک کامپیوتر، CATI[2] وارکردن اطلاعات باکمک کامپیوتر، CAPI[3]

1) Computer Assited Coding
2) Comouter assisted Data Entry
3) Computer assisted personal interviewing

(۱)

مصاحبه شخصی باکمک کامپیوتر ، CATI مصاحبه تلفنی باکمک کامپیوتر بعنوان نمونه ای از دستاوردهای تکنولوژی مدرن اشاره کرد . کارایی روش اخیر مخصوصا " درآمارگیریهای اقتصادی دربرخی ازکشورها تجزیه شده است بهره گیری از کامپیوتر به کمک مصاحبه تلفنی مبتنی بر برنامه ریزی تماسهای تلفنی وشماره گیری اتوماتیک ، انجام پرسش وپاسخ ومقایسه جوابها با آمار قبلی وبالاخره ورود اطلاع به فایل درحین صحبت وهمزمان باآن ادیت وحذف خطا می‌باشد . بدیهی است پیاده کردن چنین روشی علاوه برنیاز به تکنولوژی ونیروی انسانی ماهر مستلزم کارایی بخش های دیگر مانند وسایل ارتباطی است . هرکدام ازروشهای فوق مزایای قابل توجهی دربهبود کیفیت و ارائه بهنگام آمار می‌تواند داشته باشد . امروزه سازمانهای تولید کننده آمار کشور به مزایای این گونه روشها آگاهی یافته ودرجهت استفاده از آنها کوششهایی بعمل آورده‌اند . کدگذاری به کمک کامپیوتر یکی از روشهای فوق است که باامکانات موجود کشور قابل پیاده شدن بوده وهم اکنون استفاده ازآن مرحله آزمایشی خودرا طی می‌کند بهمین منظور مزایای این روش بعنوان نمونه ای از استفاده های تکنولوژی درتولید آمار شرح داده می‌شود .

کد گذاری یکی از فعالیتهای تولید آمار است که معمولا " به زمان زیادی نیاز دارد بطوریکه مثلا " کد گذاری فرمهای یک سرشماری مانند سرشماری عمومی نفوس وممکن به صورت دستی بیش از ۶ ماه وقت نیاز دارد وفعالیت کد گذاری به این ترتیب است که کد گذار برای یافتن کد مورد نظر خود درزمینه فعالیتهای اقتصادی ، مشاغل وغیره می‌بایست جداول کد را که درواقع کتاب قطوری‌است جستجو کند وپس از انتخاب کد فعالیت ، آنرا درمحل پیش بینی

---

1) Computer Aasisted Telephone Interveiwing

شده دربرنامه درج نماید این کار علاوه بر وقت گیر بودن آن باتوجه به پیچیدگی کدها معمولا " دارای خطای نسبتا " زیادی نیز می‌باشد .

درروش CAC ماشین بااستفاده از برنامه های داده شده ،عمل تطابق ،عبارت را باکد انتخاب شده از طریق جدول کد مربوط انجام میدهد . ابتدا اپراتور به توسط دستگاه ورودی،عبارت را وارد کامپیوتر می‌کند سپس کامپیوتر با جستجو درجدول کد مربوطه ،کلیه عبارتی که نزدیکترین تطابق با عبارت ورودی را داشته باشد بهمراه کدهای وابسته شان لیست می‌کند سپس کد گذار می‌تواند یکی از عبارات لیست شده را که مناسبترین می‌باشد انتخاب نموده ودر صورت ناکافی بودن ،بخشی از عبارت را وارد کامپیوتر نماید تا عملیات تطابق دوباره صورت پذیرد. اگر اطلاعات وعبارات لیست شده برای کد گذار باز هم ناکافی باشد یک امکان گسترش می‌تواند به سیستم CAC اضافه گردد تا کدگذار را قادر به نمایش یک لیست از عبارات بایک کد معین نماید. درحال حاضر تلاش زیادی بعمل می‌آید تا روش CAC بصورت یک سیستم تمام اتوماتیک درآید ودربرخی ازکشورها نیز این کار عملی شده است .

مشکلات مربوط به تهیهٔ چارچوب آماری

صدیقه صالحی

مرکز آمار ایران

ساختن چارچوب آماری از اهمیت بسیاری برخوردار است. بهتر این است که وقت بیشتری برای ساختن چارچوب آماری گذاشته شود و با بررسی و تفکر چارچوب مطلوبی با توجه به هدفهای آمارگیری نمونه ای انتخاب گردد، تا هر گونه خطایی که از نمونه ها بدست می آید دقیق تر و با واقعیتهای جامعه نزدیکتر باشد.

تعریف چارچوب آماری

در کتابهای آماری تعاریفی از چارچوب وجود دارد که چند نمونه از آن در زیر آمده است.

۱- چارچوب یعنی اسامی آمارگیری نمونه ای که عبارتست از فهرست مجموعه کلیه واحدهای جامعه مورد مطالعه که از آن تعداد معین نمونه انتخاب میشود.

۲- اگر بخواهیم یاد آوری فهرست کلیه افراد جامعه به نمونه گیری بپردازیم در اینصورت مجموعه اطلاعاتی را که در این فهرست ثبت شده است، چارچوب آمارگیری می نامند.

۳- مجموعه کاملی از واحدهای نمونه که بایستی نمونه مورد نظر از بین آنها ارائه شود، چارچوب جامعه نمونه گیری نامیده میشود.

چارچوب نمونه گیری ممکن است عبارتی از اشخاص یا واحدهای مشخص یا یک پرونده یا یگانی، مجموعه ای از کارتهای منفرد زنی الفبایی از یک منطقه جغرافیایی تقسیم بندی شده و یا اطلاعاتی از تمام افراد در یک سال معین ثبت باشد.

۴- کلمه چارچوب آمارگیری عبارتست از فهرست کاملی از عناصری که می خواهیم در مورد آنها مطالعه انجام دهیم مثلا فهرست کامل ساختمانهای یک شهر.

۵- نمونه دیگری از تعاریف ارائه شده برای چارچوب آماری در زیر آمده است.

به کار بردن روشهای نمونه گیری برای جمع آوری اطلاعات مستلزم داشتن فهرست کاملی از واحدهای نمونه گیری می باشد. مجموعه واحدهای نمونه گیری تشکیل جامعه مورد نظر را میدهند. یکچنین فهرستی را چارچوب نمونه گیری می نامند. چارچوب نمونه گیری ممکن است شامل مشخصاتی درباره یک منطقه جغرافیایی به حدود واحدهای شمارش تحویل گیرنده را برای چارچوب در نظر می گیریم و به ادامه بحث می پردازیم.

۶- هرنوع تصویرمناسب ازواحدهای آماری جامعه موردبررسی را یک چارچوب آماری گویند بطوری که به ازاء هرفرد جامعه ما یک تصویرمناسب از آن در اختیار داشته باشیم . این تصویر ممکن است تشخیص یک نماره به هرفرد جامعه باشد و یا یک کارت جایگاه در روی نقشه و یا یک نام و یا مجموعه ای از یک نقشه به هر واحد جامعه اختصاص یافته باشد .

شناسائی جامعه آماری

در تمام تعاریفی که برای چارچوب آماری و یا چارچوب نمونه گیری ارائه شده است، مشخص شد که اصولا قبل از معرفی چارچوب آماری و تعریف آن ، باید جامعه آماری را بشناسیم و آنرا تعریف کنیم .

در کتابهای درسی و دانشگاهی برای جامعه آماری نیز تعاریفی آمده است که چندنمونه از آن درزیر آمده است .

۱- جامعه عبارتست از مجموعه کلیه واحدهای مورد بررسی که در واقع، هدف آمارگیری، برآورد مشخصات آن می باشد .

۲- اگر اشیاء یا نمونه ها را نسبت به یک یا چند صفت مشترک یکجا در نظر گرفته شوند، گوئیم یک جامعه آماری ... یا عبارتساده تر یک جامعه تشکیل شده است .

۳- مجموعه افراد یا اشیائی که دارای یک یا چند صفت مشترک باشند، جامعه آماری و یا بطور ساده جامعه نامیده می شوند .

۴- جامعه آماری به مجموعه ای از افراد یا اشیاء یا به هر جامعه دیگری اطلاق می گردد که افراد و یا عناصر آن جامعه، صفت یا صفتهای مشخص کننده درزمینه تحقیق مورد نظر را دربرداشته باشد مثل جامعه ساکنان یک شهر، جامعه کارگاههای خریدوفروش کالا .

۵- ملاحظه کنید در تعاریف بالا هم هرچند مفهوم را نمی رسانند ولی خالی از نقص نیز نمی باشند ، اکنون تعریف زیر را مورد جامعه آماری درنظر می گیریم .

جامعه آماری عبارتند از مجموعه واحدهائیکه دارای یک صفت مشترک باشند وضمنا صفات متغیر دیگری در آنها باشد که موضوع بررسی آماری قرار می گیرد .

چه جوامعی موضوع چارچوب سازی آماری قرار می گیرند :

باتوجه به تعاریفی که از چارچوب جامعه آماری آمد به این نتیجه می رسیم که چارچوب آماری باید از جوامعی تهیه گردد که محدود و یا قابل شمارش باشد پس به این ترتیب از جوامع نامحدود نمی توان چارچوب آماری تهیه نمود ، چون نمی توان محدوده این نوع جوامع را مشخص کرد .

خصوصیات لازم برای ساختن یک چارچوب آماری مطلوب چیست :

به منظور برآورد مقدار یا صفاتی از یک جامعه باید از آن جامعه نمونه گیری بعمل آوریم و برای آنکه نمونه گیری خوب انجام باید چارچوبی از واحدهای آماری جامعه مورد بررسی در اختیار داشته باشیم ، چارچوبی با خصوصیاتی که حداقل نیازمندیها را تامین نماید .

یک چارچوب را کامل گوئیم ، اگر هر عنصر جمعیت کامل یکبار و فقط یکبار در آن بطور جداگانه آمده باشد . عناصر دیگر در آن نشان داده نشوند و اطلاعات دقیق باشد .

در کار نمونه گیری واقعی ، چارچوبهای کامل کم هستند و همواره نواقصی وجود دارد . چارچوب ممکن است شامل بعضی از واحدهای جامعه نباشد که در این صورت چارچوب ناقص است و یا ممکن است بعضی از واحدهای جامعه بیش از یکبار در آن وارد شوند که در آن صورت جامعه دارای واحدهای تکراری است و اگر بعضی از اطلاعات موجود در چارچوب صحیح نباشد چارچوب دقیق نخواهد بود .

اگر از وقتی که چارچوب ساخته شده است در آن تجدید نظر نشده باشد ممکن است برای زمان حاضر قابل استفاده نباشد و یا برای استفاده کافی نباشد و واضح است که نقائص چارچوب باعث ایجاد خطا در نتایج حاصل میشود . بنابراین چارچوب باید قبل از استفاده به دقت بررسی شود و قبوه کافی باید بکار رود که نقائص موجود در چارچوب تا حد امکان نتایج حاصل از بررسی را تحت تاثیر قرار ندهند . تهیه چارچوب خوب برای برنامه های نمونه گیری بسیار مشکل است .

بطور خلاصه یک چارچوب آماری مطلوب باید تا حد امکان این خصوصیات را داشته باشد .

۱ـ کامل باشد ، یعنی چارچوب آماری را کامل گوئیم هرگاه هر عنصر آن یکبار و فقط یکبار در آن بطور جداگانه آمده باشد یعنی تمام عناصر جامعه در آن برگزار شده باشد . اگر چارچوب حاوی بعضی از عناصر جامعه نباشد کامل نیست و ناقص تلقی میشود در این حالت باید این واحدها را پیدا کرده و در کار نمونه گیری ، معمولا - چارچوبهای کامل نادرستند و همواره در چارچوبها خطا که برای آمار گیری نمونه ای تهیه میشود نواقصی وجود دارد .

۲ـ بهنگام باشد ، چارچوب آماری را بهنگام گوئیم هرگاه اطلاعات آن نسبت به زمانیکه مورد استفاده نمونه گیری قرار می گیرد کهنه و قدیمی نباشد .

۳ـ در چارچوب واحدهای تکراری نباشد ، هر عنصر آن یکبار و فقط یکبار در آن منظور شده باشد .

۴ـ در چارچوب واحدی وارد شده باشد که در جامعه کاری مربوط نباشد .

۵ـ صحیح و دقیق باشد ، منظور هنگام تهیه چارچوب صحت اطلاعات مربوط به آن ، شرایط دیگری برای در اختیار داشتن یک چارچوب آماری مطلوب است .

چارچوب آماری نامناسب چه عمومیاتی دارد ؛

۱- ممکن است چارچوب اصولا - به جمعیت مورد نظر در آمارگیری مربوط نباشد که این مورد نامناسبترین صورت استفاده از چارچوب است .

۲- کهنه بودن اطلاعات در چارچوبهای آماری قدیمی و کهنه لزوم آوری است که موجب اشتباهات نمونه ای مهمی در نتیجه نتایج حاصل از بررسی های آماری میشود .

۳- پوشش نادرست : مانند مناسب بودن منظور داشتن و اشخاص غیر مرتبط به جامعه آماری در چارچوب است یعنی نباید از واحدهائی که اصولا - در ای صفت یا مشخصات مشترک افراد جامعه آماری نیستند آمارگیری بعمل آید .

امکان بوجود آمدن اشتباه در مراحل مختلف آمارگیری وجود دارد ، ولی بالاترین احتمال مربوط به نادرست انتخاب کردن چارچوب می باشد که در نتیجه نمونه های انتخابی اشتباه خواهند بود .

نمونه ای از یک چارچوب سازی آماری و مراحل مختلف تهیه آن :

باتوجه به اینکه تاکنون اقداماتی را که در رابطه با چارچوب سازی برای طرح آمارگیری از قیمت کالاها و خدمات مورد مصرف خانوارهای روستائی سال ۱۳۶۶ صورت گرفته است به عنوان مثالی برای چارچوب سازی یک طرح آماری از نظر می گذرانیم .

بررسی و مطالعاتی که بر روی اسناد و مدارک مربوط به سابقه طرح شاخص قیمت خرده فروشی کالاها و خدمات روستائی بعمل آمد ، مشخص شد که مشکل اصلی و اصولی در عدم دسترسی مطلوب به اطلاعاتی که در رابطه شاخص در زمینه های مربوط می باشد عدم وجود یک چارچوب آماری خوب بر اساس فرم مورد نظر طرح است . بدین جهت پیشنهاد شد که از اجرای پروژه عظیم سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ بهره گیری شود و در رابطه با طرح مذکور تحولات اساسی گنجانیده شود تا با استفاده از نتایج سرشماری ، چارچوب آماری لازم تهیه گردد .

این پیشنهاد با استقبال مواجه شد و به طور اصولی تائید گردید و بخش A فرم ۳ سرشماری به آن اضافه یافت و سوالی به شرح زیر در پرسشنامه آبادی گنجانیده شد .

• آیا اگر ساکنان آبادی به احتیاج اولیه خود را از همین آبادی خریداری می کنند ؟ بلی ☐ نه ☐

• اگر بلی ، آیا آبادیهای دیگر هم برای خرید به احتیاج اولیه به این آبادی مراجعه می کنند ؟ بلی ☐ نه ☐

اگر ساکنان آبادی برای خرید احتیاج خود به شهر یا شهرهائی مراجعه می کنند حداکثر نام دو شهر را بنویسید .

۱- ..... ۲- .....

پس از اجرای عملیات سرشماری با توجه به اطلاعات طرح آمارگیری از قیمت کالاها و خدمات مورد مصرف خانوارهای روستائی جامعه آماری طرح ، متشکل از دو زیرجامعه ، یکی زیرجامعه مربوط به مراکز خرید روستائی روستائیان و دیگری زیرجامعه مربوط به مراکز خرید شهری روستائیان حاصل گردید . بنا به تعریف و بر اساس اهداف آمارگیری ، مراکز خرید روستائی ، آبادیهائی هستند که در بخش A فرم ۳ سرشماری به کدها ۱ و ۲ یا ۱ و ۳ داده شده است و مراکز خرید شهری ، شهرهائی هستند که در قسمت دوم بخش A فرم ۳ سرشماری ، اسامی آنها ذکر شده باشد .

برای تهیه چارچوب آماری مطلوب از دو زیرجامعه بالا، ابتدا شکل کلی جدول مربوط به نظریه فرهنگ آبادیهای حاصل از نتایج سرشماری عمومی نفوس و مسکن سال ۱۳۴۵ در نظر گرفته شد و بر اساس اطلاعات مندرج در فرم ۳ سرشماری مذکور، جدول مورد نیاز از اطلاعاتی که برای تهیه چارچوب طرح لازم و یا مفید تشخیص داده شده بود، به شرح زیر طراحی شد.

۱ـ شماره ردیف آبادی در نمونه.

۲ـ نام آبادی مسکونی دائمی و کد اختصاصی آن.

۳ـ راه یا راههائی که به آبادی منتهی میشود، شامل راه زمینی و نوع آن، راه آهن و راه آبی.

۴ـ مطی جهت خروج مقنیات آبادی در روی نقشه.

۵ـ تعداد خانوار معمولی ساکن و جمعیت معمولی ساکن در آبادی.

۶ـ مطی جهت خروج کد دهگانهای مربوط به اینکه آبادی مرکز خرید محسوب میشود یا نه؛ یکی از کدهای ۱، ۳ یا ۴، ۲.

۷ـ مطی جهت نسبت اسامی شهرهای محل خرید شهری ساکنین آبادی، که باید از فرم ۳ استخراج و به صورت مختصر به جدول اضافه شود.

۸ـ وجود امکانات تعاونی و اداری آبادی، شامل ۳ سوال.

۹ـ وجود امکانات تبلیغ و درمانی آبادی، شامل ۱۱ سوال.

۱۰ـ وجود امکانات مربوط به مظاهر فرهنگ و ارتباطات آبادی، شامل ۳ سوال.

۱۱ـ امکان مربوط به وجود برق در آبادی.

۱۲ـ وجود امکانات تولیدی و بازرگانی و خدماتی آبادی، شامل ۱۲ سوال.

با توجه به زیرجامعه های طرح و بر اساس اطلاعات بالا، دو جدول دستگل برای هریک از استانهای کشور طرح شده بدین ترتیب که جدول اول تمام آبادیهای مسکونی دائمی دشت و نیمه دشت بخش، شهرستان و کل استان نشان میدهد و جدول دوم شامل آبادیهای مسکونی دائمی استان است که در قسمت اول از بخش ۸ فرم ۳ سرشماری به آنها کد ۱ و ۳ یا ۱ و ۴ داده شده باشد.

با استخراج ماشینی اطلاعات سرشماری، از اولین اطلاعاتی که کار انتقال اطلاعات فرم ۳ آنها به روی نوار خاتمه یافته بود، بررسی چارچوبها مد نظر و عمل قرار گرفت و استان ایلام به عنوان نماینده ای از استانهای منطقه جنگی و استان یزد به عنوان نماینده استانهایی که در آنها ایل صحرائی قرار ندارد انتخاب گردید.

بررسی جداگانه و مستقلی که از جدول دوم چارچوب مربوط به طرح، هریک از استانهای ایلام و یزد بعمل آمد معلوم کرد که بین کدهای مربوط به وضعیت صنعتی آبادی و شاخصهای ساکن در آن و همچنین بین کدهای مربوط به وضعیت تجدید حیات ساکنان آبادی با وجود امکانات خدماتی و بازرگانی آبادی، علامتهائی وجود دارد لذا تنی توان چارچوبهای تامین را که به این ترتیب بدست آمده است تغییر داد یا تغییر داد.

مثلا در استان ایلام، ۶۱/۵ و در استان یزد ۷۵/۲ آبادیهائی که بپرسش سوال ۸ـ۸ بیکی از کدهای ۱ و ۳ یا ۱ و ۴ مشخص شده بودند یا اصولا کارگاهی نداشتند و یا فاقد یکی از گروههای تعاونی بودند؛ این دو کارگاه از کارگاههای مهم و تاثیری برای داد و حد فاصل تولید روستائیان در آبادیهای مسکونی دائمی داشتند و اصولا خانوار و جمعیت اندکی در آنها ساکن بود. مثلا از هرقدیلر در استان ایلام ۱۲ و در استان یزد ۲۲ آبادیهائی که در پاسخ سوال ۸ـ۸ با کد ۲ مشخص شده بودند، دارای تعداد زیادی از کارگاههای مورد نظر طرح از جمله بقالی و گروههای تعاونی مصرف بودند و در

عین حال خانوار و جمعیت‌خانوار نیز در آنها ساکن بود . به عبارت‌دیگر این بررسی نشان داد که به تفکیک‌دهستان اندیکاتورهای و ماشینی اطلاعات‌شرح ۳ ، دسته ای از آبادیهای قرارگرفته درچارچوب‌فاقد خصوصیت‌جامعه آماری مربوط هستند، یعنی به جوامع آماری دیگر تعلق دارند . و دسته مهمی از آبادیها که در چارچوب‌قرار نگرفته اند، دارای شرط مرکز خرید روستائی بودن هستند و باید به چارچوب‌اضافه شوند .

بر اساس‌نتایج حاصل از بررسی روی جداول ماشینی دو استان ایلام و یزد تصمیم گرفته شد که چارچوب‌آماری متشکل از آبادیهائی باشد که :

۱ـ مسکونی دائمی و دارای خانوار معمولی ساکن باشد .

۲ـ درسوال ۱ـ۸ کد ۱ علامت‌خورده باشد .

۳ـ اگردر سوال ۱ـ۸ کد ۲ علامت‌خورده باشد آبادی دارای هردوکارگاه یقه‌آبی و شرکت‌خدماتی معرفی‌باشد .

در بررسی بخش‌آمده ازدو استان ایلام و یزد معلوم گردید که آبادیها به جای نقاط شهری به عنوان مراکز خرید شهری از آبادیهای هم جوار نام برده اند. یعنی این بخش ازبررسی نیز نشان داد که چارچوب‌نهائی شهری طرح نیزخالی از اشکال نیست و باید برای اصلاح و کامل شدن این دفعتاکافی مبذول شود و از اطلاع بدست‌آمده برای مرکز شهری که به اشتباه روستاها را معرفی کرده بودند برای اصلاح و تکمیل چارچوب‌روستائی مراکز خرید روستائیان بهره گیری شد .

در مرحله دوم ، شهرستان‌شناسی اطلاعات‌مورد نظر از آبادیهای مسکونی دائمی دریافت‌شد وبا استفاده از فرم ۳ سرشماری عمومی نفوس و مسکن ۱۳۶۵ برای حدود ۵۵۰۰ آبادی مسکونی دائمی نام مراکزخرید شهری روستائیان مشخص‌شد سپس‌با استفاده از این چارچوب‌درسطح دهستان مبادرت‌به تعیین نام شهر و قراوانی آبادیهائی گردید که اهالی آنها جهت‌تامین اقلام‌مورد به این شهرها مراجعه داشتند و جداول در سطح ۱۵۴۴ دهستان برحسب‌اولویت‌اول و دوم بدست‌آمد . در مرحله سوم از جداول تهیه شده برای مراکز خریدشهری هر دهستان خلاصه‌ای تهیه شد و جدول توزیع توأمان قراوانی مراکزخرید شهری درسطح شهرستانهای هر استان که به تفکیک‌اولویت‌اول و دوم خرید روستائیان قرار دارند، تهیه و تنظیم شد و نتایج در ۲۴ جدول استانی که در برگیرنده نام مراکز خرید شهری ۱۹۲ شهرستان (معنای ۳ شهرستان خرمشهر ، آبادان ، قصر شیرین که از مناطق جنگی بوده اند) میشود، تهیه و ارائه شد و چارچوب‌آماری مراکز خرید شهری روستائیان پایان یافت .

در این مرحله اسامی ۱۴۴۹ مرکزخرید دیگرهم که بوسیله معتمدین آبادیها معرفی شده بود که می بایست وضعیت‌آنها نسبت به چارچوب‌شهری و یا روستائی مشخص می گردید . این کار با همکاری گروه پرداری کافی ، که لازمه تهیه هرچارچوب‌خوب‌نمونه صورت‌گرفتنی در اجرای آن بنابه ضرورت‌از امکانات و مدارک‌زیر استفاده بعمل آمد .

الف ـ نشریات‌مربوط به فایل فرم ۱۴۲ سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ ، شامل فهرست‌اسامی ، کد اقتصادی و خانوار کلیه آبادیهای کشوردرسطح دهستان ، بخش ، شهرستان و استان .

ب ـ نشریات‌مربوط به فایل حوزه بندی آبادیها ، شامل فهرست‌اسامی ، کد اقتصادی و شماره حوزه کلیه آبادیهای کشور که پس ازسرشماری عمومی نفوس و مسکن سال ۱۳۶۵، اطلاعات‌ناشی از اجرای سرشماری درکدام‌بخش آمده است .

ج ـ فهرست‌شناسی کلیه آبادیهای سرشماری شده درسال ۱۳۶۵ (فایل کامل فرم ۳ سرشماری) ، شامل آبادیهای مسکونی دائمی (کد ۱) ، آبادیهای مسکونی موسمی (کد ۲) ، آبادیهای خالی ازسکنه موسمی (کد ۳) ، آبادیهای خالی ازسکنه دائمی (کد ۴) .

۴ـ فهرست جداول ماشینی اطلاعات فرم ۳ کلیه آبادیهائی که سرشماری عمومی نفوس و مسکن سال ۱۳۶۵، مسکونی دائمی معرفی شده بودند (فهرستی که چارچوب نهائی مراکز خریدشهری روستائیان را دربردارد) .

۵ـ فهرست جداول ماشینی اطلاعات فرم ۳ کلیه آبادیهائی که بافواصل تعیین شده بعنوان مراکز خرید روستائی روستائیان مشخص شده بودند (چارچوب نهائی زیرجامعه روستائی طرح) .

۶ـ نظریه تقسیمات کشوری جمهوری اسلامی ایران درسال ۱۳۶۴ ، با اعمال تغییرات مربوط به سال ۱۳۶۵ .

۷ـ فهرست ماشینی اسامی و مشخصات کلیه شهرهای کشور ، درسطح شهرستان و استان .

۸ـ فهرست اسامی و مشخصات آبادیهائی که درسرشماری عمومی نفوس و مسکن سال ۱۳۶۵، به صورت شهری حوزه بندی و سرشماری شده اند (روستا ـ شهرها) .

۹ـ فهرست ماشینی کلیه آبادیهائی که درفرم ۳ سرشماری ، تعداد خانوار آنها صفر ولی نوع سکونت در آبادی ، مسکونی وبا گزارش نشده است .

۱۰ـ فرم ۳ تکمیل شده درسرشماری عمومی نفوس و مسکن سال ۱۳۶۵، کلیه آبادیهای مورد اشکال ونیاز .

۱۱ـ نشریات فرهنگ آبادیهای استانهای کشور ، مربوط به سرشماری عمومی نفوس و مسکن سال ۱۳۵۵ .

۱۲ـ نقشه های روستایی کشور ، درتقسیمات کشوری سال ۱۳۵۶ و سال ۱۳۶۵ .

۱۳ـ نشریات فرهنگ آبادیهای کشور ، مربوط به آمارگیری سال ۱۳۶۰ جهادسازندگی .

۱۴ـ نشریات حاصل از استخراج ماشینی نتایج اجرای سرشماری کشاورزی آبادیهای کشور درسال ۱۳۵۳ .

درنهایت، تحقیق برروی ۱۶۶۹ اسامی معرفی شده بوسیله معتمدین محلی آبادیهای مسکونی دائمی کشوربه شرح زیربه نتیجه رسید :

۱ـ تعداد ۱۲۰۶ نام (۷۲/۳ x) مربوط به اسامی آبادیهائی بود که درچارچوب نهائی مراکزخرید روستائی طرح قرار داشتند .

۲ـ تعداد ۱۲ نام (۱/۰ x) مربوط به اسامی شهرهائی بود که اشتباها - از آن نام برده شده بود .

۳ـ تعداد ۱۷ نام (۱/۰ x) مربوط به آبادیهائی بودند که مسکونی موسمی گزارش شده بودند که به علت غیرمستمر بودن کارگاهها نمی توانستند ازواحدهای جامعه آماری طرح باشند .

۴ـ تعداد ۷۴ نام (۴/۴ x) از اسامی معرفی شده درفهرست اول که چارچوب نهائی مراکزخرید شهری روستائیان را در برداشت قرار داشتند ولی درهیچ یک از آنها حتی یکنفر نانوائی یا فروشگاه تعاونی مصرف هم وجود نداشت بنابراین به عنوان مرکزخرید نمی توانستند محسوب شوند .

۵ـ تعداد ۸۶ نام (۵/۲ x) از اسامی باتمامی کنکاش و مراجعه به مدارک و اسناد شناسائی نشده وچنین نتیجه گیری شد که این اسامی نامهای جدیدی استکه بعد از انقلاب اسلامی ایران ، آبادیها به آنها نامیده شده اند یا اسامی دوم آبادیها می باشد .

۶ـ تعداد ۲۴۵ نام (۱۴/۷ x) از اسامی درچارچوب نهائی قرارنداشتند ولی درفهرست اول موجود بودند یعنی فرد داشتن بقالی و تعاونی مصرف شوا - ر ا نداشتند ولی با وجود این به فهرست نهائی اضافه گردیدند .

۷- تعداد ۲٤ نام (۲٤/۱ ٪) از نامی معرفی شده ، آبادیهائی بودند که در فهرست اول و دوم وجود نداشتند .
تعداد ۱۳ تا از این آبادیها روستا - شهر به حساب آمده بودند بنابر این در فهرست اول و دوم نام آنها وجود نداشتو ۱۱ تای بقیه تواما - در ای یکانی و تعاونی نبودند . بنابر این این ۲٤ آبادی نیز به فهرست چارچوب اضافه گردید .

به این ترتیب از ۱۶۲۹ آبادی که به اشتباه به جای مرکز شهر از آنها نام برده شده بود تعداد ۱۲۲۵ آبادی یعنی ۷۵/٤٪ به فهرست چارچوب روستائی اضافه شد و فقط ۲۴ آبادی حذف گردید .

بعد از بررسی بعمل آمده از کل ۳۳۹ روستا شهر کشور ۲۱۷ آبادی در فهرست جدید و مطابق با فهرست اول قرار داشتند ولی ۲۲ روستا شهر و ۱۲ آبادی بزرگ تغییر در فهرست محاسبه شده بودند که با بررسی سوابق اضافه به فهرست اضافه گردیدند .

به این ترتیب آبادیهائی که به علت فقدان اطلاع در فهرستها انعکاس نیافته بودند کشف گردیدند و پس از تکمیل اطلاعات آنان به فهرستها اضافه شدند و ضمن حذف تعدادی روستا شهرها نیز به واحدهای جامعه آماری اضافه گردیدند و در چارچوب گنجانده شدند به این ترتیب کار تکمیل چارچوب آماری روستائی طرح به پایان رسید و تعداد واحدهای آماری آن به ۱۰۷٤٤ آبادی بالغ گردید .

آمارگیری جاری جمعیت سال ۱۳۷۰ و استخراج ماشینی

عبدالحسین صدقیانی

مرکز آمار ایران

هدف اصلی از سرشماریها وآمارگیریهای نمونه ای جمع آوری اطلاعات موردنیاز جهت برنامه ریزیهای درزمینه های مختلف ازقبیل آموزشی ،اقتصادی ،بهداشتی ،کشاورزی وبطورکلی اقتصادی ـ اجتماعی می باشد تمام تلاش مسئولان ودست اندرکاران سرشماریها وطرحهای آماری این است که صحیح ترین ودقیقترین اطلاع رازجامعه وبصحت یاصحتهای موردنظر بدست آورند . بدین منظور درکلیه مراحل ، اعم از مرحله تهیه طرح ، مرحله عملیات میدانی ومرحله استخراج ، تدوین وانتشار نتایج سعی می شود که از تجارب ونظرات وترفندهایی که ممکن است میزان خطاهای غیرنمونه ای راکاهش دهند استفاده شود .

بطورکلی خطاهای غیرنمونه ای درسرشماریها وآمارگیریهای نمونه ای رابه ۳ دسته اصلی می توان تقسیم نمود .

۱ـ خطاهای غیرنمونه ای مرحله تهیه طرح شامل :

۱ـ۱ تعاریف ومفاهیم نامناسب

۱ـ۲ انتخاب روشهای نامناسب

۱ـ۳ انتخاب ابزار نامناسب

۱ـ۴ پرسشنامه های نامناسب

۱ـ۵ دستورالعمل نامناسب

۱ـ۶ چارچوب نامناسب

۱ـ۷ نقشه نامناسب

۲ـ خطاهای غیرنمونه ای مرحله عملیات میدانی شامل :

۲ـ۱ پرسش نامناسب

۲ـ۲ پرسشنامه های بدون پاسخ

۲ـ۳ پاسخهای اشتباه

۳ـ خطاهای مرحله استخراج ، تدوین وانتشار شامل :

۳ـ۱ خطای بازبینی ؟

۳ـ۲ خطای کدگذاری

۳ـ۳ خطای پانچ

۳ـ۴ خطای برنامه نویسی

۳ـ۵ خطای تدوین وانتشار

ازمیان ۳ دسته خطای نگرشده به دسته سوم یعنی خطاهای مرحله استخراج ، تدوین وانتشار وبه طور اخص به خطای کدگذاری وپانچ می پردازیم .

احتمال پیشامدهای فازی

سید محمود طاهری*- ناصررضا ارقامی

دانشگاه فردوسی مشهد

اندازه های احتمال پیشامدهای فازی[1]

نظریه احتمال و نظریه مجموعه های فازی، هر دو برای مطالعه موارد شامل عدم قطعیت و اطمینان وضع شده‌اند. اولی به عنوان الگوی عدم قطعیت آماری (منسوب به پیشامدهای تصادفی) و دومی به عنوان الگوی عدم قطعیت ناشی شده از تشخیص، عقیده و قضاوت انسانی. این دو رهیافت به عدم قطعیت نه متناقض با یکدیگرند و نه یکی، دیگری را شامل می شود. اساساً هیچ چیز برای تطبیق یا وفق این دو رهیافت وجود ندارد، چرا که طبیعت هر کدام، از دیگری کاملاً متفاوت است[۲].

اما این باعث نمیشود که نتوان در یک مسأله، از هر دو نظریه توأمان استفاده کرد. در واقع، روشهای نظریه احتمال و نظریه مجموعه های فازی میتوانندبرای مطالعه عدم قطعیت ودقت بهتر کیب شوند. گام اول در این راه توسط پروفسور زاده در ۱۹۶۸ برداشته شد[۶] وتا کنون، مطالعات در این مورد توسعه زیادی یافته است.

این مقاله شامل ارائه و تشریح رهیافت زاده در این مسأله و نیز تعمیم ها و تصحیح هایی است که دیگران در ادامه کار وی انجام داده اند.

۱ - تعاریف و مفاهیم

فرض کنید X فضایی از نقاط (اشیاء) باشد. یک زیرمجموعهٔ فازی A از X، توسط یک تابع عضویت، $\mu_A(x)$، مشخص می شود. که به هر نقطه در X یک عدد حقیقی از بازهٔ [۰ و ۱] نسبت می دهد. مقدار $\mu_A(x)$، درجهٔ عضویت x در A نامیده می شود. بنابراین نزدیکی مقدار $\mu_A(x)$ به عدد یک، به

1-Probability Measures of Fuzzy Events

معنی بالا بردن درجهٔ عضویت x در A است و بالعکس، وقتی A یک زیرمجموعهٔ معمولی فرض شود، آنگاه $\mu_A(x)$ بر طبق آنکه x در A عضو باشد یا نباشد، تنها دو مقدار یک و صفر را اختیار می‌کند.

بنابراین مجموعه‌های معمولی حالت خاصی از مجموعه‌های فازی هستند.

تعریف ۱-۱- مجموعه نقاطی از x که برای آن نقاط $\mu_A(x)>0$، را تکیه‌گاه[2] A گوییم.

تعریف ۱-۲- مقدار $M=\sup_x \mu_A(x)$ را ارتفاع[3] مجموعهٔ A یا درجهٔ ماکزیمال[4] در A گوییم.

تعریف ۱-۳- اگر x نقطه‌ای باشد که برای آن $\mu_A(x)=0/5$ آنگاه x را یک عضو A [5] گوییم.

مثال ۱-۱- فرض کنید $X=[0,100]$ و $x\in X$ به عنوان سن تلقی شود. یک زیر مجموعهٔ فازی X که پیری را نشان دهد میتواند اینگونه تعریف شود:

$$\mu_A(x)=\begin{cases}0 & \text{اگر } 0\le x\le 50\\ \left[1+\left(\frac{x-50}{5}\right)^{-2}\right]^{-1} & \text{اگر } 50\le x\le 100\end{cases}$$

مثال ۱-۲- فرض کنید $X=\{1,2,3,4,5\}$، یک زیر مجموعهٔ فازی از X مانند A که "کوچک بودن" را نشان دهد می‌تواند اینگونه تعریف شود:

$$\mu_A(x)=\begin{cases}1 & x=1\\ 0/6 & x=2\\ 0/3 & x=3\\ 0/1 & x=4\end{cases}$$

دقت کنید که در اینجا تکیه‌گاه A مجموعهٔ $\{4,3,2,1\}$ بوده و ارتفاع آن واحد است و هیچ نقطه‌ای یک عضو برای A نمی‌باشد.

نماد گذاری - برای سهولت نمادهای زیر را به کار می‌بریم [۲]، یک مجموعهٔ متناهی غیر فازی مانند $X=\{x_1,\dots,x_n\}$ را می‌توان اینگونه بیان کرد:

$$X=x_1+\dots+x_n \qquad (1-1)$$

2-Support
3-Height
4-Maximal Grade
5-Crossover Point

ویا (۲-۱)
$$X = \sum_{i=1}^{n} x_i$$

که در آن منظور از علامت + ، اجتماع است نه جمع حسابی. به عنوان یک تعمیم، زیر مجموعهٔ فازی A از X را نیز می توان اینگونه بیان کرد:

$$A = \mu_1 x_1 + \cdots + \mu_n x_n \qquad (۳-۱)$$

ویا (۴-۱)
$$A = \sum_{i=1}^{n} \mu_i x_i$$

که در آن μ_i ها، درجات عضویت x_i ها در A هستند. اگر خود x_i ها نیز عدد باشند، برای جلوگیری از ابهام می نویسیم:

$$A = \mu_1 / x_1 + \cdots + \mu_n / x_n \qquad (۵-۱)$$

ویا (۶-۱)
$$A = \sum_{i=1}^{n} \mu_i / x_i$$

وهنگامیکه تکیه گاه یک مجموعهٔ فازی پیوسته باشد می نویسیم:

$$A = \int_X \mu_A(x) / x \qquad (۷-۱)$$

با توجه به اینکه در اینجا منظور از انتگرال، اجتماع عناصر منفرد $\mu_A(x)/x$ ، $x \in X$ است.

مثال ۱-۳- فرض کنید $X=\{a,b,c,d\}$ ویا به طور معادل $X = a + b + c + d$ در این حالت یک زیر مجموعهٔ فازی A را به صورت زیر نمایش می دهیم:

$$A = 0/3\,a + b + 0/9\,c + 0/5\,d$$

واگر $X = 1 + 2 + \cdots + 100$

می نویسیم:
$$A = 0/3/25 + 0/9/30$$

مثال ۱-۴- برای مثال (۱-۱) مجموعهٔ فازی A بصورت زیر نوشته می شود:

$$A = \int_{50}^{100} \left[1 + \left(\frac{x-50}{5}\right)^{-2}\right]^{-1} / x$$

تعریف ۱-۴- مجموعهٔ فازی A را نرمال[6] گوییم اگر ارتفاع آن واحد باشد. یعنی اگر $\sup_x \mu_A(x) = 1$ در حالات دیگر A را زیرنرمال[7] گوییم.

متذکر می شویم که هر مجموعهٔ زیر نرمال A را می توان با تقسیم μ_A بر $\sup_x \mu_A(x)$ نرمال کرد.

مثال ۱-۵- مجموعهٔ فازی A در مثال (۱-۲) نرمال است ولی زیرمجموعهٔ B آنقدر خیلی کوچک و نه خیلی بزرگ، از $X = 1 + 2 + \cdots + 100$ به صورت زیر نرمال است.

$$B = 21/2 + 0/3/3 + 0/5/4 + 0/8/5 + 0/8/6 + 0/5/7 + 0/3/8 + 0/1/9$$

6-Normal
7-Subnormal

تذکر - در حالاتی که تکیه گاه یک مجموعهٔ فازی گسسته است از نماد دیگری نیز برای نشان دادن یک مجموعهٔ فازی استفاده می کنند. برای مثال مجموعهٔ B مثال فوق را اینگونه می نویسند :

$$B = \left\{ \frac{0.1}{9} , \frac{0.3}{8} , \frac{0.5}{7} , \frac{0.8}{6} , \frac{0.8}{5} , \frac{0.5}{4} , \frac{0.3}{3} , \frac{0.1}{2} \right\}$$

ما از این به بعد در بعضی موارد از این نوع نماد گذاری استفاده خواهیم کرد. همچنین در بعضی موارد و برای اختصار به جای $\mu_A(x)$ خواهیم نوشت $A(x)$.

تعریف ۱-۵- برای مجموعه های فازی بجای بحث از تعداد عناصر آن مفهوم توان یک مجموعهٔ فازی به صورت زیر تعریف میشود :

$$|A| = \sum_i \mu_A(x_i) \qquad (۸-۱)$$

که یک تعمیم طبیعی برای تعداد عناصر است .

مثال ۱-۲- فرض کنید X مجموعهٔ ساکنان یک شهر و A مجموعهٔ فازی انسانهای بیکار در آن شهر باشد . اگر $\mu_A(x)$ بیانگر درجهٔ عضویت شخص x در زمرهٔ بیکارها باشد (مثلاً $\mu_A(x)=0/5$ اگر x نیمه وقت کار کند .)، آنگاه $|A|$ را می توان تعداد بیکاران تمام وقت تعبیر کرد.

تعریف ۱-۶- یک مجموعهٔ فازی را تهی گوییم اگر و فقط اگر تابع عضویت آن برای هر $x \in X$ صفر باشد .

تعریف ۱-۷- دو مجموعهٔ فازی A و B را مساوی گوییم، $A = B$، اگر و فقط اگر به ازای هر $x \in X$:

$$\mu_A(x) = \mu_B(x) \quad (\mu_A = \mu_B) .$$

تعریف ۱-۸- مکمل یک مجموعهٔ فازی A یک مجموعهٔ فازی A' است که برای آن

$$\mu_{A'}(x) = 1 - \mu_A(x) .$$

تعریف ۱-۹- گوییم A زیر مجموعهٔ B است یا B شامل A است، اگر و فقط اگر به ازای هر $x \in X$ ،

$$\mu_A(x) \leqslant \mu_B(x) .$$

تعریف ۱-۱۰- اجتماع دو مجموعهٔ فازی با توابع عضویت $\mu_A(x)$ و $\mu_B(x)$، یک مجموعهٔ فازی C است که برای آن

$$\mu_C(x) = \max \left[\mu_A(x) , \mu_B(x) \right] \qquad (۹-۱)$$

و یا به زبان ساده تر

$$\mu_C = \mu_A \vee \mu_B \qquad (۱۰-۱)$$

تعریف ۱-۱۱- اشتراک دو مجموعهٔ فازی A و B ، با توابع عضویت $\mu_A(x)$ و $\mu_B(x)$، یک مجموعهٔ فازی C است که برای آن:

$$\mu_C(x) = \min \left[\mu_A(x) , \mu_B(x) \right] \qquad (۱۱-۱)$$

و یا به بیان ساده تر

$$\mu_C = \mu_A \wedge \mu_B \qquad (۱-۱۲)$$

یک راه استیضاف شهودی تعریف های اجتماع و اشتراک این است :

اجتماع A و B عبارت است از کوچکترین مجموعهٔ فازی که شامل A و B باشد ، و اشتراک آن دو عبارت است از بزرگترین مجموعهٔ فازی که هم در A و هم در B شامل شود . بررسی معادل بودن این دو تعریف با تعاریف فوق بسیار ساده است برای حالت اجتماع گوییم :

از رابطهٔ (۱-۹) داریم:

$$\mu_C = \max[\mu_A, \mu_B] \geq \mu_A, \mu_B$$

حال اگر D یک مجموعهٔ فازی باشد که شامل هر دوی A و B باشد آنگاه

$$\mu_A \leq \mu_D \quad , \mu_B \leq \mu_D$$

پس

$$\mu_D \geq \max[\mu_A, \mu_B] = \mu_C$$

و بنا به تعریف

$$C \subset D$$

تعریف (۱-۱۲)- مانند حالت معمولی : A و B را جدا از هم گوییم اگر $A \cap B$ تهی باشد .

بنا به تعریف (۱-۱) ، A و B جدا از هم خواهند بود اگر و فقط اگر ، اشتراک تکیه گاههای آنها تهی باشد .

مثال (۱-۷) فرض کنید

$$X = \{ ۱۰, ۲۰, ۳۰, \ldots, ۱۰۰ \}$$

و زیر مجموعهٔ فازی A از X نشان دهندهٔ ویژگی «کوچک» بصورت زیر باشد :

$$A = \left\{ \frac{۱}{۱}, \frac{۰/۸}{۲}, \frac{۰/۶}{۳}, \frac{۰/۳}{۴} \right\}$$

و زیر مجموعهٔ فازی B از X نشان دهندهٔ ویژگی «نزدیک به ۵» بصورت زیر باشد :

$$B = \left\{ \frac{۰/۴}{۳}, \frac{۰/۸}{۴}, \frac{۱}{۵}, \frac{۰/۸}{۶}, \frac{۰/۴}{۷} \right\}$$

و C زیر مجموعهٔ فازی نشان دهندهٔ ویژگی «خیلی کوچک» بصورت زیر باشد :

$$C = \left\{ \frac{۱}{۱}, \frac{۰/۶}{۲}, \frac{۰/۴}{۳}, \frac{۰/۱}{۴} \right\}$$

آنگاه داریم:

$$A' = \left\{ \frac{۰/۲}{۲}, \frac{۰/۴}{۳}, \frac{۰/۷}{۴}, \frac{۱}{۵}, \frac{۱}{۶}, \frac{۱}{۷}, \frac{۱}{۸}, \frac{۱}{۹}, \frac{۱}{۱۰} \right\}$$

$$A \cup B = \left\{ \frac{1}{1}, \frac{۷۸}{۲}, \frac{۷۱}{۳}, \frac{۷۸}{۴}, \frac{1}{۵}, \frac{۷۸}{۶}, \frac{۷۴}{۷} \right\}$$ و

$$A \cap B = \left\{ \frac{۷۳}{۳}, \frac{۷۴}{۴} \right\}$$

و

$$|A| = ۲,۷ \quad , \quad |B| = ۳,۴ \quad , \quad |C| = ۲,۱ \quad , \quad C \subset A$$

مجموعه

<u>بعضی خواص مربوط به $\cup$ و $\cap$ و مکمل</u>

۱) خاصیت جابجایی $A \cup B = B \cup A$ ، $A \cap B = B \cap A$

۲) خاصیت شرکت پذیری $A \cup (B \cup C) = (A \cup B) \cup C$ ؛ $A \cap (B \cap C) = (A \cap B) \cap C$

۳) خاصیت توزیع پذیری $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ ، $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

۴) قوانین دمورگان

الف) $(A \cup B)' = A' \cap B'$

ب) $(A \cap B)' = A' \cup B'$

تعریف ۱-۱۳- اگر A یک زیر مجموعهٔ فازی از X باشد، آنگاه یک مجموعهٔ تراز[۵] α یا A_α از A، یک مجموعهٔ غیر فازی به صورت زیر است.

$$A_\alpha = \{ x \mid \mu_A(x) \geq \alpha \} \qquad (۱-۱۳)$$

یک مجموعهٔ فازی A را میتوان به مجموعه های تراز آن تجزیه کرد، یعنی:

$$A = \int_0^1 \alpha A_\alpha \qquad (۱-۱۴)$$

یا

$$A = \sum_\alpha \alpha A_\alpha \qquad (۱-۱۵)$$

که در آن $\int$ و $\sum$ اجتماع A_α را نشان می دهد.

گزاره ۱-۱۶- اگر A و B دو مجموعهٔ فازی دلخواه باشند داریم

$$(A \cup B)_\alpha = A_\alpha \cup B_\alpha \qquad (۱-۱۶)$$

$$(A \cap B)_\alpha = A_\alpha \cap B_\alpha \qquad (۱-۱۷)$$

برای مثال، در مورد (۱-۱۶) داریم

$$(A \cup B)_\alpha = \{ x \mid \mu_{A \cup B}(x) \geq \alpha \} = \{ x \mid \max [\mu_A(x), \mu_B(x)] \geq \alpha \}$$

۵-Level Set

$$= \left\{ x \mid \mu_A(x) \geq \alpha \text{ یا } \mu_B(x) \geq \alpha \right\} = \left\{ x \mid \mu_A(x) \geq \alpha \right\} \cup \left\{ x \mid \mu_B(x) \geq \alpha \right\}$$

$$= A_\alpha \cup B_\alpha$$

مثال ۱-۱۳- اگر

$$X = \left\{ 1, 2, \ldots, 10 \right\}$$

و

$$A = \left\{ \frac{0.1}{2}, \frac{0.3}{1}, \frac{0.5}{7}, \frac{0.9}{6}, \frac{1}{9} \right\}$$

با استفاده از (۱-۱۳) داریم :

$$A = 0.1 \left\{ \frac{1}{2}, \frac{1}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{9} \right\} + 0.3 \left\{ \frac{1}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{9} \right\} + 0.5 \left\{ \frac{1}{7}, \frac{1}{6}, \frac{1}{9} \right\} + 0.9 \left\{ \frac{1}{6}, \frac{1}{9} \right\} + 1 \left\{ \frac{1}{9} \right\}$$

که به فرم (۱-۱۵) است با مجموعه های تراز زیر :

$$A_{0.1} = \left\{ 2, 1, 7, 6, 9 \right\}$$

$$A_{0.3} = \left\{ 1, 7, 6, 9 \right\}$$

$$A_{0.5} = \left\{ 7, 6, 9 \right\}$$

$$A_{0.9} = \left\{ 6, 9 \right\}$$

$$A_1 = \left\{ 9 \right\}$$

تعریف ۱-۱۴- مجموعهٔ فازی A را محدب گوییم اگر و فقط اگر هر مجموعهٔ تراز A_α برای $\alpha > 0$ محدب باشد .

۲ـ اندازه های احتمال پیشامدهای فازی (رهیافت زاده)

اندازه های احتمال معمولی و لزوم تعمیم ـ در نظریه احتمال ، یک پیشامد A ، یک عضو از میدان سیگمایی $\mathcal{F}$ حاصل از زیر مجموعه های فضای نمونه Ω است ، و یک اندازه احتمال P ، یک اندازه تعریف شده بر یک فضای اندازه پذیر $(\Omega$ و $\mathcal{F})$ است ، یعنی یک تابع حقیقی مقدار که به هر $A \in \mathcal{F}$ یک احتمال $P(A)$ نسبت می دهد به قسمی که :

$$\begin{aligned} &1)\ P(A) \geq 0 \\ &2)\ P(\Omega) = 1 \\ &3)\ P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \end{aligned} \qquad (۱ـ۲)$$

(A_iها مجدا از هم)

مفهوم یک پیشامد و احتمال متناظر با آن از اساسی ترین مفاهیم در نظریه احتمال اند . در آنجا یک پیشامد ، مجموعه ای معمولی از نقاط فضای نمونه است واحتمال متناظر با آن نیز با جمع(انتگرال)مقادیر احتمالات متناظر با اعضای مجموعهٔ منظور حاصل میشود . مثلاً برای مجموعهٔ A

$$P(A) = \int_A dP \qquad (۲ـ۲)$$

دقت کنید که با در نظر گرفتن تابع نشانگر مجموعهٔ A ، یعنی

$$\mu_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \qquad (۳ـ۲)$$

میتوان (۲ـ۲) را اینگونه بیان کرد

$$P(A) = \int_{\Omega} \mu_A(x)\, dP = E(\mu_A) \qquad (۴ـ۲)$$

معادلهٔ (۲ـ۴) رابطهٔ بین احتمال یک پیشامد و امید ریاضی تابع نشانگر آن پیشامد را بیان می کند زاده [۹] همین رابطه را بسادگی برای مجموعه های فازی تعمیم داده است.

بیان وتشریح رهیافت زاده ـ گفته شد که در نظریهٔ احتمال ، یک پیشامد مجموعه ای از نقاط فضای نمونه است. با وجود این ما در گفتار روزمره گزاره هایی را پیشامد می نامیم که دقیقا مجموعه ای از نقاط نیستند . مثلا پیشامد اینکه امروز هوا گرم است یا تقریبا برابر ۵ است یا در ۲۰ بار پرتاب این سکه ، معمولا دفعات شیر بیشتر از خط است و مانند اینها .

این گزاره ها ، گزاره هایی مبهم(فازی) هستند بااستفاده از مفاهیم مجموعه های فازی ،می توان مفاهیم یک پیشامد واحتمال مربوطه را بطور طبیعی تعمیم داد باساس این تعمیم جایجایی تابع نشانگر به کار برده شده در (۲ـ۴) با تابع عضویت یک مجموعه فازی است .

برای سادگی فرض کنید که Ω فضای n بعدی اقلیدسی R^n و $\mathcal{F}$ میدان سیگمایی حاصل از مجموعه های

(غیرفازی) بورل در $\mathbb{R}^n$ و P یک اندازه احتمال بر $\mathbb{R}^n$ باشند. پس فضای احتمال در اینجا $(\mathbb{R}^n, \Pi, P)$ است. یک نقطه در $\mathbb{R}^n$ را اختصاراً به x نشان خواهیم داد.

تعریف ۲-۱- فضای احتمال $(\mathbb{R}^n, \Pi, P)$ را در نظر بگیرید. یک پیشامد فازی g در $\mathbb{R}^n$ عبارت است از یک زیر مجموعهٔ فازی μ_A از $\mathbb{R}^n$ که تابع عضویت آن، $\mu_A : \mathbb{R}^n \to [0,1]$، یک تابع اندازه پذیر بورل باشد.

تعریف ۲-۲- احتمال یک پیشامد فازی توسط انتگرال لبگ-استیلتیس زیر تعریف می شود.

$$P(A) = \int_{\mathbb{R}^n} \mu_A(x)\, dP = E(\mu_A) \qquad (۲-۵)$$

بنابراین همانند (۲-۴) احتمال یک پیشامد فازی برابر با امید ریاضی تابع تشانگر آن است. وجود انتگرال فوق نیز با فرض اندازه پذیر بورل بودن μ_A تضمین شده است.

تعاریف زیادی از نظریه مجموعه های فازی برای بحث فعلی لازم است، در اینجا بطور خلاصه به آنها اشاره می کنیم. برای هر x داشتیم

$A \subset B$	$\Leftrightarrow$	$\mu_A(x) \leq \mu_B(x)$ (۱)
$A = B$		$\mu_A(x) = \mu_B(x)$ (۲)
$A^c =$ متمم A		$\mu_{A^c}(x) = 1 - \mu_A(x)$ (۳)
$A \cup B =$ اجتماع A و B		$\mu_{A\cup B}(x) = \max[\mu_A(x), \mu_B(x)]$ (۴)
$A \cap B =$ اشتراک A و B		$\mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)]$ (۵)
$AB =$ ضرب A در B		$\mu_{AB}(x) = \mu_A(x) \cdot \mu_B(x)$ (۶)
$A \oplus B =$ جمع A و B		$\mu_{A\oplus B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\cdot\mu_B(x)$ (۷)

حال برای نتیجه گرفتن بعضی روابط مقدماتی از (۱) تا (۷) آماده ایم. ابتدا به عنوان یک نتیجه فوری از (۱) و (۲-۵) داریم :

$$A \subset B \Rightarrow P(A) \leq P(B) \qquad (۲-۶)$$

و بطور مشابه از (۴) تا (۷) داریم

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (۲-۷)$$

$$P(A \oplus B) = P(A) + P(B) - P(AB) \qquad (۲-۸)$$

با استفاده از استقراء از (۲-۷) و (۲-۸) روابط زیر را، بطور مشابه با حالت مجموعه های معمولی، برای مجموعه های فازی داریم

9-Fuzzy Event

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \cdots + (-1)^n P\left(\bigcap_i A_i\right) \qquad (۹\text{-}۲)$$

$$P(A_1 \oplus \cdots \oplus A_n) = \sum_i P(A_i) - \sum_{i<j} P(A_i A_j) + \cdots + (-1)^n P(A_1 \cdots A_n) \qquad (۱۰\text{-}۲)$$

همچنین (۷-۲) و (۸-۲) نامساوی بول را برای مجموعه های فازی نتیجه می دهند .

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} P(A_i) \qquad (۱۱\text{-}۲)$$

$$P(A_1 \oplus A_2 \oplus \cdots) \leqslant \sum_{i=1}^{\infty} P(A_i) \qquad (۱۲\text{-}۲)$$

تعریف ۲-۳. گیریم A و B دو پیشامد فازی در فضای احتمال $(\mathbb{R}^n, \mathcal{B}, P)$ باشند A و B را مستقل گوییم اگر

$$P(AB) = P(A) \cdot P(B) \qquad (۱۳\text{-}۲)$$

توجه می کنید که در تعریف استقلال عبارت AB را بجای $A \cap B$ به کار برده ایم. گزاره ۲-۱- فرض کنید $\Omega_1 = \mathbb{R}^{n_1}$ و $\Omega_2 = \mathbb{R}^{n_2}$ و P_1 و P_2 اندازه احتمالهای دلخواه به ترتیب بر $\mathcal{B}_1 \times \mathcal{B}_2$ باشد که P_1 و P_2 اندازه های احتمال بر Ω_1 و Ω_2 و نیز A_1 و A_2 دو پیشامد فازی در $\Omega_1 \times \Omega_2$ باتوابع عضویت زیر باشند $\mu_{A_1}(x_1) = \mu_{A_1}(x_1, x_2) d P_2(x_2)$ و $\mu_{A_2}(x_2) = \int \mu_{A_2}(x_1, x_2) d P_1(x_1)$ آنگاه A_1 و A_2 به بیان (۱۳-۲) مستقل اند [۷]. توجه کنید که دقتی در تعریف استقلال AB را نیز بکار ببریم این حکم برقرار نیست .

تعریف ۲-۴. احتمال شرطی A به شرط B طبق رابطه زیر تعریف می شود .

$$P(A|B) = \frac{P(AB)}{P(B)} \qquad \text{به شرط آنکه} \quad P(B) > 0 \qquad (۱۴\text{-}۲)$$

مانند حالت معمولی اگر A و B مستقل باشند داریم .

$$P(A|B) = P(A) \qquad (۱۵\text{-}۲)$$

بسیاری از مفاهیم نظریه احتمال مانند میانگین ، واریانس و ... برحسب توابعی از توزیع های احتمال بیان می شوند . در اینجا مفهوم یک پیشامد فازی اقتضا می کند که مفاهیم مربوطه برحسب یک پیشامد فازی و نیز یک اندازه احتمال بیان شوند . برای مثال مفاهیم میانگین و واریانس یک پیشامد فازی A با اندازه احتمال P بر $\mathbb{R}^n$ را می توان به صورت های زیر تعریف کرد (در حالت معمولی این مفاهیم چندان شناخته شده نیستند) .

$$E_P(A) = \frac{1}{P(A)} \int_{\mathbb{R}^n} x \mu_A(x) \, dP \qquad (۱۶\text{-}۲)$$

$$\mathrm{Var}_P(A) = \frac{1}{P(A)} \int_{\mathbb{R}^n} (x - E_{P(A)})^2 \mu_A(x) \, dP \qquad (۱۷\text{-}۲)$$

که در آنها $P(A)$ عامل استاندارد کردن است .

مثال۲-۱- یک تاس نا اریب را پرتاب می کنیم ، در اینصورت

$$X = \{1, 2, 3, 4, 5, 6\}$$

$$P(x_i) = \frac{1}{6} \qquad i = 1, \ldots, 6$$

فرض کنید پیشامد فازی A، یک زیر مجموعه فازی از X، که نشان دهنده "خیلی کوچک بودن عدد رو شده" و نیز پیشامد فازی B که نشان دهنده "تقریبا پنج" است اینگونه تعریف شوند .

$$\mu_A(x) = \begin{cases} 1 & x=1 \\ 0/7 & x=2 \\ 0/4 & x=3 \\ 0/2 & x=4 \\ 0/1 & x=5 \\ 0 & x=6 \end{cases} \qquad \mu_B(x) = \begin{cases} 0 & x=1 \\ 0/1 & x=2 \\ 0/3 & x=3 \\ 0/8 & x=4 \\ 1 & x=5 \\ 0/8 & x=6 \end{cases}$$

آنگاه داریم: (توجه کنید که در اینحالت در روابط (۲-۵) و (۲-۱۶) و (۲-۱۷) علامت انتگرال به علامت سیگما تبدیل می شود.)

$P(A) = 0/40$	$E_p(A) = 2/04$	$\sqrt{\text{var}_p(A)} = 1/29$
$P(B) = 0/50$	$E_p(B) = 4/70$	$\sqrt{\text{var}_p(B)} = 1/14$
$P(A \cup B) = 0/78$	$P(A \cap B) = 0/12$	$P(AB) = 0/08$
$P(A \oplus B) = 0/82$	$P(A \mid B) = 0/16$	$P(B \mid A) = 0/20$

۳- احتمال پیشامدهای فازی(رهیافت باگر)

در قسمت قبل مشاهده کردیم که تعریف ارائه شده توسط زاده منجر به یک مقدار عددی معین و دقیق برای احتمال یک پیشامد فازی می شود. از دیدگاه شهودی، این دقت و تعیین یا مفهوم وطبیعت مبهم یک پیشامد فازی خود به نوعی احتمال فازی (یعنی یک زیر مجموعه فازی از یک واحد باشد. در تلاش برای برقراری و صورتبندی این احساس شهودی، باگر[۳]، براساس اصل توسیع ، تعریف دیگری برای احتمال یک پیشامد فازی مطرح کرده است .

اصل توسیع [۸] - فرض کنید R و S دو مجموعه باشند. همچنین فرض کنید f یک زیر مجموعه فازی از R بوده و G مجموعه زیر مجموعه های معمولی R باشد یعنی $G = P(R)$ آنگاه اگر f یک نگاشت بصورت

زیر باشد :

$$f : G \longrightarrow S \qquad (۱-۳)$$

با توجه به اتحاد تجزیه(۱-۱۵) یعنی $A = \bigcup_{\alpha \in I} A_\alpha$ می توانیم فلسروf را بصورت زیر بر زیر مجموعه های فازی R گسترش دهیم:

$$f(A) = \bigcup_{\alpha \in I} \alpha f(A_\alpha) \qquad (۲-۳)$$

مثال ۳-۱- فرض کنید $R=\{1,2,3\}$ و $S=\{0,1,4,9\}$ که در اینصورت

$$G = \mathcal{P}(R) = \{\phi, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$$

وتابع بصورت زیر تعریف شده باشد : $f: G \to S$

$$f(B) = \begin{cases} 0 & \text{اگر } B = \phi \\ 1 & B = \{1\}, \{2\}, \{3\} \\ 4 & B = \{1,2\}, \{1,3\}, \{2,3\} \\ 9 & B = \{1,2,3\} \end{cases}$$

حال اگر زیر مجموعه فازی A از R را بصورت زیر در نظر بگیریم .

$$A = \left\{ \frac{0.2}{1}, \frac{0.5}{2}, \frac{0.8}{3} \right\}$$

آنگاه داریم

$$f(A_{0.2}) = f(\{1,2,3\}) = 9$$
$$f(A_{0.5}) = f(\{2,3\}) = 4$$
$$f(A_{0.8}) = f(\{3\}) = 1$$

و در نتیجه

$$f(A) = \left\{ \frac{0.2}{9}, \frac{0.5}{4}, \frac{0.8}{1} \right\}$$

تعریف ۳-۱- در استفاده از اصل تعمیم، به عنوان یک حالت خاص، فرض کنید A یک پیشامد فازی از X و P یک تابع احتمال تعریف شده بر زیر مجموعه های بورل X باشند . آنگاه با استفاده از اصل تعمیم، احتمال یک پیشامد فازی A از X را اینگونه تعریف می کنیم .

$$P_Y(A) = \bigcup_{\alpha \in I} \alpha P(A_\alpha) \qquad (۳-۳)$$

که در آن $P(B_\alpha)$ احتمال پیشامد A_α (یعنی یک زیر مجموعه معمولی X) است .

باتوجه به آنکه تمام A_α ها مجموعه های بورل هستند ،تابع(۳-۳) خوش تعریف است .بعلاوه چون $P(A_\alpha) \in [0,1]$ ، پس $P_Y(A)$ یک زیر مجموعه فازی از $[0,1]$ است .

متذکر می شویم که اگر A یک زیر مجموعه غیر فازی(معمولی) از X باشد ، آنگاه تعریف جدید منطبق بر تعریف معمولی $P_Y(A)$ می باشد .

مثال ۳-۲- فرض کنید

$$X = \{x_1, x_2, x_3, x_4\}$$

و

$$B = \left\{ \frac{1}{x_1}, \frac{1}{x_2}, \frac{0}{x_3}, \frac{1}{x_4} \right\}$$

آنگاه چون برای هر α ، $B_\alpha = B$ و در نتیجه $P(B_\alpha) = P(B)$ داریم :

$$P_Y(B) = \bigcup_{\alpha \in I} \alpha \left\{ P(B_\alpha) \right\} = \left\{ \frac{1}{P(B_1)} \right\} = \left\{ \frac{1}{P(B)} \right\} = \left\{ P(B) \right\}$$

مثال ۳-۳- مجموعه X مثال قبل و زیر مجموعه فازی A از X بصورت $A = \left\{ \frac{1}{x_1}, \frac{0.3}{x_2}, \frac{0.5}{x_3}, \frac{1}{x_4} \right\}$ را در نظر بگیرید.

اگر P یک تابع چگالی احتمال تعریف شده بر X باشد بطوریکه $P(x_1) = 0.3$ و $P(x_2) = 0.2$ و $P(x_3) = 0.1$ و $P(x_4) = 0.4$ و آنگاه چون

$$\alpha_1 = 0.3 \qquad A_{\alpha_1} = \{x_1, x_2, x_3, x_4\} \quad , \quad P(A_{\alpha_1}) = 1$$

$$\alpha_2 = 0.5 \qquad A_{\alpha_2} = \{x_1, x_3, x_4\} \quad , \quad P(A_{\alpha_2}) = 0.7$$

$$\alpha_3 = 0.7 \qquad A_{\alpha_3} = \{x_1, x_4\} \quad , \quad P(A_{\alpha_3}) = 0.6$$

$$\alpha_4 = 1 \qquad A_{\alpha_4} = \{x_4\} \quad , \quad P(A_{\alpha_4}) = 0.4$$

بنابر این

$$P_Y(A) = \bigcup_{\alpha \in I} \alpha P(A_\alpha) = \bigcup_{i=1}^{4} \alpha_i P(A_{\alpha_i}) = \left\{ \frac{0.3}{1}, \frac{0.5}{0.7}, \frac{0.7}{0.6}, \frac{1}{0.4} \right\}$$

ویژگیهای $P_Y(A)$

قضیه ۳-۱-[۲]- فرض کنید X یک مجموعه و P یک تابع چگالی احتمال تعریف شده بر تمام زیر مجموعه های بورل X باشد .همچنین A و B دو زیر مجموعه فازی X باشند که دارای توابع عضویت اندازه پذیر بورل باشند .حال اگر $B \subset A$ آنگاه

$$P_Y(A) \not\subset P_Y(B) \qquad (۳-۴)$$

و اگر $P_Y(A)$ و $P_Y(B)$ هر دو تکیه گاه های $P_Y(A)$ و $P_Y(B)$ باشند ، آنگاه

$$\mu_{P_Y(B)}(x_0) \leq \mu_{P_Y(A)}(x_0) \qquad (۳-۵)$$

نتیجه ۳-۱- تحت شرایط قضیه فوق اگر بعلاوه تکیه گاههای $P_Y(A)$ و $P_Y(B)$ یکی باشند آنگاه $P_Y(B) \subset P_Y(A)$.

مثال ۳-۴- فرض کنید

$$X = \{x_1, x_2, x_3, x_4\}$$

$$A = \left\{ \frac{\cdot/8}{x_1}, \frac{\cdot/5}{x_2}, \frac{\cdot/3}{x_3}, \frac{\cdot/2}{x_4} \right\}$$

$$B = \left\{ \frac{\cdot/4}{x_1}, \frac{\cdot/4}{x_2}, \frac{\cdot/2}{x_3}, \frac{\cdot/1}{x_4} \right\}$$

(توجه کنید که $B \subset A$)

آنگاه مجموعه های سطح A عبارتند از :

$$A_{\cdot/4} = \{x_1, x_2, x_3, x_4\}$$

$$A_{\cdot/3} = \{x_1, x_2, x_3\}$$

$$A_{\cdot/5} = \{x_1, x_2\}$$

$$A_{\cdot/8} = \{x_1\}$$

و برای B نیز

$$B_{\cdot/1} = \{x_1, x_2, x_3, x_4\}$$

$$B_{\cdot/2} = \{x_1, x_2, x_3\}$$

$$B_{\cdot/3} = \{x_1, x_2\}$$

$$B_{\cdot/4} = \{x_1\}$$

حال اگر $P(x_1)=\cdot/2$ و $P(x_2)=\cdot/3$ و $P(x_3)=\cdot/4$ و $P(x_4)=\cdot/1$ آنگاه

$$P_Y(A) = \left\{ \frac{\cdot/2}{1}, \frac{\cdot/3}{\cdot/9}, \frac{\cdot/5}{\cdot/5}, \frac{\cdot/8}{\cdot/2} \right\}$$

$$P_Y(B) = \left\{ \frac{\cdot/1}{1}, \frac{\cdot/2}{\cdot/9}, \frac{\cdot/3}{\cdot/6}, \frac{\cdot/4}{\cdot/3} \right\}$$

و

مشاهده می کنیم که در این مثال $P_Y(B) \not\subset P_Y(A)$ لذا این به دلیل آن است که تکیه گاههای $P_Y(B)$ و $P_Y(A)$ یکسان نیستند.

قضیه زیر یک شرط کافی برای یکسان بودن تکیه گاههای $P_Y(A)$ و $P_Y(B)$ بیان می کند.

قضیه ۳-۲-[۲] - فرض کنید A و B دو زیر مجموعه فازی X باشند و $B \subset A$. آنگاه $P_Y(A)$ و $P_Y(B)$ برای تکیه گاههای یکسان هستند اگر

$$A(x_i) > A(x_j) \quad \text{اگر و فقط اگر} \quad B(x_i) > B(x_j)$$

اگر و فقط اگر

$A(x_i) = A(x_j)$ اگر و فقط اگر $B(x_i) = B(x_j)$ (۶-۳)

تحت شرایط قضیه فوق طبق نتیجه (۱-۳) داریم $P_Y(B) \subset P_Y(A)$.

قضیه ۷-۳. فرض کنید A و B دو پیشامد فازی از X با احتمالهای $P_Y(A)$ و $P_Y(B)$ باشند ، آنگاه

$$P_Y(A \cup B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A_\alpha) + P(B_\alpha) - P(A_\alpha \cap B_\alpha) \right\} \qquad (۷-۳)$$

اثبات ـ داریم $A \cup B = \bigcup_{\alpha \in I} \alpha \left\{ (A \cup B)_\alpha \right\}$ و $P_Y(A \cup B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A \cup B)_\alpha \right\}$

اما بنابه (۲-۳) داریم $(A \cup B)_\alpha = A_\alpha \cup B_\alpha$ بنابراین

$$P_Y(A \cup B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A_\alpha \cup B_\alpha) \right\}$$

و چون $P(A_\alpha \cup B_\alpha) = P(A_\alpha) + P(B_\alpha) - P(A_\alpha \cap B_\alpha)$

پس

$$P_Y(A \cup B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A_\alpha) + P(B_\alpha) - P(A_\alpha \cap B_\alpha) \right\}$$

قضیه ۸-۳. فرض کنید A و B دو پیشامد فازی از X با احتمالهای $P_Y(A)$ و $P_Y(B)$ باشند ، آنگاه

الف)

$$P_Y(A \cap B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A_\alpha \cap B_\alpha) \right\} \qquad (۸-۳)$$

ب) اگر A و B مستقل باشند

$$P_Y(A \cap B) = \bigcup_{\alpha \in I} \alpha \left\{ P(A_\alpha) \cdot P(B_\alpha) \right\}$$

اثبات ـ قسمت (الف) با توجه به اینکه طبق (۱-۱۷) داریم $(A \cap B)_\alpha = A_\alpha \cap B_\alpha$ و قسمت (ب) با توجه به اینکه اگر A و B مستقل باشند داریم $P(B_\alpha \cap B_\alpha) = P(A_\alpha) \cdot P(B_\alpha)$ ثابت می شود

احتمال فازی و عدد اصلی ـ اکنون نشان می دهیم که تعریف زاده و تعریف ارائه شده در این بخش برای احتمال یک پیشامد فازی ، با دو تعریف متفاوت از عدد اصلی یک مجموعه فازی مرتبط اند . بدینوسیله تفاوت ماهوی این دو تعریف واضح تر می شود .

بنابه (۸-۱) چنانچه A یک زیر مجموعه فازی با تکیه گاه متناهی از X باشد ، آنگاه توان A ، $|A|$ عبارت است از

$$|A| = \sum_{x_i \in X} \mu_A(x_i) \qquad (۹-۳)$$

زاده در [۸] پنده توان را به مفهوم عدد اصلی فازی تعمیم می دهد . به اینصورت که فرض کنید A یک زیر مجموعهٔ فازی از X با تکیه گاه متناهی باشد . آنگاه با استفاده از اصل تعمیم ، عدد اصلی فازی $|A|_a$ ، با بینگونه تعریف می شود .

$$|A|_a = \bigcup_{\alpha \in I} \alpha \left\{ |A_\alpha| \right\} \qquad (۱۰-۳)$$

که در آن $|A_i|$ عدد اصلی مجموعه غیر فازی A_α است .توجه می کنید که $|A|_\alpha$ یک مجموعه فازی از اعداد صحیح غیر منفی است .

بطور مختصر تعریف اول یعنی $|A|$ جمع درجات عضویت اعضای X را می دهد در حالی که دومی یعنی $|A|_\alpha$ برای هر درجه بانتدی $\alpha \in I$ عدد اصلی مجموعه عناصری از X که درجه عضویت آنها حداقل α باشد را اندازه می گیرد.

اکنون فضای نمونه متناهی زیر را در نظر بگیرید .

$$X = \{x_1, \ldots, x_n\}$$

ونیز فرض کنید x_i ها ، هم احتمال باشند .اگر E یک زیر مجموعه معمولی از X باشد ،آنگاه احتمال E عبارت است از

$$P(E) = \frac{|E|}{|X|} = \frac{Card\, E}{n} \qquad (۱۱-۳)$$

حال فرض کنید F یک زیر مجموعه فازی از X باشد .اگر بخواهیم احتمال F را با استفاده از تعریف بالا بدست آوریم ، می توانیم توان را به عنوان عدد اصلی F به کار ببریم یعنی

$$P(F) = \frac{|F|}{|X|} = \frac{\sum_i \mu_F(x_i)}{n} = \frac{1}{n}\sum_i \mu_F(x_i) \qquad (۱۲-۳)$$

که این مبنای تعریف زاده برای احتمال یک پیشامد فازی است .اما اگر از عدد اصلی $|F|_\beta$ در فرمول فوق استفاده کنیم داریم

$$P_\gamma(F) = \frac{|F|_\beta}{|X|} = \frac{\bigcup_{\alpha \le 1}\{|F_\alpha|\}}{n} = \bigcup_\alpha \left\{\frac{|F_\alpha|}{n}\right\} = \bigcup_\alpha \{P(F_\alpha)\} \qquad (۱۳-۳)$$

که در آن $P(F_\alpha)$ ، احتمال معمولی مجموعه تراز F_α است .مشاهده می کنیم که این ،متناظر با تعریف یاگر برای احتمال یک پیشامد فازی است .

۳.۴ احتمال‌های پیشامدهای فازی (تصحیح و اضافه یاگر توسط کلمنت)[10]

مفهوم احتمال یک پیشامد فازی که توسط یاگر تعریف شده ، دارای چند نقص است. طبق این تعریف اولا : چنانچه یک پیشامد فازی زیر مجموعه یک پیشامد فازی دیگر باشد لزوما احتمال اولی زیر مجموعه احتمال دومی نخواهد بود. ثانیا : ویژگی جمع پذیری (رابطه ۷-۳) ناشی شده از تعریف یک ویژگی خوبی نیست [۱]. کلمنت [۱] ضمن تأیید صحت شهودی روش یاگر با یک تغییر تکنیکی جزیی در

10-Klement

رهیافت وی ؛ تعریف پیراسته تری را برای احتمال یک پیشامد فازی ارائه کرده است .

یاد آوری می کنیم که تعریف احتمال یک پیشامد فازی تَوَسط یا گربصورت زیراست :

$$P_Y(A) = \bigcup_{\alpha} \alpha \left\{ P_{A_\alpha}(A) \right\} \qquad (۱\text{-}۴)$$

این تعریف ، در حالت کلی فاقد یک ویژگی اساسی است یعنی در حالت کلی الزام (۴-۲) یعنی

$A \subset B \Rightarrow P_Y(A) \subset P_Y(B)$ برقرار نیست.

برای ارائه روش پیشنهادی کلمنت و مقایسه باروش یاگرمثال (۴-۳) را مجدداً یاد آوری می کنیم.

فرض کنید $X = \{x_1 , x_2 , x_3 , x_4\}$

و اندازهٔ احتمال P نیز اینگونه باشد .

$$P(x_1) = ۰/۲ \qquad P(x_2) = ۰/۳ \qquad P(x_3) = ۰/۴ \qquad P(x_4) = ۰/۱$$

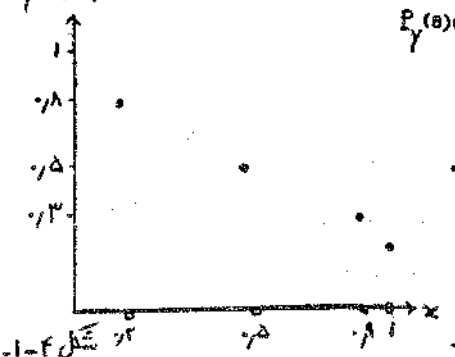
آنگاه برای مجموعه های فازی A و B بصورت زیر

$$A = \left\{ \frac{۰/۸}{x_1} , \frac{۰/۵}{x_2} , \frac{۰/۳}{x_3} , \frac{۰/۲}{x_4} \right\}$$

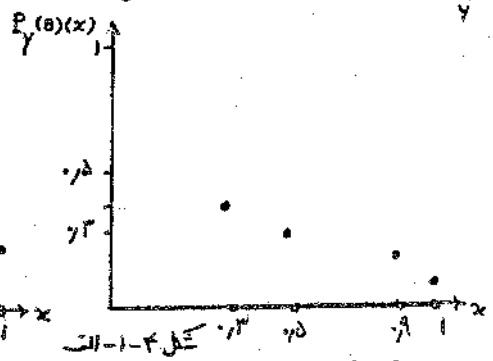
$$B = \left\{ \frac{۰/۳}{x_1} , \frac{۰/۴}{x_2} , \frac{۰/۲}{x_3} , \frac{۰/۱}{x_4} \right\}$$

$$P_Y(A) = \begin{cases} ۰/۸ & \text{اگر } x = ۰/۲ \\ ۰/۵ & x = ۰/۵ \\ ۰/۳ & x = ۰/۹ \\ ۰/۲ & x = ۱ \\ ۰ & \text{سایرنقاط} \end{cases} \qquad P_Y(B) = \begin{cases} ۰/۴ & \text{اگر } x = ۰/۳ \\ ۰/۳ & x = ۰/۵ \\ ۰/۲ & x = ۰/۹ \\ ۰/۱ & x = ۱ \\ ۰ & \text{سایرنقاط} \end{cases}$$

ملاحظه می کنید که برای این مثال $A \subset B$ اما $P_Y(B) \not\subset P_Y(A)$.دلیل این عدم یکنوایی P_Y این است که دو زیر مجموعه فازی $P_Y(A)$ و $P_Y(B)$ تکیه گاههای متفاوتی دارند . یعنی زیر مجموعه ای از بازه واحد که در آن $P_Y(A)$ و $P_Y(B)$ صفر نمی شوند ، متفاوت است .(شکل ۴-۱)



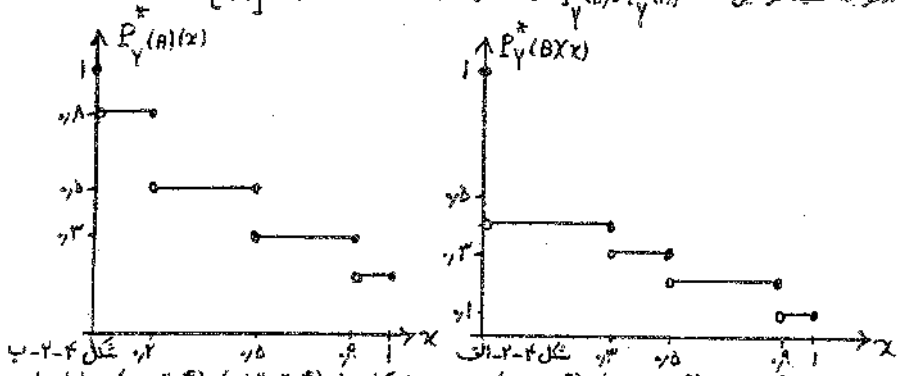
شکل ۴-۱-ب



شکل ۴-۱-الف

این مشکل (عدم یکنوایی) با یک تغییر کوچک در نمودار $P_Y(A)$ و $P_Y(B)$ برطرف می شود (شکل

۴ـ۲).(توجه کنید در این حالت $P^*_Y(A)$ و $P^*_Y(B)$ تکیه گاههای یکسان دارند ، هر دو $[0,1]$.)



تصحیح شکل های (۴ـ۱ـ الف) و (۴ـ۱ـ ب) بصورت شکل های (۴ـ۲ـ الف) و (۴ـ۲ـ ب) معادل با

تعریف زیر برای احتمال یک پیشامد فازی است [۱]:

$$P^*_Y(A)(\alpha) = \sup_{\alpha} \{\alpha \mid P_Y(A_\alpha) \geq \alpha\} \quad , \alpha \in I \qquad (۴ـ۴)$$

ویژگی های P^*_Yـ ابتدا می گوییم که $P_Y(A)$ و $P^*_Y(A)$ در نقاط اساسی بر هم منطبق هستند .

گزارهٔ ۴ـ۱ـ فرض کنید A یک پیشامد فازی باشد . اگر به ازای $\alpha \in I$ ، عدد $\alpha \in I$ وجود داشته باشد به

قسمی که $P_Y(A_\alpha) = \alpha$ آنگاه داریم

$$P_Y(A)(\alpha) = P^*_Y(A)(\alpha) \qquad (۴ـ۳)$$

$P^*_Y(A)(\alpha) \leq \alpha_0$

اثبات ـ با قرار دادن $\alpha_0 = \sup\{\alpha \mid P_Y(A_\alpha) = \alpha\}$ داریم $P_Y(A)(\alpha_0) = \alpha_0$. بعلاوه می توان ثابت کرد که $P^*_Y(A)(\alpha) \geq \alpha_0$ و حکم ثابت است .

گزاره ۴ـ۲ـ یکنوایی P^*_Y ـ فرض کنید A و B دو پیشامد فازی باشند طوری که $A \subseteq B$ آنگاه

$$P^*_Y(A) \subseteq P^*_Y(B) \qquad (۴ـ۴)$$

اثبات ـ این قضیه با توجه به این نکته که اگر $A \subseteq B$ آنگاه برای هر α داریم $A_\alpha \subseteq B_\alpha$ واضح است .

گزارهٔ ۴ـ۳ـ پیوستگی P^*_Y ـ فرض کنید $\{A_n, n \in N\}$ یک دنباله افزایشی از پیشامدهای فازی باشند به

قسمی که $A_n \uparrow A$ ، آنگاه

$$P^*_Y(A_n) \uparrow P^*_Y(A) \qquad (۴ـ۵)$$

اثبات ـ با ملاحظه اینکه اگر $A_n \uparrow A$ آنگاه $P(A_n) \uparrow P(A)$ ، حکم نتیجه می شود.

بعلاوه خود تابع $P^*_Y(A)$ نیز چند ویژگی خوب دارد.

گزارهٔ ۴ـ۴ـ برای تمام پیشامدهای فازی A داریم

$$P^*_Y(A)(\alpha) = 1 \quad \text{و} \quad P^*_Y(A)(\alpha) = 0 \qquad \text{برای همه } \alpha \qquad (۴ـ۶)$$

$$P^*_Y(A)(x) = \inf\left\{ P^*_V(A)(y) \mid y \in [0, x] \right\} \qquad (۷\text{-}۴)$$

اثبات ـ گزاره های (۴ـ۶) از رابطهٔ (۴ـ۳) بسادگی نتیجه می شود. (۴ـ۷) نیز با توجه به غیر صعودی بودن $P^*_Y(A)$ نتیجه می شود.

کلمنت بر مبنای مفاهیم اعداد فازی و نیز آنچه وی آن را اندازه های فازی، فازی ـ مقدار [۱] می نامد، یک ویژگی جمع‌پذیری را برای تعریف خود مطرح می کند که خوانندگان علاقمند را به [۱] ارجاع می‌دهیم. وی همچنین یک تعبیر (۴ـ۳) را بر اساس عدد اصلی مجموعهٔ فازی در [۱] ارائه کرده است.

۵ـ احتمال پیشامدفازی (رهیافت دوم یاگر)

در بخش قبل تصحیحی که کلمنت در رهیافت یاگر برای تعریف احتمال یک پیشامد فازی پیشنهاد کرده است و همچنین چندویژگی مطلوب ناشی شده از آن تصحیح شرح داده شد. در عین حال یاگر خاطرنشان می کند که نقصی در تعریف کلمنت وجود دارد که پایه نگرانی است:

فرض کنید A یک زیر مجموعه معمولی از X باشد در این صورت (۵ـ۱)

$$A_\alpha = \begin{cases} A & \alpha > 0 \\ X & \alpha = 0 \end{cases}$$

بنابر این از (۳ـ۳) داریم (شکل ۵ـ۱)

$$P_Y(A) = \left\{ \frac{1}{P(A)} \right\} \qquad (۵\text{-}۲)$$

اما بر اساس (۴ـ۳) داریم (شکل ۵ـ۲)

$$P^*_Y(A)(x) = \begin{cases} 1 & 0 \leq x \leq P(A) \\ 0 & P(A) < x \end{cases} \qquad (۵\text{-}۳)$$

به بیان دیگر، روش یاگر در حالت یک مجموعه معمولی منجر به تعریف معمولی احتمال می شود (درجه عضویت یک برای $P(A)$ و صفر برای هر مقدار دیگر) ولی در این وضعیت روش کلمنت می گوید که به هر عدد کوچکتر از $P(A)$ (حتی مثلا صفر) عدد عضویت یک را منتاظر کنیم و این به هرحال یک نقص برای این روش است.

یاگر در مقاله ای که در آن مجدداً[11] به موضوع احتمال یک پیشامد فازی پرداخته است [۴]، ضمن بیان اشکال فوق، تعریف جدیدی برای احتمال یک پیشامدفازی مطرح کرده است که در اینجا به آن می پردازیم.

مجدداً تعریف کلمنت برای احتمال یک پیشامد فازی را در نظر بگیرید:

11-Fuzzy-Valued Fuzzy Measures

$$P^*_\gamma(A)(\alpha) = \sup_{\alpha} \{\alpha \mid P_\gamma(A_\alpha) \geq \alpha\} \quad , \alpha \in I \qquad (۴\text{-}۵)$$

یک تعبیر $P^*_\gamma(A)(\alpha)$ این است که برای هر $\alpha \in I$، مقدار $P^*_\gamma(A)(\alpha)$ را به عنوان درجه راستی گزاره احتمال A حداقل α است، در نظر بگیریم. (وقتی A یک مجموعه معمولی باشد، تعبیر فوق به آسانی قابل درک است.) حال مجموعه فازی A'، یعنی مکمل مجموعه فازی A را در نظر بگیرید.

$$\mu_{A'}(\alpha) = 1 - \mu_A(\alpha) \qquad \forall \alpha \in X$$

فرض کنید A'_α ها، مجموعه های تراز A' باشند، آنگاه برای A' داریم

$$P^*_\gamma(A')(\alpha) = \sup_{\alpha} \{\alpha \mid P_\gamma(A'_\alpha) \geq \alpha\} \quad , \alpha \in I$$

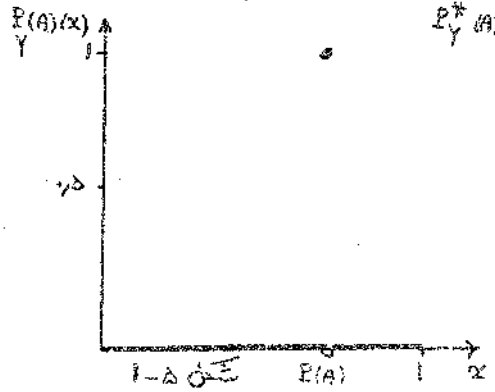
که در آن $P^*_\gamma(A')(\alpha)$ درجه راستی این گزاره است: احتمال عدم وقوع A حداقل α است. اکنون $P_{*\gamma}(A)$ را اینگونه تعریف می کنیم.

$$P_{*\gamma}(A) = 1 - P^*_\gamma(A') \qquad (۵\text{-}۵)$$

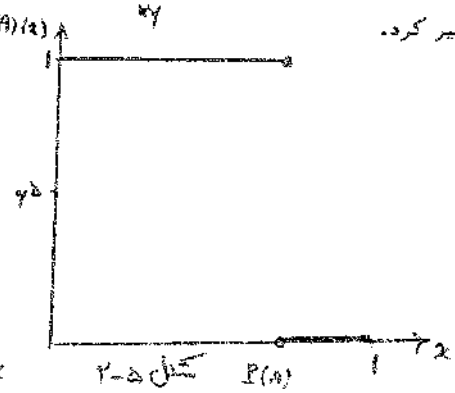
بنابر این

$$P_{*\gamma}(A)(\alpha) = P^*_\gamma(1-\alpha) = \sup_{\alpha} \{\alpha \mid P_\gamma(A'_\alpha) \geq 1-\alpha\} \qquad (۶\text{-}۵)$$

و مشابه تعابیر قبلی $P_{*\gamma}(A)(\alpha)$ را می توان درجه راستی گزاره احتمال A حداکثر α است تعبیر کرد.



شکل ۵-۱



شکل ۵-۲

حال می توان با ترکیب دو احتمال (۵-۴) و (۵-۵)، تعریف پیراسته تری برای احتمال یک پیشامد فازی بصورت زیر بدست آورد.

$$\hat{P}_\gamma(A) = P^*_\gamma(A) \cap P_{*\gamma}(A) \qquad (۷\text{-}۵)$$

و یا

$$\hat{P}_\gamma(A)(\alpha) = \min\left[P^*_\gamma(A)(\alpha) \ , \ P_{*\gamma}(A)(\alpha)\right] \qquad (۸\text{-}۵)$$

در اینجا برای هر $\alpha \in I$، مقدار $\hat{P}_\gamma(A)(\alpha)$ عبارت است از درجه راستی گزاره احتمال پیشامد A دقیقاً

برابر با α است."

تذکر ۱ـ همانگونه که در میان مطالب فوق ذکر شد ، یا گر تعابیری برای $\hat{P}^{*}_{\gamma}(A)(\alpha)$ و $\hat{P}^{*}_{\gamma}(A)(\alpha)$ در نظر گرفته است ، به اینصورت که $\hat{P}^{*}_{\gamma}(A)(\alpha)$ را تعبیر می کنند به درجه راستی گزاره "احتمال پیشامد A" ، حداقل α است" و $\hat{P}^{*}_{\gamma}(A)(\alpha)$ را به درجه راستی گزاره "احتمال پیشامد A حداکثر α است" با کمی دقت در تعریف های فوق و شکل نمودارهای توابع عضویت این دو مجموعه فازی تعابیر فوق واضح اند. برای واضح شدن این تعبیر که : $\hat{P}_{\gamma}(A)(\alpha)$ عبارت است از درجه راستی گزاره "احتمال پیشامد A دقیقا (=برابر α است" می گوییم:

اگر دو گزاره

۱ـ احتمال پیشامد A حداقل برابر α است.

۲ـ احتمال پیشامد A حداکثر برابر α است ـ کاملا راست باشند آنگاه می توانیم نتیجه بگیریم که گزاره

احتمال پیشامد A دقیقاً برابر α است . نیز کاملا راست است . به عبارت دیگر درجه راستی دو گزاره اول اعدادی بین صفر و یک باشند ،مسلما ارزش درستی گزاره نتیجه ، برابر با حداقل این دو عدد است .به عبارت دیگر درجه تعلق α به مجموعه $\hat{P}_{\gamma}(A)$ برابراست با مینیمم درجات تعلق α به دو مجموعه $\hat{P}^{*}_{\gamma}(A)$ و $P_{*\gamma}(A)$.

برای مثال فرض کنید درجه راستی گزاره ۱ برابر۷/۰ و درجه راستی گزاره ۲ برابر ۴/۰ باشد . حال استنتاج زیر را در نظر بگیرید .

ـ درجه راستی گزاره "احتمال پیشامد A حداقل x است" برابربا۷/۰است .

ـ درجه راستی گزاره "احتمال پیشامد A حداکثر x است" برابربا۴/۰است .

∴ ـ درجه راستی گزاره "احتمال پیشامد A دقیقا برابر x است" برابر با۴/۰است .

که این نتیجه گیری همان چیزی است که رابطه (۸ـ۵) بیان می کند .

تذکر ۲ـ در جملات وتعابیر نظیر "احتمال پیشامد A حداقل αاست" . ویا "احتمال پیشامد A دقیقاً برابر α است" . اندکی تسامح ، ملحوظ است . آیا منظور از "احتمال پیشامد A" در این نوع جملات دقیقا چیست ؟ اگر منظور همان مفهوم معمولی احتمال باشد که این مفهوم برای یک پیشامد فازی تعریف نمی شود .اگر تعریف زاده برای احتمال یک پیشامد فازی مد نظر است که در صورتبندی یا گر این مفهوم، به کار نرفته است. اگر منظور ،مفهومی است که خود یا گر مطرح کرده است که در این صورت یک عدد فازی را نمی توان برابر با یک عدد تصور کرد . با این همه ما می توانیم از لفظ احتمال به کار برده شده در این نوع عبارات، همان مفهوم شهودی ای که از احتمال داریم را مد نظر بگیریم .

مثال ص۱۰۵ـ فرض کنید

$$X = \{x_1, x_2, x_3, x_4\}$$

$$\underline{P}(x_1) = 0.2 \quad \underline{P}(x_2) = 0.3 \quad \underline{P}(x_3) = 0.4 \quad \underline{P}(x_4) = 0.1$$

و

و

$$A = \left\{ \frac{1}{x_1}, \frac{0.7}{x_2}, \frac{0.5}{x_3}, \frac{0.1}{x_4} \right\}$$

آنگاه ، شکل(ص۳۰۵)

$$\underline{P}^*_Y(A)(x) = \begin{cases} 1 & 0 \leq x \leq 0.1 \\ 0.7 & 0.1 < x \leq 0.5 \\ 0.5 & 0.5 < x \leq 0.9 \\ 0.1 & 0.9 < x \leq 1 \\ 0 & \text{سایر نقاط} \end{cases}$$

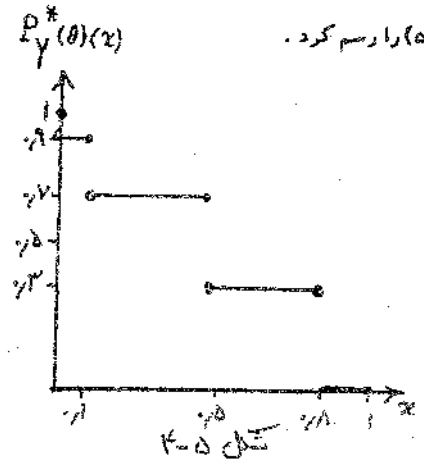
وچون

$$A^1 = \left\{ \frac{0}{x_1}, \frac{0.3}{x_2}, \frac{0.5}{x_3}, \frac{0.9}{x_4} \right\}$$

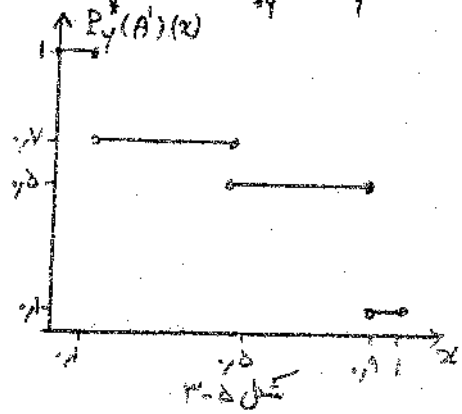
پس مجموعه های تراز A^1 و احتمالات متناظر عبارتند از شکل (ص۴۰۵) ،

$\underline{P}^*_Y(A^1_\alpha)=1$	$A^1_\alpha = \{x_1, x_2, x_3, x_4\}$	$\alpha = 0$
$\underline{P}^*_Y(A^1_\alpha)=0.8$	$A^1_\alpha = \{x_2, x_3, x_4\}$	$0 < \alpha \leq 0.3$
$\underline{P}^*_Y(A^1_\alpha)=0.5$	$A^1_\alpha = \{x_2, x_4\}$	$0.3 < \alpha \leq 0.7$
$\underline{P}^*_Y(A^1_\alpha)=0.1$	$A^1_\alpha = \{x_4\}$	$0.7 < \alpha \leq 0.9$
$\underline{P}^*_Y(A^1_\alpha)=0$	$A^1_\alpha = \phi$	$0.9 < \alpha \leq 1$

چون $\underline{P}^*_Y(A^1) = \underline{P}(A)$ ، به آسانی می توان شکل (ص۵۰۵) را رسم کرد .



شکل ۴-۵



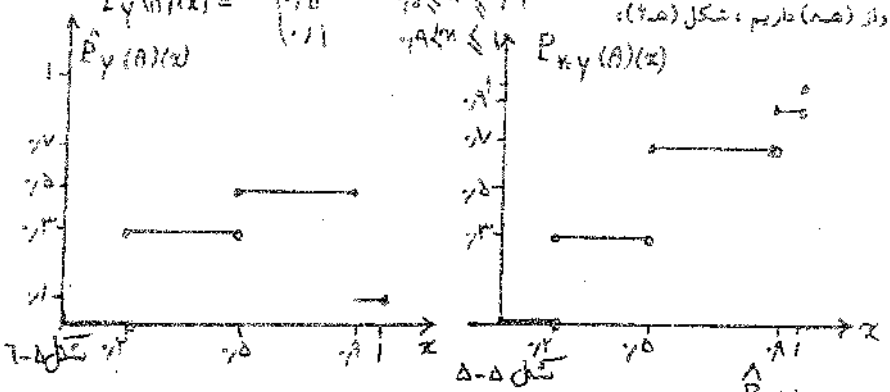
شکل ۳-۵

زیرا:

$$\hat{P}_{*\gamma}(A)(x)=\begin{cases}0 & 0\le x<0.2\\ 0.3 & 0.2\le x<0.5\\ 0.7 & 0.5\le x<0.9\\ 0.9 & 0.9\le x<1\\ 1.0 & x=1\end{cases}$$

$$\hat{P}_{\gamma}(A)(x)=\begin{cases}0 & 0\le x<0.2\\ 0.3 & 0.2\le x<0.5\\ 0.5 & 0.5\le x<0.9\\ 0.1 & 0.9\le x\le 1\end{cases}$$

واژ (صـ۸) داریم ، شکل (صـ۲):



شکل ۵-۵ شکل ۵-۶

ویژگیهای $\hat{P}_{\gamma}(A)$:

یک ویژگی مهم $\hat{P}_{\gamma}$ این است که برای پیشامد معمولی A ،مقدار Pr(A) یعنی تعریف معمولی احتمال یک پیشامد را نتیجه میدهد .

قضیه صـ۱ـ برای هر مجموعهٔ معمولی A از X داریم

$$\hat{P}_{\gamma}(A)=Pr(A) \qquad (۹ـصـ)$$

که در آن Pr(A) kاحتمال معمولی Aاست.یعنی

$$Pr(A)=\sum_{x\in A}P(x) \qquad (۱۰ـصـ)$$

اثبات ـ فرض کنید $X=\{x_1,x_2,\dots,x_n\}$. بدون از دست دادن کلیت ،فرض کنید :

$$A=\{x_1,\dots,x_j\}$$

در این حالت $P(A)=\sum_{i=1}^{j}P(x_i)$ چون A مجموعه ای معمولی است پس

$$A_\alpha=\begin{cases}\{x_1,\dots,x_j\} & \alpha>0\\ X & \alpha=0\end{cases}$$

و همچنین

$$\hat{P}^{*}_{\gamma}(A)(x)=\begin{cases}1 & 0\le x\le P(A)\\ 0 & P(A)<x\end{cases}$$

بطور مشابه برای مجموعه معمولی $P(A') = 1 - P(A) = \sum_{i=j+1}^{n} P(x_i)$ که برای آن $A' = \{x_{j+1}, \ldots, x_n\}$

داریم

$$A'_\alpha = \begin{cases} \{x_{j+1}, \ldots, x_n\} & \alpha > 0 \\ X & \alpha = 0 \end{cases}$$

وهمچنین

$$P^*_Y(A')(\alpha) = \begin{cases} 1 & 0 \le \alpha \le 1 - P(A) \\ 0 & 1 - P(A) < \alpha \end{cases}$$

و چون $P_{*Y}(A)(\alpha) = P^*_Y(A')(1-\alpha)$

پس

$$P_{*Y}(A)(\alpha) = \begin{cases} 1 & P(A) \le \alpha \le 1 \\ 0 & \alpha < P(A) \end{cases}$$

وبالاخره با توجه به رابطه (۵-۷)

$$\hat{P}_{*Y}(A)(\alpha) = \begin{cases} 1 & \alpha = P(A) \\ 0 & \text{سایر نقاط} \end{cases}$$

قضیه ۵-۲- فرض کنید زیر مجموعه فازی A از X طوری است که حداقل یک $x \in X$ وجود دارد که

$$\mu_A(x) \neq 0 \text{ و } 1 \quad , \quad P(x) \neq 0$$

آنگاه برای هر $u \in I$

$$\mu_{\hat{P}_Y(A)}(u) < 1 \qquad (۵-۱۱)$$

یعنی $\hat{P}_Y(A)$ زیر نرمال است و بنابراین هیچ $u \in I$ وجود ندارد که برای آن، درجه راستی گزاره (احتمال A، دقیقاً برابر u است.) یک شود.

اثبات - اثبات بوسیله برهان خلف، به آسانی انجام می شود.

اکنون وضعیتی را بررسی می کنیم که $A \subset B$ و نشان می دهیم که، بطور اساسی، احتمال در نقطه ای که در آن، $\hat{P}_Y(B)$ بزرگترین مقدار راستی را اتخاذ می کند بزرگتر است از احتمال در نقطه ای که $\hat{P}_Y$ (A) در آن بیشترین مقدار راستی را کسب می کند. (توجه دارید که طبق قضیه قبل برای A و B فازی، توابع $\hat{P}_Y(A)$ و $\hat{P}_Y(B)$ زیر نرمال هستند.)

لم ۱- برای هر زیر مجموعهٔ فازی A از X، تابع $\hat{P}_Y(A)$ یک تابع بطور یکنواخت نزولی بر I است.

اثبات - با توجه به تعریف $\hat{P}_Y(A)$ و ویژگی یکنوایی احتمال و اینکه برای $\alpha_1 < \alpha_2$ داریم $A_{\alpha_2} \subset A_{\alpha_1}$ حکم ثابت می شود.

لم ۲- برای هر زیر مجموعه فازی A از X، تابع $P_{*Y}(A)$ یک تابع بطور یکنواخت صعودی بر I است.

اثبات - با استفاده از لم قبل و نیز رابطهٔ $P_{*Y}(A)(x) = P^*_Y(A')(1-x)$ حکم ثابت می شود.

قضیه ۵-۳- برای هر زیر مجموعه فازی A از X، $\hat{P}_\gamma(A)$ یک زیر مجموعهٔ فازی محدب از I است.

اثبات - فرض کنید $P^*_\gamma(A)=G$، بنا به ام ۱، G بر آبطور یکنواخت نزولی است، پس برای هر u_1 و u_2 از I و برای هر $\lambda \in I$

$$G(\lambda u_1 + (1-\lambda)u_2) \geq \min\left[G(u_1), G(u_2)\right]$$

بنا بر این $P^*_\gamma(A)=G$ یک زیر مجموعهٔ فازی محدب از I است.

بطور مشابه به دلیل اینکه $P_{*\gamma}(A)=H$ بر آبطور یکنواخت صعودی است، یک زیر مجموعه فازی محدب از I است و بنا بر قضیه (۱-۴)، $\hat{P}_\gamma(A)=H\wedge G$، نیز یک زیر مجموعه فازی محدب از I خواهد بود.

نتیجه ۵-۱- از محدب بودن $\hat{P}_\gamma(A)$ نتیجه می شود که $\hat{P}_\gamma(A)$ بالاترین درجهٔ عضویت خود را در یک نقطه یا یک بازهٔ بسته $[a,b]$ اختیار می کند.

قضیه ۵-۴-

الف) اگر $A\subset B$ آنگاه $P^*_\gamma(A) \subset P^*_\gamma(B)$ $\quad P(A)\ P(B)$

ب) اگر $A\subset B$ آنگاه $P_{*\gamma}(A) \supset P^*_\gamma(B)$ $\quad P(A)\ P(B)$

اثبات - قسمت (الف) همان گزاره (۳-۴) است. قسمت (ب) نیز با استفاده از قسمت (الف) و با توجه به اینکه در اینجا $B^1\subset A^1$، بسادگی ثابت می شود.

با استفاده از مطالب فوق می توان قضیه زیر را نتیجه گرفت.

قضیه ۵-۵- اگر $A\subset B$ و $\hat{P}_\gamma(A)$ و $\hat{P}_\gamma(B)$ مقادیر ماکزیمم خود را بترتیب در بازه های $[a1,a2]$ و $[b1,b2]$ اختیار کنند، آنگاه $a1\geq b1$ و $a2\geq b2$.

اکنون مشاهده می کنیم که اگر $A\subset B$، آنگاه مقداری که توسط آن بیشترین درجه راستی B حاصل می شود، همیشه بزرگتر است از مقداری که توسط آن بیشترین درجهٔ راستی A بدست می آید.

قضیه ۵-۶- فرض کنید $A\subset B$. اگر $\hat{P}_\gamma(B)$ و $\hat{P}_\gamma(A)$ مقادیر ماکزیمم خود را بترتیب در بازه های $[a1,a2]$ و $[b1,b2]$ اختیار کنند آنگاه

(ص-۱۲)

الف) برای هر $u>b2$ داریم $\quad \hat{P}_\gamma(B)(u) \geq \hat{P}_\gamma(A)(u)$

ب) برای هر $u>a1$ داریم $\quad \hat{P}_\gamma(A)(u) \geq \hat{P}_\gamma(B)(u)$

اثبات - چون برای $u>b_\gamma$ داریم

$$P_{*\gamma}(A)(u) \geq P^*_\gamma(A)(u)$$

$$\Rightarrow P_{*\gamma}(B)(u) \geq P^*_\gamma(B)(u)$$

بنابراین

$$\hat{P}_\gamma(A)(u) = \min\left[P_{*\gamma}(A)(u), P^*_\gamma(A)(u)\right] = P^*_\gamma(A)(u)$$

$$\hat{P}_Y(B)(w) = P_Y^*(B)(w)$$

و به طور مشابه

و با توجه به یکنواییِ P_Y^*، (الف) حاصل می‌شود. به روش مشابه (ب) نیز ثابت می‌شود.

مثال ۵-۱- فرض کنید X فضای نمونهٔ حاصل از پرتاب یک تاس نااریب باشد، یعنی

$X=\{1,2,3,4,5,6\}$

و $P(1)=P(2)=\ldots=p(6)=1/6$

همچنین فرض کنید زیرمجموعهٔ فازی A از X که نشان‌دهندهٔ «خیلی کوچک» بودن است، اینگونه تعریف شود:

$$\mu_A(x)=\begin{cases}1 & \text{اگر } x=1\\ 0/7 & \text{اگر } x=2\\ 0/4 & \text{اگر } x=3\\ 0/2 & \text{اگر } x=4\\ 0/1 & \text{اگر } x=5\\ 0 & \text{اگر } x=6\end{cases}$$

حال فرض کنید تاس را یک دفعه پرتاب کنیم، در اینصورت احتمال پیشامد فازی A، یعنی احتمال روشدن یک عدد خیلی کوچک (به بیان فوق)، طبق چهار روشی که مطرح کرده‌ایم به صورت زیر بدست می آید.

۱- براساس رهیافت زاده $P_Z(A)=\frac{1}{n}\sum_{i=1}^{6}\mu_A(x_i)=\frac{1}{6}\times 2/4=0/4$

یعنی احتمال روشدن یک عدد «خیلی کوچک» تقریباً برابر ۰/۴ است.

۲- براساس رهیافت اول یا گر

در اینجا مثلاً احتمال ۰/۶ دارای درجهٔ عضویت ۰/۷ در $P(A)$ است:

$$P_Y(B)(x)=\begin{cases}1 & \text{اگر } x=1/6\\ 0/7 & \text{اگر } x=2/6\\ 0/4 & \text{اگر } x=3/6\\ 0/2 & \text{اگر } x=4/6\\ 0/1 & \text{اگر } x=5/6\\ 0 & \text{اگر } x=6/6\end{cases}$$

به بیان دیگر مابه گزاره «احتمال آنکه عدد خیلی کوچک رو شود ۰/۶ است» درجه‌ٔ راستی ۰/۷ را نسبت

می‌دهیم.

۳ـ براساس رضایت کلمنت

$$\overset{*}{P}_{\gamma}(A)(x)=\begin{cases}1 & \text{اگر } 0<x\leq 1/2\\ 0/7 & \text{اگر } 1/2<x\leq 2/2\\ 0/4 & \text{اگر } 2/2<x\leq 3/2\\ 0/2 & \text{اگر } 3/2<x\leq 4/2\\ 0/1 & \text{اگر } 4/2<x\leq 5/2\\ 0 & \text{اگر } 5/2<x\leq 6/2\end{cases}$$

در این حالت می گوییم درجه‌ٔ راستی گزارهٔ «احتمال وقوع پیشامد A حداقل ۲/۶ است» برابر ۰/۷ است.

۴ـ براساس رضایت دوم یاگر

$$\hat{P}_{\gamma}(A)(x)=\begin{cases}0 & \text{اگر } 0<x\leq 1/2\\ 0/1 & \text{اگر } 1/2<x\leq 2/2\\ 0/2 & \text{اگر } 2/2<x\leq 3/2\\ 0/2 & \text{اگر } 3/2<x\leq 4/2\\ 0/1 & \text{اگر } 4/2<x\leq 5/2\\ 0 & \text{اگر } 5/2<x<6/2\end{cases}$$

در این حالت می گوییم درجه‌ٔ راستی گزارهٔ «احتمال وقوع پیشامد A دقیقاً برابر ۲/۶ است» برابر ۰/۱ است.

در این مثال ملاحظه می‌کنیم که تابع $\hat{P}_{\gamma}(A)$ شباهت زیادی با تابع $\overset{*}{P}_{\gamma}(A)$ دارد. در حقیقت این دو تابع در بسیاری از نقاط با خود یعنی در [۱و۲/۱] برابر بوده و در نقطهٔ دیگر یعنی در [۲/۱و۰] نیز تابع $\hat{P}_{\gamma}(A)$ حالت متغیر را نسبت به نقطهٔ دیگر دارد. این به‌دلیل مساوی بودن احتمالات اعضاء مجموعهٔ مرجع X است. مثلهای که خواهد آمد این وضعیت را به حالت کلی تعمیم می‌دهد.

نتیجه ـ فرض کنیم مجموعهٔ مرجع X دارای n عضو هم‌احتمال باشد یعنی

$$X=\{x_1,\dots,x_n\} \quad ; \quad P(x_1)=\dots=P(x_n)=1/n$$

و نیز فرض کنید تابع عضویت $\overset{*}{P}_{\gamma}(A)$ بصورت زیر باشد. (صعودی)

آنگاه طبق اینکه n زوج یا فرد باشد تابع عضویت $\hat{P}_{\gamma}(A)$ بصورتهای زیر خواهد بود.

$$P_Y^*(A)(x) = \begin{cases} p_1 & 0 < x \le 1/n \\ p_2 & 1/n < x \le 2/n \\ \vdots & \vdots \\ p_n & (n-1)/n < x \le 1 \end{cases}$$

$$\hat{P}_Y(A)(x) = \begin{cases} p_n & 0 < x \le 1/n \\ p_{n-1} & 1/n < x \le 2/n \\ \vdots & \vdots \\ p_{n-k+1} & k/n < x \le 1/2 \\ p_{n-k+1} & 1/2 < x \le (k+1)/n \\ \vdots & \vdots \\ p_{n-1} & (n-2)/n < x \le (n-1)/n \\ p_n & (n-1)/n < x \le n/n = 1 \end{cases}$$

الف) $n = 2k$

$$\hat{P}_Y(A)(x) = \begin{cases} p_n & 0 < x \le 1/n \\ p_{n-1} & 1/n < x \le 2/n \\ \vdots & \vdots \\ p_k & k/n < x \le (k+1)/n \\ \vdots & \vdots \\ p_{n-1} & (n-1)/n \le x < (n-2)/n \\ p_n & (n-1)/n < x \le 1 \end{cases}$$

ب) $n = 2k+1$

References:

1-E.P.Klement,Some remarks on a paper by R.R.Yagev,inform.SCi.27:211-220(1984)

2-A.L.Ralescu and D.A.Ralescu,Probability and Ruzziness,Inform.SCi.34:85-92 (1984)

3-R.R.Yager,A note on Probilities of fuzzy events,Inform.SCi.18:113-129 (1979)

4- --------,A representation of probability of a fuzzy sabset , Fuzzy sets and systems 13:273-283 (1965)

5- L.A.Zadeh,Fuzzy Set,Inform.cotrols 8:338-353 (1965)

6- --------,Probability measures of fuzzy events,J.Math.Anal.Appl.23:421-427 (1968)

7- --------,The concept of a linguistic variable and its application to approximate reasoning.part I,Inform.SCi.8:199-249;part II,Inform.SCi.8:301-357;part III,Inform.SCi.9:43-80.(1975)

8- --------,A theory of approximate reasoning, University of California at Berkley, 1977.

نقش مراکز اطّلاعاتی در پژوهشهای آماری

پرویز طه

مرکز آمار ایران

مقدمه

درجامعه فرا صنعتی‌امروز ، اطلاعات و نظامهای اطلاعاتی ، جایگاه ویژه و رفیعی در فعالیتهای بشری از قبیل برنامه ریزی ، سیاستگذاری ، تحقیق وتوسعه وبالاخره تصمیم گیری کسب کرده‌است . جهان‌غرب از مدتها پیش ، این پدیده را شناسایی کرده و برای بهره‌گیری موثر و بهینه آن ، سرمایه‌گذاریهای چندین میلیون دلاری به عمل آورده ونیروی انسانی کار آزموده و تکنولوژی‌پیشرفته‌ای رابرای شناسایی منابع اطلاعاتی و نظارت کامل برآنها به خدمت گرفته‌است .

قلمروی جامع‌اطلاعات با تاثیر پذیری ازتحولات عصرحاضر ، درجریان یک تحول سریع پیش‌می‌رود . ظرفیت رایانه ها افزایش می‌یابد ، بسته های نرم‌افزاری پیش ساخته قابلیتهای بیشتری کسب‌می‌کنند ، ارتباطات دوربرد توسعه می‌یابدوتمام‌اینها عوامل موثر برکارآیی وکاربرد اطلاعات بو شمارمی‌روند .

امروزه‌اطلاعات تبدیل به حربه‌ای برای اخذ امتیازات اقتصادی‌واجتماعی شده‌است . گسترش‌ودامنه گردش اطلاعات و سهم‌موثر آن در پیشرفتهای اجتماعی واقتصادی وانجام پژوهشهای علمی موجب‌بروز تحول بنیادی درجوامع‌شده‌است . اطلاعات به همان نسبتی که درتصمیم‌گیریها و برنامه ریزیها و روند پژوهش نقش اساسی دارد ، موجب بالارفتنِ بینشهای فردی وتغییررفتاراجتماعی می‌شود وبالاخره پدیده انفجار اطلاعات جوامع را مجبور می‌کند تاگردش تبادل اطلاعات را بااستفاده از تکنولوژی نوین سهارکنند وبه نحو مطلوب به‌جریان اندازند .

درچنین شرایطی ارزش حیاتی اطلاعات برای برنامه‌ریزی ، تحقیق و توسعه و رشدتکنولوژی‌کاملا" شناخته‌شده‌است و برای بهبود نظامهای اطلاعاتی وخدماتی که دراین راستاارائه می‌شود ، لزوم‌انجام سرمایه‌گذاریهای کلان مورد حمایت و تاکید قرارمی‌گیرد .

درفرایند تولید واستفاده ازاطلاعات ، عواملی چند دارای نقش بنیادی ومرکزی هستند . این عوامل رامی‌توان به چندگروه تقسیم کرد : ۱ ـ منابع اطلاعاتی که انواع اطلاعات مربوط به گستره دانش بشری وحاصل فعالیتهای مختلف اجتماعی ، اقتصادی و فرهنگی و پژوهشهای علمی و فنی را طبق نیاز جویندگان اطلاعات تامین و فردسترس می‌گذارند . ۲ـ نظامهای اطلاعاتی که با شیوه هـــای علمی به شناسایی ، گردآوری ، سازماندهی ، ذخیره و بازیابی اطلاعات و منابع اطلاعاتی می‌پردازند . ۳ ـ واحدهای اطلاعاتی یامراکز اطلاعات که درقالب کتابخانه ها ، مراکز مدارک واسناد ، مراکـــــز ارجاعی ، بانکهای اطلاعاتی وغیره براساس نظامهای متداول اطلاعاتی بهامر گردآوری ، سازماندهی ، بازیابی و اشاعه اطلاعات می‌پردازند . برای آشنایی بیشتر باماهیت و ویژگیها وکاربرد این عوامـــل ابتدا مفهوم اطلاعات وکاربردهای آن و سپس نظام اطلاعاتی را مورد بحث قرارمی‌دهیم . سپس بـــا اشاره به واحدهای اطلاعاتی و شرح انواع آن ، خدمات عمده اطلاعاتی را که این واحدهاارائه‌می‌کننـــد بیان می‌کنیم و سپس با توجه به نقش منابع اطلاعاتی وکاربرد وسیع آنها ، شمه‌ای ازآنهارا کـــه در انجام پژوهشهای آماری اهمیت بیشتری دارند به اجمال معرفی می‌کنیم .

مفهوم اطلاعات

اطلاعات ، پدیده ای بسیار پیچیده وگسترده‌است که تاکنون تعریف جامع ، روشن وهمه پذیـــری برای آن بدست نیامده‌است . گروهی آن را مترادف دانش ، آگاهی یا بصیرت و خبر تعـــریف‌کرده‌اند . عده‌ای ، اطلاعات را دانسته هایی می‌دانند که درتصمیم‌گیری بکار می‌رود و درواقع نوعی ارتبـــاط دوسویه میان تصمیم گیری واطلاعات قائلند . بدین ترتیب‌که ازنظرآنان ، اطلاعات فقط درتصمیم‌گیری بکارمی‌رود و تصمیم‌گیری نیز تنها بادراختیارداشتن اطلاعات امکان پذیراست . درواقع این‌تعریف بیشتر ناظر برکاربرد اطلاعات است تاماهیت آن .

دسته‌ای دیگر ، اطلاعات را داده ها یا دانسته هایی طبقه بندی شده می‌دانند که دریک‌نظـــام ، به شیوه‌ای خاص با یکدیگر ارتباط یافته و معنی و مفهوم خاصی به دست آورده‌است . بنابراین ، میان داده و اطلاعات تفاوتی وجود دارد . بدین‌گونه که داده ها ، بیانگر واقعـــیاتی با گستره‌وسیع ، خـــام وسازمان نیافته‌است ، مانند داده های کامپیوتری که به شکل اعداد ، کد ، حروف وکمیتهای قیاســـی وارد کامپیوتر شده و پس از داده آمایی یا پردازش ، به شکل اطلاعات ، قابل استفاده‌می‌شود . باایـــن تعریف ، اطلاعات ، زیرمجموعه ای از داده هاست که ویژگیهایی دارد . بدین صورت که سازمان‌یافتـــه و طبقه بندی شده‌است و می‌تواند صفحا و مفهومی خاص را انتقال دهد . قابل تبیین و تفسیـــراست و

مبنای تصمیم‌گیری قرارمی‌گیرد.

دراصطلاحنامه کتابداری ، اطلاعات به صورت " حقیقت یا مفهومی که از طریق رسانه هسته‌ای اطلاعاتی ارائه می‌شود" و نیز به عنوان اجزا وعناصراصلی هرنظام کنترل شده " تعریف شده‌است .

باتوجه به تعاریف و مفهومهای گوناگونی که ازسوی صاحبنظران و متخصصان برای اطلاعات‌ارائه شده‌است ، یک وجه‌اشتراک را از خلال تعریفها وتفسیرهای مختلف می‌توان برای اطلاعات بازشناخت وآن عبارت ازاین است که اطلاعات پدیده‌ای است که باایجاد آگاهی ، تصمیم‌گیریها را تحت تاثیر قرارمی‌دهد.

ویژگی دیگر اطلاعات که درحقیقت شاخص آن به شمارمی‌رود ، این است که باورود به یک‌ارگانیزم، ساخت آن را دگرگون می‌سازد وهرچه‌باراطلاعاتی‌بیشتر باشد، تاثیر آن برارگانیزم مربوط ، بیشتر خواهد بود . درحقیقت می‌توان گفت که اطلاعات ازغیراطلاعات باهمین‌ویژگی تمایز می‌یابد.

گفتنی است دراین مبحث ، اصطلاح اطلاعات به مفهوم‌گسترده‌آن به‌کاررفته‌است . بنابراین ، نمودکمی آن یعنی آمار نیز مطمح نظرقرارداشته‌است .

نظام اطلاعاتی

نظام اطلاعاتی درمفهوم کلی خود ، وسیله یامجموعه وسایلی‌است که برای‌ارتباط میان فردیاافراد با فرد یاافراد دیگر به کارمی‌رود واز ارتباطات شفاهی و تلفنی تاروشهای کاملا " خودکارکامپیوتری گردآوری ، ذخیره ، پردازش واشاعه اطلاعات و انتقال دوربردآن را دربرمی‌گیرد.

درمفهوم خاص نظام اطلاعاتی‌شامل‌مجموعه‌رویه‌هاوتجهیزاتی‌است‌که‌کارگردآوری،پردازش،ذخیره‌سازی وارائه اطلاعات را باشیوه‌ای سازمان یافته وباکارآیی بالاانجام می‌دهد وآن را از طریق واحدهای اطلاعاتی دردسترس جویندگان اطلاعات قرارمی‌دهد. به طورکلی نظام اطلاعاتی رامی‌توان چنین تعریف کرد : " نظامی که تبادل و داده آمایی اطلاعات را میسر می‌سازد ویاروندی نظام‌یافته برای جمع آوری ، آماده‌سازی ، ذخیره سازی و بازیابی اطلاعات به نحوی که پاسخگوی نیازهای‌مختلف اطلاعاتی باشد.

نیاز به نظامهای اطلاعاتی با افزایش سریع حجم اطلاعات مورد تقاضا به طور فزاینده‌ای رو به گسترش است ، به‌گونه‌ای که درعصر حاضر ، این نظامها ، جزء جدایی ناپذیر نظامهای مدیریت جوامع

پیشرفته وداده پردازی خو دکارشده‌اند .

ازانواع نظامهای اطلاعاتی می‌توان به عنوان نمونه به‌نظام اطلاعاتی تحقیقات جاری کشاورزی یا "کاریس" (Current Agricultural Research Information System) نظام‌بازیابی مدارك پزشکی یا (مدلار)Medical Literature Analysis and Retrieval System( ) ، نظام شبکه اطلاعاتی (Information Network System) ، که درمراکز اطلاعاتی نظیر کتابخانه های تخصصی مراکز اسناد و مدارك و مراکز پخش واشاعه‌اطلاعات کاربرد دارندو یا به‌نظامهای طبقه بندی مورد استفاده درفعالیتهای آماری واقتصادی نظیر طبقه بندی استاندارد بین‌المللی‌کلیه‌فعالیتهای‌اقتصادی ، طبقه بندی استاندارد بین‌المللی کالا ، طبقه‌بندی فعالیتهای آماری سازمان ملل متحد و طبقه بندی فعالیتهای آماری اسکاپ اشارەکرد .

واحدهای اطلاعاتی

واحدهای اطلاعاتی را شاید بتوان با ارزشترین و باثبات ترین سرمایه نظام اطلاعاتی محسوب داشت . جایی که بیشترین دانش و اطلاعات بالقوه موجود، جمع آوری ، سازماندهی ، ذخیره‌و برای بازیابی واشاعه آماده می‌شود . واحدهای اطلاعاتی شامل انواع کتابخانه ها ، مراکز اسنادومدارك مراکز اطلاعاتی ، مراکزارجاعی ، مراکز پخش و اشاعه اطلاعات و دیگر مخازن اطلاعاتی‌است که همگی‌اینها ازارکان برجسته نظام اطلاعاتی موجود یا برنامه‌ریزی شده به شمارمی‌روند .

الف . کتابخانه ها :

این واحدهای‌اطلاعاتی ، شناخته شده‌ترین و قدیمیترین انواع واحدهای اطلاعاتی‌هستند که به‌امر سازماندهی وذخیره مدارك دست اول ومنابع مرجع می‌پردازند . تاکید این واحدها بیشتربر روی مجموعه ها وامانت مدارك‌است .

ب . مراکز اسنادومدارك :

این مراکز علاوه براینکه به جمع‌آوری و نگهداری‌وذخیره مدارك می‌پردازند ، تاکید آنهابیشتر روی توزیع و اشاعه اطلاعات است . فعالیت عمده این مراکز شامل گردآوری ذخیره و بازیابی مدارك معین ، تهیه خبرنامه ، چکیده نامه ، ترجمه و نسخه برداری ازاصل مدارك و بالاخره ارزیابی

مدارك دست اول و دست دوم علمی و فنی است .

ج - مراکز پخش واشاعه اطلاعات :

این مراکز به عنوان یك نظام پردازش اطلاعات به شمار می‌روند که وظیفه اصلی آنها ، عملیات انتقالی روی اطلاعات است . این وظیفه را یا ازطریق ارجاع به منابع مناسب و یا به عنوان مرکزی که خود به جمع آوری ، طبقه بندی واشاعه اطلاعات تخصصی می‌پردازد ، انجام می‌دهند . وظایف این مراکز به‌طورعمده عبارتند از : کتابداری ، دکومانتاسیون ، تولید مواد سمعی و بصری ، انتشارات ، ارتباطات سازمانی و پردازش داده ها .

خدمات این مراکز شامل موارد زیراست :

جستجوی داده ها وجمع آوری مواد چاپی ، رابط بین گروههای شاغل دربخشهای تحقیقات ، چاپ و انتشار مواد چاپ نشده و جستجوبرای نیازهای استفادهکنندگان بدون آنکه درانتظار رسیدن آنان باشد و بالاخره ارزشیابی اطلاعات قابل دسترس وپردازش آنها .

د - مراکز ارجاعی :

این مراکز معمولا " وظیفهارجاع درخواستها به منابع اطلاعاتی را برعهده دارند . وظایف آنها عبارت است از جمع آوری اطلاعات درباره داده ها و منابع آنها دریك موضوع خاص ، تهیه فهرستهای جامع ازانواع داده ها و اطلاعات و خدمات قابل دسترس وراهنمایی استفاده کنندگان به منابع اطلاعاتی مناسب از طریق فهرستهای موضوعی .

هـ - مراکز اطلاعاتی

این مراکز ، غالبا " کار انتخاب ، سفارش ، ذخیره ، بازیابی واشاعه اطلاعات رادریك زمینه خاص موضوعی ویابك رسالت خاص برعهده دارند . این مراکز معمولا " خبرنامه یابولتنهای نمایه و آگهیهای تحقیقاتی منتشر می‌کنند و خدمات ترجمه رانیز ارائه می‌دهند . دراین مراکز داده ها در یك زمینه خاص ، به طور جامع گردآوری و سپس توسط متخصصان موضوعی مورد تحلیل وارزیابی قرار می‌گیرد و بعدازطریق خدمات آگاهی رسانی جاری ، خدمات اشاعه اطلاعات گزینشی ، انتشارات و پاسخگویی به پرسشها اشاعه می‌یابد .

و - مراکز داده ها :

این نوع مراکز ، به پردازش داده های خام یانتایج داده ها می‌پردازند - مثل داده های حاصل از آمارگیریها و نمونه‌گیریهای موضوعی و داده های مربوط به علوم محض و علوم اجتماعی - خدمات این مراکز به‌طور عمده‌شامل: ارزشیابی داده ها ، تالیف ، اشاعه اطلاعات و خدمات ارجاعی است .

غیر از واحدهای اطلاعاتی که به آنها اشاره شد ، منابع اطلاعاتی بخش دولتی و غیردولتی‌ازقبیل سازمانها و نهادهای اجرایی ، قوه قانونگذاری و قضایی ، واحدهای‌اداری محلی ، سازمانهای‌بین‌المللی و موسسات پژوهشی به سبب دارا بودن داده های آماری واداری ، به عنوان واحدهای اطلاعاتی‌قلمداد می‌شوند که بخش مهمی اززیربنای ملی اطلاعات را به صورت مخازن داده هادرخود جای داده‌اند .

خدمات اطلاعاتی

کلیه واحدهای اطلاعاتی ازانواع کتابخانه های سنتی گرفته تامراکز اطلاعاتی کاملا" پیشرفته ، هرکدام درقالب وظایف خود ، براساس امکانات موجود به تامین نیازهای اطلاعاتی مراجعان خاص خود می‌پردازندنوع خدمات این واحدها را برحسب ویژگی دنیای اطلاعات و نوع نیازاستفاده کنندگان ، می‌توان در ۸ گروه عمده به شرح زیر مشخص‌کرد:

۱ - زمانی که استفاده کننده ، از اطلاعات و آمار تولید شده دریك زمینه خاص آگاهی ندارد ، به عبارت دیگر درمواردی که بین تولیدکنندگان واستفاده کنندگان آماروااطلاعات فاصله‌بسیاری وجود دارد . دراین موارد ، خدمات آگاهی رسانی جاری ، خدمات ترویجی وانتشارات وتبلیغات کارسازهستند .

۲ - تراکم اطلاعات در مقیاس وسیع ، مراجعه کننده را باانبوهی عظیم ازاطلاعات درزمینه مورد نظر مواجه می‌سازدکه درنتیجه ، نیاز به‌انتخاب اصلح مطرح می‌شود . نظامهای‌نمایه سازی ، طبقه بندی اطلاعات وارائه خدمات مرجع ، راهگشای این مشکل هستند .

۳ - رشد سریع وروزافزون اطلاعات و درنتیجه ، افزایش مداوم میزان کهنگی اطلاعات ، لزوم دسترسی سریع به اطلاعات روز آمد درزمینه موضوعی‌خاص را مرکد می‌سازد . دراین مورد،نقدها ومرورهای موضوعی ، گزارشها و مطالعات مربوط به وضع موجود دریك زمینه تخصصی،تهیه خلاصه مطالب به صورت چکیده‌های تفصیلی ، تجزیه و تحلیل وارزشیابی اطلاعات از اهم خدمات قابل ارائه محسوب می‌شوند .

۴ ـ ماهیت میان رشته‌ای پراکندگی اطلاعات به‌ویژه درمواردی که استفاده کننده ، فقط در محدوده‌یک موضوع خاص و معین تخصص دارد . روشهای نمایه سازی برای موضوعهای میان رشته‌ای ، تجزیه و تحلیل اطلاعات و خدمات جستجو ، دراین زمینه کارسازهستند .

۵ ـ گوناگونی زبانها و وجود اطلاعات مورد نظر درمتون خارجی از یک سو وعدم آشنایی کافی با زبانهای خارجی ازسوی دیگر ، مشکلی است فراراه استفاده کنندگان ازاطلاعات . ارائه خدمات ترجمه درواحدهای اطلاعاتی ، رافع‌این مشکل‌است .

۶ ـ دامنه وسیع استانداردها وروشهای‌ارائه آنرا و عقاید که درمقابل آن ، فقط تعداد محدودی از استانداردها وطرحها برای بعضی ازاستفاده کنندگان مناسب‌است . دراین زمینه ، انتخاب وارائه یا تنظیم مجدد اطلاعات و تلفیق آن براساس نیاز استفاده کنندگان ، خدمت مناسب قابل ارائه‌است .

۷ ـ وجود اختلاف درکیفیت واعتبار داده های آماری درزمینه موضوعی مورد نظر وکمبود وقت یا عدم توانایی برای‌ارزشیابی وانتخاب داده های معتبر ، مشکل دیگر استفاده کنندگان‌از منابع اطلاعاتی به ویژه داده های آماری به شمارمی‌رود . خدمت قابل ارائه دراین زمینه ، تجزیه وتحلیل وارزشیابی داده ها واطلاعات است .

۸ ـ درمواردی که نسخه‌ای ازمدارک مورد نیازاست ، تهیه‌نسخه هایی ازمدرک بااستفاده‌ازوسایل تکثیر ونسخه برداری وانتقال دوربرد ، ازخدمات قابل ارائه به شمار می‌رود .

منابع اطلاعاتی

به طور کلی ، نوع منابع اطلاعاتی مورد استفاده درپژوهش وبرنامه ریزی ، منابعی‌است که‌صرف ـ نظر ازکیفیت اطلاعات ، برای استفاده کننده ، قابل حصول و مشخص‌باشد . ازجمع استفاده کنندگانی که به طورفعال درجستجوی اطلاعات‌اند دراین راه کوشش به خرج می‌دهند ، تعداد اندکی‌از منابع حاوی اطلاعات مختلف و سودمند و ویژگیهای آنها آگاهی دارند . شناخت منابع اطلاعاتی درارتباط با نیازهای دقیق جویندگان اطلاعات ، ضمن اینکه ، تحصیل اطلاعات دقیق و مطلوب را فراهم می‌آورد ، دروقت ، انرژی و تلاش مربوط به جستجوی اطلاعات نیز صرفه جویی قابل ملاحظه‌ای ایجاد می‌کند .

درواحدهای اطلاعاتی ، اطلاعات درزمینه های گوناگون ، به شکل مواد چاپی و غیر چاپی ودراشکال

مختلف فیزیکی نگهداری وارائه می‌شود. برخی‌ازاین اطلاعات که جنبه بنیادی دارند و به شیوه‌ای‌خاص تنظیم وارائه شده‌اند، منابع بنیادی نامیده می‌شوند. این قبیل منابع به سبب کاربرد وسیع وتخصصی آنها، از اهمیت ویژه‌ای برخوردار ند و به طورکلی ازنظر شکل و محتوا به منابع دسته‌اول، منابع دست دوم و منابع دست سوم تقسیم می‌شوند. منابع دست اول شامل هرنوع مواد بنیادی جدیدی است که برای نخستین بارارائه می‌شود،مانندگزارش نتایج اجرای طرحهای آماری، گزارشهای‌پژوهشی، انتشارات رسمی، رساله‌ها، صورتجلسه‌ها،کنفرانسها وکارپروژه ها. منابع دست دوم دنابعی‌هستند که عموما" درتهیه و تدوین آنها، از منابع دست اول استفاده می‌شود و براساس نظامهای مشخصی چون نمایه نامه ها، چکیده نامه ها و واژه نامه ها تنظیم می‌شوند. منابع دست سوم، منابعی است که حاوی اطلاعاتی درباره منابع دست اول و دست دوم باشد و جوینده را دراستفاده‌ازاین منابع یاری دهد، مانندکتاب شناسی کتاب شناسیها.

اینک به طور اجمال به برخی‌ازمنابع فوق که دراغلب پژوهشهای آماری اهمیت بسیاری دارند اشاره می‌کنیم:

۱ـکتاب شناسیها

با توجه به انبوه عظیم مواد منتشر شده درخلال سالهای اخیر و روند سرسام آور انتشارمنابع اطلاعاتی، جستجوی دقیق و تخصصی دراین مجموعه عظیم ازاطلاعات، به سبب گسترش سریع‌مرزهای دانش و تکنولوژی، تنها از طریق ابزارهای کتاب‌شناختی که هدف اصلی آنها، شناسایی، تعیین و انتخاب موادی است که از پیش دردسترس قرارداشته‌است، امکان پذیرمی‌شود. ابزارهای‌کتاب‌شناختی ازجمله کتاب شناسیها و مقاله نامه ها درکنترل سیل اطلاعات و تعیین و شناخت مطالب موردنظر نقش مهمی‌راایفا می کنند.

۲ـنمایه ها

نمایه عبارت از راهنمای نظام یافته مندرجات هرنوع نوشته‌ای‌است که درآن‌شناسه یا واژه های کلیدی‌به ترتیب خاص وبارجاع به محل معینی مانند صفحه یا مدخل مورد نظر تنظیم می‌شود. هدف نمایه کمک به جوینده‌ای‌است که مایل است اطلاعات مورد نظر را به طورسریع ومطمئن دردسترس‌داشته باشد. نمایه شامل انواع مختلف است که مهمترین آنها نمایه کتاب‌ونمایه‌نشریات ادواری‌یا پیوسته‌هاست نمایه کتاب، متداولترین نوع نمایه است که حاوی نمایه عنوان و نمایه موضوع است و معمولا" درانتهای

کتابهای علمی یا درمجلدهای جداگانه منضم به کتاب و نشریات علمی درج می‌شود. نمایه نشریات ادواری ، نمایه مقاله‌های پیایندهاست که اطلاعات کتاب شناختی یا خصوصیات مقاله رادردسترس جویندده قرارمی‌دهد. نمایه مقاله ها درنشریات‌ادواری کاری را انجام می‌دهد که کتاب شناسیها درمدیای کتاب‌انجام‌گردیدش‌آن‌اند فهرست مندرجات نیز یک نوع نمایه مقاله های‌نشریات ادواری به شمار می‌رود.

۳ ـ چکیده نامه ها

همان طورکه گفتیم شما یه ها غالبا" شامل اطلاعات کتاب شناختی مقاله هایی است که در فهرست وارد می‌شود. ولی نمایه ، محتوای مقاله ها را مورد بحث قرارنمی‌دهد. گاهی‌اوقات نیز، عنوان مقاله گمراه کننده‌است وممکن‌است محتوای دقیق مقاله را مشخص نکند. چکیده نامه ها برای رفع این مشکل و برای پاسخگویی به‌این گونه مسائل و ارائه خلاصه و بازنمودی صحیح از محتوای مقالهها تهیه و تدوین می‌شوند. چکیده ها نشان می‌دهند که چه چیزی کشف شده ، چه چیزی درحال‌اکتشاف است‌یاچه‌چیزی تحت پژوهش علمی قرارندارد. به عبارت دیگرچکیده ها کمک می‌کنند تاخواننده ، اطلاعات خود را درمورد پیشرفتهای حاصل دررشته مورد نظر ، روز آمد نگهدارد. همچنین چکیده ها علاوه برجلوگیری از اتلاف وقت ، باارائه محتوای خلاصه نوشته ها موجب می‌شوند که جوینده‌اطلاعات ، نسبت به ادامه مطالعه و بررسی اصل مقاله تصمیم‌گیری کند.

۴ ـ منابع نقد و بررسی

این منابع ، وظیفه استخراج ، ارزیابی وگردآوری‌اطلاعات را برعهده دارد و پژوهشگران رادربهره‌برداری از داده های اطلاعاتی وآماری یاری می‌دهد. این منابع ، اطلاعات وآمار ارزیابی شده‌ای را که به‌رشته های موضوعی خاص ، ارتباط دارد ، به پژوهشگران عرضه می‌کند وازاین نظر ، درانتقال آراوعقاید بسیار موثر است. ازمهمترین منابع نقد و بررسی ، پیایندهای نقد وبررسی است که به‌انتشار بررسیهای وضع موجود ، بررسیهای انتقادی ، تفسیر ی وارزیابی شده می‌پردازند. نام‌این منابع درمتون خارجی معمولا" باعناوینی چون Annual Review , Advances , Comments و Essays شروع می‌شود.

۵ ـ صورتجلسه‌های کنفرانسها

این قبیل منابع را می‌توان به عنوان بررسی وضع موجود دریک موضوع اختصاصی قلمدادکرد.

سازمان دهندگان کنفرانسها ، معمولا " نقدهایی‌را به عنوان معرفی هرجلسه از کنفرانس مطرح می‌کنند که عموما " نقدهای خوبی به شمارمی‌روند .

۶ـ فرهنگهای جغرافیایی

این منابع شامل اسامی جغرافیایی وتقسیمات کشوری واطلاعاتی درباره آبادیها ، دهستانها ، بخشها ، رودخانه ها ، کوهها ، جمعیت ، خانوار وخصوصیات اجتماعی واقتصادی واحدهای کوچک جغرافیایی درتقسیمات کشوری است که همراه با نقشه های موضوعی ، نیاز جویندگان اطلاعات دراین زمینه را بر طرف می‌سازند .

۷ـ دستنامه ها و دستورنامه ها (Handbooks and Manuals)

دستنامه مرجعی است که موضوعات مهم و بنیانی یک رشته را به طور اختصار بیان می‌کند و خواننده را دریادگیری جنبه های خاص یک موضوع یاری می‌دهد . تاکید دستنامه ها دراکثر موارد بر بنیادهای ثابت و مورد قبول دانشمندان یک رشته استواراست.تابیان آخرین پیشرفتهای حاصل دریک زمینه . دستورنامه حاوی دستورالعملهایی است مبنی براینکه‌مراحل انجام کار دریک زمینه خاص چگونه باید باشد . دستورنامه که به نام کتاب آموزشی ودستورالعمل‌نیز نامیده می‌شود ، می‌کوشد به سوالاتی که به یک یا چند فعالیت خاص مربوط می‌شود ، پاسخ مقتضی دهد . به طور خلاصه هدف دستورنامه ، هدایت افراد به هدف وروش شناسی اموروسازماندهی است .

۸ـ سالنامه ها

سالنامه به طورکلی ، تلخیصی است از داده ها وآمارهای مختلف مربوط به رویدادهای گذشته در یک کشور ، معمولا سالنامه نشان دهنده روند رشد و توسعه درزمینه های گوناگون‌است . تغییرات حاصل دررشد وتوسعه را ازطریق جداول آماری که به‌صورت سریهای زمانی‌تنظیم شده‌است به سهولت می‌توان مورد بررسی قرارداد . ازآنجا که سالنامه های معتبر درنقل اطلاعات خود به سایر منابع‌اطلاعاتی استناد می‌کنند ، می‌توان آنهارا به عنوان فهرستهای غیر رسمی نیز مورد استفاده قرارداد .استفاده‌های خاص به منابع آماری بخش دولتی ، خواننده سالنامه را به مواد و منابع عمده درزمینه های موردنظر راهنمایی می‌کند . ازانواع سالنامه ها می‌توان به سالنامه آماری اشاره کرد . این نوع سالنامه ها ، نشانگر اوضاع اجتماعی ـ اقتصادی و فرهنگی یک کشور دریک دوره زمانی خاص به شمار می‌روند .

۹ ـ انتشارات دولتی

این منابع ، منابعی قابل اطمینان ، معتبر وروز آمد هستند و هیچ واحد اطلاعاتی هرچند غنی از نظر مجموعه، بدون دسترسی به انتشارات دولتی ، قادر به ارائه خدمات مرجع کافی نیست . این منابع تقریباً " درکلیه زمینه ها انتشار می‌یابند وازنظر بررسیهای علمی وکسب اطلاعات مرجع منابع کاملا " مفیدی به شمار می‌روند . دراین منابع ، گزارشهای دولتی واطلاعاتی وجود دارد که در سایر منابع نمی‌توان به آنها دسترسی پیداکرد . این منابع به صورت کتاب ، جزوه ، کاربرگه ، مجله ، گزارش به ویژه گزارشهای آماری و دولتی وجود دارد .

شبکه ملی اطلاعات

حال که با واحدهای اطلاعاتی ، ویژگیها وکاربرد آنها آشنا شدیم ، ضرورت دارد به انتظاراتی که جامعه استفاده کننده از آماروااطلاعات ، درحال حاضر ازاین واحدها دارد ، اشاراتی داشته‌باشیم .

کتابخانه ها بنابرشیوه سنتی خود بیشتر به‌امر تهیه ، رده‌بندی وذخیره مدارک تاکید دارند و به تنظیم مجدد اطلاعات و اشاعه آن رغبتی نشان نمی‌دهند . حال آنکه دراثر رشد تصاعدی اطلاعات که به حق‌انفجار اطلاعات لقب یافته‌است ، وتوسعه سریع نرم‌افزارهای پیش ساخته و وسایل ارتباط دوربرد و نیز نیازهای نوین جامعه استفاده کننده ازآماروااطلاعات ، این دو وظیفه یعنی تنظیم مجدد اطلاعات واشاعه آن ، به طور موکدی مورد توجه قرارگرفته است .

به طور کلی از واحدهای اطلاعاتی انتظار می‌رود که یک سلسله خدمات اطلاعاتی کارآمد وجامع درتمام‌زمینه های موضوعی ارائه دهند واین نیازی‌است که به طور روزافزون از جانب پژوهشگران ، برنامه ریزان و سایر استفاده کنندگان‌ازاین‌واحدهامطرح می‌شود .

ازسوی دیگر ، تفاوتهای موجود دربین سازمانهای مختلف تولیدکننده‌آمار واطلاعات ازلحاظ چگونگی تولید اطلاعات و نحوه ارائه آن ، لزوم ایجاد ارتباط مستمر و هماهنگی بین این سازمانها و پیروی از تعاریف ، طبقه بندیها و روشهای استاندارد مشترک را مورد تاکید قرارمی‌دهد . تامین‌این دونیاز به صورت مطلوب ، از طریق ایجاد شبکه ملی اطلاعات با پیش بینی تمهیدات وانجام‌اقداماتی به صورت مرحله‌ای و تدریجی امکان پذیرخواهد بود .

به طور کلی ایجاد شبکه ملی اطلاعات را می‌توان به سیر تولید یک کالا تشبیه کرد که‌ازمرحله تولید تامصرف ادامه دارد . کالای مورد نظر ما اطلاعات و داده های آماری است که نتیجه یامحصول

شبایی فعالیتهای آماری و اطلاعاتی کشور به شمار می‌رود . و سیر تولید ، رشته فعالیتهایی است که شامل ایجاد طبقه بندیها و تعاریف ، روشهای اجرای طرحهای آماری ، پردازش اطلاعات ، تامین و آموزش نیروی انسانی متخصص ، تامین بودجه وامکانات لازم وبالاخره اشاعه اطلاعات خواهد بود . این شبکه کلیه سازمانهای تولید کننده آمار ، واحدهای اطلاعاتی و موسسات پژوهشی را دربرمی‌گیرد .

اطلاعاتی که دراین شبکه ازمرحله تولید تامصرف ، مورد تبادل قرارمی‌گیرد ، درزمینه هایی خواهد بود که به طورعمده ، نیازهای آماری کشور را منطبق با اهداف توسعه اجتماعی ، اقتصادی و فرهنگی تامین کند . برای تحقق این امر ، انجام اقداماتی اساسی و فراگیر ضرورت دارد که به اهم آنها فهرست واراشاره می‌کنیم .

۱ ـ اطلاعات به عنوان یکی از عوامل توسعه مورد توجه کافی قرارگیرد وسیاستهای ملی اطلاعات هماهنگ با برنامه‌های توسعه اجتماعی ـ اقتصادی کشور تنظیم شود .

۲ ـ اتخاذ سیاستهای کارا برای ارتقای فعالیتهای سازمانهای تولید کننده آمار واطلاعات و واحدهای اطلاعاتی ، شامل توسعه زیر بنایی فعالیتهای تولید آمار واطلاعات وگسترش خدمات اطلاع رسانی .

۳ ـ مشارکت درمنابع و فعالیتها برای کنترل انفجار اطلاعات ازطریق فعالیتهای مشترک وسازمان یافته درسطح ملی .

۴ ـ زمینه های موضوعی وفعالیتهایی که برای برنامه ریزیهای توسعه ازاولویت برخوردارنسبت ، مورد توجه خاص قرارگیرند و اطلاعات وآمارمربوط به آنها با همکاری منظم وسازمان یافته واحدهای تولید کننده آمار ، ازطریق شبکه های اطلاعات تخصصی فراهم و دردسترس گذاشته شود .

۵ ـ برای کنترل اطلاعات وآمار تولید شده ، روش هماهنگ برای گردآوری و پردازش داده ها اعمال شود .

۶ ـ ایجاد پایگاههای داده ها درسطح هر واحد اطلاعاتی و درسطح ملی .

۷ ـ استفاده از تکنولوژی جدید درامر ذخیره‌سازی ، پردازش ، بازیابی ، انتقال دوربرد واشاعه اطلاعات .

۸ ـ تهیه وبکارگیری استانداردهای مشترک برای جنبه‌های گوناگون فعالیتهای تولید آمار وخدمات اطلاع رسانی .

ازجمله اقداماتی که درزمینه های فوق باید به صورت هماهنگ ودر قالب استاندارد انجام شود می‌توان به موارد زیر اشاره کرد :

الف) تعاریف ومفاهیم طبقه‌بندیها وروشهای آماری

ب) رده‌بندی موضوعی آمارهای تولیدی یا مورد نیاز

ج) شکلهای مشترک ارتباطی برای انتقال داده‌ها از طریق نظامهای کامپیوتری

د) تهیه کتاب شناسی ، چکیده نامه و نمایه برای گزارشها وضابع آماری

هـ) طراحی پایگاههای داده ها ،نظامهای ارتباطی ،شیوه های ارتباط ،نرم افزارها وزبان ارتباطی مشترک

و) خدمات امانت ،مرجع واشاعه اطلاعات گزینشی

بدیهی است تحقق این هدف بستگی به مشارکت فعال و همکاری همه جانبه سازمانهای ذی ربط دارد که درقالب اصولی مشخص ومدون ودرچارچوب نظامهای ملی ،این امر خطیر رابه انجام رسانند .

منابع وماخذ :

۱ ــ آترتون ،پاولین . مبانی نظامها وخدمات اطلاعاتی . تهران : وزارت فرهنگ وآموزش عالی . مرکز اسناد و مدارک علمی ،۱۳۶۳ . فصلهای اول ،پنجم ، ششم و هشتم .

۲ ــ احسان اغلو ، اکمل الدین ، و دیگران . شبکه اطلاع رسانی در کشورهای اسلامی . ترجمه عبدالحسین آذرنگ و دیگران . تهران : دفتر پژوهشهای فرهنگی ،۱۳۶۹ .

۳ ــ حری ، عباس . " بانک اطلاعات فرهنگی " . فرهنگ وسینما . شماره ۲ (اردیبهشت ۱۳۶۹) . ص ۱۶ ــ ۱۴ .

۴ ــ مهراد ، جعفر . مقدمه‌ای بر خدمات مرجع عمومی . شیراز : دانشگاه شیراز ، ۱۳۶۷ .

۵ ــ یونسکو . اطلاعات علمی و فنی در آسیا و اقیانوسیه . بررسی کلی ، نیازها و پیشنهادها . ترجمه سیمین بردبار . تهران : وزارت فرهنگ وآموزش عالی . مرکز اسناد و مدارک علمی ، ۱۳۶۵ .

۶ ــ یونسکو . یونیسیست : راهنمای ایجاد و توسعه مراکز ارجاعی . ترجمه محمدنقی میهدوی . تهران : وزارت فرهنگ و آموزش عالی . مرکز اسناد و مدارک علمی ،۱۳۶۳ .

کاربرد آمار و اطّلاعات در مدیریّت و اقتصاد از طریق طرّاحی و اجرای سیستم اطلاعات مدیریّت

جعفر عسگری

مرکز آمار ایران

۱ـ اطلاعات واهمیت آن

در صنعت تولید اطلاعات ، کوچکترین ذره اطلاعاتی را BIT ، ترکیب چند Byte را با هم و بالاخره ترکیب چند Byte را داده یا DATA می نامند .

ایجاد ارتباط بین چند DATA را اطلاعات یا INFORMATION می نامند . به ترتیب مزبور اطلاعات دارای دو یا چند بعد است و قابل تعبیر می باشد . به عبارت دیگر داده‌ها از طریق پرورش به اطلاعات تبدیل میشود . برای اطلاعات تعاریف مختلفی شده است " اطلاعات فرآیندی است که از طریق آن به داده‌های مفید انسان افزوده میشود و یا اطلاعات دریافت و تعبیر هر پیامی است که موثر در تصمیم گیری باشد " پایه اصلی هر تصمیمی اطلاعات است و بدون آن نمی توان تصمیم درستی اتخاذ کرد .

انسان در طول عمر طولانی خود ، در ارتباط با اطلاعات واستفاده از آن چهار عصر را پشت سر گذاشته است . در عصر شکار و در عصر کشاورزی برای اداره کردن انسانها وانجام تولید از چماق و شلاق استفاده شده است . در عصر صنعتی تفکر مدیریتی برای شناخت انسان وانگیزه‌های آن برای تولید بیشتر ، مدیریت علمی و اطلاعات را مطرح می کند .

در این عصر سعی شده که اطلاعات به دانش تبدیل شود وکاربرد آن به کمال سر پرستی و مدیریت انجام گیرد .

در عصر چهارم ، یعنی عصر اطلاعات ، که از سال ۱۹۵۷ آغاز میشودو درسال ۲۰۰۰ پایان خواهد یافت ، تولید اطلاعات واستفاده از آن در تصمیم گیری ، بتدریج ، در اولویت قرار می گیـــرد . در این عصر تصمیم خوب و مفیـــد به تصمیمی گفته میشود که ۹۰ درصد آنرا اطلاعات و ۱۰ درصد آنرا احساس و یا بینش مدیرتشکیل دهـــد . روش مدیریتـــی این عصر روش مدیریت مشارکتـــی یا "مدیریـــت توام با مفاهمـــه" اســـت . دراین روش چنین فرض میشود که اگر مردم اطلاعات صحیح در اختیارشـــان قرارگیـــرد می توانند تصمیمات درست اتخاذ کنند و درنتیجه درست کار کننـــد . در این عصـــر برای ایجاد تفاهم و تبادل اطلاعات با گروه کثیر وگســـترده از تکنولـــوژی اطلاعات ، که مهمترین ابزار آن کامپیوتـــر است ، استفاده میشود که هم سرعـــت و هم دقـــت را بهمراه دارد .

درحال حاضر حجم اطلاعات علمی در دنیا بهحدی است که کمترمحقق و دانشمندی می‌تواند بـــه مجموعه اطلاعات تخصصی خود اشراف و یا دسترسی پیدا کنـــد . حجم مجلات علمی هر ۱۵ ســـال یکبار ، تعداد محققیـــن هر ۲۰ سال یکبار وجمعیت جهان هر ۵۰ سال یکبار دو برابر میشود . طبق برآوردهای موجود ، هزینه دوبارهکاری در تحقیقات برابر است با ده درصـــد مجموع هزینههایـــی که صرف تحقیقات می شود که دلیل عمده آن عدم اطلاع ازجامعه اطلاعات است ضمن آنکه حـــدود ۱۲۰۰ بانك اطلاعات علمی در سطح بین‌المللـــی برای تبادل اطلاعات فعال می باشند ، درصـــد اشتغال در صنعت تولید اطلاعات در ایالات متحده آمریکا ، اکنون بر بخشهای دیگـــر پیشی گرفته اســـــت .

عصر پنجم ، که به آن عصر " فرااطلاعات " نیز گفته میشود ، ازسال ۲۰۰۰ آغاز میشود . در این عصر اطلاعات به عنوان کلید اصلی تحقیق ،توسعه وانتقال تکنولوژی محسوب میشود . تکنولوژی بیش از پیش خصلت اطلاعاتی می یابد وطبق پیش‌بینی دانشمندان جامعه شناسی وعلوم ارتباطات شتاب در پیشرفت تکنولوژی اطلاعاتی باعث انفجار دانشها واطلاعات می شود واز آن طریق تکامل مغز بشری جای تکامل زیستی را میگیرد . بشر به‌مرحله تازه‌ای ازتحول تمدن پا خواهدگذاشت که برمبنای دانش و اطلاعات شکل گرفته است .

آ- آمار و اهمیت آن

آن بخش از اطلاعات پرورده شده که جنبه کمی دارد اطلاعات آماری نامیده میشود ، هر مجموعه ای ازارقام واعداد (اطلاعات پرورده شده) مربوط به طبیعت وفعالیتهای اقتصادی و اجتماعی جامعه رااطلاعات آماری یا آمار می نامند . مفهوم دیگر آمار علم‌آمار است که اصول و روشهای جمع‌آوری اطلاعات آماری ،تابش دادن آنها ،تجزیه وتحلیل و استنتاج آماری رامورد بحث قرار می دهد .

تا قبل از ۱۹۵۰ برای جمع‌آوری ،تجزیه و تحلیل و تعبیرو تفسیر اطلاعات آماری کمتر از تکنیکهای آماری استفاده می شد وکلمه آمار به اندازه ، مفهوم "آمار" توسعه وگسترش نیافته بود .

در چند دهه اخیر جنبه‌های نظری و تحلیلی آمار در ارتباط با طراحی ،جمع‌آوری ، بازبینی و بالابردن و استخراج ،تجزیه وتحلیل وتعبیر آمارها دچار تحول اساسی شده‌است .

تحولات وتوسعه آماردرسالهای اخیر نتیجه استفاده وکاربرد اطلاعات‌آماری در زمینه‌های مختلف اقتصادواجتماعی توسط دولتها و بخش خصوصی بوده است .

فکرکردن به آمار و استفاده ازآن از جمله کاربردهای روزانه ماست، و درسایه آن ضمن کسب تجربه ، برای بسیاری ازموضوعات زندگی خودمان رامعمولا کسب میکنیم ، ماعمولا " در امورات شخصی، مسایل خود را در قالب فرمولهای پیچیده علمی مطرح وحل میکنیم امادرمسایل وامورات زندگی

اجتماعی استفاده از آمار تسهیلات فراوانی را برای حل مسایل مزبور فراهم می کند زیرا :

- آمار امکان میدهد که افراد موضوعات ، و ضمیمت‌ها و پدیده‌ها را براساس واقعیتهای آنها تعریف و توصیف نماییم .

- آمار امکان میدهد که تشخیص ومقایسه و سنجش موضوعات و رفتارها صورت گیرد .

- آمار امکان میدهد که با تشخیص صحیح از ناصحیح محتمل از غیر محتمل و مربوط از نامربوط تصمیمات عاقلانه ای اتخاذ شود .

- آمار اساس تحقیقات و مطالعات وزبان بین‌المللی است وامکان میدهد که مکالمه بین ما و جهان خارج انجام پذیرد .

- آمار امکان میدهد که اتخاذ تصمیمات و فرموله کردن اقدامات بگونه ای صورت گیرد که از حداقل خطرات و زیانها برخوردار شود .

۳ - تولید کنندگان اطلاعات آماری

تولید کنندگان اطلاعات آماری را در هرکشور به شرح زیر می توان طبقه بندی نمود :

الف - پژوهشگران فردی

آمارهای تولید شده توسط پژوهشگران فردی نتیجه تحقیقات و بررسیهای تجربی است که معمولا " با هدف کشف وقایع علمی صورت میگیرد . چنین اطلاعاتی در علوم طبیعی واجتماعی در زمینه هایی مانند شیمی ،زیست شناسی ،روانشناسی ، اقتصاد، کشاورزی و صنعت مطرح میشود واین اطلاعات معمولا " به یک یا چند مسئله مشخص و در سطوح کوچک مربوط میشود . آمارهای تولید شده توسط پژوهشگران فردی ویا موسسات علمی تحقیقاتی ، در صورتیکه چاپ شوند، معمولا " در گزارشها ی فنی و نشریات تخصصی آنها انعکاس می یابند .

ب - سازمانهای فردی

سازمانهای بازرگانی وصنعتی ،ارگانهای محلی نظیر شهرداریها ، واحدهای حمل ونقل نظیر

شرکتهای کشتیرانی و هواپیمایی و سایر سازمانها در ارتباط با فعالیتهای خود تولید آماری کنند.
اکثریت این اطلاعات از فرمها و اوراق اداری و ثبتی این سازمان ها نظیر لیست حقوق، حساب
مشتریان و فروش، گزارش پیشرفت کار و صورتحساب ها قابل استخراج میباشند.
این سازمانها در مواردی نیز با انجام آمارگیری (تکمیل پرسشنامه) اطلاعاتی را از داخل و یا خارج
سازمان جمع آوری میکنند. این پرسشنامه‌ها ممکن است برای پاسخگویی به سئوالات مشخصی
در زمانهای متناوب و یا آمارگیری موردی برحسب نیاز سازمان صورت گرفته شده باشد.
همچنین در مواردی اتحادیه‌های مربوط به چنین سازمانهایی (مانند سازمان صنایع ملی ایران،
اتحادیه تولید کنندگان وسایل گازسوز و ...) در جهت نیاز جمعی اعضاء وابسته خود و یا
همکاری آنها آمارهائی را جمع آوری میکنند. این اطلاعات برمبنای اطلاعات آماری دست اولی
که توسط اعضاء تولید میشود جمع آوری میشود. بخشی از این اطلاعات که معمولا" جنبه محرمانه
ندارد، برای تامین نیازهای داخلی روزنامه‌ها، اتاق بازرگانی و دولت در اختیار آنها قرار میگیرد.

ج - دولت

بزرگترین تولید کننده آمار و اطلاعات در هر کشوری دولت است. اطلاعات آماری که توسط
دولت تولید میشود در ارتباط با رفاه کشور و اقتصاد ملی میباشد. بخش قابل توجهی از آمارهای
دولت از اوراق و اسناد ثبتی و اداری که روز به روز برمبنای فعالیتها و وظایف جاری سازمانهای
دولتی انجام میگیرد، و بعضا" بموجب قوانین و مقررات می باشد صورت میگیرد.
مثال بارز این موضوع در کشورمان آمارهای تولیدی ثبتی مربوط به صادرات و واردات است که
برمبنای اطلاعات ثبتی مندرج در برگ سبز گمرکی صورت میگیرد. نمونه دیگر اطلاعات
مربوط به مشخصات پروانه‌های ساختمانی صادره میباشد که متقاضیان احداث ساختمان
ناچارند آنرا تکمیل نموده و در اختیار شهرداریها قرار دهند. سایر آمارهای مربوط به دولت
براساس نیازهای آماری موجود با اجرای آمارگیریهای نمونه‌ای و یا سرشماری سالانه یا هر چند سال
یکبار، جمع آوری میشود. اجرای این آمارگیریها ممکن است مستقیما" توسط دولت و یا واحدهایی

که از طرف دولت برای اینکار تعیین شده است صورت بگیرد .

آمارهای تولید شده توسط دولت که ماهیتا "جنبه محرمانه ندارد، در گزارشهای آماری (نشریات بولتن‌های آماری، مقالات آماری، سالنامه‌های آماری) منتشر میشود .

آمارهای تولید شده توسط دولت که غالبا " به عنوان آمارهای رسمی نامبرده میشود گستره وسیعی از فعالیتهای اقتصادی واجتماعی را در پوشش قرار میدهد . طبقه‌بندی توصیه شده توسط سازمان ملل در این زمینه در نمودار شماره ۱ آورده شده است .

در کشور جمهوری اسلامی ایران مرکز آمار ایران بعنوان سازمان مرکزی تولید آمار، بانک مرکزی جمهوری اسلامی ایران، وزارت کار و امور اجتماعی، وزارت کشاورزی، وزارت بهداری، وزارت صنایع ووزارت کشور ازجمله مهمترین سازمانهای تولید کننده آمار در بخش دولت می باشند .

د- سازمانهای بین‌المللی

سازمان ملل متحد و آژانسهای وابسته به آن نظیر کمیسیون اقتصادی واجتماعی آسیا واقیانوسیه (ESCAP)، موسسه بین‌المللی کار (ILO)، سازمان بهداشت جهانی (WHO)، سازمان خواربارو کشاورزی جهانی (FAO)، صندوق بین‌المللی پول (IMF)، بانک آسیایی (ADB) از جمله سازمانهایی هستند که در زمینه‌های مربوط به فعالیتهای خود آمارهای بین المللی را تولید می کنند . این موسسات با ارسال پرسشنامه به سازمانهای آماری کشورها وجمع‌آوری اطلاعات لازم، آمارهای تطبیقی کشورها را در نشریات خود منتشر می کنند . این موسسات همچنین روشها و استانداردهای آماری که کشورها را در تولید آمار یاری می کنند و باعث قابلیت مقایسه آمارها در سطوح بین‌المللی شود توسعه و گسترش میدهند .

۴ - استفاده کنندگان آمار

استفاده کنندگان اطلاعات آماری را نیز می‌توان به چهار طبقه پژوهشگران و محققین فردی شرکتها، موسسات و سازمانهای فردی، دولت وموسسات بین‌المللی تقسیم بندی نمود که

در این‌میان دولت همواره به‌عنوان مهمترین وبزرگترین استفاده کننده اطلاعات محسوب میشود .

منابع مزبور اطلاعات آماری را عموماً " باهدف :

- تصمیم گیری در مسایل روز

- ارزشیابی برنامه‌ها و فعالیتهای گذشته

- برنامه ریزی و پیش بینی فعالیتهای آینده

مورد استفاده قرار میدهند .

نقش اطلاعات آماری صحیح وبهنگام درسلامت وکیفیت تصمیم گیریها توسط مدیران ، در ارزشیابی برنامه های گذشته باهدف تشخیص نقاط قوت و ضعف عملکرد برنامه‌ها و در سیاستگزاری وپیش‌بینی برنامه‌های آینده برای توسعه اقتصادی کشور درسطح خردوکلان غیرقابل انکار است .

توسعه استفاده از آمار و اطلاعات صحیح وبهنگام برای اهداف فوق‌الذکر مخصوص در پرتوپیشرفت فوق العاده تکنولوژی کامپیوتری ، ضمن افزایش درصد دقت ارزشیابی‌ها وپیش‌بینی برنامه‌ها، باعث توسعه‌فعالیتهای آماری‌توسط‌کشورهای مختلف دراکثر زمینه‌های اقتصادی و اجتماعی‌شده است . امروز در ایران بیش‌از ۸۰ درصد سازمانهای مهم دولتی وخصوصی کشوردارای فعالیت تهیه آمار می باشند وسالانه حدود ۳۰ میلیارد ریال برای جمع آوری ، استخراج و انتشاراطلاعات آماری صرف هزینه میشود . این درحالی است که از یک طرف توجه دولت و مسئولان آمار کشور هرسال در جهت توسعه فعالیتهای آماری ، تکمیل خلاءهای آماری موجود ، تهیه چارچوب های آماری دقیق با پوشش کامل و انجام آمارگیریهای نمونه‌ای و سرشماریها در تزاید می باشد و از طرف دیگر توجه دانشگاهها و موسسات آموزشی در جهت توسعه و اشاعه علم آمار و تربیت نیروی انسانی متخصص آماروافزایش دوره‌های دکتری،کارشناسی ارشدوکارشناسی وکاردانسی بسیار محسوس میباشد . طراحی واجرای سرشماری فعالیتهای اقتصادی ، برای اولین بار در ایران ، درسال ۱۳۷۲ توسط‌مرکزآمارایران باهدف شناسایی دقیق فعالیتهای اقتصادی در ایران و تهیه چارچوب آماری از آنها وتهیه اطلاعات اقتصادی لازم مصداقی برتوجه دولت

در سرمایه گذاری در بخش آمار و توجه مسئولان آماری کشور برای استحکام واستفنای منابع آمـــار موجود میباشد. همچنین برگزاری اولین کنفرانس آمار در اصفهان و نیز تاسیس انجمن آمـــار ایـــران در سال جاری نشانه دیگری ازتوسعه واشاعـــه علم آماروکاربردهای آن نزدمسئولیـــــن و دانشگاهیان بشمار میرود.

یکی ازمصادیق مهم استفادهازاطلاعات آماری در اقتصاد تهیه جدول داده - ستانده اقتصاد ایـــران می باشد که درحال حاضـــر توسط مرکز آمار ایران در دست محاسبه می باشد. در تهیه این جـــدول که دارای ۷۹ سطر و ستون می باشد ، با اجرای آمارگیریهای نمونه ای در ۴۸ زمینه از فعالیتـهای اقتصادی (کشاورزی معادن ، صنعت ،آب و برق و گاز ، ساختمان ، بازرگانی وهتلداری و رستوران، حمل‌ونقل و ارتباطات وخدمات) آمار تولید و هزینه‌های بخشهای اقتصادی جامعه جمع آوری شده و درحال حاضر محاسبات‌فنی آن در دست انجام می باشد . نتایج حاصل از محاسبات این جـــدول ضمن ارائه ضرایب فنی‌روابط‌بین بخشهای اقتصادی ، شاخصهای کلان اقتصاد جامعه مانند تولید ناخالص داخلـــی ، هزینه ملی ، ارزش افزوده ، تشکیل سرمایه و ... را به تفکیک بخشهای اقتصادی در اختیار برنامه ریزان جامعه قرار می دهـــد.

۵ - عرضه وتقاضای اطلاعات آماری

تقاضای اطلاعات آماری و یا به عبارت دیگـــر نیازهای آماری جامعه در سطح خردوکـــلان روز بـــروز درتزایـــد می باشدوحال آنکه ، در طـــرف مقابل ، عرضه وتولید اطلاعات آماری بـــا محدودیتهای فنـــی ، اجرائـــی ومالی مواجه می باشد . در مقابل عرضه‌و تقاضای اطلاعات آمـــاری چه در سطح کلان وچه درسطح خرد چند نکته زیر قابل بررسی اســـت :

الـــف - اطلاعات آماری ، در حال حاضـــر ، به عنوان یک " خدمت اقتصادی " می تواننـــد تلقـــی شـــود چراکه طراحی ، جمع آوری ، استخراج وانتشار، ذخیره سازی‌و به هنگام سازی آن‌توام با هزینه‌های بسیار زیاد می باشـــد.

ب - هر " خواسته آماری " مدیران ، مسئولان ، پژوهشگران و سازمانها لزوما " نمایندگی بمنظور " نیازهای آماری " واقعی برای برنامه ریزی و تحقیق تلقی شود چرا که تامین بعضی از خواسته‌های آماری نه‌مفید ، نه امکان پذیر و نه اقتصادی و متناسب با امکانات مالی و اجتماعی می باشد .

ج - تولید هرچه بیشتر اطلاعات واطلاعات آماری و ارائه آن به مدیران و مسئولان جامعه خدمت بیشتر و بهتر به آنها و به امر برنامه ریزی وتصمیم گیری محسوب نمی شود زیرا این امر نه‌تنها هزینه کردن اقتصادی امکانات را به همراه ندارد بلکه باعث تضییع وقت مدیران و بعضا " خطای تصمیم گیریهایی شود .

د - توجه و اقدام به موضوع مطرح در بندهای ب و ج ضرورت بررسی و تشخیص نیازهای آماری واقعی را مطرح می سازد بطوریکه آمارشناسان در تماس و تبادل نظر با برنامه ریزان ، مدیران و تصمیم گیران بایستی "نیازهای آماری " راز " خواسته‌های آماری " اعلام شده تفکیک و پالایش نمایند .

هـ - توجه و اقدام برای‌تامین "نیازهای آماری واقعی " ضرورت طبقه بندی نیازها ، اولویت آنها و برنامه ریزی اجرائی آنها را در قالب نظام آماری کشور و برخورداری ازامکانات آماری موجود ایجاب می‌نماید . برنامه ریزی مورد اشاره تولید اطلاعات آماری مورد نیاز را در قالب امکانات موجود و یابرخورداری از تعاریف و مفاهیم و طبقه‌بندیهای استاندارد بروشهای علمی آماری کیفیت و دقت اطلاعات آماری وحداقل فاصله بین‌تولید وانتشار آمار امکان پذیر می نماید .

در قالب نکات فوق‌الذکر ، عرضه‌اطلاعات‌آماری مورد‌نیازدرسطح جامعه ازطریق نظام‌آماری کشور درسطح خرد (یک سازمان یک‌شرکت ، یک دانشگاه ، یک‌موسسه و ...) از طریق طراحی و اجرای سیستم اطلاعات آماری مطرح می شود . در هریک از دوحالت فوق الذکر ، بااستخراج نیازهای آماری واقعی ، تولیدو عرضه اطلاعات‌آماری هدفمند مرنیازهای مدیران و محققین‌میتنی

می شـــود . در صورتیکه هدف تامین اطلاعات مورد نیاز مدیران سازمان یاشرکت باشد موضـــــوع طراحی واجرای " سیستم اطلاعات مدیریت " MANAGEMENT INFORMATION SYSTEM مطـــرح می شـــود . و در صورتیکه هدف تامین اطلاعات آماری مورد نیاز مدیران سازمان یا شرکت باشـد در آنصورت لزوم طراحی و اجرای " سیستم اطلاعات آماری مدیریت " یا MANAGEMENT STATISTICAL INFORMATION SYSTEM مطرح می شـــــود .

در طراحی (MIS) و یا (MSIS) مراحل مشروحه زیر مورد بررسی و اقدام قرار می گیـرد .

الف ـ شناسایی اهداف ، نمودار سازمانی ووظایف مدیریتی سازمان (شناسایی وضع موجود)

ب ـ شناسایی نیازهای واقعی مدیران سازمان به اطلاعات وآمار از طریق یافته‌های بند الف و نیز با طراحی واجرای طرح شناسایی نیازها (مصاحبه با مدیران سازمان)

ج ـ شناسایی امکانات آماری سازمان باطراحی و اجرای طرح امکانات آماری

د ـ تهیه تعاریف ومفاهیم و طبقه بندیهای مورد نیاز بر اساس اهداف و نیازهای اطلاعاتی و آمـــــاری

هـ ـ طراحی طرحهای لازم به منظور جمع‌آوری اطلاعات و آمار

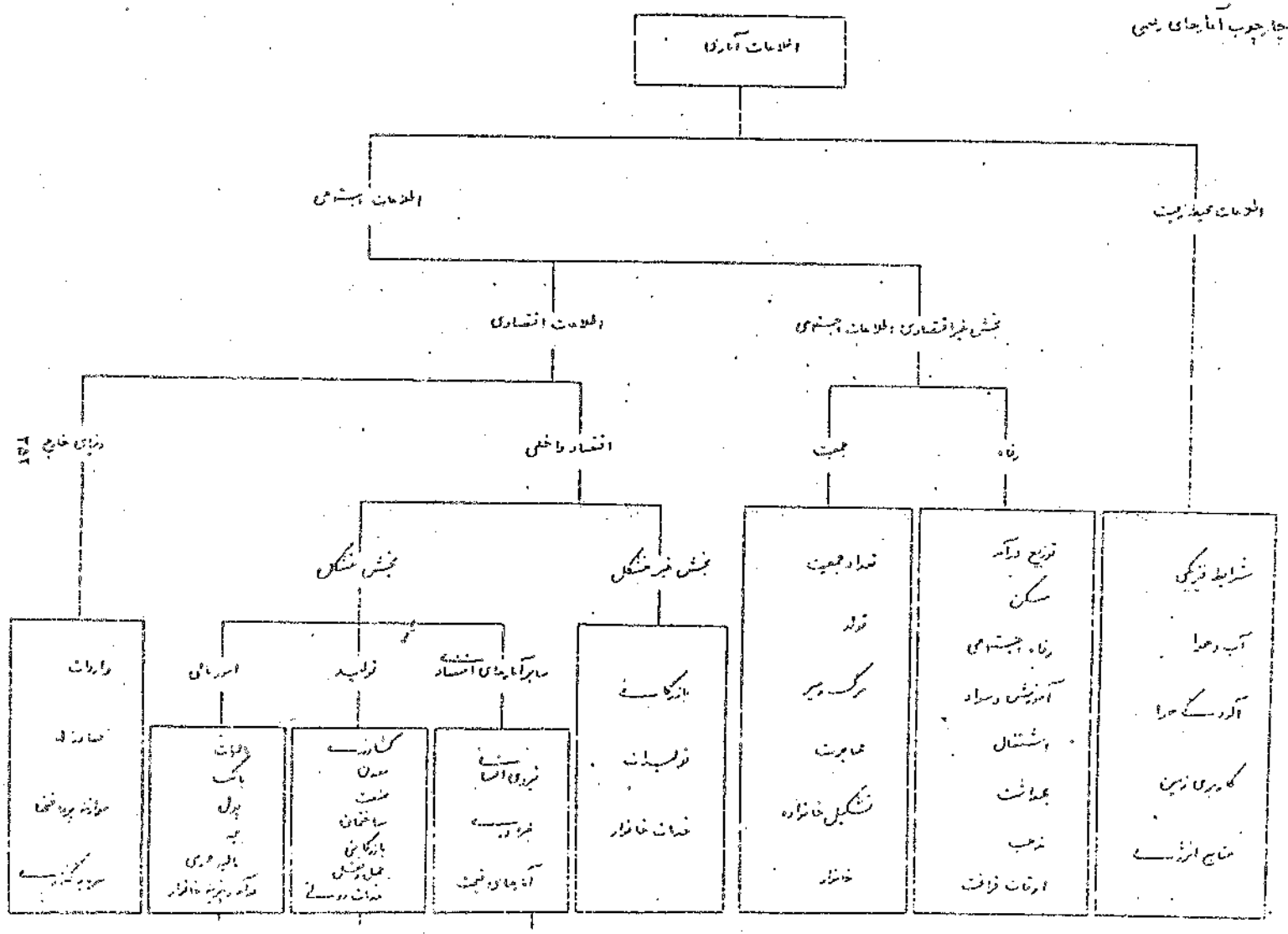
و ـ طراحی وپیش‌بینی نرم افزارهاو سخت افزارهای لازم برای استخراج و پرورش اطلاعـــات وآمـــــار

ز ـ پیش بینی روش بـهنگام سازی اطلاعات و آمار بر اساس نیازهای جدید سازمان

ج ـ سازماندهـــی واحدهای اداری برای اجرای طرحهای تهیه شده و برآورد بودجه و نیـروی انسانی سالانـــــه .

درپایان ضمن تاکید براهمیت اطلاعات وآمار واستفاده ازآن درتصمیم‌گیری وتحقیقات ، امید است تقاضا و عرضه اطلاعات آماری در کشورمان در سطح خرد وکلان در قالب تفکر سیستمـــی قرارگیـــرد تا تولیدات هدفمنـد سیستم ، نیازهای آماری از قبـل شناسایی شده جامعـــــــه را تغذیه نماید .

نمودار ۱ - چارچوب آمایهای رسمی



تجزیه و تحلیل الگوی مصرف خانوارهای شهری در استان تهران

علی عسگری

سازمان برنامه و بودجهٔ استان تهران

۵شنبه ۹:۰۰ تالار ۲

بخش اعظم تولیدات هر کشوری توسّط جمعیّت و یا خانوارهای آن کشور به مصرف می‌رسد. در اقتصاد میزان مصارف خانوارها، رفتار و گرایش آنها در انتخاب کالاها و خدمات برای مصرف، حساسیّت خانوارها برای تغییرات مصارف خود در قبال تغییرات درآمد، مجموعاً تحت عنوان الگوی مصرف مورد مطالعه قرار می‌گیرد. در مقاله پیشنهادی سعی شده است تا با توجّه به اهمیّت مبحث مزبور در اقتصاد از یک طرف و نتایج آمارگیری از هزینه و درآمد خانوارهای شهری استان تهران از طرف دیگر، تصویری از وضع و الگوی مصرف خانوارهای شهری استان تهران در سالهای قبل و بعد از پیروزی انقلاب اسلامی ارائه شود و تغییرات حاصل از الگوی مزبور در ارتباط با تحوّلات اقتصادی و اجتماعی مورد مطالعه و بررسی قرارگیرد.

مطالعهٔ مزبور جنبه آماری-کاربردی دارد و با استفاده از اطّلاعات آماری موجود، عملکرد قانون انگل در الگوی مصرف خانوارهای استان تهران و نیز تأثیر تحوّلات مربوط به جنگ تحمیلی، سیاست سهمیه‌ای‌کردن کالاهای مصرفی، افزایش قیمت کالاهای اساسی و ... مورد بررسی قرارگرفته است.

کاربرد نقشه در آمار

نجم‌الدین علامه

مرکز آمار ایران

مقدمه :

با توجه به پیشرفت تکنولوژی درعصرحاضروامکان سرعت بخشیدن به کارهای اجرایی توسط ماشین های حسابگر ،لازم است اجراکنندگان پروژه های عمرانی وتعاونی کسانی که دارای تخصص های مختلف هستندبا مطالعه دستآوردهای جدید وآشنایی با متدهای پیشرفته ،دانسته های خودرا بهنگام نگه داشته ودررا ه اعتلای سطح دانش ونوآوری درتعاونی زمینه های تخصصی ،مدیریت ،اجرا و... بکوشند .

یکی ازدستآوردهای ماشین های حسابگرامکان تهیه نقشه ،ترسیم واصلاح نقشه های موضوعی ازجمله نقشه های آماری توسط P.C. یا PERSONAL COMPUTER است،که دراین رابطه عامل برای ذخیره سازی اطلاعات موردنظر باید ابتدا با توجه به مختصات جغرافیایی ،نقشه جامع وکامل کشور ،که قبلا" تهیه وشبکه بندی شده است را به ماشین ،هدایت کندوسپس اطلاعات مربوط به کلیه نقاط شهری وروستایی در داخل شبکه ها را تعیین نماید .نقشه های آماری اعم ازشهری وروستایی نیز با توجه به شبکه های ایجاد شده وذخیره سازی اطلاعات لازم برای این نقشه های مبنا ،به راحتی قابل ارائه وتهیه می باشد .دراین مقاله درمورد چگونگی تهیه نقشه مبنا ، نقشه های آماری شهری وروستایی،واحدجغرافیایی وکاربرداین نقشه ها در آمارگیری ها وسرشماریها بحث خواهیم گفت .

به‌جهت شناخت اهمیت نقشه‌در آمار ،لازم است آشنایی مختصری با کره‌زمین و عملیاتی که منجر به‌تهیه‌نقشه‌یک منطقه می‌شود داشته‌باشیم‌وسپس به‌انواع نقشه‌ها و از آن جمله‌نقشه‌های آماری بپردازیم‌وچگونگی تهیه‌این نقشه‌ها را درگذشته و حال وآینده‌بیان کنیم.

به‌خدمت گرفتن ماشین های حسابگر (COMPUTER) در رابطه‌با تهیه نقشه‌های آماری درآینده‌ضمن یاری رساندن به‌نقشه‌برداران برای ترسیم ،تکثیر وتعیین محدوده‌کار ماموران آمارگیری به‌عملیات زمینی(اصلاح ،تکمیل وبرداشت نقشه) کمک خواهدکرد ونقشه‌برداران برای اصلاح نقشه‌ها کما فی السابق باید با دقت وصحت بیشترکار خودرا ادامه‌داده‌وهرساله‌با برنامه‌ریزی مناسب نقشه‌ها را بهنگام نگه‌دارند.

تعریف نقشه : نقشه‌عبارتست از ترسیم‌تصویرقاشم‌عوارض طبیعی ومصنوعی روی زمین برروی کاغذبا مقیاس های مختلف‌منظوراز عوارض طبیعی عبارتست از تشخیص‌اراضی که‌بدون دخالت انسان خودبخوددرروی زمین بوجودآمده‌اندمانند کوهها ،روخانه‌ها ،تپه‌ها ،گودالها ،بریدگیها و... که‌هرکدام‌با علائم ویژه‌از نقشه‌ها مورداستفاده‌قرارمی‌گیرد .عوارض مصنوعی تشخیص‌اراضی هستندکه به‌دست انسان درروی زمین احداث‌شده‌اندمانندساختمانها ،دیوارها ،کوچه‌ها ، خیابان ها ومیادین که‌با مقیاس‌مشخصی روی نقشه‌نشان داده‌می‌شوند.

برای تهیه‌نقشه‌یک منطقه‌بطورکلی سه‌سری عملیات مشخص ومتمایز از یکدیگرانجام‌می‌گیرد.

۱ـ عملیات ژئودزی

۲ـ عملیات توپوگرافی(نقشه‌برداری)

۳ـ عملیات کارتوگرافی (ترسیم‌نقشه)

دراین مجموعه‌عملیات توپوگرافی که‌به‌عبارت دیگرمی‌توان تحت عنوان نقشه‌برداری از آن نام‌بردموردنظرمی‌باشدوچون دراین موردوبرای انجام بارهای عملیات ازعکسهای هوایی هم‌استفاده‌می‌شود،اشاره‌مختصری نیزبه‌آن خواهدشد.

عملیات ژئودزی مربوط به تعیین نقاطی است که به عنوان نقاط پایه از آنها در عملیات نقشه برداری استفاده می‌شود. در عملیات ژئودزی نقطه‌ای را به عنوان نقطه اصلی در یک نقطه انتخاب کرده و مختصات آنرا به طریقه نجومی تعیین می‌کنند. سپس نقاطی را به فاصله ۴۰ تا ۵۰ کیلومتری نقطه اصلی انتخاب نموده و پس از تعیین مختصات آنها، نقاط دیگری بین این نقاط که فاصله آنها ۱۵ تا ۲۰ کیلومتر است انتخاب و مختصات آنها را تعیین می‌کنند. سپس نقاط دیگری که بین نقاط دسته آخر واقع شده و فاصله آنها از یکدیگر ۸ تا ۱۵ کیلومتر است انتخاب کرده و بهمان طریق، مختصات نقاط را بدست می‌آورند و بالاخره در مرحله چهارم نقاطی به فاصله ۳ تا ۵ کیلومتر بین نقاط دسته سوم انتخاب نموده و مختصات تعداد زیادی از نقاط در منطقه مورد نظر به دست می‌آید. پس از انجام این مرحله از کار، عملیات نقشه برداری شروع می‌شود. در نقشه برداری از این نقاط به عنوان نقاط ژئودزی استفاده می‌نمایند. به عبارت دیگر عملیات نقشه برروی محل و موقعیت تمام عوارض زمین باید به طریق ترسیمی نشان داده شود. عوارض زمین به دو دسته تقسیم می‌شود. به عوارضی که در سطح خاک و کمی پایین تر از سطح خاک وجود داشته و به آسانی حد آنها معلوم و مشخص می‌شود، عوارض مسطحاتی (پلانیمتری) گفته می‌شود. این عوارض ممکن است مانند رودخانه طبیعی بوده و یا مانند ساختمانها، راهها، ارتباطی، مزارع و غیره که دست انسان در ساختن و بوجود آوردن آن دخالت دارد مصنوعی باشند. به این دسته از عوارض می‌توان خطوط موضوع مانند، استانها، شهرستانها و ... را اضافه کرد که برای مشخص کردن آنها نیازی به نقشه برداری نمی‌باشد.

به دسته دیگر از عوارض که مربوط به شکل و فرم زمین بوده و پستی و بلندیها را نشان میدهد، عوارض ارتفاعی (آلیمتری) می‌گویند. عملیات نقشه برداری به منظور برداشت عوارض مسطحاتی را ارتفاعی زمین است که معمولاً روی ورقه‌ای به نام پرگه برداشت انجام می‌گیرد. همانطوریکه قبلاً گفته شد عملیات نقشه برداری متکی به نقاط ژئودزی بوده و در داخل شبکه‌ای که از این نقاط تشکیل شده انجام می‌گیرد. واضح است که این نقاط وسیله کنترلی برای عملیات نقشه برداری منطقه

مورد نظر خواهد بود. ولی چون تعداد این نقاط که فاصله آنها در مرحله چهارم عملیات ژئودزی بین ۳ تا ۵ کیلومتر است برای نقشه برداری بزرگ مقیاس از نظر کنترل نقشه کافی نیست. نقشه بردار مجبور است برای زیاد تر کردن نقاط کنترل بنحوی تعداد نقاط را زیاد تر کند بطوریکه در هر کیلومتر مربع از منطقه برداشت لااقل یک نقطه با مختصات معلوم داشته باشد این دسته از نقاط را می‌توان نقاط درجه ۵ دانست. با توجه به این نقاط منطقه مورد نظر را با روش مثلث بندی یا کشیرالاضلاع بندی نقشه برداری می‌کنند و کلیه عوارض واقع در این مثلث بندی یا ... را با توجه به مختصاتشان روی برگه برداشت انتقال می‌دهند و بعداً با استفاده از روش ترسیمی عوارض زمین را روی برگه برداشت نشان می‌دهند.

پس از تعیین و محاسبه نقاط ژئودزی چنانچه عوارض و جزئیات زمین با اندازه‌گیریهای لازم و استفاده از روش ترسیمی روی برگه برداشت برده شود طریقه عمل را برداشت مستقیم می‌گویند ولی اگر از عوارض و جزییات زمین عکسبرداری هوایی بشود و از دستگاههای مخصوص تبدیل عکس به نقشه برای برداشت عوارض و جزییات استفاده شود این طریقه را برداشت فتوگرامتری می‌گویند. باید در نظر داشت که همیشه برداشتهای فتوگرامتری دارای نواقصی است که باید با روش برداشت مستقیم تکمیل گردد. مثلاً موقعی که از یک منطقه جنگلی عکسبرداری هوایی بشود راههای مواصلاتی در روی عکس بخوبی ظاهر نمی‌شود یا مناطقی که وجود بر ما نع عکسبرداری هوایی شده است که در هر حال تکمیل نواقص فوق به طریق برداشت مستقیم انجام می‌گیرد.

امروزه طریقه برداشت گرامتری جایگزین برداشت مستقیم شده است و فقط در کارهای محلی کوچک و یا در مناطقی که از آن عکسبرداری نشده باشد، همچنین برای تعلیم مبتدیان از روش برداشت مستقیم استفاده می‌شود. بنابراین می‌توان گفت برداشت مستقیم هنوز دارای اهمیت فوق العاده است.

کار توگرافی به کلیه مراحل ترسیم نقشه از آوردن مختصات نقاط روی برگه برداشت، ترسیم مقدماتی و ترسیم نهایی عوارض و چاپ نقشه گفته می‌شود.

انواع نقشه‌ها : پس از تهیه نقشه از طریق گرامتری ، هر سازمانی بنا بر روش کار خود و هدفی را که دنبال می‌کند از نقشه‌های فوق به عنوان نقشه پایه استفاده کرده و احتمالاً با تغییراتی که در آن ایجاد می‌کند نقشه‌ها را برای انجام فعالیت خود آماده می‌کند .

از جمله انواع نقشه‌ها می‌توان به نقشه‌های زیر اشاره کرد :

۱ـ نقشه‌های مسطحاتی یا پلانیمتری PLANIMETRY که موقعیت سطحی نقاط مختلف را نشان میدهد .

۲ـ نقشه‌های توپوگرافیک TOPOGRAPHIC که موقعیت افقی و ارتفاعی نقاط و عوارض مختلف را نشان میدهد .

۳ـ نقشه‌های برجسته که شکل سه‌بعدی دارد روی پلاستیک و یا مقوا یا گچ درست می‌شود که مقیاس ارتفاعات آن معمولاً برای بهتر جلوه نمودن سه برابر مقیاس افقی آن می‌باشد .

۴ـ نقشه‌های عکسی یا PHOTO MAP : اگر عکس هوایی منطقه را مورتیفیک کرده و روی کاغذ چاپ کنیم و شبکه‌بندی اطلاعات لازم و نام عوارض و آبادیها را روی آن بنویسیم نقشه عکسی بدست می‌آید .

۵ـ نقشه‌های شهری که در آن تمامی کوچه‌ها و خیابانها و عوارض ساخته شده شهری نشان داده می‌شود . و معمولاً در مقیاسهای ۱:۲۰۰۰ یا ۱:۲۵۰۰ تهیه می‌شود .

۶ـ نقشه راهها ROAD MAP که خطوط مواصلاتی و تابلیت عبور راهها را نشان میدهد .

۷ـ نقشه آماری یکی از انواع نقشه‌های مسطحاتی است که در مقیاسهای ۱:۲۰۰۰ ، ۱:۲۵۰۰ شهری و ۱:۱۰۰۰۰۰ ، ۱:۵۰۰۰۰۰ روستایی تهیه می‌شود .

نقشه‌های آماری که برای آمارگیریها و سرشماریها مورد استفاده قرار می‌گیرد بر دو نوعند : شهری و روستایی

ـ نقشه‌های آماری شهری ، نقشه‌هایی هستند که طبق تعریف مرکز آمار ایران برای شهرهای دارای شهرداری تهیه می‌شوند . و مقیاس آنها ۱:۲۰۰۰ ، ۱:۲۵۰۰ می‌باشند .

ـ نقشه‌های آماری روستایی ،نقشه‌هایی هستندکه عمدتا "راهها وموقعیت آبادیها وخطوط شهری را با مقیاس ۱:۵۰۰۰۰ یا ۱:۱۰۰،۰۰۰ نشان میدهند.

درحالی حاضربلوک، احدآماری شهری است وبرای آشنایی،به تعریف آن میپردازیم.

تعریف بلوک : بلوک عبارتست ازمجموعه ساختمانهای ساخته‌شده ،باغ وزمینهای ساخته‌نشده‌که به یکدیگرچسبیده واز اطراف بوسیله معابرعمومی یا عوارض طبیعی احاطه‌شده‌باشد ،بطوریکه اگرازیک نقطه‌آن حرکت کنیم پس ازطی دوربلوک بازهم به همان نقطه برسیم.

نقشه‌های آماری یکی ازنقشه‌های موضوعی است که‌لازم است به‌جهت بکارگیری درطرحهای آماری همواره‌بهنگام باشد .تهیه این نقشه‌ها ازچندطریق ممکن میشود

۱ـ برای تهیه‌نقشه‌های آماری ازعکسهای هوایی که‌بوسیله‌هواپیما گرفته میشودوزمان تهیه‌آنها کمی مانده‌به‌اجرای طرحهای آماری وسرشماریها است استفاده‌میشود .این عکسها پس ازچاپ موزاییک ورفع خطاها ،توسط نقشه‌برداران زمینی گویا میشود ،بدین معنی که‌بلوکها واسامی کوچه ،خیابان و.... را برروی عکسها منعکس وپس ازترسیم ،حوزه عمل آمارگیربرروی آنها تعیین وآمارگیری انجام میشود.

۲ـ تهیه‌نقشه‌های آماری ازروی عکسهای هوایی که‌هرپنج یا ده‌سال یکبار عکسبرداری میشودوبه‌جهت استفاده‌درطرحهای آماری وهمانگونه‌که‌درقسمت یک گفته‌شد ،پس ازچاپ وموزائیک عکسی ورفع خطاهای موجوددرعکسها با امکانات موجوددرآتلیه‌ها وبدون مراجعه به محل عکسبرداری شده ،تهیه میشود.

۳ـ تهیه‌نقشه‌های آماری ازروی عکسهای هوایی که‌کلیه مراحل فتوگرامتری ازقبیل تعیین مختصات نقاط زمین برروی هرعکس ،تبدیل عکسها بوسیله دستگاههای فتوگرامتری ودرنتیجه تهیه‌نقشه ازطریق کارتوگرافی خواهدبود .لازم به‌تذکراست که‌مراحل ظهورعکس ،موزائیک (پهلوی هم‌چیدن عکسها) تبدیل وتهیه‌نقشه‌های دقیق بهداز عملیات فتوگرامتری وکارتوگرافی نیازبه‌توضیح

زیادی دارد که به جهت تخصصی بودن هریک از این مراحل از ذکر آنها در این مقاله خودداری می‌کنیم.

ـ تهیه نقشه‌های آماری از طریق روش اول برای کشورهایی که وسعت خاک آنها زیاد می‌باشد پرهزینه است، لکن از پوشش کامل برخوردار می‌باشد و تصویر کلی و همگامی از منطقه مورد آمارگیری را نشان میدهد.

ـ روش دوم از روش اول کم‌هزینه‌تر می‌باشد زیرا عکسبرداری هر پنج یا ده سال یکبار انجام می‌شود و نقشه‌هایی که از روی این چنین عکسهای هوایی حاصل می‌شود باید توسط نقشه‌برداران تا زمان سرشماریها بهنگام نگه داشته شود.

ـ روش سوم نظیر روش دوم است، با این تفاوت که فاصله زمانی عکسبرداری مشخص نیست و بهمین لحاظ لازم است پایه و اساس نقشه‌ها که بر مبنای چنین عکسهایی بنا شده و از آنها پس از طی مراحل مختلف نقشه‌های دقیق آفست تهیه میشود، نقشه‌برداران با دقت زیاد و با برنامه زمانی تنظیم هر ساله نقشه‌ها را اصلاح و تکمیل نمایند.

در اولین سرشماری بعد از پیروزی انقلاب اسلامی و با توجه به اهمیت نقشه‌های آماری در سرشماری عمومی نفوس و مسکن مرکز آمار ایران طبق قراردادی که با سازمان نقشه‌برداری منعقد کرد، سازمان مذکور از تمامی شهرهای بزرگ و تعدادی از نقاط روستایی عکسبرداری کرد و با ظهور عکسها و موزائیک آنها و رفع مختصری از خطاهای موجود یک نسخه از فیلم عکسها را در اختیار مرکز آمار ایران قرار داد. مرکز آمار ایران نیز با همکاری گروههای نقشه‌استانها، ضمن از تکمیل و آماری کردن نقشه‌ها و نوشتن اسامی کوچه، خیابان بر روی آنها و ترسیم نقشه‌ها را برای سرشماری آماده کرد. پس از تهیه، ترسیم و تکثیر نقشه‌ها و نوشتن تعداد تقریبی خانوارها در روی بلوکهای شهری و آبادیها محدوده کار هریک از ماموران آمارگیری را که اصطلاحاً "حوزه آماری" نامیده می‌شد بر روی نقشه‌ها مشخص کردند و سرشماری را با موفقیت و بدون هیچ گزارشی مبنی بر عدم دسترسی به واحدهای آماری در شهر و روستا به مرکز آمار ایران تحویل نمودند.

با توجه به مطالبی که گذشت لزوم آگاهی ماموران آمارگیر از نقشه و شناخت

موقعیت جغرافیایی خود در محل آمارگیری محرز می‌شود. مسلم است اگر کسی طبیعت اطراف خود را بخوبی نشناسد امکان آگاهی از اطلاعاتی که می‌خواهد کسب کند خیلی کمتر از موقعی است که به نقشه و نقشه‌خوانی آشنایی داشته باشد. ما برای آگاهی بیشتر ماموران آمارگیری و راهنمایی بیشتر آنان در جمع آوری آمار و اطلاعات لازم است با ساده‌ترین روش و همانگونه که در راهنمای آمارگیری به او یاد می‌دهیم که چگونه به اصلاح و تکمیل نقشه‌های شهری بپردازد، در رابطه با شناخت نقشه و نقشه‌خوانی هم به او آموزش می‌دهیم. در راهنمای آمارگیری طریقه و روش بلوکسی که به دو بلوک تقسیم شده یا چند بلوک که به یک بلوک تبدیل شده و نقشه بسرداری از بلوکهای جدید را یادآور می‌شویم. لازم است برای شناخت بهتر ماموران آمارگیری بلوک که در حال حاضر واحد آماری شهری را معرفی می‌کند به صورت سه بعدی در راهنمای ماموران آمارگیری ترسیم شود، تا ماموران آمارگیری بتوانند به راحتی شکل بلوک را در ذهن خود مجسم نمایند.

تجربه نشان داده است که انجام سرشماریها و آمارگیریها بدون نقشه امکانپذیر نیست.

در سرشماریها به ویژه سرشماری نفوس که در مدت زمان مشخص انجام می‌گیرد محدوده کار یک مامور آمارگیری و حدود و ثغور کار او باید بر روی نقشه مشخص شود تا او بتواند با توجه به خانوارهای ساکن در آن، آمارگیری در محدوده کار خود را به اتمام برساند. نقشه، جامعه آمارگیری را پوشش می‌دهد و اگر قبل از آمارگیری بخوبی و با دقت اصلاح شده باشد سازمان اجراکننده با اطمینان کامل و با پوشش کل جامعه آماری اطلاعات مورد نیاز را بدست خواهد آورد.

اگر نقشه‌های آماری با دقت تهیه شده باشد، آمارگیریهای نمونه‌ای هم با موفقیت انجام می‌گیرد و رسیدن به واحدهای نمونه سریعتر انجام می‌شود.

با توجه به گسترش شهرهای کشور و لزوم داشتن نقشه‌های بهنگام و دقیق از تمامی شهرها و تهیه نقشه‌های بلوکه از آبادیها جهت آگاهی بهتر و بیشتر ماموران آمارگیری از مکانهای واقع در آبادیها در آینده و با توجه به تکنولوژی جدید و پیشرفته جهان در تهیه پکیج های (برنامه‌های از پیش نوشته شده) مختلف، مرکز

آمار ایران ناگزیر است که برای تهیه، تکمیل و بهنگام کردن نقشه‌های آماری از تکنیک های پیشرفته یاری بگیرد. در این راستا و به‌جهت بررسی در مورد وضعیت موجود و آینده نقشه‌ها، گروهی تحت عنوان گروه فایل جغرافیایی در مرکز آمار ایران تشکیل و بحثهای مقدماتی در رابطه با بخدمت گرفتن ماشین های حسابگر جهت اخذ اطلاعات جغرافیایی و نقشه‌های آماری انجام شده و در آینده ادامه خواهد یافت. جمع بندی بحثها در این خصوص را می‌توان در چند بند خلاصه کرد:

۱ـ بحث در رابطه با تعیین واحد جغرافیایی و واحد آماری به‌گونه‌ای که این دو یا برهم منطبق و یا به‌قسمی باشند که از طریق واحد جغرافیایی بتوانیم به واحد آماری دست یابیم و این واحد در تمامی طرحهای آماری و سرشماریها مورد استفاده و استناد قرار گیرد.

۲ـ برای شناخت واحدهای آماری، نیاز به داشتن نقشه مبنا می‌باشد و این نقشه باید دارای مختصات جغرافیایی باشد تا بتوانیم ارتباط منطقی و صحیحی با قسمتهای مختلف کشور چه در نقاط روستایی و چه در نقاط شهری موجود آوریم. برای دست یابی به این هدف پس از بررسی و تحقیق برروی نقشه‌های مختلف کشور به این نتیجه می‌رسیم که بهترین نقشه موجود که در حال حاضر بتواند جوابگوی نیازهای آینده نقشه‌های آماری باشد، نقشه‌های ۱:۵۰۰۰۰ اداره جغرافیایی ارتش است که این اداره در سالهای اخیر به بازنگری قسمتهای مختلفی از این نقشه‌ها پرداخته است. از جمله ویژگیهای این نقشه به مختصات جغرافیایی و نشان دادن دقیق بستی و بلندیهای و بطور کلی توپوگرافی کشور می‌توان اشاره کرد.

شبکه‌ای که برروی نقشه‌های ۱:۵۰۰۰۰ ترسیم شده براساس سیستم تصویر استوانه‌ای یا UNIVERSAL TRANSVER MERCATOR (U.T.M.) است که توسط گرجاردمرکاتور دانشمند هلندی پیشنهاد شد. در این سیستم کره زمین را در امتداد خط استوا درون استوانه‌ای مماس می‌کنند و عمل تصویر را از نقطه‌ای مانند مرکز زمین انجام می‌دهند. سپس استوانه تصویر را در امتداد یکی از یالهای برش و بصورت صفحه هموار و مسطحی می‌گسترانند. صفحه استوانه‌ای، زمین همانند خط مستقیم و افقی در وسط صفحه تصویر می‌شود و مدارها نیز بصورت خط مستقیم و موازی

طوری ظاهر می‌شوند که هرچه از خط استوا دورتر شوند بر فاصله آنها از یکدیگر افزوده می‌شود و نصف‌النهارها نیز همانند خطوطی موازی و متساوی‌الفاصله عمود بر خط استوا تصویر می‌شوند. بدیهی است مناطقی که پیرامون خط استوا واقع شده‌اند پس از تصویر شکل واقعی خود را تا اندازه‌ای حفظ می‌کنند و هرچه از استوا دورشویم از دقت تصویر کاسته می‌شود تا جایی که در نقطه قطب زمین که در بی‌نهایت تصویر گردیده‌اند بکلی شکل واقعی خود را از دست داده و به یک خط مستقیم مبدل می‌گردند.

در این سیستم کره زمین را به ۶۰ قاچ شش درجه‌ای تقسیم می‌کنند و استوانه تصویر را به نوبت بر نصف‌النهار مرکزی هریک از قارچ ها مماس می‌سازند و چهره هر قاچ را روی استوانه مزبور تصویر می‌کنند. کشور ایران بین قاچ یا ZONE ۳۸ تا ۴۲ (۵ قاچ ۶ درجه‌ای) قرار دارد و سطح کشور ایران ۳۰ درجه کره زمین را بخود اختصاص داده است.

با توجه به مطالبی که گفته شد، اگر واحد آماری تعریف شود و نقشه پایه با توجه به ویژگیهایی که گفته شد مشخص شود، به جهت ایجاد ارتباط صحیح بین نقاط شهری و روستایی و دسترسی به واحدهای آماری در آینده لازم است کلیه نقشه‌های روستایی مرکز آمار ایران که تاکنون مورد استفاده قرار می‌گرفت، بر مبنای نقشه‌های ۱:۵۰۰۰۰ تغییر یابد و نقشه‌های شهری نیز با برخورداری از مختصات جغرافیایی که از طریق نقشه‌های ۱:۵۰۰۰۰ محاسبه و تعیین می‌شود، به واحدهای آماری شهری منتقل گردد. در این صورت کلیه مناطق شهری و روستایی بوسیله یک شبکه مختصات جغرافیایی بهم دیگر اتصال خواهد یافت...

برای ارائه اطلاعات و نشریات آماری در قالب تقسیمات کشوری به نظر میرسد بهترین شبکه‌بندی در تقسیم‌بندی کشور برای رسیدن به واحدهای آماری شهری و روستایی تقسیمات اداری کشور است، که کشور را از طریق مرزهای استان شهرستان، بخش و دهستان تقسیم‌بندی کنیم و سپس از تلفیق مختصات هریک از این تقسیمات از طریق دهستان به بلوک یا آبادی و یا به واحدهای آماری برسیم. در صورتیکه این پیشنهاد مورد توجه قرار گیرد و نقشه‌های آماری و شبکه‌بندی آنها با توجه به روش مذکور تهیه شود، نقشه‌های تهیه شده قابل ارائه به ماشین های حساب

خواهد شد و از این پس کلیه تغییرات حاصل در نقشه‌های شهری و روستای و تقسیم
نقشه‌ها برای اجرای عملیات میدانی به سادگی میسر می‌شود.

مروری بر نمونه‌گیری سیستماتیک

علی عمیدی

مرکز نشر دانشگاهی

۲شنبه ۱۷:۰۰ تالار ۷

دقیقترین راه تعیین مقدار مشخصه‌ای از یک جامعهٔ آماری، بررسی تمام واحدهای آن جامعه است. بدیهی است که یک بررسی فراگیر مستلزم صرف وقت و هزینهٔ زیاد است که گاهی نیز به از بین رفتن اشیایی منجر می‌شود که واجد آن مشخصه‌اند. لذا به جای تعیین مقدار دقیق مشخصهٔ مطلوب، با نمونه‌گیری از جامعهٔ تحت بررسی به برآورد مشخصه اکتفا می‌کنیم. برای انجام نمونه‌گیری، روشهای کلاسیک و غیرکلاسیک متعددی وجود دارند. ساده‌ترین این روشها، روش نمونه‌گیری تصادفی است که در واقع پایهٔ اساسی تمام روشهای دیگر نمونه‌گیری است. ولی خود این روش به دلیل اینکه در مقایسه با سایر روشها غالباً دارای دقتی کمتر است کاربرد زیادی ندارد. روشهای دیگر نمونه‌گیری، نظیر نمونه‌گیری با طبقه‌بندی، نمونه‌گیری با احتمال متغیر، نمونه‌گیری خوشه‌ای یک، مرحله‌ای و چند مرحله‌ای و نمونه‌گیری سیستماتیک و غیره برحسب نوع جامعه برای برآورد مشخصه‌ها و استنباطها بکار می‌روند. ما در این مقاله مروری بر نمونه‌گیری سیستماتیک خواهیم داشت. ضمن تعریف نمونه‌گیری سیستماتیک در حالتهای $N = nk$، $N \neq nk$، و در حالت دوری، برآوردکننده‌ای تاریب برای میانگین جامعه ارائه می‌دهیم و دربارهٔ برآورد چگونگی تغییر واریانس این برآوردکننده صحبت می‌کنیم، و دقت آن را با نمونه‌گیری تصادفی ساده مقایسه می‌کنیم. اطلاع از این نمونه‌گیری، به دلیل سرعت کار، کمی هزینه و کاربرد آن در بررسی اکثر مسائل آموزشی، برای معلمین ضروری است.

منابع آماری

شهلا غفرانی
مرکز آمار ایران

حاصل نظریه‌های علم آمار، کوششهای آمارشناسان و گردآورندگان وتدوین کنندگان آمار، درمنابع آماری متجلی می گردند که این منابع پاسخگوی بخش قابل توجهی از پژوهشگران و مراجعان به کتابخانه‌ها ومراکز اسناد و اطلاعات است .

هر پرسشی در خصوص کمیت یا آمار در موضوعی خاص ، مراجعه کننده را برای یافتن پاسخ به کتابخانه رهنمون می شود که پاسخ این قبیل پرسشها در منابع مرجع مختلف ، از قبیل دائره المعارفها ، منابع جغرافیائی ... و خصوصا " در منابعی که اطلاعات آنها کلا " وباعمدتا " آماری است ، یافت می شوند که به این نوع منابع مرجع که از منابع ردیف اول است ، " منابع آماری " اطلاق می شود.

منابع آماری در مجموعه منابع مرجع هرکتابخانه ، از اهمیت زیادی برخوردارند واطلاعاتی که از این نوع منابع بدست می آید، در نتیجه گیریهای تحقیقات ، نقش تعیین کننده دارد .

برای ورود به بحث ، ابتدا به تعریفی از آمار تاریخچه آن ، روشهای مختلف گردآوری و ... اشاره می شود.

۱ - تعریف آمار : تعریف آمار در طول زمان با توجه به کاربردهای آن تغییر کرده است ، لیکن جوهر اصلی تمام آنها این است که آمار علم ارائه واقعیات به صورت کمی است .

۲ - تاریخچه آمار : استفاده از آمار در اخذ تصمیم همواره مد نظر تصمیم گیرندگان امور در دوره‌های مختلف بوده است . محققان در آثار خود به آماربرداریهائی که انگیزه آنها جمع آوری مالیات ، تدارکات نظامی

و غیره بوده ، اشاره کرده‌اند .

۳-کاربردهای آمار : تاریخچه طولانی آمار نشان دهنده کاربرد آن در تصمیم گیری های مهم است - در جوامع امروزی ، استفاده از آمار نیز ابعاد وسیعتری یافته است و از ساده ترین برنامه ریزیها تا برنامه‌ریزیهای درازمدت و تحقیقات علمی و اجتماعی را شامل می شود .

بیشترین استفاده از آمار را دولتها جهت برنامه ریزی وتعیین سیاستهای آینده ، ارزیابی وضع موجود و عملکرد فعالیتهای گذشته کرده‌اند و نقش آن در برنامه ریزیهای توسعه آشکار است و در واقع مدیران جامعه برای تصمیم گیری نیازمند اطلاعات ، خصوصاً " از گویاترین نوع آن ، یعنی اطلاعات آماری " هستند .

۴-سازمانهای آماری : تولید آمار عملی فنی است که نیاز به دانش و تجربه و سازماندهی خاص دارد . سازمانهای آماری تبلور چنین نیازی هستند که بار عمده تولید و ایجاد هماهنگی در آمار ، بر دوش آنهاست . وظیفه سازمانهای عمده آماری در سطح بین المللی ، هماهنگ ساختن فعالیتهای آماری در سطح جهانی است . تا از این طریق امکان انجام مقایسه های بین المللی بوجود آید .

سازمانهای آماری از نظر عملکرد به سه گونه متمرکز ، نیمه متمرکز و غیرمتمرکز تقسیم می‌شوند . در کشورهائی با نظام آماری متمرکز ، سازمان مسئول آمار وظیفه انجام آمارگیریهای مورد نیاز مملکت را بر عهده دارد .

در ایران نظام آماری سه دوره مشخص را پشت سر گذارده است و در حال حاضر نظام آماری ایران نظامی نیمه متمرکز است که در آن سازمانهای مختلف بر حسب نیاز ، دارای واحدهای آماری هستند .

۵-تولید آمار : چنانکه ذکر شد ، جمع آوری و تولید آمار فرایندی فنی است که اعتبار آن ناشی از دقت و کاربرد صحیح روشها در تمام مراحل تولید و جمع آوری و تدوین است .

به طور کلی آمارها با یکی از سه شیوه سرشماری ، آمارگیری نمونه ای و روش ثبت آمار تولید می شوند .

در سرشماری کلیه اعضاء یک جامعه آماری مورد پرسش قرار می گیرند و لذا بخاطر گستردگی کار و سطوح جغرافیائی ، هزینه فراوان و ... در مقاطع زمانی خاص انجام می شود . دقت آن در مقایسه با سایر روشها بیشتر است . اطلاعات بدست آمده از این روش وسیع ، ولی کم عمق است و تصویری کلی از وضع موجود ارائه می کند .

در شرایط محدودیت زمان و بودجه، در صورت تهیه دقت در اندازه گیری برخی از خصوصیات جامعه آماری ، می توان از روش آمارگیری نمونه‌ای استفاده کرد. اطلاعات بدست آمده از این روش عمیق تر از سرشماری است .

آمارهای ثبتی حاصل ثبت اطلاعات فعالیتهای جاری سازمانها هستند که بخش عمده‌ای از اطلاعات آماری را تشکیل می دهند. با این روش می توان اطلاعاتی با کیفیت مطلوب حتی در پائین‌ترین سطوح جغرافیائی و سازمانی به صورت مستمر گرد آوری کرد. سالنامه‌های آماری ، محصول مشخص و نهائی این نوع گرد آوری اطلاعات هستند .

۶. منابع آماری : منابع آماری که در اینجا مراد از آن منابع چاپی است ، یکی از مهمترین منابع مرجع در کتابخانه ها و مراکز اسناد محسوب می شوند که پاسخگوی بسیاری از پرسشها در مورد کمیت هستند. این منابع در کتابخانه جایگاه ویژه‌ای دارند و بعد از فرهنگها و دائره المعارفها ، پرمراجع ترین مواد هستند .

منابع آماری از نظر شکل دارای انواع گوناگون هستند و از یک جدول یک صفحه‌ای تا سالنامه‌های آماری را شامل می شوند .

این منابع از نظر دوره انتشار و موضوع به شرح ذیل قابل تقسیم بندی می باشند :

- از نظر دوره انتشار
  - ۱. ادواری
  - ۲. تک نگاشت
- از نظر موضوع
  - ۱. مجموعه‌ها
  - ۲. موضوعات خاص

۶-۱. نشریات آماری ادواری : این نشریات از نظر دوره معمولا " یک ماهه، فصلی، شش ماهه، نیمه سالانه و سالانه هستند. گزارشهای سالانه معمولا " گزارش نهائی محسوب می شوند .

نوع دیگر نشریات ادواری نشریاتی هستند که حاصل انجام طرحهای جاری آماری هستند و نتایج

آنها هر ساله منتشر می شود .

۶-۲ . تك نگاشت ها : این قبیل نشریات از نظر نحوه تهیه به دو گروه عمده تقسیم می شوند :

۶-۲-۱ . نشریات حاصل از انجام سرشماریها که معمولا " در فواصل زمانی مشخص اجرا می شوند .

۶-۲-۲ . نشریات حاصل از انجام طرحهای موردی که بر حسب نیاز بـــه آمار در زمینه ای خاص،بـــه اجرا در می آید . بعضا " این نوع منابع حاصل سرشماریهای موردی هستند .

۶-۳ . مجموعه ها :اطلاعات این نوع منابع معمولا " درزمینه‌های مختلف است واین‌نوع‌منابع‌ازنظرنوع‌موضوعی برخوردارند .مهمترین این‌منابع سالنامه‌های‌آماری‌کشورهاهستندکه‌آیینه تمام نمـای فعالیتهای اجتماعی، اقتصادی ،صنعتی ، . . . هستند . اطلاعات بخشهای مختلف دراین نوع منابع کلی است

۶-۴ .موضوعات خاص :این نوع منابع یك موضوع را از ابعـــاد مختلف بررسی می کنند و بـــــــا توجه به محدود بودن موضوع امکان عمیق شدن اطلاعات را بوجود می آورنـــد . این نشریات از نظـــــــــر دوره انتشار معمولا " ادواری هستند،اما بعضا " نیز ممکن است به صورت تك نگاشت منتشر شوند .

۷ .ارزیابی منابع آماری : برای ارزیابی یك منبع آماری باید کلیه عوامل دخیل در تهیه آن را مـــد نظر قرار داد ، هر چند که ممکن است در بررسی اجمالی یك منبع امکان اشراف به کلیه آنها وجود نداشتـــه باشد . ایـــن عوامل عبارتند از :

۷-۱ . سازمان تولید کننده آمار : در اینجا باید اعتبار سازمان تولید کننده آمار مورد بررســـــــــــــی قرار گیرد تا بر اساس آن بتوان میزان اعتبار آمار ارائه شده را سنجید . داشتن واحد و تشکیلات آمـــــاری و افراد متخصص می تواند یك معیار برای سنجش این اعتبار باشـــد .

۷-۲ . زمان انتشار : سعی م ر بهنگام نگاهداشتن آمار یکی از مهمترین مسائل دست اندرکاران سازمانهای آماری است . آمارهائی که با تاخیر منتشر می شوند، معمولا " ارزش اطلاعاتی خود را از دست می دهنـــد . لذا مقایسه پوشش زمانی‌منبع و تاریخ انتشار آن ، بهنگام بودن اطلاعات مشخص می کند .

۷-۳ . منبع آماری : در ارزیابی فیزیکی یك منبع آماری عوامل مهمی باید مورد بررسی قرار گیرند کـــه اهم آنها به شرح ذیل انـــد :

۷-۳-۱ . عنـــــوان : عنوان باید بیانگر محتوای نشریه باشد . ذکر پوشش زمانی و مکانی بسیار ضروری است .

۷-۳-۲. مقدمـــــه : دریک منبع آماری مانند هر کتاب دیگری مقدمه باید خواننده را به دنیای کتاب وارد کنـد . در مقدمه معمولا " مسائل کلی درباره ضرورت و نحوه تدوین بیان می شود .

۷-۳-۳ . تعاریف و مفاهیم : ارائه تعاریف و مفاهیم از ضروری ترین بخشهای یک منبع آماری است که نشان دهنده مفاهیم واژه ها و عبارات ـ که الزاما " با معانی متداول آنها یکی نیست ـ می باشـد .

۷-۳-۴ . علائم اختصاری : ارائه این علائم در صورتی ضروری است که این علائم در متن نشریه مورد استفاده قرار گرفته باشـد .

۷-۳-۵ . نقشه : وجود نقشه نیز در صورتی ضروری است که آمار مربوط به منطقه خاص جغرافیایی باشـد .

۷-۳-۶ . خلاصه یافته ها : با توجه به اینکه آمارها نشان دهنده کیفیت یا کمیتی خاص از یک موضوع هستند ، ارائه مختصری از آنها در قالب یک متن بازبانی ساده خواننده را با وضعیت کلی جامعه مورد آمارگیری آشنا می سازد .

۷-۳-۷ . جداول : مهمترین بخش یک منبع آماری جداولی است که یافته ها در آن منعکس شده اند و ارائه آنهـا بر پایه قواعدی است که آمار شناس نسبت به آنها آگاهی کامل دارد .

جداول دارای انواع گوناگونی هستند که ذکر آنها در حوصله این نوشته نیست ، لیکن خصوصیات عمده ای که یـک جدول آماری باید دارا باشد به شرح ذیل اسـت :

۷-۳-۷-۱ . عنوان جدول : که معمولا " به صورت عبارت نوشته می شود و باید خلاصه ، کافی و رسا باشد و بیانگـر محتوی محدوده جغرافیایی ، زمانی و ... باشد .

۷-۳-۷-۲ . سری زمانی (Time Series) : داشتن سری زمانی در یک جدول معمولا " به غنـای آن می افزاید و این در صورتی امکان پذیر است که آمارگیریهای قبلـی نیز در یک چارچوب و قالب انجام شده باشد . سری زمانی امکان مقایسه در یک دوره مشخص را برای جوینده با مراجعه به یک منبع ، میسر سازد .

۷-۳-۷-۳ . یکسان بودن پوشش زمانی و مکانی جداول : فکر این مساله به این منظور است که برای مثال در یـک نشریه آماری کشور باید با مقایسه چند جدول با یکدیگر به این مساله توجه کرد که آیا پوشش زمانی و مکانی جداول در یک سطح است و یا با یکدیگر اختلاف دارند .

۷-۳-۸ . نمودار : " مهمترین روش نمایش آمار پس از جدول ، استفاده از نمودار است که روز به روز توسعه و اهمیت بیشتری پیدا کرده است " [۱] . با آنکه نمودار از ضروریات یک منبع آماری نیست ، لیکن با استفاده از آن که نوعی نمایش

هندسی آمار است می توان امکان مقایسه بصری را بین آمارهای مختلف ایجاد کرد . این وسیله به هیچ وجه جای جدول را نمی گیرد و معمولا " به عنوان عامل کمکی مورد استفاده قرار می گیرد .

نمودار دارای انواع مختلفی است که نوع آن باید متناسب با توانائی استفاده کننده ، انتخاب شود .

۷ـ۳ـ۹ - جامعه استفاده کننده : اصولا " در تهیه هر منبعی مخاطبان نقش تعیین کننده در نحوه ارائه مطالب دارند و هدف اصلی تهیه آنها محسوب می شوند . در تهیه یک منبع آماری نیز باید به این مساله توجه داشت و امکانات این جامعه را در استفاده از منبع ، نحوه ارائه جداول ، نمودارها و ... در نظر داشت .

۷ـ۳ـ۱۰ - مآخذ مورد استفاده : در ارزیابی یک منبع آماری ، بخصوص منابعی که با روش ثبت آمار تهیه می شوند ، باتوجه به مآخذ آن می توان اعتبار و ارزش آمار را سنجید .

۷ـ۳ـ۱۱ - نمایه ـ که متاسفانه اکثر نشریات معتبر آماری فاقد آنند ـ از اهمیت زیادی برخوردار است . ازآنجا که نشریات آماری معمولا " بر اساس طبقه بندی های خاصی ـ که الزاما" استفاده کننده با آنها آشنائی ندارد ـ تدوین می گردند ، وجود نمایه ای که بر اساس اصول علمی تهیه شده باشد ، در آنها بسیار ضروری است .

۷ـ۳ـ۱۲ - همخوانی آمار (consistency) : مراد از همخوانی آمار این است که اطلاعات ارائه شده در منابع مختلف ، با یکدیگر دارای همخوانی باشند و مراجعه کننده با مراجعه به منابع مختلف دچار سردرگمی نگردد . در اینجا می توان آمار ارائه شده در منبع مورد بررسی را با چند منبع معتبر دیگر مقایسه کرد تا همخوانی آمارها مشخص گردد .

۸ـ جایگاه آمار در طبقه بندی های جهانی : آمار ، به دلیل اهمیت آن در رده بندی های معمول جهانی جایگاهی خاص دارد . در نظام رده بندی دهدهی دیویی ، علم آمار در علوم خالص و داده های آماری درباره نفوس ، آمارهای عمومی ، نیروی انسانی و ... در علوم اجتماعی و آمار سایر رشته های خاص ، در کنار سایر متون آن رشته ها با تقسیمات فرعی قرار می گیرند .

در نظام رده بندی کنگره آمریکا نیز علم آمار در علوم و داده های آماری در خصوص نفوس ، آمارهای عمومی ، نیروی انسانی و ... در علوم اجتماعی قرار می گیرند .

منابع :

۱ - سادات میرمحمدی ، ابوالقاسم و دیگران . مبانی تدوین نشریات آماری . تهران : مرکز آمار ایران ، ۱۳۶۹ ، ص ۵۹ .
۲ - مصاحبه با برخی از کارشناسان مرکز آمار ایران ، خصوصا " جناب آقای فاضل رضوی .

چگونگی آمار خدمات در ایران

فرزانه فرزین

مرکز آمار ایران

مقدمه :

عملکرد چند دهه اخیر بسیاری از کشورهای در حال توسعه که برنامه ریزی های خود را برپایه رشد بیشتر تولید و درآمد ملی قرار داده بودند چندان موفق نبوده و در اکثر موارد این رشد همراه با افزایش نابرابریهای درآمد و امکانات، تشدید استقراض، وابستگی، مصرف زیاد و کاهش سریع منابع و ذخایر طبیعی محدود و غیرقابل تولید مجدد، و بخصوص افزایش کاذب و سریع بخش خدمات و بالا رفتن سهم آن در تولید ملی آنها بوده است.

عدم موفقیت برنامه ریزی های مزبور، در سالهای اخیر مفهوم جامع و چندین بعدی "توسعه" یا "توسعه اجتماعی - اقتصادی" را بجای مفهوم ساده و صرفا کمّی "رشد" تولید یا درآمد ملی مطرح کرده و لزوم تهیه و تنظیم برنامه های توسعه اجتماعی - اقتصادی و اجرای آنها را قطعی ساخته است.

امروزه در بیشتر کشورهای جهان، حسابهای اقتصادی به عنوان ابزار اصلی برنامه ریزی توسعه و معیار اساسی اندازه گیری رشد اقتصادی بکار گرفته می شود. زیرا برای تعیین اهداف برنامه و تعیین اولویت آنها و نیز جهت نیل به اهداف ملی برنامه توسعه، لازم است اطلاعات آماری مطمئن، دقیق و منظم از فعالیتها و جریانهای اقتصادی کشور و چگونگی روابط بین این جریانها در چارچوبی جامع، گویا و قابل مقایسه تولید و ارائه شوند.

چارچوب حسابهای ملی و اقتصادی و از جمله جدول داده ستانده، قالبی دقیق و علمی برای تهیه آمارهای اساسی در زمینه های مختلف مربوط به توسعه اجتماعی و اقتصادی است. حسابهای اقتصادی هم راهنمای مناسبی برای جمع آوری آمار و اطلاعات است و هم وسیله ای برای رفع نواقص و اشتباهات

آماری و ارتقاء کیفیت آمار و افزایش ضریب اطمینان و دقت در تصمیم گیریهای و سیاست گذاریهای اقتصادی ـ اجتماعی می باشد. اقدام برای تهیه حسابهای مزبور خلاء های مهم آماری در بخش های مختلف اقتصادی را نشان داده و لذا اولویت تهیه، تدوین و اجرای طرحهای آماری جهت کسب آمار مورد نظر در زمینه های مختلف را نیز تعیین می کند.

بخش خدمات با وجود آنکه بخصوص در سالهای اخیر سهم قابل ملاحظه و عمده ای در مجموعه فعالیتهای اقتصادی دارد، از جهت آمار و اطلاعات دارای خلاء های چشمگیر و مهم است بطوریکه در موارد متعددی حتی فاقد آمار و اطلاعات جاری نیز می باشد.

طی مقاله حاضر کلا" به آمار خدمات پرداخته شده است. ابتدا در قسمت "الف"، تعریف و مفهوم خدمات در اقتصاد کلان و حسابهای ملی بخصوص در سیستم حسابهای ملی، تفکیک کالاها از خدمات و معیارهای این تمایز و تعریف خدمات و مسائل مورد خاصی ارائه شده است.

در قسمت "ب" طبقه بندی فعالیتهای اقتصادی و خدمات آمده است.

در قسمت "ج" شمول خدمات و انواع فعالیتهای خدماتی برحسب کدهای مربوطه در طبقه بندیهای مورد عمل و توصیه شده حسابهای ملی تشریح شده است.

در قسمت "د" سهم بخش خدمات در تولید ناخالص و درآمد ملی کشور با کمک جداولی مشخص شده است.

قسمت "ه" به چگونگی آمار بخش خدمات در کشور اختصاص دارد و آمارهای خدمات در ۲ قسمت آمارهای پایه ای و آمارهای مربوط به تولید ناخالص، هزینه های واسطه، ارزش افزوده و درآمد و تشکیل سرمایه مورد بررسی قرار گرفته است.

قسمت "و" به شرح تهیه، تدوین و اجرای طرحهای مختلف آماری گیری از کلیه فعالیتهای اقتصادی به ویژه فعالیتهای خدماتی برای تهیه و تنظیم حسابهای ملی و پاره ای از جداول داده ها ـ ستانده های اقتصاد کشور اختصاص دارد.

در قسمت "ز" نتایج و پیشنهادات ارائه شده است.

در آخر منابع و مآخذ مورد استفاده در تهیه این مقاله درج شده اند.

الف ـ تعریف و مفهوم خدمات در اقتصاد کلان و حسابهای ملی :

در بررسی ها و تحلیل های اقتصاد کلان بخصوص در دهه های اخیر ، اشاره به خدمات متداول شده است و طی آنها از مطالبی نظیر سهم فزاینده خدمات در اقتصاد ، افزایش نسبتا " کمتر در بهره وری خدمات (در مقایسه با صنعت و ...) ، یا بالارفتن نسبتا " سریع تر سطح قیمت خدمات ، یا سهم نسبتا " کمتر مواد و کالاهای مصرفی در اراشه خدمات و یا درصد بالای سود در ارزش افزوده خدمات سخن به میان آمده است . لیکن جای تعجب است که هنوز تعریف بین المللی جا افتاده و مورد توافقی برای خدمات وجود ندارد . تا همین اواخر بسیاری از اقتصاددانان واژه خدمات را گویا دانسته و تعریفی برای آن ارائه نداده بودند . بعضی دیگر صرفا " برداشت های خود را به جای تعریف خدمات ارائه می داشتند . بدین وسیله نتیجه این امر آن بوده است که خدمات به صورتی که در آمارهای مختلف تهیه شده توسط کشورهای مختلف مورد اشاره واقع شده است غالبا " دارای محتوا و پوشش متفاوت بوده و تفاوت آنها اکثرا " بسیار چشمگیر است .

در سالهای اخیر نیاز به داشتن تعریف استاندارد و توافق شده ای برای خدمات یا لااقل آگاهی کافی نسبت به تفاوتهای تعاریف و مفاهیم و پوشش و شمول آن که در آمارهای گوناگون بکار می رود به میزان قابل ملاحظه ای افزایش یافته است . یکی از علل این امر ، افزایش علاقه به تسریع در توسعه آمار خدمات بوده است . عامل مهم دیگر ، مذاکرات تجاری GATT (موافقتنامه عمومی تعرفه و تجارت) است که طی آنها صراحتا " از دفتر آمار سازمان ملل متحد و سایر سازمانهای بین المللی خواسته است تا آمار مربوط به مبادله خدمات از جمله روشن ساختن مفهوم خدمات و بخصوص مفهوم آن در زمینه سازگاری بین المللی را بهبود بخشند .

بی شک هر تعریفی که برای خدمات پذیرفته و تائید شود باید در قالب سیستم های آماری حسابهای ملی و اقتصادی موجود قابل تطبیق باشد . بدین

ابتدا ، تعریف وشمول خدمات باید حتی الامکان با سیستم حسابداری ملی همانگ باشد .چون این سیستم بعنوان مهم ترین وسیله ادغام ومقایسه و ارائه استاندارد ومنظم وقابل مقایسه بین المللی آمارهای اقتصادی محسوب می شود .

الف ـ تعریف ومفهوم خدمات در سیستم حسابهای ملی (S.N.A) :

سیستم حسابهای ملی خدمات را تعریف نمی کند ، لیکن دارای تعریف دقیقی درمورد " کالاها وخدمات" است وبرحسب چگونگی ونظورازتولید آنها را طبق طبقه بندی مشخصی دسته بندی می کند .منتها استفاده ازاین تعریف و طبقه بندی کالاها وخدمات ، وقتی که حفظ سازگاری میان سیستم حسابهای ملی وآمارخدمات موردنظرباشد دارای اشکالاتی است که عمده آنها به قرارزیر است :

۱ـ سیستم حسابهای ملی مرزتولیدرابرای تعریف کالاها وخدمات تعیین کرده است وبموجب آن محصولات خارج ازاین مرز "خدمات" محسوب نمی شوند .به این ترتیب محصول کلیه فعالیتهای انجام شده درداخل خانوارکه برای فروش نباشد (بجزخدمات خانگی خریداری شده ازخارج خانوار) ازقبیل آشپزی ، نظافت ، مراقبت ازکودک وکلا" خانه داری را ازمفهوم وحوزه خدمات خارج می سازد .

۲ـ با توجه به اینکه واحد آماری درآمارتولیدسیستم حسابهای ملی "کارگاه" است ،خدمات انجام شده داخل کارگاه به عنوان خدمات مجزا محسوب نمی شودوفقط خدماتی که به واحدهای خارج ازکارگاه یعنی به کارگاههای دیگرارائه می شود ثبت می گردد .بعنوان مثال فعالیتهای واحدتعمیرات وسایل نقلیه یک کارگاه خدماتی به عنوان خدمات مجزا نمی محسوب نمی شود ، مگراینکه تعمیرگاه ، قسمتی ازخدمات خودرا به کارگاههای دیگرارائه کرده وبفروشد .لذا میزان تولید خدمات وسهم آن درکلی تولیدبدلیل ارزش افزوده نامحدودی بستگی به قرارروابط فنی کارشناسان وحدومرزپوشش فعالیتها دارد .

ج- نظام حسابهای ملی مشخصاً " پرداخت بابت خدمات را از پرداخت بابت عوامل تولید (مزد و حقوق، بهره و ...) مجزا می کند و فقط پرداخت بابت خدمات را تحت پوشش "کالاها و خدمات" قرار می دهد و پرداختی به عوامل تولید را به عنوان توزیع اولیه درآمد محسوب می کند.

د- سیستم حسابهای ملی پرداختهایی را که بابت خدمات صورت گرفته از انتقالات جاری جبران نشده مانند پرداختهای اجباری، جرایم و غرامات مجزا می کند. اساس این تفکیک قراردادی یا اختیاری بودن این پرداخت‌ها است. مثلاً" عوارض بزرگراه، پل، ورود به پارکینگ به عنوان پرداختهای بابت خدمات طبقه‌بندی می‌شوند، درحالیکه عوارض گذرنامه، عوارض فرودگاه و پرداختهای مربوط به گواهینامه رانندگی به عنوان انتقالات جاری جبران نشده منظور می شوند و در طبقه‌بندی کالاها و خدمات قرار نمی گیرند.

الف-۲- تفکیک خدمات از کالاها :

همانطورکه قبلاً" گفته شد، سیستم حسابهای ملی کالاها و خدمات را از سایر گروهها تفکیک می کند ولی کالاها را از خدمات متمایز نمی سازد.

بسیاری از اقتصاددانان کوشیده‌اند تا مرز بین خدمات و کالاها را تعیین کنند و با توجه به نظرات و پیشنهادات آنان بنظر می رسد هیچ معیار منحصر به فردی وجود ندارد که براساس آن بتوان کالاها را از خدمات متمایز نمود. لذا برای این تفکیک باید از معیارهای متعددی استفاده شود. مهم ترین معیارهای مزبور عبارتند از:

ا- کالاها ملموس و خدمات غیرملموس هستند. این معیار در مورد اکثر کالاها و خدمات صادق است ولی در مواردی نیز کالا و یا خدمات دارد. مثلاً" برق که مجموعه محسوب یک کالا طبقه‌بندی می شود غیرملموس است. یا خدماتی که محصول نهایی آنها به صورت اشیاء ملموس می باشد مانند عکس (که محصول نهایی خدمات عکاسی

است) یا نوار کامپیوتر (که محصول نهایی خدمات کامپیوتری است) .

۲ـ خدمات درهمان زمان تولید (ارائه) مصرف می شوندوغیرقابل ذخیره می باشند. این امر الزاما " درکلیه موارد صادق نیست .زیرا امکان دارد مدت زمان قابل ملاحظه‌ای بین تهیه (تولید) وبهره‌برداری (استفاده) ازیک خدمت مشاوره‌ای وجودداشته باشد .

۳ـ خدمات قابل حمل ونقل نیستندویا بایددرهمان محل تولیدمصرف شوند . کهاین مسئله درچنددههقبل کاربردداشتهودراین سالهابا توسعه وپیشرفت تکنیکهای مدرن اطلاعاتی بنظرنمی رسدصادق باشد .

۴ـ درخدمات مصرف داده‌ها یا نهاده‌ها(INPUTS) نیروی کار بیشتر و مصرف داده (نهاده یا هزینه INPUT) های واسطه کمتراز کالاهاست کهاین مشخصه بیش ازآنکه معیاری درتشکیک کالا ازخدمات باشداز خصوصیات آنهاست .

الف ـ۳ـ تعریف خدمات وموارددتداخلی :

هیل (HILL) درسال ۱۹۷۷ خدمات راچنین تعریف کرد :

"خدمت عبارت است ازتغییری که‌دروضعیت یک فردیا یک کالای متعلق به‌یک واحداقتصادی به‌وجودمی آیدواین تغییرنتیجه فعالیت یک واحداقتصادی دیگر با موافقت قبلی شخص یا واحداقتصادی اولیه است" .مشخصه این تعریف این بود که‌بیشترازخصوصیات ظاهری (مانندقابل حمل بودن یا نبودن)،براساس خصوصیات اقتصادی استوار بود ، منتها اختلاف دربرداشت های افرادازتغییردروضعیت شخص یا کالاها را بدنبال داشت .به‌عنوان مثال خدمات انتظامی وایمنی (ازقبیل خدمات پلیس ها یا مأمورین آتش نشانی) درواقع برای جلوگیری ازبروزتغییردروضعیت وشرایط شخص یا کالاصورت می گیرد .

هیل (HILL) درسال ۱۹۸۷ بااشاره‌به‌یک ویژگی مهم اقتصادی خدمات ، یعنی تماس بین تولیدکننده (ارائه‌کننده) ومصرف کننده (استفاده‌کننده)

خدمات ، تعریف قبلی خودرا چنین تکمیل کرد :

"... درتصور از خدمات این برداشت مستتر است که خدمات باید به برخی از واحدهای اقتصادی ارائه شود ، و این مشخصه در مورد تولید کالا که تولید کننده معمولا" اطلاعی از مصرف کننده کالای خود ندارد و وجود ندارد . یک کتاب ورزیا یک متصدی معمولا" جدا از مشتریان (مصرف کنندگان کالای خود) محصولاتی را تولید می کند ، لیکن یک معلم نمی تواند بدون شاگرد درس کند" .

مشخصه فوق جهت درک خصوصیات اقتصادی خدمات (مثلا" کوچکتر بودن واحدهای تولیدکننده خدمات نسبت به واحدهای تولیدکننده کالاها) وتفکیک مواردمرزی (مثلا" طبقه بندی کارت پستال به عنوان کالا و عکس به عنوان خدمات) مفید است ولی برای تفکیک همه کالاها از خدمات قابل استفاده نیست .

درتفکیک خدمات از کالاها هنوز مشکل دیگری وجود دارد . بسیاری از خدمات (نظیر خدمات رستوران) حاوی مقدار زیادی کالا است و کلیه کالاها عملا" حاوی خدمات هستند . برای تفکیک کالاها از خدمات معیار تصمیم گیری باید تا حد نهایی انجام معامله باشد بدون اینکه نسبت به جدا کردن عناصر کالا از عناصر خدمات اقدام شود . لذا اجاره دادن خانه و ارائه غذا در رستوران علیرغم محتوای کالایی زیاد آنها باید به عنوان خدمات محسوب شوند و در مقابل ، کتاب و صفحه گرامافون علیرغم محتوی خدمات زیاد آنان ، کالا منظور گردند .

هیل درتکمیل تعریف خود مشخصه دیگری را هم ذکر کرد که می تواند تصمیم گیری را در خیلی جهت تفکیک کالاها از خدمات باشد و آن اینست : "... اگر محصول نهایی جهت یک استفاده کننده ویژه تولید شود یک خدمت است ، ولی محصول نهایی عبارتست از آن که برای استفاده کننده تمام شود (و تولیدکننده آن از قبل نمی داند که چه کسی استفاده کننده آن خواهد بود) کالا محسوب می شود . لذا نرم افزارهای (Package) عمومی کالاست ولی نرم افزار سفارشی برای مشتری مطابق نظر و خواست و اهداف او یک خدمت است . یک عکس کدرنما نمای از شخصی برای الصاق درگذرنامه اش ، که

توسط یک عکاس برداشته شده یا خدمت است ولی عکس از یک انسان مشهور که
می تواند مثل یک کارت پستال توسط هرفردی درمغازه ها خریداری شود یک
کالاست .

با مساوی رفوق ، کتاب نیز کالاست ولی نه اینکه تمام فعالیتهای تولید کتاب
فقط تولید کالا باشد . زیرا وقتی که نویسنده ای کتاب می نویسد یعنی نسخه
دستنویس خود را به ناشر می دهد ، یک خدمت تولید کرده و تولید کتاب به عنوان
کالا از مرحله انتشار آغاز می شود .

با توجه به مساوی رها و شاخص فوق ، بخش ساختمان از نظر اینکه کالاهای
غیرمنقول ایجاد می کند ، به گفته برخی اقتصاددانان ، تولیدکننده خدمات
است . ولی نظریه غالب این است که محصول نهایی رشته فعالیت ساختمان
(ساختمانها ، پل ها ، جاده ها ، استکله ها و ...) کالاست زیرا ملموس است و می تواند
ذخیره شود .

درمقابل اجاره داراییهای ملموس ، مثلا" اجاره دادن خانه ، بدون تغییر
درماهیت آنها ، به عنوان خدمت درنظرگرفته می شود . نکته مهم دیگر این است
که منظور کردن کالاها و خدمات در آمار و تولید و هزینه مشابه یکسان نیست . مثلا"
درحالی که خرده فروشی در آمار و تولید یک فعالیت خدماتی است (که افزوده
بازرگانی ، تولید ناخالص آنست) ، درآمار هزینه (مصرف) ، اجناس خریداری
شده از خرده فروشی به عنوان کالا محسوب می شوند که شامل افزوده بازرگانی
آنها نیز می باشد .

معیار دیگر در تفکیک کالا از خدمت ، چگونگی معامله و مبادله است . مثلا" یک بطری نوشابه
خریداری شده از یک فروشگاه کالاست ، درصورتی که اگر این بطری در رستوران سرو شود به
عنوان جزئی از یک خدمت محسوب میشود . بهمین ترتیب دندان مصنوعی
که برای دندانپزشک در لابراتوار ساخته میشود کالاست ، لیکن همین دندان مصنوعی
وقتی توسط دندانپزشک در دهان بیمار کارگذاشته شود قسمتی از خدمت دندانپزشکی است .

ب ـ طبقه‌بندی فعالیتهای اقتصادی وخدمات :

طبقه‌بندی استاندارد بین المللی فعالیتهای اقتصادی (I.S.I.C. Rev.3) که تقریبا " در انطباق کامل با سیستم حسابهای ملی سازمان ملل (S.N.A. ۱۹۶۸) می باشد دارای ۱۰ کد ۱ رقمی برای بخش های عمده فعالیتهای اقتصادی بشرح زیر می باشد :

بخش عمده ۱ـ کشاورزی وشکار

بخش عمده ۲ـ معدن وحفاری

بخش عمده ۳ـ صنعت

بخش عمده ۴ـ برق ،گاز وآب

بخش عمده ۵ـ ساختمان

بخش عمده ۶ـ عمده‌فروشی ،خرده‌فروشی ،رستوران وهتلداری

بخش عمده ۷ـ حمل ونقل ،انبارداری وارتباطات

بخش عمده ۸ـ خدمات مالی ،بیمه ، املاک ومستغلات وکسب وکار

بخش عمده ۹ـ خدمات عمومی ،اجتماعی وشخصی

بخش عمده ۱۰ـ فعالیتهای نامشخص

همانطورکه ملاحظه می شوددرطبقه‌بندی فوق ازبخش عمده ۶ ببعد به خدمات اختصاص داده شده است .

درتجدیدنظرسوم طبقه‌بندی بین المللی فعالیتها (I.S.I.C. Rev.3) جمعا" ۱۷ طبقه از طبقات عمده ۱۷ گانه آن به خدمات اختصاص داده شده که شامل ۱۲۵ دسته فعالیت با کد ۴ رقمی است .

ج ـ شمول خدمات وانواع فعالیتهای خدماتی برحسب گدهای مربوطه درطبقه‌بندیهای موردعمل حسابهای ملی :

بخش های عمده ۶ و ۷ و ۸ و ۹ طبقه‌بندی بین المللی فعالیتها (I.S.I.C.Rev.3)

بترتیب شامل ۵۸ گروه فعالیت خدماتی با کد ۴ رقمی به شرح زیر است :

۶۱۰۰- عمده فروشی

۶۲۰۰- خرده فروشی

۶۳۱۰- رستوران ها ، کافه ها و سایر اماکن صرف غذا و نوشیدنی

۶۳۲۰- هتل ها ، مسافرخانه ها ، اردوگاهها و سایر اماکن مشابه

۷۱۱۱- راه آهن

۷۱۱۲- حمل مسافر در شهر ، حومه و بین شهرها

۷۱۱۳- سایر وسایل حمل زمینی مسافر

۷۱۱۴- حمل و نقل زمینی بار

۷۱۱۵- حمل و نقل به وسیله خطوط لوله

۷۱۱۶- خدمات کمکی حمل و نقل زمینی

۷۱۲۱- حمل و نقل دریایی و آب های ساحلی

۷۱۲۲- حمل و نقل آبی داخلی

۷۱۲۳- خدمات پشتیبانی حمل و نقل آبی

۷۱۳۱- وسایل نقلیه هوایی

۷۱۳۲- خدمات پشتیبانی حمل و نقل هوایی

۷۱۹۱- خدمات فرعی حمل و نقل

۷۱۹۲- انبارداری

۷۲۰۰- ارتباطات

۸۱۰۱- موسسات پولی

۸۱۰۲- سایر موسسات مالی

۸۱۰۳- خدمات مالی

۸۲۰۰- بیمه

۸۳۱۰- مستغلات داری

۸۳۲۱ـ خدمات حقوقی

۸۳۲۲ـ خدمات حسابداری، ممیزی ودفترداری

۸۳۲۳ـ خدمات کامپیوتری

۸۳۲۴ـ خدمات مهندسی مشاوره‌ای، معماری وفنی ومهندسی

۸۳۲۵ـ خدمات تبلیغاتی

۸۳۲۹ـ خدمات کسب وکار طبقه‌بندی نشده‌درجای دیگر

۸۳۳۰ـ اجاره وکرایه‌دادن ماشین آلات وتجهیزات

۹۱۰۰ـ اداره عمومی ودفاع

۹۲۰۰ـ خدمات سینمایی ونمایشی

۹۳۱۰ـ خدمات آموزشی

۹۳۲۰ـ خدمات تحقیقاتی وپژوهشی

۹۳۳۱ـ خدمات پزشکی، دندانپزشکی وسایرخدمات بهداشتی ودرمانی

۹۳۳۲ـ خدمات دامپزشکی

۹۳۴۰ـ خدمات خیریه

۹۳۵۰ـ اتحادیه‌های کسب وکار، حرفه‌ای وکارگری

۹۳۹۱ـ خدمات مذهبی

۹۳۹۹ـ خدمات اجتماعی طبقه‌بندی نشده‌درجای دیگر

۹۴۱۱ـ تولیدفیلم سینمایی

۹۴۱۲ـ توزیع ونمایش فیلم سینمایی (سینماها)

۹۴۱۳ـ پخش برنامه‌های رادیوتلویزیون

۹۴۱۴ـ تولیدکنندگان خدمات تئاتری وسرگرمی

۹۴۱۵ـ نویسندگان، آهنگسازان وسایرهنرمندان مستقل طبقه‌بندی نشده درجای دیگر.

۹۴۲۰ـ کتابخانه‌ها، موزه‌ها، باغهای گیاهشناسی، باغهای وحش وسایر

خدمات فرهنگی طبقه بندی نشـــده درجای دیگـر .

۹۴۹۰- خدمات ورزشی ،سرگرمی وتفریحی طبقه‌بندی نشده‌درجای دیگر

۹۵۱۱- تعمیرکفش وکالاهای چرمی

۹۵۱۲- تعمیروسایل برقی

۹۵۱۳- تعمیروسایل نقلیه‌موتوری وموتورسیکلت

۹۵۱۴- تعمیرساعت وجواهرآلات

۹۵۱۹- سایرخدمات تعمیری طبقه‌بندی نشده‌درجای دیگر

۹۵۲۰- خدمات لباسشویی وتعمیرورنگرزی پوشاک وکالاهای پارچه‌ای وفرش

۹۵۳۰- خدمات خانگی

۹۵۹۱- خدمات پیرایش وآرایش

۹۵۹۲- استودیوهای عکاسی

۹۵۹۹- خدمات شخصی طبقه‌بندی نشده‌درجای دیگر

۹۶۰۰- ارگانهای بین المللی وسایرسازمانهای خارج ازکشور

بدین ترتیب از ۱۶۰ کد ۴ رقمی موجوددرطبقه‌بندی I.S.I.C. Rev2 حـــدود ۶۰ کدآن به‌خدمات اختصاص داده‌شده‌است .

درتجدیدنظرسوم طبقه‌بندی بین المللی فعالیتها (۱۹۹۰ I.S.I.C Rev3) ۱۱ طبقه‌از طبقات عمده ۱۷ گانه‌فعالیتها به‌خدمات اختصاص داده‌شده‌است که‌شامل حدود ۱۲۵ دسته‌فعالیت با کد ۴ رقمی است .بدلیل اینکه‌استفاده‌ازتجدیدنظـــر سوم طبقه‌بندی مزبورهنوزدرکشورما شروع نشده‌وطبقه‌بندی مزبورا صولا"درانطبا ق با سیستم حسابهای ملی تجدیدنظرشده (Revised S.N.A.) باید پیاده شودکـــه سیستم مزبورقرار است از ۱۹۹۳ برای اجراء به‌کشورها توصیه‌گردد . لذا ازبیـــان وشرح فعالیت‌های ۱۲۵ گانه‌خدمات با کد ۴ رقمی طی طبقه‌بندی مزبـــــور خودداری می گردد .

د ـ سهم بخش خدمات درتولیدناخالص ودرآمدملی کشور :

جدول ۱ سهم خدمات را درتولیدناخالص داخلی ، تولیدناخالص ملی و درآمدملی کشور درسالهای ۶۸-۱۳۶۱ به قیمت های جاری نشان می دهد :

جدول ۱ـ سهم خدمات درتولیدناخالص داخلی ،تولیدناخالص ملی ودرآمدملی ایران درسالهای ۶۸-۱۳۶۱ به قیمت های جاری

(درصد)

شرح	۱۳۶۱	۱۳۶۲	۱۳۶۳	۱۳۶۴	۱۳۶۵	۱۳۶۶	۱۳۶۷	۱۳۶۸
سهم خدمات درتولیدناخالص داخلی	۵۰/۰	۵۲/۸	۵۷/۰	۵۵/۹	۵۹/۴	۵۷/۸	۶۱/۳	۶۱/۰
سهم خدمات درتولیدناخالص ملی	۴۹/۴	۵۲/۷	۵۵/۸	۵۳/۹	۵۷/۵	۵۶/۱	۶۰/۰	۵۹/۶
سهم خدمات دردرآمدملی	۵۲/۶	۵۸/۹	۶۱/۵	۶۱/۱	۶۴/۷	۶۲/۴	۶۷/۴	۶۷/۰

مآخذ : بانک مرکزی جمهوری اسلامی ایران .

جدول ۲ـ سهم خدمات درتولیدناخالص داخلی،تولیدناخالص ملی ودرآمدملی ایران درسالهای ۶۸-۱۳۶۴ به قیمت های ثابت سال ۱۳۶۱

(درصد)

شرح	۱۳۶۴	۱۳۶۵	۱۳۶۶	۱۳۶۷	۱۳۶۸
سهم خدمات درتولیدناخالص داخلی	۵۱/۸	۵۰/۰	۴۸/۲	۴۷/۷	۴۷/۰
سهم خدمات درتولیدناخالص ملی	۵۰/۰	۴۸/۲	۴۶/۸	۴۶/۶	۴۵/۸
سهم خدمات دردرآمدملی	۵۷/۰	۵۸/۸	۵۵/۵	۵۷/۱	۵۶/۱

مآخذ : بانک مرکزی جمهوری اسلامی ایران .

جدول ۳- سهم ارزش افزوده بخشهای عمده اقتصادی در مجموع ارزش افزوده‌های کشور به قیمت های ثابت سال ۱۳۶۱

(درصد)

بخش های عمده اقتصادی	۱۳۶۴	۱۳۶۵	۱۳۶۶	۱۳۶۷	۱۳۶۸
جمع :	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
کشاورزی	۱۷/۵	۲۰/۲	۲۱/۱	۱۹/۵	۱۹/۵
نفت	۱۲/۸	۱۲/۱	۱۴/۰	۱۵/۹	۱۶/۵
صنایع و معادن، آب و برق و گاز و ساختمان	۱۸/۱	۱۷/۹	۱۶/۶	۱۶/۹	۱۷/۰
خدمات	۵۱/۱	۴۹/۸	۴۸/۳	۴۷/۷	۴۷/۰

مآخذ: بانک مرکزی جمهوری اسلامی ایران - اداره حسابهای اقتصادی .

همانطورکه در ۳ جدول فوق ملاحظه می شود سهم خدمات در تولید ناخالص داخلی ،تولید ناخالص ملی و درآمد ملی در سالهای اخیر همواره با بیش از ۶۰٪ و حتی نزدیک به ۷۰٪ (به قیمتهای جاری) بیشترین بوده و به قیمتهای ثابت سال ۱۳۶۱ همواره با بیش از ۵۰٪ و یا کمی کمتر از آن بیشترین سهم را نسبت به سایر بخش های اقتصادی داشته است .

بهمین ترتیب سهم ارزش افزوده خدمات در مجموع ارزش افزوده های کشور با اینکه در سالهای اخیر روبه کاهش بوده است معهذا کماکان بیشترین سهم را همه ساله داشته است و بیش از ۵۰٪ یا کمی کمتر از آن بوده است .

۵ - چگونگی آمار بخش خدمات در کشور :

متاسفانه با وجود سهم عمده و تقریبا " همیشه چندین برابر خدمات در کلیه کشورهای

کلان اقتصادی نظیر تولید ناخالص داخلی ، تولید ناخالص ملی ، درآمد ملی ، و ارزش افزوده کشور، آمار دقیق ، بهنگام و جامعی از آن در کشور در دست نیست و بطور کلی آمار بخش خدمات در ایران دارای سابقه و تاریخ مهم ، کوتاه و ناقص است .

در جایی که آمار مربوط به فعالیتهای کشاورزی ، صنعت ، معدن و نفت از سالها قبل (مثلاً" در مورد صنعت از حدود ۳۰ سال قبل) تهیه شده و در دسترس بوده است ، در مورد خدمات و بخصوص بخشهای بسیار مهم و عمده آن نظیر بهداشت و درمان ، آموزش ، فرهنگ و هنر و ورزش ، حمل و نقل و ... تاکنون هیچگونه آمار جامع ، دقیق و مبتنی بر اصول علمی آماری تهیه نشده است .

مسأله بسیار مهم در کلیه آمارهای مربوط به تولید ناخالص داخلی ، تولید ناخالص ملی و درآمد ملی و نظایر آن این است که آمار مربوط به خدمات بطور تقریبی و بخصوص از روش باقیمانده تخمین و برآورد شده است . در حالیکه کل هزینه و سرجمع هلنیز همیشه (و در غیاب آمار کامل و محیّر از بخشهای مختلف بخصوص خدمات) تقریبی و با برآوردی همراه با فروض و احتمالات متعدد بوده است .

هرگاه آمار مربوط به خدمات در دو قسمت زیر بررسی شود مسائل و کمبودها روشن تر بنظر خواهد رسید :

۱ـ آمار پایه ای بخش خدمات اعم از شبکتی یا آمارگیری :

آمار پایه ای مربوط به همه بخش های خدماتی بطور شبکتی وجود ندارد و فقط در سوار دهها فعالیتهای خدماتی که دولت عهده دار تأمین مالی و کنترل آنها می باشد، نظیر آموزش دولتی ، بهداشت و درمان دولتی و اداره عمومی و دفاع ، نظم و امنیت عمومی و بعضاً " فعالیتهای خدماتی مالی و پولی که با شکها عمدتاً " عامل آنها هستند ، آمار پایه ای مربوط به تعداد و توزیع کارگاهها در مناطق مختلف شهری و روستایی کشور ، تعداد کارکنان آنها و مجموع های هزینه و درآمد آنها

درنتایج بودجه سنتی دستگاههای دولتی مطرح شده و از سالها قبل وجود دارد.

در مورد تهیه آمارهای پایه ای خدمات از طریق آمارگیری و فهرست برداری نیز بجز سرشماری های عمومی نفوس و مسکن کشور توسط مرکز آمار ایران (که آنها نیز هم عمدتاً" برای تهیه آمارهای پایه ای کارگاه طراحی و اجراء نشده اند) تنها چند مورد آنهم در مناطق شهری کشور توسط وزارت کار و امور اجتماعی و با همکاری مرکز آمار در سالهای ۱۳۵۲ و ۱۳۶۳ طرحهایی تهیه و اجراء شده است.

لذا از آنجا که جدآمارهای پایه ای ثبتی و یا آمارهای پایه ای برمبنای آمارگیری و فهرست برداری فاقد تعاریف و مفاهیم و طبقه بندیها و حجم و اصول استاندارد یکسانی می باشند، می توان گفت آمارهای پایه ای در مورد خدمات بطور کلی و جامع تا کنون تهیه و ارائه نشده است تا بتوان مستقیماً" و با اطمینان از آنها استفاده کرد و یا بر پایه آنها طرحهای آمارگیری تهیه و اجراء نمود.

۲- آمار مربوط به تولید ناخالص، هزینه های واسطه، ارزش افزوده، درآمد و تشکیل سرمایه بخش خدمات، اعم از ثبتی یا آمارگیری:

گفتیم آمارهای کامل در مورد تعداد کارگاههای خدماتی و تعداد کارکنان آنها در سطح کل کشور وجود ندارد. در مورد اطلاعات مربوط به تولید یا هزینه، درآمد و تشکیل سرمایه همه کارگاههای خدماتی نیز به طریق اولیٰ آماری بصورت ثبتی موجود نیست. مگر در مورد خدمات دولتی و یا تعدادی خدمات خصوصی که در قوانین و لوایح بودجه دولت و یا پرونده های مالیاتی چرّف و مشاغل و کارگاهها در وزارت دارایی و یا وزارت اقتصاد و دارایی آمارو ارقام ارائه شده است.[1] آمارو ارقام مزبور نیز سهم بودجه بطور مستقیم برای کارشناسان و برنامه ریزان و استفاده کنندگان قابل استفاده نیست، زیرا از تعاریف، طبقه بندی، شمول و اصول استاندارد و یکسانی برخوردار نمی باشد.

۱- مشخص وزارت دارایی سالنامه وزارت اقتصاد و دارایی فصلی در مورد برخی از خدمات آمارهایی حاوی تعداد و نام و نشانی مکانی کارگاهها و نوع فعالیت آنها بعنوان ارقام تخمینی درآمدشان برای ارزیابی وضعیت مالی کارگاههای مزبور و تخمین مالیات آنها دارد.

بانک مرکزی از سالها قبل و بدنبال برآورد تولید و درآمد داخلی و ملی کشور و به منظور داشتن آمار مربوط به تولید ، هزینه‌ها واسطه ، درآمد و تشکیل سرمایه ثابت ، در مورد برخی خدمات و آنهم از بخش خصوصی ، نظیر خدمات آرایش و پیرایش ، تاکسی ، لباسشویی ، گرمابه ، رادیولوژی ، سینما و ... آمارگیریهایی در تعدادی از شهرهای کشور انجام داده است . منتها طرح فنی ، روش های نمونه گیری یا آمارگیری (بطور کلی و تشریحی) و اصولاً آمار پایه ای و نرمی آمارگیریهای مزبور مشخص نیست .

دفتر آمار وزارت اقتصاد نیز در سالهای دهه ۴۰ بنا به پیشنهاد سازمان برنامه و تقاضای بانک مرکزی ، برای برآورد درآمد ملی در بخش خدمات خصوصی همزمان با اجرای آمارگیریهای صنعتی در شهرهای کشور از کارگاههای خدماتی نظیر موسسات آموزشی خصوصی ، گرمابه ها و ... در شهرها و به روش نمونه گیری آمارگیری کرد و از فعالیت آنها در سال ۱۳۴۹ اطلاعاتی انتشار داد .

مرکز آمار ایران بدنبال تصویب قانون سال ۱۳۵۳ و تشکیل مدیریت آمارهای مالی و محاسبات ملی از سال ۱۳۵۴ در مورد تعدادی از فعالیتهای خدماتی نظیر آموزش ، بهداشت و درمان ، خدمات خیریه و مذهبی ، بانکداری ، مسکن ، بیمه و ... اقدام به محاسبات و برآوردهایی نمود که حاصل آنها در بیش از ۳۰ جلد نشریه حاوی آمار و اطلاعات مربوط به تولید ناخالص ، هزینه های واسطه ، ارزش افزوده ، درآمد ، تشکیل سرمایه و تعداد کارکنان و پرداخت به عوامل تولید در این فعالیتها بود . گزارشات مزبور عمدتاً " تحقیقاتی " بوده و برپایه آمارگیریهای مشخص و موردی که آنها هم خود براساس آمارچارچوبی مستحکم و پیشگامی طراحی و تهیه نشده بوده اند .

به عنوان مثال برای کار عظیم تهیه جدول داده ستانده های اقتصاد کشور در سال ۱۳۵۲ ، مدیریت آمارهای مالی و محاسبات ملی مرکز آمار در سالهای ۵۵ و ۱۳۵۶ تعدادی آمارگیری در زمینه بسیاری رشته فعالیتهای خدماتی در

مناطق شهری کشور، تهیه و اجرا کرد. که چارچوب و قالب آنها را عمدتا " آمسار با بهای وزارت کار و مرکز آمار در سال های ۱۳۵۲ و ۱۳۵۳ تشکیل می داد و شکیل آمار و اطلاعات مزبور برای پوشش همه خدمات و در سطح کل کشور (اعم از شهری و روستایی) بر عهده کارشناسان و با کارگاه های شناسی و برآوردهای مبتنی بر نظرات تخصصی و تجربی آنان انجام گرفت .

تهیه و ارائه منتشریه برآورد تولید و ارزش افزوده خدمات در مناطق شهری کشور در سال ۱۳۵۶ توسط اینجانب یکی از همان اقدامات و گزارشات فوق الذکر می باشد .

و- تهیه ، تدوین و اجرای طرحهای مختلف آماری گیری از کلیه فعالیتهای اقتصادی به ویژه فعالیتهای خدماتی برای تهیه و تنظیم حسابهای ملی و بالاخص جدول داده ها - ستانده های اقتصاد کشور :

از سال ۱۳۶۷ برای تهیه و تنظیم جدول داده ستانده های اقتصاد کشور و طی مراحل اولیه آن خلاء های آماری عمیق و مهم در مورد انواع فعالیتهای خدماتی مشخص شد و برای رفع نیاز و پر کردن خلاء های مذکور تهیه و اجرای طرحهای آماری برای آمارگیری از قلمات لازم افتاد ، بطوریکه همزمان با تهیه و اجرای طرحهای آماری در مورد انواع فعالیتهای اقتصادی (صنعت ، معدن ، کشاورزی ، ساختمان و ...) ، تهیه و تنظیم و اجرای طرحهای متعدد آمارگیری از کلیه فعالیتهای خدماتی در اولویت مناسب قرار گرفت ، در نتیجه طی سالهای ۱۳۶۸ تا کنون حدود ۳۰ طرح به شرح زیر در زمینه آمارگیری از فعالیتهای خدماتی توسط مرکز آمار ایران تهیه و تنظیم و با همکاری سازمان برنامه و بودجه ، ستانها ، وزارتخانه های آموزش و پرورش ، کار و امور اجتماعی ، فرهنگ و ارشاد اسلامی و سازمانهای مختلف نظیر سازمان دامپزشکی ، در مناطق شهری و روستایی کشور

اجرا شده است . استخراج نتایج طرحهای مزبور در مراحل بایانی کار می باشد .

نکته مهم در مورد طرحهای مزبور فهرست (فریم) مورد استفاده آنها است که عمدتا اطلاعات حاصل از فهرست برداری سرشماری عمومی نفوس ومسکن ۱۳۶۵ کشور توسط مرکز آمار بوده است .

اطلاعات فریمی مزبور دارای اشکالاتی بود که پس از کار زیاد کارشناسی و جهت حفظ وکاربرد آن به عنوان یک چارچوب واحد برای کلیه طرحهای آماری (وبه ویژه خدمات) قابلیت استفاده یافت که خود یکی از ویژگیهای مثبت وسهم طرحهای مزبور را تشکیل می دهد واین همین استفاده از فریم وچارچوب واحد برای طرحهای مختلف آماری است .

دومین ویژگی طرحهای مزبور جامعیت وگستردگی اطلاعات مأخوذه طبق اجرای آنها است بخصوص نظر به اینکه اطلاعات حاصل از طرحهای مزبور می بایستی درخانه های مربوط به خود در جدول داده ها ـ ستانده ها ظاهر شود وجدول فرایب فنی را هم تنذیه کند ، لذا طی آمارگیریهای مزبور اطلاعات تفصیلی در مورد انواع هزینه های جاری وسرمایه ای آنها به تفکیک گدهای ۴ رقمی (براساس طبقه بندی بین المللی فعالیتها) منبع تهیه وارائه آنها پرسش واخذ گردید .

سومین مشخصه مهم طرحهای مذکور پرسش واخذ اطلاعاتی در باره مهم واردات هزینه های واسطه آنها برحسب نوع هزینه بوده است .

چهارمین وشایدتازه ترین وپیچیده ترین ویژگی طرحهای مذکور پرداختن به مسأله فعالیت اصلی وفعالیت یا فعالیتهای فرعی در واحد مشاهداتی ومورد مراجعه بعنی کارگاه بوده است که این مورد برای اولین بار در ایران وشاید در بسیاری از کشورها مورد ملاحظه قرار گرفته است .

پنجمین مشخصه طرحهای مزبور مجموعه بودن وکامل بودن آنها است بطوریکه کل آنها کل فعالیتهای خدماتی با گدهای ۴ رقمی در طبقه بندی بین المللی فعالیتها را در کل کشور تحت پوشش قرار داده اند .

ویا بلاخره آخرین (ویا اولین) شعبه مهم این طرحها روش یکسان آمارگیری از طریق مراجعه و مصاحبه و تماس مستقیم آمارگیران با مشمولان کارگاهها بوده است .

فهرست طرحهای تهیه و تنظیم شده ، آموزش داده شده ، و اجرا شده بدقرار زیر است :

۱ـ طرح آمارگیری از عمده فروشی و خرده فروشی

۲ـ طرح آمارگیری از محل های صرف غذا و شوشبندی

۳ـ طرح آمارگیری از اقامتگاههای عمومی

۴ـ طرح آمارگیری خدمات حمل و نقل جاده ای مسافر

۵ـ طرح آمارگیری خدمات حمل و نقل جاده ای بار

۶ـ طرح آمارگیری خدمات کلی و جنبی حمل و نقل (شامل پارکینگ و باسکول داری و پاکسس)

۷ـ طرح آمارگیری خدمات حمل و نقل آبی

۸ـ طرح آمارگیری خدمات حمل و نقل هوایی

۹ـ طرح آمارگیری خدمات انبارداری

۱۰ـ طرح آمارگیری خدمات مشاوره ای ساختمان و زمین (شامل سنگاههای معاملات املاک)

۱۱ـ طرح آمارگیری صنوقهای قرض الحسنه

۱۲ـ طرح آمارگیری خدمات کسب و کار

۱۳ـ طرح آمارگیری از نمایندگیهای بیمه

۱۴ـ طرح آمارگیری خدمات وسایل نقلیه موتوری

۱۵ـ طرح " " پیرایش و آرایش

۱۶ـ " " " بیمارستانی

۱۷ـ " " " غیربیمارستانی

۱۸ـ طرح آمارگیری خدمات آموزشی (آموزش عالی)

۱۹ـ " " " " (آموزش عمومی)

۲۰ـ " " " " (آموزش فنی وحرفه‌ای وسایرخدمات آموزشی)

۲۱ـ " " " پژوهشی وتحقیقی

۲۲ـ " " " دامپزشکی

۲۳ـ " " " ارشادی ـ مذهبی وخیریه

۲۴ـ " " " فرهنگی ـ ورزشی

۲۵ـ " " " اداری ـ قضایی

۲۶ـ " " " شهری وبهسازی محیط

۲۷ـ " " سایرخدمات اجتماعی شخصی وخانگی

ز ـ نتایج وپیشنهادات :

تهیه واجرای طرحهای متعددآماری درزمینه‌کلیه‌فعالیتهای اقتصادی بدوبیژه‌خدمات (برای اولین بار با مشخصات قبلاً گفته‌شده‌درکشور) نتایج قابل توجه وبعضاً بسیار راهنما وراهگشا داشته‌است که عمدتاً از برخورد با مسائل ومشکلات وحل ورفع آنها ویا حتی داشتن آنها وعدم حل ورفعشان حاصل آمده‌است . عمده مسائل ومشکلات مزبورعبارتند از :

۱ـ نداریف ومفاهیم غیراستانداردومتغیردرموردا کثرکمّیت ها وفعالیتها

۲ـ عدم استفاده‌کامل ازطبقه‌بندی استانداردبین المللی فعالیتهادر موردآماربایه‌ای

۳ـ نارسایی ونقص درتعریف مفاهیم اصلی وکلیدی بخصوص "کارگاه" ، " فعالیت اقتصادی" ، " فعالیت اصلی" یا "فعالیت عمده" ، " فعالیت فرعی" ، هزینه" ، "درآمد" ، " جاری " ، " سرمایه‌ای" و...

۶ـ مشکلات حین اجرا و آمارگیری طرحهای فهرست برداری (ناشی از مأمور، ناشی از فرم ها، ناشی از راهنماها، ناشی از آموزش، ناشی از امکانات و محدودیتهای زمانی و مکانی و مالی آمارگیری و ناشی از برخورد با مشکوبی و شرایط اجتماعی ـ اقتصادی جامعه و ...).

۷ـ مشکلات خاص اطلاعات فهرست برداری و تشخیص نادرست کارگاه و فعالیت و فعالیت اصلی.

۸ـ مسائل ناشی از تخصیصی بودن اطلاعات خواسته شده در مواجهه با ارتباط سیستم حسابداری اکثریت قریب به اتفاق کارگاهها.

۹ـ مشکلات ناشی از زنده شمار آماربایه ای بینگام و حداقل در مورد فعالیتهای اقتصادی غیرمتشکل، یا غیرمشهود با مکان (داخل خانوار) وبخصوص فعالیتهای اقتصادی بدون مکان (از قبیل، خدمات تحمل ونقل، آموزش، بهداشت و درمان، خدمات خصوصی وشخصی و خانگی).

با توجه به حاصل سالها بررسی، مطالعه، تجربه و کارگاه شناسان متعدد در زمینه آمار خدمات وبخصوص براساس نتایج طرحهای آماری مذکور و در برخورد با مسائل و مشکلات فوق الذکر موارد زیر عمده و مطالعه و بررسی و کار آماری در راستای آنها اساسی بنظر می رسد:

۱ـ ایجاد و ارائه تعاریف و طبقه بندیهای مناسب و پذیرش تعاریف استاندارد قابل مقایسه بین المللی توسط وزارتخانه ها، نهادها، سازمانها و ارگانهای دستاندرکار مسائل و امور اقتصادی ـ آماری.

۲ـ پذیرش طبقه بندی استاندارد بین المللی (I.S.I.C) و انطباق فعالیتهای اقتصادی کشور بستة آن توسط وزارتخانه ها، سازمانها و ارگانهای فوق الذکر

۳ـ بطور کلی بررسی و پذیرفتن نظام حسابهای ملی توصیه شده سازمان ملل (S.N.A) توسط وزارتخانه ها و سازمانهای فوق در تهیه و ارائه آمار و اطلاعات اقتصادی.

۴ ـ تهیه و تنظیم و اجرای مطلوب یک طرح فهرست برداری برای تهیه و ارائه آمارهای قریبی و پایه ای در باره کلیه فعالیتهای اقتصادی در کل کشور اعم از کارگاهی و غیرکارگاهی ساکن (داخل خانوار) یا بی مکان .

۵ ـ ارزشیابی طرحهای تهیه و اجرا شده و یافتن موارد ضعف و نارساییهای آنها

۶ ـ ارزشیابی علمی ، عملی و واقعی از امکانات و محدودیتها و بخصوص در وهله اول ارزشیابی اضافه کلی و تفضیلی طرحها در چارچوب اصول اولیه برنامه ریزی (تعیین هدف تعیین اولویت) با توجه به زمان و مکان و منابع مالی و نیروی انسانی و نیازهای واقعی و اساسی کشور .

۷ ـ تهیه و اجرای طرحهای آماری در مورد خدمات بر حسب اولویت های تعیین شده و با توجه به موارد فوق بخصوص بندهای ۵ و ۶

۸ ـ انجام سیستماتیک و منظم و با برنامهٔ کار تهیه ، تنظیم و اجرای طرحها ، نظارت بر اجرای آنها ، بازبینی و استخراج و انتشار بهنگام نتایج آنها .

منابع و مآخذ :

۱ـ سالنامه‌های آماری کشور ، مرکز آمار ایران .

۲ـ گزارشات مختلف آماری درباره برآورد تولید ، ارزش افزوده و هزینه‌های واسطه فعالیتهای مختلف خدماتی ـ مرکز آمار ایران .

۳ـ مجموعه طرحهای مختلف آمارگیری خدمات (منجمله مجموعه طرح آمارگیری خدمات وسائل نقلیه موتوری ، آموزش و ...) .

۴ـ گزارشات بانک مرکزی درباره درآمد ، هزینه و ارزش افزوده فعالیتهای خدماتی (انتشارات ها ، رادیولوژی ها و ...) .

۵ـ گزارش درآمد ملی ایران در سالهای ۵۶-۱۳۳۸ ـ بانک مرکزی ایران .

۶ـ طبقه‌بندی استاندارد بین المللی فعالیتهای اقتصادی (I.S.I.C. Revis).

۷ـ " " " " " (I.S.I.C. Rev.3).

۸ـ نظام حسابهای ملی (A System of National Accounts 1968, U.N.).

۹ـ ترجمه تعاریف و مفاهیم و اصطلاحات نظام حسابهای ملی، مدیریت آمارهای مالی و محاسبات ملی ـ مرکز آمار ایران .

۱۰ـ ترجمه فصل ششم نظام حسابهای ملی (تولید ، مصرف ، تشکیل سرمایه و ...) ، مدیریت آمارهای مالی و محاسبات ملی ـ مرکز آمار ایران .

۱۱ـ ترجمه فصل هشتم نظام حسابهای ملی (حسابها و جداول استاندارد) ، "مدیریت آمارهای مالی و محاسبات ملی ـ مرکز آمار ایران .

۱۲ـ ترجمه راهنمای نظام حسابهای ملی، " مدیریت آمارهای مالی و محاسبات ملی ـ مرکز آمار ایران .

۱۳ـ فهرست مجموعه کدهای مربوط به طبقه‌بندی مورد عمل در سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ کشور ـ مرکز آمار ایران .

۱۴ـ نتایج آمارگیری از مشخصات پایه‌ای کارگاههای مناطق شهری کشور در سالهای ۱۳۵۲ و ۱۳۵۳ و ۱۳۶۲ ـ وزارت کار و امور اجتماعی با همکاری مرکز آمار ایران .

۱۵ـ گزارش "مفهوم خدمات در آمار" ـ سازمان ملل متحد .

شاخص‌های قیمت

محمدحسین فرنودی

بانک مرکزی جمهوری اسلامی ایران

شاخص‌های قیمت وآخرین تجدیدنظرآنها در سال ۱۳۶۱

تهیه شاخص های قیمت برای مطالعه اوضاع اقتصادی وتحقیق درتغییرات وضع زندگی مردم هرکشورامری ضروری است . سابقه نشان میدهدکه تهیه شاخصهای قیمت درجهان درآغاز توسط سازمانهای خصوصی وافراد شروع شدولی بعدا " مقامات دولتی آنهارابرای مقاصدمختلف بشکل کاملتری محاسبه وبورداستفاده قراردادند .

بطورکلی شاخصهای قیمت تغییرات قیمت تعدادمعینی ازکالاهاوخدمات رابصورت درصدنشان میدهند .برای تهیه شاخص معمولا" سال معینی رابعنوان سال پایه انتخاب میکنند و قیمت کالاهاوخدمات موردنظررادرماهها وسالهای بعد نسبت به قیمتهای سال پایه مقایسه میکنند .

متداول ترین شاخص های قیمت یکی شاخص بهای کالاهاوخدمات مصرفی ودیگری شاخص بهای عمده فروشی کالاهاست که ذیلا " بذکرموارد استفاده هرکدام ، چگونگی محاسبه آنهـا و آخرین تجدیدنظردرآنهادرسال ۱۳۶۱ ، پرداخته میشود .

۱ـ شاخص بهای کالاها وخدمات مصرفی

۱ـ۱ـ مقدمه : شاخص بهای کالاها وخدمات مصرفی معیار سنجش تغییرات قیمت‌ها است ثابت وعمده‌ترین کالاها وخدمات است که توسط خانوارهای شهرنشین ایران بمصرف میرسد. این شاخص بعنوان وسیله ای برای اندازه گیری سطح عمومی قیمتهای خرده فروشی یکی‌ازبهترین روابط سنجش قدرت خرید داخلی‌پول کشوربشمارمیرود. شاخص مزبوربعنوان میزان الحرارهای جهت سنجش قیمتهادورداستفاده قرارمیگیرد.

۱ـ۲ـ تاریخچه : شاخص بهای کالاها وخدمات مصرفی اولین باردرایران درسال ۱۳۱۵ تهیه شد وتاسال ۱۳۲۸ درآن‌تجدیدنظری بعمل‌نیامد. شاخص مزبور برمبنای ضرائبی کلیه‌از بررسی هزینه ۱۷ خانوارنمونه در۷شهربزرگ کشور بدست آمده بود، برپایه ۱۰۰= ۱۳۱۵ محاسبه گردید.

در تجدید نظری که در سال ۱۳۲۸ صورت گرفت ، زمینه‌های مختلف تهیه شاخص ابعاد تازه ای یافت . تعداد خانوارهای نمونه ای که ترکیب هزینه آنها مبنای ضرائب شاخص قرار گرفت به ۳۲۳۷ خانوار افزایش یافت وتعداد شهرهای نمونه وکالاها وخدمات مورد قیمت گیری بمراتب بیشتر شد. بعلاوه در مفاهیم ، تعاریف و روش محاسبه برخی از گروههای شاخص تغییرات اساسی بوجود آمد. محاسبه این شاخص تا سال ۱۳۴۸ ادامه یافت .

معمولا" در غالب کشورها رسم متداول بر این است که هر ده سال یکبار در شاخصها تجدید نظر بعمل می‌آورند. بهمین لحاظ در سال ۱۳۴۸ با توجه به تغییرات قابل ملاحظه ای که در سطح درآمد و ترکیب مصرف خانوارها پدید آمده بود، در این شاخص تجدید نظر بعمل آمد.

درسالهای ۱۳۵۰ به بعد بسبب حرکات داخلی جمعیت ومهاجرت تعداد کثیری از خانوارهای روستائی‌به شهرها، رشد جمعیت شهری بسیار سریع گردید و در نتیجه پیدایش تعدادی شهر را موجب شد. بعلاوه با افزایش نسبتا" سریع درآمدها وبالا رفتن قدرت خرید مصرف کنندگان همراه با به بازار آمدن محصولات جدید وتبلیغات وسیعی که برای فروش کالاهای گوناگون ازطریق رسانه‌های گروهی صورت می‌گرفت ، تغییرات قابل ملاحظه ای را در ترکیب هزینه‌های مصرفی خانوارهای شهری پدید آورد. افزایش واردات کالا در نتیجه بالا

رفتن در آمد ارزی کشور و تغییر ترکیب آنها، نوع و مقدار اقلامی از کالاهای فعلی را که غالباً در دوره‌های کوتاهی از سال در بازار یافت میشد، زنبیل خرید خانوارها را تا حدودی دگرگون ساخت و در نتیجه در تغییر هزینه مصرفی خانوار اثر گذاشت . مجموعه این عوامل موجب شد تا تجدید نظر جامع دیگری در شاخص مزبور در سال ۱۳۵۳ بعمل آید . در این سال تعداد شهرها از ۳۵ شهر نمونه تجدیدنظر قبلی یکباره به بیش از دو برابر و به تعداد ۷۴ شهر افزایش پیدا کرد و تعداد کالاها و خدمات مورد قیمت گیری از ۱۶۹ کالا و خدمت به ۲۷۲ قلم کالا و ۲۲ خدمت و رویهم به ۳۱۵ قلم کالا و خدمت افزایش داده شد .

قرار بود تجدیدنظر دیگری در سال ۱۳۵۸ یعنی اولین سال بعد از انقلاب مردم ایران در شاخص بهای کالاها و خدمات مصرفی بعمل آید ولی بعلت اینکه ساختار الگوی مصرف خانوارها هنوز ماهیت خود را پیدا نکرده بود تصمیم گرفته شد تا در سال ۱۳۶۱ این تجدیدنظر بعمل آید .

۳-۱- پوشش شاخص در آخرین تجدیدنظر در سال ۱۳۶۱ : پوشش جغرافیائی شاخص خصوصیات جمعیتی که این شاخص برای آنان محاسبه میشود و نیز دزینده‌هائی که در شاخص مورد توجه قرار می‌گیرد، درجه اعتبار آنرا برای کاربردهای مختلف مشخص میسازد .

۱-۳-۱- پوشش جغرافیائی : درجه صحت و اعتبار شاخص در ارتباط مستقیم با میزان پوشش جغرافیائی آن است، زیرا تنوع شرایط اقلیمی بدون شک در خصوصیات اقتصادی و اجتماعی مثل عادات مصرف، سلیقه‌ها و گرایشهای افراد منعکس میشود . با عتبار این دیدگاه در تجدیدنظر سال ۱۳۶۱ تصمیم گرفته شد تا شاخص در سطح هریک از ۲۴ استان کشور بصورت مستقل محاسبه گردد و در نهایت از ترکیب شاخص استانهای مختلف شاخص کل در مناطق شهری کشور بدست آید . در تجدیدنظر اخیر تعداد شهرها از ۷۴ شهر تجدیدنظر سال ۱۳۵۳ به ۷۸ شهر افزایش پیدا کرد و تعداد ۶ شهر هم جابجا شد تا امکان تهیه شاخص در سطح هریک از استانهای کشور میسر گردد . تعریف مناطق شهری بکار برده شده در تجدیدنظر سال ۱۳۶۱ همان ضوابط بکار برده شده توسط وزارت کشور بودکه در تجدیدنظرهای قبلی استفاده میشد . طبق این تعریف کلیه نقاطی که در سال ۱۳۶۱ جمعیتی معادل ۵۰۰۰ نفر و بیشتر داشتند شهر تلقی گردید .

۱-۱-۳-۱- شهرهای مورد بررسی در تجدیدنظر سال ۱۳۶۱ : در تجدیدنظر سال ۱۳۶۱ ابتدا فکر تجدیدنظر در کلیه شهرهای نمونه قبلی مورد توجه قرار گرفت . زیرا با افزایش شهرهای

جدیدمیبایستی مراحل علمی نمونه گیری ازابتدا دنبال شود .اما چون تغییرشهرهای قبلی علاوه برمشکلات اداری ازقبیل نقل وانتقالات آمارگیران بومی ازشهرزادگاهشان به شهرهای دیگر، موجب ازبین رفتن تجربیاتی میشدکه طی سالیانکار درمورداوضاع واحوال اقتصادی واجتماعی شهرمحل خدمت خود کسب کرده بودند. همچنین ازآنجاکه شهرهای نمونه قدیم کلیه شهرهای مهم کشوررادربرداشت وحدود $\frac{2}{5}$ جمعیت شهری کشوردرآنهاساکن بود و در هر نمونه گیری جدید ناچار قسمت اعظم آنهامجددا " انتخاب میگردید، ازاینروتصمیم گرفته شد که ازتغییرشهرهای موجود موردبررسی خودداری شود وبرای گسترش دامنه پوشش جغرافیائی شاخصهاتنها ۴ شهرجدید به آنها اضافه شود و۶ شهرهم جابجا گردد.

۲-۱-۲-۱- مراحل انتخاب شهرهادرتجدیدنظرسال ۱۳۶۱ : کلیه شهرهائی که درسال ۱۳۶۱ بیشتر از یکصد هزار نفر جمعیت داشتند و نیز مراکز استانها با قاطعیت (شانس صد درصد) انتخاب شد. این تعداد برابر ۳۲ شهر گردید . دراین تجدید نظر تنها ۴ شهر به شهرهای نمونه قبلی اضافه شد و ۶ شهر هم جابجا گردید.

چون یکی ازمزایای سریهای زمانی مقایسه آمارهای گذشته وآینده میباشد، بنابراین درتجدید نظر سال ۱۳۶۱ ضمن آنکه برای اولین بار برای هر یك از ۲۴ استان کشور بصورت جداگانه شاخص تهیه میشود، برای هر یك از ۱۲ شهر بزرگ کشور که در تجدید نظر سال ۱۳۵۳ برای هر کدام شاخص جداگانه محاسبه میشد وهمچنین برای سایر شهرها که تعدادشان در شاخص قبلی ۶۲ شهر بود و در شاخص جدید ۶۶شهر است ، نیز جداگانه شاخص محاسبه میگردد.

۲-۳-۱- پوشش جمعیتی : شاخص بهای کالاهاوخدمات مصرفی نشان دهنده تغییرات سطح عمومی تعداد ثابت ومعینی کالا وخدمت است که توسط خانوارهای شهرنشین کشوربمصرف میرسد. چون با توجه به وسعت کشور وکثرت تعداد شهرها وخانوارهای ساکن در آنها، امکان دسترسی به کلیه خانوارها درکلیه مناطق شهری مقدور نیست ، لذا درسال ۱۳۶۱ تعداد ۱۵۰۰۰ خانوار نمونه در کل کشور توسط مرکز آمار ایران انتخاب گردید که روش انتخاب آنها سه مراحله ای و از طریق انتخاب حوزه ، بلوك و خانوار نمونه بوده است . خانوارهای نمونه در سطح ۲۴ استان کشور بطریقی توزیع گردید تا تعمیم نتایج در سطح استان معنی دار باشد. با بررسی خانوارهای نمونه در هر استان ، فراشب اهمیت کالاها وخدمات مشمول شاخص بهای

کالاها و خدمات مصرفی بتفکیک در هر استان و در نتیجه در کل کشور انتخاب گردید. چون تعداد خانوارهای نمونه در هر یک از ۱۲ شهر بزرگ کشور که در شاخص قبلی برای هر یک از آنها بصورت جداگانه شاخص تهیه میشد، کافی بود، بنابراین در سال ۱۳۶۱ نیز برای حفظ سریهای زمانی کافی السابق برای این تعداد از شهرهای بزرگ بطور جداگانه علاوه بر استانها، بتفکیک، شاخص تهیه میشود.

برای آنکه کلیه نوسانات فصلی مصرف در بررسی هزینه خانوارها نشان داده شود، تعداد کل خانوارهای نمونه در هر شهر و یا استان بین ۱۲ ماه سال بطور مساوی تقسیم گردید. علاوه بر توزیع زمانی، توزیع جغرافیائی نیز تا حد امکان در هر شهر مورد توجه قرار گرفت.

۲ - ۳ - ۱ - خانوارهای مشمول شاخص: فکر محدود کردن خانوارهای مشمول شاخص مبتنی بر دست یافتن به روند عمومی و واقعی قیمتها بوده، به حذف تغییرات استثنائی و اتفاقی استوار است. بعبارت دیگر همگن ساختن خانوارهای مشمول شاخص، روند تغییرات آن را برای اکثریتی که ترکیب مصرف یکسان و یا نزدیک به هم دارند، مشخص میکند. در تجدید نظر سال ۱۳۶۱، خانوارهای مشمول شاخص تقریبا" شامل کلیه خانوارهای ساکن در مناطق شهری ایران میشد. اما چون الگوی مصرف خانوارهائی که هزینه سالانه آنها بیشتر از ۶۰۰۰۰۰ ریال در سال بود، با الگوی مصرف جامعه تفاوت فاحشی داشت، این قبیل از خانوارها کنار گذاشته شد. این خانوارها از نظر تعداد تنها ۲ / ۰ درصد و از نظر هزینه کمی کمتر از ۲ / ۰ درصد کل هزینه ها را بخود اختصاص می دادند.

۲ - ۳ - ۱ - پوشش از نظر هزینه: شاخص بهای کالاها و خدمات مصرفی کلیه مخارج مصرفی خانوارهای مشمول آن را در بر میگیرد. در غالب موارد، تعیین مرزی میان هزینه های سرمایه ای مشخص است ولی در بارهای موارد قراردادی و اختیاری است. برای مثال در حالیکه کلیه اقلام پوشاک مشمول شاخص هستند اما لباس کار خریداری شده توسط کارفرمایان را باید خارج کرد یا در حالیکه هزینه های اتومبیل را باید منظور داشت ولی هزینه اتومبیلهای مورد استفاده خانوار در بازرگانی و تجارت باید خارج شود.

۲ - ۱ - کالاها و خدمات مشمول شاخص: مهمترین هدف در هر تجدید نظر در شاخصهای قیمت استفاده از ضرایب اهمیت به هنگام تر و تکمیل و تصحیح فهرست مشخصات مورد قیمت گیری است. در تجدید نظر سال ۱۳۶۱ برای محاسبه ضرایب و انتخاب کالاها و خدمات مورد قیمت گیری در سطح هر

یکدام از ۲۴ استان کشور وهمچنین ۱۲ شهربزرگ به تفکیک وشهرهای کوچک رویهم ازنتایج بررسی بودجه خانوار دااستفاده بعمل آمد. تجدیدنظراخیرنتایج تجربیات ارزنده عملی تجدیدنظرهای قبلی رابپیش روداشت که خود متضمن بهبود شیوه‌های عملی وافزایش دقت وصحت بیشتر در درانتخاب راه کارهابود. انتخاب کالاهاوتعیین شرائب آنهاضمن آنکه مبتنی برروش علمی بود، تجارب ودست آوردهای عملی‌صاحب نظران مسائل قیمت راسورد توجه قرارداد. لذابس از استخراج فهرست مقدماتی کالاهاوخدمات، نظرات وپیشنهادات کارشناسان امر، هیات‌نماینده آماری وآمارگیران اقتصادی موردتوجه قرارگرفت وباتوجه به کلیه امکانات ومحدودیتهای موجود، فهرست نهائی کالاهاوخدمات منتخب تهیه شد.

۱-۳-۱- تعدادکالاهاوخدمات منتخب: درتجدیدنظرسال ۱۳۶۱ درمجموع درسطح کشور تعداد ۳۰۲ قلم کالا و خدمت انتخاب شدتاقیمت آنهااز ۷۸شهرمورد بررسی بانک مرکزی جمهوری‌اسلامی‌ایران جمع آوری گردد. درانتخاب کالاهاوخدمات مشمول شاخص نکات ذیل موردتوجه قرارگرفت.

۱- مهمترین عاملی‌که درانتخاب هرکالاویاخدمت مشمول‌شاخص درتجدیدنظرسال ۱۳۶۱ موردتوجه قرارگرفت، اهمیت نسبی آن کالایاخدمت درمجموع هزینه خانوارهابود.

۲- عامل‌دیگر امکان جمع آوری قیمت کالاویاخدمت منتخب بود. گرچه تعدادی‌ازکالاها و خدمات دارای اهمیت نسبی‌قابل‌توجه‌ای درهزینه‌خانوارهابودند، اماچون امکان‌جمع آوری‌قیمت آنها بسیار مشکل‌بود، ازانتخاب آنهادرنظر گردید.

۳- علاوه بردوعامل فوق، انتخاب پاره‌ای ازکالاهاوخدمات مشمول‌شاخص درسال ۱۳۶۱ بر مبنای‌بیش بینی روند مصرف آنهادرآتیه بوده‌است. گرچه اهمیت نسبی کمی رادرمقایسه با دیگر کالاهاوخدمات منتخب بخوداختصاص داده بودند.

۴- حفظ قابلیت مقایسه شاخص‌جدیدباشاخص های‌گذشته عامل‌دیگری بود که درانتخاب کالاهاوخدمات مشمول شاخص درسال ۱۳۶۱ مورد توجه قرارگرفت تاکالاهاوخدمات منتخب جدید بااقلام شاخصهای قبلی‌تجانس بیشتری داشته باشند.

۵- عامل‌دیگری که درتجدیدنظر سال ۱۳۶۱ درانتخاب کالاهاوخدمات مشمول‌شاخص

مورد توجه قرارگرفت ، درمورد کالاهای فصلی بود که فقدان قیمت آنها مشکلات عدیده‌ای را در محاسبات شاخص بوجود می‌آورد . دراین تجدیدنظر سعی بعمل آمد تا کالاهائی انتخاب شود که مدت زمانی بیشتر درطول سال عرضه گردند .

۲-۳-۱- مشخصات کالاها وخدمات منتخب : جزئیات دقیق هریک از کالاها وخدمات منتخب مشمول شاخص درتجدیدنظر سال ۳۶۱ قلم کالا بطور دقیق مشخص گردید ودر پرسشنامه‌های مخصوص که برای قیمت گیری درنظر گرفته شد درج وبرای هر یک از ۲۸ شهر مورد بررسی ارسال گردید .

۲-۳-۱- طبقه بندی کالاها وخدمات مشمول شاخص : طبقه بندی کالاها وخدمات مشمول شاخص بر اساس پیشنهاد سازمان ملل طبقه بندی بین المللی محاسبات ملی (S.N.A*) میباشد که بنا بمقتضیات اقتصاد ایران تغییرات جزئی بشرح زیر در آن ضروری تشخیص داده شد .

۱- خوراک خارج ازخانه که درطبقه بندی سازمان ملل درگروه سایر کالاها وخدمات قرار گرفته است ، در طبقه بندی شاخص جدید درگروه خوراکیها ودخانیات ، طبقه بندی شده است .

۲- لوازم التحریر از گروه سایر کالاها وخدمات درطبقه بندی سازمان ملل به گروه تحصیل وسائله منتقل شد .

درذیل اسامی ۸ گروه اصلی و شماره‌ی گروه فرعی شاخص طبق این طبقه‌بندی همراه با اهمیت نسبی هر گروه نسبت به شاخص کل داده شده است .

گروه	اهمیت نسبی
شاخص کل	۱۰۰/۰۰
۱- خوراکیها ، آشامیدنیها ودخانیات	۴۱/۵۱
خوراک	۳۲/۰۶
غذا درخانه	۲۵/۵۷
نان ، برنج ، آرد وفرآورده‌های آن	۵/۵۱
حبوبه	۰/۲۲

* United Nation A System of National Accounts, Serial, No, 2, Rev. 3, Ny, 1969

لبنیات وتخم مرغ	۴/۸۰
گوشت قرمز، مرغ وماهی	۱۲/۱۶
میوه ها وسبزیها	۹/۹۱
روغن حیوانی ونباتی	۰/۵۱
قند وشکر وفرآورده‌های آن	۰/۸۷
چای	۰/۸۰
سایر کالاهای خوراکی	۰/۴۹
غذا درخارج ازخانه	۱/۴۹
آشامیدنیها	۰/۳۵
دخانیات	۲/۱۰
۲-پوشاک	۸/۰۵
پوشاک آماده	۳/۷۳
کالاهاوخدمات مربوط به تهیه پوشاک	۲/۴۰
انواع کفش وتعمیر آن	۱/۹۲
۳- مسکن وسوخت وروشنائی	۲۶/۵۱
مسکن	۲۴/۴۵
آب ، برق وسوخت	۲/۰۶
۴ اثاثه وکالاهاوخدمات مورداستفاده درخانه	۷/۳۹
اثاثه خانه	۶/۴۰
کالاهاوخدمات مورداستفاده درخانه	۰/۹۹
۵- درمان وبهداشت	۳/۳۵
دارو	۱/۲۴
خدمات درمانی	۲/۷۱

۰۶ حمل ونقل و ارتباطات	۸/۲۶
حمل ونقل	۷/۶۷
ارتباطات	۰/۵۹
۰۷ تفریح، تحصیل ومطالعه	۱/۵۹
تفریحات	۰/۷۸
تحصیل ومطالعه	۰/۸۱
۰۸ کالاها وخدمات متفرقه	۴/۲۴

۵-۱ جمع آوری اطلاعات

۱-۵-۱-منابع اطلاع : پوشش مطلوب نوسانات قیمت یک کالا درسطح شهر ویا شهرهای نمونه مستلزم جمع آوری تعداد مظنه کافی از نقاط مختلف آن شهر است ۰ زیرا غالبا" یک کالا در نقاط مختلف شهر با قیمتهای متفاوتی بفروش میرسد ونوسان قیمت آن لزوما" یکی نیست باینجهت تعداد منابع اطلاعی که قیمتها از آن جمع آوری وچگونگی پراکندگی آنها درسطح شهر از جمله عوامل موثری است که می‌تواند در نشان دادن تغییرات صحیح سطح عمومی قیمتها بسیار موثر واقع شود۰ با توجه به این امر در تجدید نظر اخیر کوشش بعمل آمد تا نتایج حاصل از تجارب گذشته در تعیین و انتخاب منابع اطلاع نمونه بشرح ذیل بکار گرفته شود۰

۱-۱-۵-۱-تعداد مظنه های مورد لزوم در هر شهر : فقدان چار چوب کامل و بهنگام از تعداد منابع اطلاع یعنی فروشگاهها وموسساتی که کالا یا خدمت خاصی را برای فروش عرضه میکنند و عدم دسترسی به اطلاعات مربوط به حجم فروش هر کالا در هر فروشگاه تعیین تعداد منابع اطلاع نمونه لازم رابسیار مشکل ساخته است ۰ لذا بر اساس تجارب قبلی و با توجه به امکانات پرسنلی ، تعداد مظنه لازم در شهرهای مختلف برای کالاهای خوراکی و غیرخوراکی بشرح ذیل تعیین گردید۰

تعداد منابع اطلاع برای جمع‌آوری قیمت	تهران	شهرهای بزرگ	۵۱ شهر کوچک
مواد خوراکی	۲۰	۱۰	۵
مواد غیرخوراکی وخدمات	۱۰	۵	۳

بااینترتیب مجموع منابع اطلاعی که در ۷۸شهر نمونه مورد مراجعه آمارگیران اقتصادی بانک مرکزی جمهوری اسلامی ایران قرار میگیرد بیش از ۲۱ هزار مغازه خرده فروشی میباشد . البته تعداد اقلام مورد قیمت گیری در منابع اطلاع مختلف یکسان و برابر نیست ، زیرا در حالیکه مثلا "از مطب یک پزشک فقط اجرت یک بار و قلم خدمت پرسش و ثبت میگردد ، تعداد قیمتهائی که از یک سوپرمارکت و یا یک فروشگاه بزرگ جمع آوری می شود در برخی موارد ممکن است به بیش از ۲۰۰ منطقه بالغ گردد .

۲-۵-۱- روش انتخاب منابع اطلاع : انتخاب منابع جدید توسط آمارگیران اقتصادی خبره و قدیمی صورت میگیرد . باین ترتیب که در اولین تماس ، هدفها ی بانک مرکزی جمهوری اسلامی ایران از جمع آوری قیمتها برای صاحب منبع تشریح میگردد و ضمن اعتماد ، همکاری وی جلب میشود . این انتخاب براساس ضوابط زیر صورت میگیرد .

الف - پراکندگی در سطح شهر : توزیع جغرافیائی منابع اطلاع باید بترتیبی باشد که منابع اطلاع در کلیه مناطق شهر پراکنده باشند تا کلیه نوسانات قیمت ناشی از تفاوت محلهای فروش در حوزه عمل شاخص قرار گیرد .

ب - موقعیت محلی منبع اطلاع : حتی الامکان سعی میشود منابع اطلاع در مراکز خرید، بازارهای بزرگ و همسایگی خانه ها واقع باشند .

ج - حجم فروش منابع اطلاع از عوامل موثر در انتخاب آنها میباشد اما بسبب فقدان اطلاعات در مورد حجم فروش هر فروشگاه در هر محل، سعی میشود تا از فروشگاههای موجود ، فروشگاهی برای قیمت گیری کالاهای مورد نظر انتخاب شود که مشتریان زیادتر و تحرک و فروش بیشتری داشته باشند .

د - میزان همکاری و صمیمیت و صداقت صاحبان منابع اطلاع در پاسخ به سوالات آمارگیران از عوامل موثر در انتخاب آنها میباشد .

۲-۵-۲- نحوه جمع آوری اطلاعات : جمع آوری اطلاعات (قیمت کالاها و خدمات) در کلیه شهرها از طریق مراجعه مستقیم آمارگیران بانک مرکزی جمهوری اسلامی ایران و مصاحبه با صاحبان منابع اطلاع با استفاده از پرسشهای دقیق صورت میگیرد . بدیهی است که قیمت ماه جاری نسبت به ماه قبل افزایش یا کاهش داشته باشد، لذا این تذکر به دقت پرسش و ثبت

می‌گردد. در جمع آوری اطلاعات از آمارگیران اقتصادی مواکدا" خواسته شده است تا تنها قیمتهای عملی فروش حاکم بربازار هر شهر جمع آوری و گزارش نمایند.

در سالهای اخیر که تولید و واردات بیشتر کالاها در مقایسه با سالهای قبل شدیدا" کاهش‌یافته و تعدادی از کالاها از طریق کوپن بین خانوار ها توزیع میشود و بیشتر کالاها بویژه کالاهای اساسی با چندین نرخ خرید و فروش میشود، در نحوه قیمت گیری از کالاها در شاخص بهای کالاها و خدمات مصرفی نیز تغییراتی در مقایسه با قبل بوجود آمده است. تا قبل از کوپنی شدن تعدادی از کالاها مظنه‌های مورد نظر در هر شهر مظنه‌هائی بود که اکثریت قریب باتفاق خانوارها کالاهای مورد مصرف خود را با آن نرخها خریداری می‌کردند و معمولا" این مظنه‌ها همان نرخهای واقعی و مورد مصرف غالب خانوارها بود که اکثرا" هم با نرخهای تعیین شده توسط مقامات مملکتی وشهری در هر شهر چندان تفاوتی نداشت و اگر هم تفاوتهائی دیده میشد از چند درصد تجاوز نمی‌کرد. با کوپنی شدن تعدادی از کالاها تفاوت بین قیمتهای رسمی کالاها که توسط مقامات مملکتی تعیین میگشت و نرخهای واقعی کالا در بازار آزاد در غالب شهرها شدیدا" افزایش پیدا کرد. در بادی امر قیمت این قبیل از کالاها که از طریق کوپن در اختیار خانوارها در هر شهر قرار می‌گرفت، همان قیمتهای رسمی و اعلان شده بود. اما از نیمه دوم سال ۱۳۶۶ که بمرور فاصله بین کوپنهای اعلان شده در غالب شهرها زیاد شد و تقریبا" کلیه شهرها قیمتهای بازار آزاد چنین کالاهارا بمرکز گزارش کرده جهت دقیق ترشدن نتایج تصمیم گرفته شد تا قیمتهای بازار آزاد کالاهای کوپنی در محاسبات شاخص منعکس گردد. ترکیب قیمتهای بازار آزاد و قیمتهای رسمی کالاهای کوپنی باتوجه به بررسی بودجه خانوار که هزینه‌های پرداختی خانوار را برای هر کالا و یا خدمت چه در بازار آزاد و چه رسمی بدست می‌دهد، تعیین می‌گردد.

بطور کلی در حال حاضر قیمت کالاها و خدمات مشمول شاخص بهای کالاها و خدمات مصرفی به سه طریق جمع آوری میشود.

۱- بر اساس قیمتهای رسمی که تعداد آنها بسیار کم است. این قیمتها در سراسر کشور یکسان و همان نرخهای دولتی است.

۲- ترکیبی از قیمتهای رسمی و بازار آزاد که تقریبا" کمی کمتر از نیمی از کالاها و

خدمات را شامل میشود .

۳- از طریق جمع آوری قیمتهای بازار آزاد که شامل بیش از نیمی دیگر از کالاها وخدمات میشود .

مظنه‌های مورد نظر در هرشهر در مورد نیمی از کالاها ازقبیل انواع گوشت ، سیگار ، برنج وارداتی ، بعضی از مواد لبنی ، تعدادی اندك کالاوخدمات موجود در گروه پوشاك و اثاثه خانه وتعدادی دیگر از کالاها که هم با نرخهای رسمی واعلان شده توسط دولت وهم با نرخهای بازار آزاد بدست مصرف کنندگان میرسد ، ترکیبی از نرخهای رسمی وبازار آزاد است . انواع میوه وسبزی ، تقریبا " کلیه خدمات ، بیشتر اقلام موجود در گروه پوشاك ولوازم خانه وبسیاری ازکالاها وخدمات که برای آنها نرخ رسمی وجود ندارد ویا اگر وجود دارد کمتر اعمال میشود ، از طریق بازار آزاد جمع آوری میشود .

دربعضی موارد برای اندازه گیری تغییرات واقعی قیمت کالا اقدام به خرید کالا میشود . مثلا " انواع نان دربسیاری از شهرها بصورت نانه‌ای وبدون عمل توزین بفروش میرسد . دراین قبیل موارد گرچه قیمت یکعدد نان درماههای متوالی ثابت است ولی به سبب تغییرات احتمالی در وزن یکعدد نان ، قیمت عملی فروش مرتبا " تغییرمی کند . بنابراین قیمت عملی فروش نان ازطریق خرید عملی وتوزین نانهای خریداری شده بدست می آید .

۲ - ۵ - ۱ - زمان قیمت گیری : دوره قیمتگیری درهرماه ازبیست وپنجم ماه قبل شروع میشود وتابیستم آن ماه ادامه پیدا میکند . آمارگیران اقتصادی موظفند درهرماه دردوره هسته‌ای معینی به هریك ازمنابع اطلاع درطول سال مراجعه وقیمت عملی فروش کالاوخدمات مشمول شاخص راجمع آوری ودرپرسشنامه های مخصوص ثبت نمایند . توزیع زمانی منابع اطلاع درطول ماه طوری است که جمع آوری قیمت یك کالا درطول مدت قیمتگیری انجام میشود . مثلا قیمت میوه از کلیه میوه فروشیها دریك روز جمع آوری نمیشود ، بلکه تعداد منابع اطلاع مورد نظر برای میوه در طول بیست وپنج روزماه توزیع میشود وازآنها نمونه قیمت بعمل می آید .

۲ - ۵ - ۲ - فرمول شاخص وروش محاسبه آن : شاخص بهای کالاها وخدمات مصرفی که بوسیله بانك مرکزی جمهوری اسلامی ایران محاسبه میگردد ، شاخص موزونی از نوع تغییرقیمتهاست تعداد ثابت ومعینی ازکالاها وخدمات مورد مصرف خانوارهای شهرنشین کشوراست . شرایط

اشیائی که برای موزون کردن این شاخص بکار میرود ، عبارت ازمتوسط هزینه های سالانه خانوارشهری مشمول شاخص برای کالاها وخدمات مصرفی میباشد .

درمحاسبه شاخص ازشاخص های موزون قیمتهای نسبی کالاها وخدمات بافروش کالاها و خدمات درسال پایه به شکل :

$$I_{m/o} = \frac{\sum_{i=1}^{k} p_o^{(i)} q_o^{(i)} \left[p_m^{(i)} / p_o^{(i)} \right]}{\sum_{i=1}^{k} p_o^{(i)} q_o^{(i)}} (100) = \frac{\sum_{i=1}^{k} p_m^{(i)} q_o^{(i)}}{\sum_{i=1}^{k} p_o^{(i)} q_o^{(i)}} (100)$$

استفاده میشودکه درحقیقت فرمول تغییر شکل یافته لاسپیرز اقتصاددان المانی میباشد . در فرمول بالا عبارت $\left[p_o^{(i)} / p_{(m-1)}^{(i)} \right] p_o^{(i)} q_o^{(i)}$ برای ماه $(m-1)$ ام محاسبه شده و بنا محاسبه کردن $p_m^{(i)} / p_{m-1}^{(i)}$ بآسانی میتوان شاخص ماه m ام را محاسبه کرد . بدیهی است این روش برای کلیه ماههای قبل از ماه m ام نیز دنبال شده است زیرا :

$$\left\{ \left[\frac{p_{m-1}^{(i)}}{p_{m-2}^{(i)}} \right] \left[\frac{p_{m-2}^{(i)}}{p_{m-3}^{(i)}} \right] \cdots \left[\frac{p_2^{(i)}}{p_1^{(i)}} \right] \left[\frac{p_1^{(i)}}{p_0^{(i)}} \right] \right\} p_o^{(i)} q_o^{(i)} \frac{p_{m-1}^{(i)}}{p_o^{(i)}} = p_o^{(i)} q_o^{(i)}$$

میباشد .

$\frac{p_m^{(i)}}{p_{m-1}^{(i)}}$ درواقع قیمت نسبی کالاها وخدمات i در ماه m ام نسبت بماه $(m-1)$ ام میباشد .

درفرمول بالا $p_o^{(i)} q_o^{(i)}$ برابر است با :

$$p_o^{(i)} q_o^{(i)} = \sum_{j=1}^{m} \lambda_o^{(i,j)} \lambda_o^{(i,j)}$$

دراین رابطه m تعداد شهرهای مورد قیمت گیری ، j بترتیب نماینده قیمت کالا و یا خدمت i ام در شهر j ام ، $\lambda_o^{(i,j)}$ و $\lambda_o^{(i,j)}$ بترتیب متوسط قیمت ومقدارکالا ویاخدمت i در شهر j در سال پایه میباشد . بنابراین برای $\frac{p_m^{(i)}}{p_{m-1}^{(i)}}$ میتوان نوشت :

$$\frac{P_{n}^{(t)}}{P_{n-1}^{(t)}} \cdot \frac{\left[\sum_{j=1}^{m} q_{0}^{(t,j)} p_{0}^{(t,j)}\right] \Big/ \left[\sum_{j=1}^{m} q_{0}^{(t,j)} p_{n}^{(t,j)}\right]}{\left[\sum_{j=1}^{m} q_{0}^{(t,j)} p_{0}^{(t,j)}\right] \Big/ \left[\sum_{j=1}^{m} q_{0}^{(t,j)} p_{n-1}^{(t,j)}\right]}$$

برای تعیین $P_n^{(t,j)}$ در شهر j ام معمولا " از قیمتهای جمع آوری شده از منابع اطلاع آن شهر میانگین حسابی ساده استخراج میشود . برای تعدادی از کالاها که هم با نرخهای رسمی و هم با نرخهای بازار آزاد در شهرها در دسترس خانوارها قرار میگیرد ، میانگین موزون قیمت بر اساس مقدار فروش محاسبه میشود .

کاربرد فرمول تغییر شکل یافته لاسپیرز که در بالا بان اشاره شد در آنست که در مواردی که کالا و یا خدمتی با مشخصاتی سپس بدلایلی از بازار خارج میشود و کالای دیگر جانشین آن میگردد، به سادگی میتوان تغییرات قیمت کالا و یا خدمت جانشین شده را بجای قیمت کالا و یا خدمت اصلی در شاخص بکار برد . این انعطاف پذیری فرمول درقبال جانشینیهای انواع کالاها با توجه به تغییرات عدیده ای که در سالهای اخیر بدلایل مختلف در مشخصات کالاها و خدمات مشمول شاخص رخ میدهد، کاملا " ضروری و مفید است .

۱-۷ - مشکلات موجود در محاسبه شاخص : یکی از مسائل مهمیکه همواره برای محاسبان شاخص مطرح است ، تغییراتی است که بمرور زمان در مشخصات کالاها ، منابع اطلاع قیمتگیری، واحدهای وزن ومقادیری که برای کالاها انتخاب شده است وحتی نوع کالاها و خدمات، رخ میدهد . با وجودیکه در محاسبه شاخص براساس فرمول لاسپیرز ضریب اهمیت کالاها وخدمات منتخب تا زمان تجدید نظر بعدی ثابت گرفته میشود، معهذا عملا " در عوامل کمی وکیفی مشکله کالاهاوخدمات تغییراتی پدید میآید که عدم توجه به آنها شاخص را از نشان دادن روند حقیقی قیمتها باز میدارد . مثلا " کالائی با مشخصات خاص از بازار خارج میشود و کالای دیگر جانشین آن میگردد ولی در این جانشینی در قیمت کالا تغییری در جهت ترقی یا تنزل رخ نمیدهد . مشکلی که در اینجا پیش میآید قضاوت در باره این است که این جانشینی را چگونه باید در شاخص منحوظ داشت و تغییر قیمت مربوطه را چگونه تعبیر نمود .

برای احتراز از این مشکل محاطور که در بالا اشاره شد در انتخاب کالاها وخدمات

برای تهیه شاخص بهای کالاها وخدمات مصرفی سعی میشود تا کالاهاو خدماتی انتخاب شوند که احتمال خروج فوری آنها از بازار کمتراست . بعلاوه برای یک کالا انواع مختلفی در نظر گرفته می‌شود وقیمت گیری ها بعمل می‌آید تا فقدان یکی از انواع نتواند به محاسبه شاخص فردی آن کالا لطمه ای وارد سازد . درمواردیکه علیرغم این پیش بینی ها اگریک‌کالا ویا یک خدمت مشمول شاخص در شهری برای فروش عرضه‌نشودبطریق ذیل قیمت آن درمحاسبات شاخص برآورد و شاخص فردی آن محاسبه می‌گردد .

۱-۷-۱- نحوه محاسبه شاخص قیمت کالاهای فصلی : با اینکه در انتخاب کالاها وخدمات مشمول شاخص بهای کالاها وخدمات مصرفی سعی‌میشود کالاها وخدماتی انتخاب گردد که مدت زمان بیشتری درطول سال در بازار شهرها عرضه شود، معهذا برای تعدادی از کالاها با دارا بودن خرایب اهمیت قابل توجه در بعضی از ماههای سال بعلت فصل امکان جمع آوری قیمت وجود ندارد و عملا " فقدان قیمت آن ها مشکلاتی را در محاسبه تغییرات آنها پدید می‌آورد . برای این قبیل از کالاهای فصلی که عموما " در گروه میوه ها وسبزیها قرار دارند ، راه عملی کار آنست که اگر کالائی بعلت فصل در یک شهر عرضه‌نشود تغییرات‌قیمت سایر کالاهای موجود در گروه مربوطه به کالای مورد نظر اختصاص می‌یابد . اینکار برای هر چند ماه که کالا در بازار آن شهر عرضه‌نشودادامه می‌یابد . در صورتیکه در گروه میوه وسبزی یکی از گروههای فرعی کلیه کالاهای موجود در آن نیز بعلت فصل فاقد قیمت باشد بهمین ترتیب عمل میشود و تغییرات قیمت گروه اصلی بالاتر از آن گروه فرعی برای تمام کالاهای موجود در آن گروه بکار گرفته میشود . در مورد سایر کالاها وخدمات که بهر دلیل امکان جمع آوری قیمت برای کالا یا خدمتی در یکی از شهرها وجود نداشته باشد، روش دیگری اعمال میشود که در ذیل بآن اشاره می‌گردد .

۲-۷-۱- جانشینی کالا ها وخدمات و برآورد تغییر قیمت کالا و یا خدمت فاقد قیمت : هر گاه یکی از منابع اطلاع ، کالا و یا خدمت مورد نظر را غیر از کالاهای فصلی که در بالا بآن اشاره شد در یک ماه در یکی از ۲۸ شهر مورد بررسی بانک برای فروش عرضه نکند، آمارگیر اقتصادی بانک مستقر در آن شهر در مرتبه اول موظف است تاحتی‌الامکان کالا و یا خدمت دیگری را که از هر جهت شبیه کالا و یا خدمت اصلی‌است ، قیمت گیری نماید . اگر این عمل

قیمت گیری با تغییر قیمت همراه بود،آمارگیر باید ضمن بررسی دقیق روشن کند چه مقدار از این تغییر قیمت ناشی از تغییر کیفیت کالا و یا خدمت جایگزین شده است و چه مقدار مربوط به افزایش یا کاهش قیمت است . اگر منبع اطلاع اساسا " هیچ یک از انواع کالا ی مورد قیمت گیری را در آن شهر عرضه نکند، باید از منبع اطلاع دیگری که کالا را برای فروش عرضه می‌کند قیمت کالا را جمع آوری نماید .

بطور کلی اگر در شهری یکی از منابع اطلاع کالا و یا خدمت مورد قیمت گیری در یکی از ماهها حذف شود، تغییر قیمت کالا یا خدمت با توجه به تغییر قیمت در سایر منابع اطلاع مورد محاسبه قرار می‌گیرد . چون در همین ماه بجای منبع اطلاع حذف شده منبع اطلاع جدیدی انتخاب میشود، در ماه بعد محاسبه نرسان قیمت کالا و یا خدمت مورد نظر با توجه به قیمت کالا یا خدمت در منبع جدید که قیمت ماه قبل آن نیز در دست است، صورت می‌پذیرد . فقدان کالا یا خدمت در یک منبع اطلاع بصورت موقت حداکثر تا دوالی سه ماه جایز است . در دو سه ماهی که کالا یا خدمت در منبع اطلاع اصلی وجود ندارد، قیمت فرضی با توجه به تغییر قیمت در سایر منابع اطلاع محاسبه میشود واین قیمت در اولین ماه عرضه کالا یا خدمت بجا قیمت واقعی که از این منبع جمع آوری میشود مورد مقایسه و سنجش قرار می‌گیرد .

اگر هیچ یک ازمنابع اطلاع مورد قیمت گیری در یک ماه یا ماههای متوالی کالا یا خدمت مورد نظر را برای فروش در یکی از شهرها عرضه نکنند، تغییر قیمت آن کالا و یا خدمت فاقد قیمت بطریق ذیل برآورد میشود .

اگر شهری که کالا و یا خدمت مورد نظر را برای فروش عرضه نکرده است ، یکسی از ۱۲ شهر بزرگ کشور باشد، تغییر قیمت کالا و یا خدمت در ۱۱ شهر بزرگ دیگر رویهم برای آن شهر فاقد قیمت در نظر گرفته میشود .

اگر شهری که کالا یا خدمت مورد نظر را برای فروش عرضه نکرده است ، یکی از شهرهای خود منتخب باشد، تغییر قیمت کالا و یا خدمت در دیگر شهرهای خود منتخب رویهم برای آن شهر فاقد قیمت در نظر گرفته میشود .

وبالاخره اگر شهری که کالا و یا خدمت مورد نظر را برای فروش عرضه نکرده است یکی از شهرهای کوچك باشد، تغییر قیمت کالا و یا خدمت در دیگر شهرهای کوچك روبهم برای آن شهر فاقد قیمت در نظر گرفته میشود.

بدیهی است این عمل برای هرچند ماه که کالا و یا خدمت مورد نظر در یك شهر و یا چند شهر عرضه نشود، تکرار میشود.

اگر کالا و یا خدمت مورد نظر در هیچیك از ۷۸ شهر مورد بررسی بانك مرکزی جمهوری اسلامی ایران برای فروش عرضه نشود، تغییرات قیمت سایر کالاها و خدمات در کل کشور برای آن کالا و یا خدمت که فاقد قیمت است در نظر گرفته میشود.

۲-۲-۱- محاسبه تغییرات گروه مسکن: شاخص گروه مسکن نمایانگر نوسانات هزینه اجاره نشینان (کرایه خانه) میباشد. محاسبه نوسانات کرایه خانه مبتنی بر بررسی وسیع و جامعی است که هرسه ماه یکبار در سطح کلیه ۷۸ شهر مورد بررسی بانك انجام میشود و رویهم حدود ۸۰۰۰ واحد مسکونی اجاری مورد پرسش قرار میگیرند. واحدهای نمونه اجاری از فهرست صورتبرداری منطقه ای شهرها استخراج گردیده و در انتخاب آنها پراکندگی جغرافیائی درسطح هر شهر کاملا" رعایت شده است.

یکی از عوامل مهم که درشبات نسبی شاخص کرایه خانه موثر است، عدم تغییرمبلغ اجاره درطول دوره ای است که مستاجر در واحد مسکنی گزیده است وتغییراجاره در اکثریت قریب باتفاق موارد مربوط بموقعی است که مستاجرنقل مکان نموده وبه محل جدید میرود والا معمولا" تازمانیکه در محل قبلی خود سکنی دارد، میزان اجاره کم وبیش ثابت باقی میماند ازطرف دیگربنابه نتایج حاصل ازبررسی مستمرمورد فیقی که هرسه ماه یکبارانجام میشود در آن ۸۰۰۰ خانوارمستاجره رمناطق شهری کشورمورد پرسش قرارمیگیرند، نسبت خانوارهائی کسم درطول سال نقل مکان نموده وازمحل قبلی خود به واحد اجاری جدید میروند تقریبا کمتراز ۱۰ درصد کل خانوارهای اجاره نشین مشمول شاخص راتشکیل میدهند. وازآنجائیکه نوسانات اجاره بهاتقریبا" به این گروه منحصرمیشود، عدم نوسان شدید شاخص اجاره بهاخلاف آنچه که مورد انتظار عموم است کاملا" قابل توجیه بنظرمیرسد.

در تجدید نظر سال ۱۳۶۱ بجای خرید خانه که یک کالای سرمایه ای است ، مبلغ ارزش اجاری خانه های شخصی در نظر گرفته شده است . چه بسا اختلافاتی که برای خرید خانه انجام میگیرد عمدتا " به عنوان سرمایه گذاری تلقی میشوند تا بر اختلاف برای کالای مصرفی ولذا بطور مستقیم مورد نظر نیستند و برای خانوارهای شخصی نشین از مبلغ اجاری استفاده گردیده است . وزن این قلم کالا از ارزش اجاری منزل مسکونی کلیه خانوارهای شخصی نشین در سال ۱۳۶۱ به کل هزینه ها بدست آمده است . بعلت مشکلاتی که در جمع آوری اطلاعات مربوط به نوسانات این ارزش در عمل بوجود آمد ، این هزینه ها در حال حاضر از کلیه شهرها کنار گذاشته شده است و در هریک از گروههای شاخص ، محاسبات شاخص کل بدون آن دنبال میشود .

۸-۱- وصل کردن شاخص ها : برای وصل کردن سری شاخصهای قدیم و شاخصهای جدید معمولا " از فاکتور مناسبی بنام فاکتور تبدیل استفاده میکنند . این فاکتور از تقسیم میانگین اعداد سالانه شاخص جدید و قدیم در سال پایه جدید برای هر سال حاصل میشود . اگر P_0 و P_1 بترتیب سال پایه شاخص قدیم و شاخص جدید فرض شود ، با بکار بردن فرمول لاسپیرز ، شاخص جدید و شاخص قدیم در زمان t (سال پایه جدید) و همچنین فاکتور تبدیل بصورت ذیل محاسبه میشود :

شاخص پر پایه قدیم در سال پایه جدید $$L = \frac{\sum P_t Q_t}{\sum P_0 Q_t} (100)$$

شاخص بر پایه قدیم در سال پایه جدید $$L' = \frac{\sum P_t Q_0}{\sum P_0 Q_0} (100)$$

فاکتور تبدیل $$\frac{L}{L'}$$

با ضرب کردن اعداد شاخص قدیم در این فاکتور ، بآسانی می توان شاخص های قدیم را بر پایه جدید محاسبه و مورد استفاده قرار داد . این عمل برای هریک از اقلام یا گروهها باید انجام داد و چنین ضریب تبدیلی را محاسبه نمود .

۲ـ شاخص بهای عمده‌فروشی کالاها در ایران

۱ـ۲ـ هدف از تهیه شاخص : شاخص بهای عمده‌فروشی کالاها برای اندازه‌گیری تغییرات سطح عمومی قیمتها محاسبه میشود و نمودار آن از تغییر قیمتها در سطح تولیدکننده و عرضه‌کننده دست اول میباشد . این شاخص نیز برای تضمین تغییرات قدرت خرید پول کشور بکار میرود .

میدان عمل قیمتهای عمده‌فروشی شامل کلیه قیمتهائی است که در معاملات تجاری مهم بر روی کالاها اعمال میشود . بنابراین در محاسبه شاخص بهای عمده‌فروشی قیمتهای جمع آوری شده عموماً قیمتهای سرمعدن (در مورد کالاهای معدنی) یا قیمت فروش محصولات در کارخانه‌ها و یا کارگاههای تولیدی (در مورد محصولات صنعتی) یا قیمت فروش عاملین دست اول (در مورد کالاهای وارداتی و یا قیمت تحویل کالا به بنادر مرزی کشور (در مورد کالاهای صادراتی) میباشد . در مورد کالاهای کشاورزی چون امکان جمع آوری قیمت در سر مزرعه وجود ندارد ، لذا برای این نوع از کالاها قیمت عمده‌فروشی از نزدیکترین بازار عمده به مزرعه جمع آوری میشود . در مورد تعدادی از کالاهای صادراتی که امکان جمع آوری قیمت از بنادر مرزی کشور وجود ندارد ، قیمتهای پرداختی تا صادرکننده در هنگام خرید کالا برای صدور کالا جمع آوری میشود .

۲ـ۲ـ تاریخچه : شاخص بهای عمده‌فروشی کالاها در ایران برای اولین بار همزمان با شاخص بهای کالاها و خدمات مصرفی در سال ۱۳۱۵ بوسیله بانک ملی ایران تهیه شد . شاخص اولیه نوسانات قیمت ۶۶ قلم کالای منتخب را که در اقتصاد کشور اهمیت داشتند مورد محاسبه قرار میداد و قیمتهای نسبی آنها را از طریق متوسط هندسی ساده برآورد می‌کرد . در سال ۱۳۱۵ بعلت در دست نبودن ارزش فروش عمده کالاها در سطح کشور به کلیه کالاهائی شامل شاخص که تنها از بازار تهران قیمت گیری میشدند ، ضریب اهمیت یکسانی داده شده بود . شاخص مزبور تا خرداد سال ۱۳۳۸ بدون هیچگونه تجدیدنظری محاسبه و منتشر گردید .

پس از گذشت حدود ۲۵ سال با توجه به اینکه شاخص سال ۱۳۱۵ دیگر نمی‌توانست نوسانات واقعی قیمتهای عمده فروشی کالاها را در سطح کشور منعکس نماید ، تهیه شاخصی موزون یعنی شاخصی که ضرایب اهمیت کالاها در آن در نظر گرفته شده باشد ، مورد توجه قرار گرفت .

در تجدیدنظری که در سال ۱۳۳۸ در شاخص بهای عمده‌فروشی کالاها در ایران همزمان با شاخص بهای کالاها و خدمات مصرفی صورت گرفت ، اقلام مختلف شاخص مورد تجدیدنظر اساسی قرار گرفت .

دراین تجدیدنظربه هریك ازکالاهای مشمول شاخص که تعدادشان بالغ بر۱۶۰کالا بودضریب اهمیتی باتوجه به‌حجم فروششان دربازارهای عمده کشوراختصاص داده شدوقیمتها علاوه‌برتهران از ۱۰ شهر بزرگ دیگرکشورکه دارای بازارعمده‌فروشی بودندجمع آوری‌گردید . فرمول‌محاسبه شاخص نیزتغییر داده شده‌و میانگین موزون قیمت کالاهاطبق فرمول لاس پیرزمحاسبه گردید . شاخص بهای عمده‌فروشی‌کالاهاسه باردیگرهمزمان باشاخص بهای‌کالاهاوخدمات مصرفی بترتیب درسالهای ۱۳۴۸، ۱۳۵۳ و ۱۳۶۱ مورد تجدیدنظرکلی قرارگرفت . درآخرین تجدیدنظرکه سال پایه‌این شاخص سال ۱۳۶۱میباشدتعدادکالاهای‌مشمول شاخص از ۶۶ قلم کالا درسال ۱۳۱۵به ۲۹۴ قلم کالاوبراکسنز قیمت گیری به بیش از ۳۰شهرافزایش پیداکرد .

۲-۲- طبقه‌بندی‌کالاهای‌مشمول‌شاخص : معمولا " کالاهابراساس هدف ونوع استفاده‌ای‌که ازشاخص بعمل‌می‌آیدطبقه بندی‌میشوند . کشورهای‌مختلف باتوجه به‌هدفهای‌خود، کالاهارابطرق مختلفی‌طبقه‌بندی‌می‌کنند . دربعضی‌ازکشورهاطبقه‌بندی‌کالاهابراساس مرحله ساخت (خام، نیمه ساخته‌وساخته‌شده) ودربعضی‌براساس رشته‌های‌اقتصادی‌انجام‌میشود . اگربدست آوردن نوسانات قیمت درهریك ازرشته‌های‌اقتصادی‌موردنظرباشد، طبقه‌بندی‌کالاهابراساس رشته های‌اقتصادی‌ترجیح دارد . ولی‌اگرهدف صرفا " اندازه گیری‌تغییرات سطح‌عمومی‌قیمتها باشد، مطلوب آنست که طبقه بندی‌طوری‌انجام گیردتاکالاهائی‌که دارای‌نوسانات قیمت ومسارف مشابه‌ای دارند، دریك گروه قرارگیرند . برای‌این گروه بندی‌میتوان ضوابطی‌نظیرموادخام‌یکه درتولیدکالاهابکارگرفته میشودیاموارداستفاده آنهارابعنوان‌معیارتشابه بکاربرد .

مناسب ترین طبقه بندی‌که برای‌مطالعه آمارهای‌قیمت پیشنهادشده ازنوع طبقه بین‌المللی استانداردبازرگانی* است که درشاخص جدیدتغییرات مختصری‌که متناسب بااحتیاجات کشور درآن داده شده‌است ، مورداستفاده قرارگرفته است . این‌طبقه‌بندی‌شامل ۸گروه اصلی‌وشعدادی‌گروه فرعی‌است که ذیلا " به آنهااشاره شده است .

---

* International Standard Trade Classification(I.S.T.C)

۰۰۰۰۰۰	موادخوراکی
۰۱۰۰۰۰	گوشت وفرآورده‌های‌آن
۰۱۲۰۰۰	گوشت قرمز
۰۱۴۰۰۰	مرغ
۰۱۵۰۰۰	ماهی وفرآورده‌های آن
۰۲۰۰۰۰	لبنیات وتخم مرغ
۰۲۳۰۰۰	لبنیات
۰۲۵۰۰۰	تخم مرغ
۰۳۰۰۰۰	حبوب
۰۴۰۰۰۰	غلات وفرآورده‌های‌آن
۰۵۰۰۰۰	میوه‌هاوسبزیها
۰۵۱۰۰۰	میوه‌های تازه
۰۵۲۰۰۰	سبزیهای تازه
۰۵۳۰۰۰	میوه‌های خشک
۰۵۶۰۰۰	میوه‌هاوسبزیهای کنسروشده
۰۵۸۰۰۰	آب میوه
۰۶۰۰۰۰	قندوشکر وفرآورده‌های آن وعسل
۰۷۰۰۰۰	قهوه وکاکائوروچای‌وادویه
۰۸۰۰۰۰	روغنهای‌حیوانی‌ونباتی
۰۹۰۰۰۰	دیگرمواد خوراکی
۱۰۰۰۰۰	مشروبات ودخانیات
۱۱۰۰۰۰	مشروبات
۱۲۰۰۰۰	دخانیات
۲۰۰۰۰۰	موادخام صنعتی‌غیرسوختنی
۲۱۰۰۰۰	پوست خام

۲۲۰۰۰۰	دانه‌ومیوه‌های روغنی
۲۴۰۰۰۰	چوب والوار
۲۶۰۰۰۰	پشم وپنبه
۲۷۰۰۰۰	موادخام معدنی
۲۸۰۰۰۰	موادمعدنی فلزی
۲۹۰۰۰۰	مواداولیه معدنی‌وگیاهی
۳۰۰۰۰۰	سوخت های‌معدنی‌وفرآورده‌های آن
۳۲۰۰۰۰	ذغال سنگ
۳۳۰۰۰۰	نفت ومشتقات آن
۳۴۰۰۰۰	گاز
۵۰۰۰۰۰	مـــــوادشیمیائی وپتروشیمی
۵۱۰۰۰۰	موادشیمیائی‌آلی
۵۲۰۰۰۰	موادشیمیائی معدنی
۵۳۰۰۰۰	انواع رنگ ومرکب
۵۴۰۰۰۰	داروهاوسموم دفع آفات
۵۵۰۰۰۰	لوازم آرایش وبهداشت
۵۶۰۰۰۰	کودشیمیائی
۵۸۰۰۰۰	محصولات پلی مریزه
۵۹۰۰۰۰	سایرمواد شیمیائی
۶۰۰۰۰۰	مصنوعات برحسب مواداولیه
۶۱۰۰۰۰	چرم
۶۲۰۰۰۰	مصنوعات لاستیکی
۶۳۰۰۰۰	مصنوعات چوبی
۶۴۰۰۰۰	کاغذومقوا
۶۵۰۰۰۰	منسوجات

۶۶۰۰۰۰	مواد معدنی غیرفلزی
۶۷۰۰۰۰	آهن وفولاد
۶۸۰۰۰۰	سایرفلزات
۶۹۰۰۰۰	مصنوعات فلزی
۷۰۰۰۰۰	ماشین آلات ووسائط نقلیه
۷۱۰۰۰۰	ماشین های مولدنیرو
۷۲۰۰۰۰	ماشین آلات تخصصی
۷۴۰۰۰۰	لوازم وماشین آلات صنعتی عمومی
۷۶۰۰۰۰	وسایل سمعی وبصری وارتباطی
۷۷۰۰۰۰	سایرماشین آلات
۷۸۰۰۰۰	وسائط حمل ونقل وقطعات آن
۸۰۰۰۰۰	کالاهای متفرقه

علاوه بر ۸گروه اصلی وگروههای فرعی فوق، شاخص بهای عمده فروشی کالاها برای سه گروه عمده کالاهای تولید ومصرف شده درکشور، کالاهای وارده به کشور وکالاهای صادراتی وهمچنین دو گروه اختصاصی محصولات کشاورزی و مصالح ساختمانی شاخصهای جداگانه ای محاسبه می گردد. گروههای فرعی دوگروه اختصاصی مذکور عبارتند از :

محصولات کشاورزی ودامپروری

زارعت

غلات

حبوب

مواد خام صنعتی

میوه‌ها وسبزیها

دامپروری وماهیگیری

حیوانات زنده وماهی

مواد خوراکی دامی

مواد غیرخوراکی دامی

مصالح ساختمانی

مصالح ساختمانی فلزی

مصالح ساختمانی غیرفلزی

۴-۲- انتخاب کالاهای مشمول شاخص وتعیین ضرایب اهمیت آن‌ها : دراانتخاب کالاهاوتعیین ضرایب اهمیت آن بمنظور تهیه شاخص بهای عمده‌فروشی کالاهادرایران درسال ۱۳۶۱ درهریک ازسه گروه عمده‌کالاهای تولیدومصرف شده درکشور، کالاهای وارده به‌کشوروکالاهای صادراتی عوامل نامبرده درذیل مدنظر قرارگرفت .

۱- اهمیت نسبی حجم فروش هرکالا درمجموعه کالاهای موجود وتولیدشده‌درسال ۱۳۶۱ (نسبت ارزش افزوده هرکالا به ارزش مجموعه کالاهای موجود درسال ۱۳۶۱)

۲- اهمیت نسبی فروش هرکالا درگروههای مربوطه

۳- دارا بودن بازارعمده فروشی دربازارهای ایران

۴- آینده نگری درموردکالاهائی‌که درآینده ممکن‌است‌به بازارهای عمده‌فروشی کشور وارد شوند ویاازاین بازارها خارج گردند .

۵- امکان جمع‌آوری قیمت عمده‌فروشی کالاها

۶- موارداستفاده شاخص فردی یک کالا

۷- مقایسه شاخصهای جدید باشاخصهای گذشته

بادرنظر گرفتن عوامل فوق برای انتخاب کالاهای مشمول شاخص بهای عمده‌فروشی کالاها ابتدا فهرست بیش از ۱۰۰۰ قلم کالاکه بنظر میرسید درسطح عمده‌فروشی دارای معاملات قابل ملاحظه‌ای درسطح‌کشور میباشد تهیه گردید ودرگروههای ۸ گانه بالاطبقه بندی‌شد ،همانطور که اشاره شد علاوه بر ۸ گروه اصلی چون شاخص درسه گروه عمده کالاهای تولیدومصرف شده‌در کشور ، کالاهای‌وارده بکشور وکالاهای صادره ازکشورنیزتهیه می‌گرددمبادرت به انتخاب کالاها و تعیین ضرایب آنهابشرح ذیل دراین سه گروه عمده شد .

۱-۴-۲- انتخاب کالاهای‌تولید ومصرف شده درکشور : دربین‌کالاهای تولیدومصرف شده درکشور ۲۲۲ قلم کالا که ازنظرتولیدنسبت به بقیه کالاهااهمیت قابل توجهی داشتند برای

قیمت گیری دراین شاخص انتخاب گردید . برای انتخاب این تعداد ازکالاها علاوه براطلاعات موجود دربانک مرکزی جمهوری اسلامی ایران ازاطلاعات وآمارهای موجود دروزارتخانه‌ها وسایر موسسات دولتی وخصوصی بویژه مرکزآمار ایران استفاده شد . دراستفاده ازاین اطلاعات سعی شد که فقط آن قسمت از حجم فروش کالاهاکه وارد بازارهای عمده فروشی کشور میشود درمحاسبات منظورگردد و کالاهائیکه بعلت خود مصرفی وپامبادله وبابعلل دیگروارد بازار عمده‌فروشی نمی‌شوند درمحاسبات منظور نگردد . حجم فروش تعدادی ازایندسته ازکالاهابااستفاده ازارقام تولیدآنهاودرنظرگرفتن صادرات و خودمصرفی ومقادیری که بازار عرضه نمیشوندوباتوجه به متوسط قیمت های عمده فروشی آنها درسال ۱۳۶۱ برآورد گردید .

۲ـ۴ـ۲ـ انتخاب کالاهای وارده بکشور : درسال ۱۳۶۱ ازمیان انبوه کالاهای وارده بکشور تنها ۵۸ قلم کالا که درمقایسه با بقیه کالاهادارای اهمیت نسبی قابل‌توجهی بودند وامکان قیمت گیری برای آنهاتاتجدیدنظر بعدی میسربود انتخاب شد . برای انتخاب این تعداداز کالاهای وارداتی ، ابتدا ازنشریه آمار بازرگانی خارجی ایران درسال ۱۳۶۱ ، مقدار و ارزش سیف* کالاهای وارداتی استخراج شدوسپس بااستفاده ازنشریه مقررات واردات وصادرات دراین سال، حقوق گمرکی وسودبازرگانی تك تك کالاهامحاسبه وبه ارزش سیف آنهااضافه گردید . ضمنا " هزینه‌های مربوط به حمل ونقل وسود واردکننده نیز برآورد وبه قیمت این اقلام اضافه گردید . دراین محاسبات نیز سعی شد تاکالاهائی که وارد بازار عمده فروشی نمیشوندویابازار عمده فروشی ندارند درمحاسبات منظور نگردد . درمورد برآورد حجم فروش کالاهائیکه بوسیله دولت وارد وباقیمت عمده فروشی معین وحتی کمتر از قیمت های خرید دراختیار عاملین فروش قرارمی‌گیرد، نظیرگندم ، گوشت و تعدادی کالای دیگر، ارزش فروش آنها مستقیما " ازسازمانهای ذیربط اخذ گردیدوملاك تعیین ضرایب اهمیت آنها قرارگرفت .

۳ـ۴ـ۲ـ انتخاب کالاهای صادره ازکشور : درسال ۱۳۶۱ دربین کالاهای صادراتی غیرنفتی کشور تنها ۱۲ قلم کالاکه به نسبت اهمیت بیشتری داشتند برای قیمتگیری درشاخص بهای

---

*قیمت سیف CIF یا Cost Insurance Freight عبارتست ازقیمت کالاهای وارداتی درمرزهای کشور بدون درنظرگرفتن حقوق گمرکی، سودبازرگانی، هزینه حمل ونقل داخلی ،سود واردکننده ودیگر حقوق مربوطه میباشد .

عمده فروشی کالاها انتخاب گردید. برای انتخاب این تعداد اندک از کالاها با استفاده از نشریه آمار بازرگانی خارجی ایران در سال ۱۳۶۱ ارزش فوب آنها مستقیماً* استخراج و ملاک تعیین ضرایب اهمیت این قبیل از کالاها قرار گرفت.

هر چند بسیاری از ۲۹۴ قلم کالای منتخب در سال ۱۳۶۱ طبقه بندی شده در سه گروه نامبرده در بالا تحت نام واحدی ذکر میشود، معهذا در اکثر موارد هر کالا شامل انواع متعددی است که از لحاظ قیمت و مشخصات با هم تفاوت دارند. از اینرو برای پوشش این اقلام از ابتدای سال ۱۳۶۱ هر ماه قیمت بیش از ۶۰۰۰ مقنه کالا از بازارهای عمده فروشی کشور جمع آوری میشود و نرخ قیمت آنها برای محاسبه شاخص مورد استفاده قرار می گیرد. این قیمتها با استفاده از پرسشنامه ای که حاوی مشخصات دقیق انواع مختلف و پر مصرف هر کالا و پرسشهائی در زمینه نوع معامله، شرایط فروش و مقدار فروش هر کالا و یا انواع مختلف آن می باشد از طریق مراجعه مستقیم و مصاحبه با صاحبان منابع اطلاع اخذ می گردد.

کالاهائی که در سه گروه نامبرده در بالا بدلایل مختلف برای قیمت گیری در این شاخص انتخاب نشدند ضرایب اهمیت هریک از آنها به نسبت اقلام منتخب میان کالاهای موجود در هر گروه که برای قیمت گیری انتخاب شده اند، توزیع گردید. در این انتخاب تقریباً در هر گروه فرعی حجم فروش کالاهای منتخب معادل ۶۰ الی ۱۰۰ درصد حجم فروش مجموعه کالاهای موجود در آن گروه را می پوشاند. این نسبت برای گروههای اصلی حدود ۸۰ تا ۱۰۰ درصد بود.

۴-۲-۲- ضرایب اهمیت کالاها و محدودیت آنها: ضرائب اهمیت کالاها در یک شاخص به هدف و نوع استفاده ای که از آن بعمل می آید بستگی کامل دارد. اگر هدف از تهیه شاخص اندازه گیری سطح عمومی قیمتهای عمده فروشی کالاها باشد، ضرایب اهمیت کالاها باید کل فروش کالاها را در کلیه مراحل فروش بجز مرحله خرده فروشی شامل گردد. در این حالت ارزش مقادیری از کالاها که بازار برنده می شود و یا برای خود مصرفی نگهداری می گردد نباید بحساب آید. ولی اگر منظور از تهیه شاخص اندازه گیری تغییرات سطح عمومی قیمت کالاها در اولین بازارهای عمده فروشی نزدیک بمراحل تولید باشد، ضرایب اهمیت کالاها باید حجمی از مبادلات را در بر گیرد که در چنین بازارهائی بفروش میرسد. در اینصورت واضح است

---

* فوب FOB یا Free on Board قیمت تحویل کالا عادر روی کشتی در مناطق بارگیری می باشد.

که ارزش مقادیری که برای خود مصرفی نگاهداری میشوند ویابعللی وارد اولین بازارعمده فروشی نمیشوندباید خارج گردد.

درشاخص جدید ضرایب اهمیت کالاها ازارقام مربوطبه سال ۱۳۶۱ استفاده شده است. ضرایب اهمیت هرگروه ازحاصل جمع ضرایب اهمیت کالاهای منتخب موجوددرگروه حاصل شده است. مشخصات کالاها ازبین انواع متعدد پرفروش کالاهادربازارهای اولیه تعیین گشته است اما ضرایب اهمیت بهکالاها داده است نه به انواع آن. ذیلا" به ضرایب اهمیت سه گروه عمده و ۸ گروه اصلی شاخص اشاره شده است:

شاخص کل	۱۰۰/۰۰
سه گروه عمده :	
کالاهای تولیدومصرف شده درکشور	۸۳/۳۳
کالاهای وارده بکشور	۱۵/۹۷
کالاهای صادره ازکشور	۰/۷۰
۸ گروه اصلی :	
موادخوراکی	۳۹/۶۲
مشروبات ودخانیات	۲/۶۶
مواد خام صنعتی غیرسوختنی	۲/۹۰
سوخت های معدنی وفرآورده های آن	۶/۷۰
موادشیمیائی وپتروشیمی	۴/۸۱
مصنوعات برحسب مواداولیه	۳۰/۷۰
ماشین آلات ووسائط نقلیه	۹/۷۳
کالاهای متفرقه	۲/۸۶

۵- نحوه جمع آوری قیمت های عمده فروشی : جمع آوری قیمت های عمده فروشی ازطریق مصاحبه مستقیم انجام میشود بدین ترتیب که آمارگیران اقتصادی بانک به مدیران کارخانه ها و تولیدی وعاملین فروش دست اول مستقیما" مراجعه وقیمت عملی فروش کالاها راجمع آوری می کنند درموقع مراجعه علاوه بر جمع آوری قیمتها، اطلاعاتی نیز درباره وضع بازاروچگونگی حجم عرضه و... نیز جمع آوری می کنند. درهنگام مراجعه حتی المقدور سعی میشود آخرین قیمت عملی فروش کالا کلا

درمعاملات معمول بوده است ، پرسش گردد ، درمواردی که کالائی درانواع مختلف وباشرائط گوناگون بفروش میرود قیمت انواعی از آن کالا جمع آوری می گردد که بیشتر معاملات روی آن انجام شده است ، بطورکلی در اکثر قریب باتفاق موارد قیمت های جمع آوری شده کالاها در این شاخص درمورد مصنوعات داخلی قیمت فروش درب کارخانه ویا کارگاههای تولیدی است ، درمواردیکه عملا " قیمت های دیگری حاکم برباز ار باشد ضمن تحقیق کافی این قیمتها جمع آوری ودرمحاسبات شاخص ملحوظ می گردد ،

قیمت انواع محصولات کشاورزی ازنزدیکترین بازار عمده فروشی به مراکز تولید جمع آوری میشود درمورد محصولات فصلی کشاورزی که در دوره های کوتاهی توسط کشاورزان ببازارهای عمده فروشی عرضه میشود وعموما " عمده فروشان بانگهداری درجاهای مناسب درماههای بیشتری ازسال آنهـــا را بفروش میرسانند، قیمت آن ها درتمام طول ۱۲ ماه سال از ۵ شهربزرگ کشور منجمله تهران جمع آوری می گردد ، برای محصولات صنعتی قیمت سرمعدن و کالاهای وارداتی قیمت فروش واردکنندگان دست اول مورد محاسبه در این شاخص قرار می گیرد ، قیمت کالاهای صادراتی ازصادرکنندگان دست اول جمع آوری میشود ،

۶ـ۲ـ مراکز قیمت گیری کالاها : چون درتهیه شاخص بهای عمده فروشی کالاها هدف بر این است که قیمت ها ازاولین بازار عمده فروشی که غالبا " محل تولید کالاها است جمع آوری شود، لذا قیمتها از ۳۱ شهر که مرکز داد وستد این نوع کالاهاست جمع آوری می گردد ، بطورکلی قیمت کلیه محصولات کشاورزی ودامی ومواد غذائی که ازحیث تولید ومصرف دردرجه اول اهمیت قرار دارند، از ۵ شهر بزرگ کشور شامل تهران ، اصفهان ، تبریز ، شیراز ومشهد جمع آوری میشود ، چون تهران مرکز تولید اکثر کالاهای صنعتی است ویا دفتر فروش بسیاری ازمحصولات صنعتی درتهران قرار دارد، لذا قسمت اعظم کالاهای مصنوع تولید شده درکشور وتقریبا " کلیه کالاهای وارداتی وهمچنین کالاهای صادراتی از تهران که مرکز عمده فروشی وداد وستد این نوع کالاهاست قیمت گیری میشود ،

۶ـ۳ـ فرمول محاسبه شاخص : درمحاسبه شاخص بهای عمده فروشی کالاها در ایران مانند شاخص بهای کالاها وخدمات مصرفی ازشاخص حسابی موزون قیمتهای نسبی کالاها با فروش کالاها در ســـال پایه ۱۳۶۱ که درقسمت ۶ـ۱ به آن اشاره شد استفاده میشود ،

نقش تعاریف و مفاهیم در طرحهای آماری

کبری قندی

مرکز آمار ایران

پس از جنگ جهانی دوم ، و به ویژه در بیست سال اخیر ، تعاریف و مفاهیم درهریک از رشته های علوم و فنون مورد استفاده تخصصهای گوناگون قرارگرفته است و در مواردی یک تعریف یا مفهوم خاص در رشته های مختلف علوم مفاهیم متفاوت داشته است . واژه آمار نیز از مواردی است که با مفاهیم گوناگون بکار می‌رود . در یکی از دانشگاههای ایران یکی از اساتید در طول چند ترم تحصیلی از دانشجویان خواسته بودند که آمار را تعریف کنند ، با جمع آوری پاسخهای دریافتی حدود ۱۶۵ تعریف گوناگون برای آمار ذکر شده بود .

در فرهنگ لغات عمید در مقابل واژه‌های تعریف و مفهوم آمده است :

تعریف = شناساندن حقیقت امری ، یا مطلبی را برای کسی بیان کردن .

مفهوم = فهمیده شده ، دانستهشده ، آنچه به فهم و ادراک در آید .

با توجه به تعریف واژه " تعریف " متوجه می‌شویم که حقیقت امر ابتدا برخود ما بایستی روشن شود .

به عبارتی برای انتقال اطلاعات ایجاد درک مشترک ضروری است . پس انتقال تصورات و مفاهیم و لغات و اصطلاحات مستلزم داشتن درک مشترک است زیرا این الفاظ حکم زبان علم‌را دارند که برای غیر اهل فن درک آن مشکل و گاهی بدون ذکر مقدمه غیر ممکن است .

بنابراین به کار بردن الفاظ مشترک یا معانی مشترک کمک موثری در تفهیم می‌کند .

تعریف مفاهیم علمی و این که روشن شود منظور ما از تعاریف کلمات و ردیف کردن منظم آنها به دنبال هم چیست خیلی اهمیت دارد . اگر به تعریف کلماتی نظیر ملیت ، سواد ،قدرت ،

تعصب و کارآیی توجه کنیم خواهیم دید که به قول مولانا " هرکسی از ظن خودشد یارمن " .

دیده شده که وقتی از فردی سئوال می‌شود آیا تعصب در قضاوت شما موثر است ؟

ابتدا ناچاریم تعصب را تعریف کنیم تا به پاسخ سئوال برسیم و یا در پرسشنامه آمساری سئوال می‌کنیم رئیس خانوار کیست ؟ بلافاصله باید با تعریف آماری که برای رئیس خانوارمشخص کرده‌ایم خود را وفق دهیم . چه بسا یک فرد مجرد ساکن یک واحد مسکونی (که بایدقبـــــلا " تعریف شده باشد) هم می‌تواند و باید طبق تعریف رئیس خانوار منظور شود . و یا تعریف واحد مسکونی . درهیک کشور توسعه یافته با داشتن حداقل امکانات رفاهی حتی‌گاهی وسایل برودتـــــی‌و حرارتی را در ذهن متبادر می‌کند . در حالیکه درکشورهای درحال‌توسعه یک زاغه حتی یک فضای غیر سقف ، بدون کوچکترین امکانات رفاهی برای یک خانوار واحدمسکونی تلقی‌می‌شود . بنابراین ابتدا هر تعریفی ظرافت خاصی را می‌طلبد که غیر قابل ایراد و مطلق نیست .

اصول تعریف :

در تعریف واژه ، نباید از خرد آن واژه استفاده کرد و این اصل معتبری است که معمولا " رعایت می‌شود . نقش تعاریف و مفاهیم در هر رشته از علوم محض و کاربردی روشن اســـت ،در طرحهای آماری الفبای طرح بعد از بیان اهداف طرح ، تعاریف و مفاهیمی است که در طرح‌بکار می‌رود . این تعاریف جملات کوتاه و گویایی است که چند کلمه کاملا " روشن آن‌راتـــکمیل می‌دهـــد . بنابراین تعریف آماری نباید حشو و زوائد داشته باشد ." منظور از حشو و زوائد آن‌کلماتی است‌که اگر از تعریف حذف کنیم از بیان حقیقت تعریف چیزی کم نمی‌کند " مفهوم یا بصورت‌جمع مفاهیم یا توضیحاتی که داده شد تا حد زیادی گویا و آن چیزی است که به فهم ادراک درآید و جـــای شک و شبهه نباشد . لذا وقتی جامعه یا نمونه یا واحد آماری را تعریف می‌کنیم آن چیزی بـــــه شنونده مفهوم می‌شودکه خواست گویندهاست . درمورد مفاهیم آماری‌دستیابی به تعریف همگن بمنظور امکان پذیرساختن مقایسه آمارهمواره موردنظربوده وبرای بوجود آوردن تعاریف بهینه که میتـــوان آنهارا تعاریفاستاندارد نامیدتلاش‌های زیادی‌انجام‌می‌گیردبرای‌سازمان‌هایی‌کهاقدام‌به جمع‌آوری آمار و آمارگیری‌می‌کنند وجودتعاریف ومفاهیم‌بعنوان‌یکاصل‌مطرح‌است .درست‌مانند تعیین‌واحدآماری‌یاجامعه‌آمـاری درفصـــل‌قبل‌طرح .

باپیشرفت‌علم‌آمارمسئلهتعاریف‌اهمیت بیشتری پیداکرده‌وفرانرازاراتعریف‌درهرطرح‌آماری‌مسئله استانداردبودن تعریف‌مطرح‌می شود بنابراین‌تمام‌طرحهای‌آماری‌کهبارعایت‌اصول‌علمی‌انجام‌می‌گیردلزوما "

با تعاریف ومفاهیم شروع می شوند همانگونه که اشاره شد خصوصیت مطلوبی که در تعاریف مربوط به آمارگیریها مورد نظر است استاندارد بودن این تعاریف است .

بطور خلاصه میتوان گفت استاندارد بودن تعریف یعنی تعیین حدود آن تعریف ، تعیین مشخصات تعریف مشخص کردن حداقل ها و در نتیجه کاهش دامنه تعریف به کمترین حد . بطور مثال اگر به تعریف شهر در سرشماری ۱۳۵۵ نگاه کنیم ممکن است با این تعریف مکررا "برخورد کنیم که شهر را نقطه ای جغرافیایی می شناسیم که حداقل دارای پنجهزار نفر جمعیت یا مرکز شهرستان باشد . این ضابطه یک حداقل "کمی" برای شهر در نظر می گیرد . حتما "شهر دارای تعریف جامعتر و وسیعتری است که عوامل کیفی فراوانی در آن در نظر گرفته می شود . اما در اینجا تنها یک حداقل و یک محدودیت در نظر گرفته می شود .

عدم استفاده از تعاریف استاندارد چه لطمه ای به کیفیت آمارگیریها و سرشماریها وارد می کند ؟ بطور خلاصه می توان گفت ،

۱- تعاریف در آمارگیریها با اهداف و شرایط یکسان ، دارای تفاوتهایی خواهند بود زیرا اینها ذهنیت کارشناسان و شرایط خاص حاکم بر هر آمارگیری نمونه ای و یا سایر سرشماری تعیین و نوشته شده اند و طبیعتا "با نظر کارشناسان مربوط با یک آمارگیری تا آمارگیری بعدی ، با سلیقه ها و شرایط متفاوتی تغییر خواهد کرد .

۲- امکان مقایسه آمار در سطح ملی بین سازمانهای مختلف تولید کننده آمار در کشور و نیز در سطح بین‌المللی با آمارهای مشابه کشورهای دیگر کاهش یافته و یا از بین می رود زیرا در هر سازمان یا در هر طرح آماری بنا به سلیقه ، امکانات و شرایط خود از یک تعریف استفاده می کند .

۳- به کیفیت آمار و نتایج بدست آمده از آمارگیریها لطمه وارد می آید زیرا ممکن است معیارهایی کمابیشی در یک تعریف وجود داشته باشد نادیده گرفته شوند .

۴- همه عوامل بالا باعث اتلاف هزینه و نیروی صرف شده برای آمارگیریها می شود ، زیرا در مجموع از وثوق و اعتبار آمار می کاهد .

شایان ذکر است که منظور از استاندارد کردن تعاریف وضع قانونهای ابدی نیست ، وضع کردن تعریف نیست ، تعاریف بایستی از عمل و تجربه ناشی شوند و تکامل پیدا کنند . در تعاریف استاندارد هم مطابق شرایط جامعه ممکن است هر چند سال تغییری داده شود و تجدید نظر گردد . بملاوه استفاده از تعاریف استاندارد بمعنی نادیده گرفتن ضرورتهای طرحهای مختلف نیست ، ممکن است واحد مسکونی در طرحی بنا بر اهداف آمارگیری دارای مشخصاتی باشد که در طرحی دیگر این مشخصات فرق کند ولی می توان حدود و حداقل استانداردی برای واحد مسکونی تعیین کرد . امکان مانور در این حدود وجود خواهد داشت .

در یک جمله می‌توان گفت استفاده از تعاریف استاندارد برای به وجود آمدن زبان مشترک آماری در سطح ملی و بین‌المللی کمک می‌کند و پیگیری از آمار را دقیق‌تر و معتبرتر می‌سازد.

به عبارت دیگر تعاریف و مفاهیم پایه و اساس تحقیق علمی است و بدون استفاده از تعاریف و مفاهیم، پیشرفت علمی امکان پذیر نیست بخصوص نتایج آماری دقیقی بدست نخواهد داد میتوان گفت که در یک طرح آماری هیچ عدد و رقمی و آماری بدون داشتن تعریف اولیه آماری قابل تفسیر و استنتاج نخواهد بود. هر تعریفی پاسخ منطقی یک پرسش است که مکان و زمان کمتر در آن نقش دارند و حتماً این پرسش انگیزه‌ای علمی برای طرح دارد.

ذکر این نکته مهم ضروری است که ارائه یک قانون کلی برای اخذ تصمیم در مورد پذیرش یا عدم پذیرش یک تعریف در تمام طرحهای آماری وجود ندارد، چرا که این تصمیم منحصراً "بستگی به موضوع خاص مورد آمارگیری و ارتباط منطقی و تک تک سئوالات پرسشنامه دارد. بدیهی است تعاریف و مفاهیم یکی از عمده‌ترین بخشهای زیربنایی در تهیه یک طرح آماری بشمار میرود. و یکی از هدفهای اساسی در گردآوری تعاریف و مفاهیم استفاده از آن در امر برنامه ریزیهای کلان می‌باشد بحث در مورد آمار و ارقام بدون دانستن تعاریف و مفاهیم به بهینه امکان پذیر نخواهد بود. ابهام در تعاریف و مفاهیم خود مشکل بزرگی در نتیجه به دست آمده ایجاد می‌کند هماهنگی داشتن تعاریف و مفاهیم در طرحهای آماری دقت و نظم و انسجام در تهیه آمار معمولاً " اولین مسئله در صحت انجام یک طرح آماری می‌باشد، و ابهام در تعاریف و مفاهیم یکی از مواردی است که اشتباهات محتوی را بوجود می‌آورد که در اثر گزارش غلط خصوصیات و صفات واحدهای مورد آمارگیری پیش می‌آید بطور کلی برای جمع‌آوری آمار مربوط به هر طرح آماری خواه از روش سرشماری خواه آمارگیری نمونه‌ای و یا آمارهای ثبتی استفاده بشود آنچه در کاربرد هر یک از این روشها اهمیت بسیار زیادی دارند استفاده از تعاریف و مفاهیم و طبقه بندی‌های استاندارد در زمینه‌های مختلف است که مورد توصیه سازمانهای بین المللی می‌باشد حائز اهمیت است.

خصوصیات مطلوب استفاده از یک تعریف، هنگامی که در طرحها از تعاریفی یکسان استفاده نشده باشد نتایجی متفاوت را نشان خواهند داد. و نوسان این نتایج گاه به حدی است که تحلیل گر را به تردید در صحت آنها وا می‌دارد تعاریف و مفاهیم و نحوه کاربرد درست آنها نیز از عواملی است که اثرات قابل توجهی در نتایج به جای خواهد گذاشت.

در اجرای طرحهای آماری استنباط یکسان پرسشگران از اصطلاحات بکار رفته در طرح اهمیت زیادی دارد زیرا اگر هریک از پرسشگران دیدگاه خاص خود را در کسب اطلاعات به کاربرد نتیجه کار یکسان

نخواهد بود وزحماتی که در این باره تحمل می شود به هدر خواهد رفت برای اخذ نتیجه مطلوب لازم است هریک از اصطلاحات مورد استفاده تعریف شود تا همه ماموران آمارگیر از اصطلاحات مزبور برداشتی هماهنگ و یکسان داشته باشند .

جدا از مسئله اهمیت تهیه تعاریف و مفاهیم از نظر قانون این وظیفه مهم مرکز آمار ایران است در ماده ۲ بند ۸ قانون مرکز آمار ایران آمده است تهیه تعاریف و مفاهیم و معیارها و طبقه بندی های آماری و ماده ۸ وزارتخانه ها و موسسات دولتی و وابسته به دولت و شرکتهای دولتی موظفند در انجام آمارگیریها از تعاریف و مفاهیم ، روشها و معیارها و طبقه بندی های آماری مرکز آمار ایران تبعیت نمایند .

بدین منظور در اجرای این وظیفه قانونی مرکز آمار ایران به جمع آوری تعاریف و مفاهیم اقدام می نماید .

تهیه فرهنگ تعاریف و مفاهیم در مرکز آمار ایران

از آغاز سال ۱۳۶۴ مرکز آمار ایران به این نتیجه رسید که تهیه تعاریف و مفاهیم استاندارد یک هدف سهل الوصول نیست و نیاز به کار اساسی و مداوم و پرزحمت و طولانی دارد . بنابراین تصمیم گرفته شد که تعاریف و مفاهیم موجود به عنوان اولین مرحله جمع آوری شده و بدون دخل و تصرف و بدون اظهار نظر کارشناسی در یک مجموعه گردآوری شود . تا مواد اولیه ای برای کاربردی و رسیدن به هدف اصلی یعنی تعاریف استاندارد فراهم آید در این مرحله مقرر شد که این کار اولا " کار اصلی یک کارشناس باشد و نه یکی از کارهای جانبی و در نتیجه جدیت و مسئولیت لازم در آن مورد نظر باشد ثانیا " آغاز کار از خانه خود یعنی مرکز آمار ایران باشد و تعاریف مرکز گردآوری شود ثالثا " تعاریف بصورت الفبایی ارائه شوند و نه موضوعی علت این تصمیم این بود که جمع آوری تعاریف همانند تدوین یک مرجع تلقی شود و در مرجع حق تفسیر نیست .

۱ـ موضوعی کردن تعاریف بنوعی ارزش گذاری تعریف محسوب می شود . بنابراین با تعاریف گردآوری شده میتوان مثل یک فرهنگ یا دائرةالمعارف برخورد کرد .

۲ـ دلیل دیگر آن سهولت دسترسی و یافتن تعاریف است .

۳ـ دلیل بعدی حجم تعاریف بود پیش بینی می شد که حجم تعاریف و مفاهیم طرحهای مرکز آمار ایران از چند صد فراتر رود در اینصورت نمی توان آنها را براحتی در یک طبقه بندی موضوعی قرار داد .

بالاخره از سال ۱۳۶۴ به بعد دفتر بررسی طرحها و محاسبات آماری مطابق شرح وظایف خود به

صورت جدی ومستمر تهیه تعاریف ومفاهیم آماری رابشکل جمع آوری تعاریف موجود ادامه داده است وتا کنون ۴ جلد نشریه که روی هم دارای چیزی حدود ۳۸۰۰ تعریف بوده از ۴۰۵ منبع استخراج ومنتشر کرده است . وبااین کاربصورت یک کار مهم وپایه ای برخوردنموده است . بطورخلاصه میتوان گفت تلاش برای تهیه تعاریف ومفاهیم آماری درمرکز آمارایران سابقه ای درحدودبیست سال داردودرسالهای قبل از ۱۳۶۴ این تلاش بصورت تبادل نظر ، نشستهای کارشناسی ومکاتبات اداری بوده است وحاصل کارمدون نشده بود .

به نظرمی رسدکه دلیل عدم دستیابی به تعاریف ومفاهیم استانداردتاکنون این بوده است که به این مسئله بصورت یک هدف سهل الوصول که در چند نشست میتوان بآن سروسامان داد نگریسته می شده اما تجربه نشان دادکه میبایستی به صورت یک هدف دوردست بااین مهم برخوردشودکه نیازبه کاری بنیادی واساسی دارد . برنامه کارمرکزآمارایران این است که باتوجه به تجارب گذشته تمام تعاریف ومفاهیم موجود سازمانهای داخل کشور راجمع آوری ومدون کند وسپس دامنه کاررا به تعاریف بین المللی تعمیم داده واین تعاریف داخلی رادرتطابق بااستانداردهای بین المللی به یک مجموعه استاندارد نزدیک نماید . امید است درآینده کاربیش ازپیش سرعت گیرد و دستیابی به هدف درزمان کوتاهتری میسرشود .

درحال حاضرمرکزآمارایران برای ساختن تعریف استاندارد که هدفمی باشد مصالح ومواد اولیه فراهم می آورد . امیداست پس از پایان کاربه دشواری بتوان واژه آماری یافت که درباره آن تعریفی جمع آوری نشده باشد واین خود به تنهایی کار بسیار ارزنده ای است .

مدل‌سازی و شبیه‌سازی تغییرات ارتفاع سطح آب دریای خزر به کمک مدل تابع انتقال درجه حرارت و میزان بارندگی

اکبر گلدسته

دانشگاه شهید بهشتی

۵شنبه ۹:۰۰ تالار ۴

مقادیر سریهای زمانی ارتفاع سطح آب دریای خزر (Y_t)، درجه حرارت (X_{1t}) و میزان بارندگی (X_{2t}) برحسب میانگین ماهانه در سالهای ۱۳۶۷-۱۳۳۰ در یک ایستگاه هواشناسی خاص موجودند. هدف ما یافتن مدل آماری است که بتواند تغییرات سری زمانی Y_t راکه در یک سیستم پویا برحسب دو سری زمانی X_{1t} و X_{2t} بیان کند و سپس مقادیر آیندهٔ سری زمانی Y_t را بکمک آن شبیه سازی کرد.

رابطه ورودی و خروجی در چنین سیستمی به شکل:

$$Y_t = \Pi_1(B)X_{1t} + \Pi_2(B)X_{2t}$$

بیان می‌شود که در آن $\Pi_1(B)$ و $\Pi_2(B)$ به ترتیب توابع انتقال مقادیر سریهای X_{1t} و X_{2t} به سری خروجی Y_t می‌باشند.

مدل با اضافه کردن نوفه N_t برای نمایش اختلالات احتمالی کامل می‌شود مدل نهایی به صورت:

$$Y = \frac{-۰/۷۵ + ۰/۵۳B}{۱ - ۰/۶B - ۰/۲۵B^۲}X_{1t} + ۰/۰۱X_{2t} + \frac{(۱ - ۰/۷۷B^{۱۲})(۱ - ۰/۱۳B)}{۱ - ۰/۹۷B}\alpha_t$$

برآورد شده است. در پایان بکمک شبیه‌سازی، مقادیر ارتفاع سطح آب برای ۵ سال آینده پیش‌بینی شده و سطح اطمینان ۹۵درصد برای آن محاسبه و ترسیم شده است.

کاربرد آزمونهای آماری در بررسی میزان امنیت دستگاههای رمزکنندهٔ مخابراتی

اکبر گلدسته*ـ هایده مهرنژاد

دانشگاه شهید بهشتی

آشنبه ۱۶:۳۰ تالار ۷

دستگاههای رمزکنندهٔ وظیفهٔ تغییر ماهیت پیامهای ارسالی از طریق یک کانال مخابراتی ناامن را برعهده دارند. واضح است که جهت حفظ امنیت ارسال پیام، توزیع آماری اطّلاعات پس از خروج از دستگاههای رمز (اطّلاعات رمزشده) باید دنباله‌ای شبه‌تصادفی از اطّلاعات باشد. در حقیقت شبه‌تصادفی بودن داده‌ها می‌بایست در هر قطعه از اطّلاعات رمزشده رعایت شود.

براین‌اساس می‌توان آزمونهای آماری را به منظور تشخیص شبه‌تصادفی بودن دنباله اطّلاعات خروجی دستگاه رمز به انجام رساند. در این مقاله انواع آزمونهای آماری که در بررسی شبه‌تصادفی بودن دنباله‌های دودوئی از اطّلاعات رمزشده به‌کار میروند بررسی و سپس آنها را برروی تنها الگوریتم استاندارد رمز داده‌ها اعمال کرده و نقاط ضعف و یا قوت آن را برمبنای نتایج بدست‌آمده مورد بحث قرار می‌دهیم.

تعیین عوامل مؤثّر در مدت زمان شیردهی بطریق آنالیز مسیری

حسین محجوب*

دانشگاه علوم پزشکی

محمود محمودی - کاظم محمد - پرویز کمالی

دانشگاه علوم پزشکی تهران

۵شنبه ۱۴:۳۰ تالار ۵

آنالیز مسیری آنالیز رگرسیون چندگانهٔ استاندارد می‌باشد که ارتباط زنجیری رگرسیونی بین متغیرها برقرار می‌باشد. مشخصه اصلی این ارتباط آن است که جریان از چپ به راست بوده و متغیری که در سمت راست ظاهر می‌شود، متغیر وابسته می‌باشد. جهت بررسی عواملی که در طول مدت زمان شیردهی مؤثر می‌باشند، نمونه‌ای تصادفی به حجم ۱۱۴۰ از کودکان ۰ تا ۶ سال شهر تهران انتخاب کرده و اطلاعات مربوط به متغیرهای سن، شغل و تحصیلات مادر و همچنین رتبه تولد کودک، نوع زایمان و سابقه بیماری مادر و مدت زمانی که مادران بچه خود را شیر می‌داده‌اند جمع‌آوری شده است. با استفاده از تکنیک آنالیز مسیری و انجام آزمونهای آماری مناسب، عواملی که در مدت زمان شیردهی مؤثرند، همراه با اثرات مستقیم و غیرمستقیم آنها در مدت زمان شیردهی بشرح ذیل می‌باشد:

اثر مستقیم رتبه تولد کودک ۰/۱۱۳، اثر سنّ مادر از طریق رتبه تولد کودک ۰/۰۷۰، اثر سنّ مادر از طریق تحصیلات مادر و رتبه تولد کودک ۰/۰۰۸، اثر سنّ مادر از طریق شغل وی و رتبه تولد کودک ۰/۰۰۰۹، اثر سنّ مادر از طریق رتبه تولد کودک ۰/۰۲۷-، اثر تحصیلات مادر از طریق شغل مادر و رتبه تولد کودک ۰/۰۱۰-، اثر شغل مادر از طریق رتبه تولد کودک ۰/۰۳۲ می‌باشد در ضمن اثر کلّ سنّ مادر ۰/۰۸۹، اثر کلّ تحصیلات مادر ۰/۰۲۷-، اثر کلّ شغل مادر ۰/۰۳۳ و اثر کلّ رتبه تولد کودک بر مدت زمان شیردهی ۰/۱۹۷ می‌باشد.

کد و کدگذاری در آمار

علی محرابی
مرکز آمار ایران

پیشگفتار

با توجه باینکه اختلافات وتشابهات معنابهی در پدیده‌های گوناگون اجتماعی -اقتصادی وجود دارد ودر این ارتباط خصوصیات فطری انسان ایجاب میکند که به این پدیده‌ها نگرشی عمیق داشته باشد لذا علاقمند است تا برای شناخت بهتر، این اختلافات وتشابهات را کشف ودریک سیستم اصولی و منطقی طبقه‌بندی نماید.

بدیهی است دراین طبقه‌بندی بشر به شناخت واقعیتها هدایت شده و از تشتت فکری که بازدارنده اخذ تصمیمات صحیح است مصون می ماند. لازم به ذکر است که یکی اززبانشناسان (OTTO JESPERSEN) پس از مطالعه زیاد که در زبانهای مختلف جهان کرده است میگوید:

- انسان حیوانی است که قادر است طبقه‌بندی کند.
- طبقه‌بندی، پایه واساس تحقیق علمی است.
- بدون استفاده از طبقه‌بندی، پیشرفت علمی امکان پذیر نیست.

حال که در این پیشگفتار اهمیت طبقه‌بندی روشن گردید شایسته است تا به نقش "کد" که کلید شمارش در هر عنوان طبقه‌بندی شده میباشد و همچنین کدگذاری پرداخته شود. ولی لازم است قبل از بیان این مطلب مقدمتا" به علت پیدایش "کد" پرداخته شود.

مقدمه

بشر نخستین پس از آشنائی با راه تامین اولین خواسته الزامی خود یعنی غذا، بدلیل داشتن خصوصیات گوناگون از جمله زیاده‌خواهی اقدام به تلاش برای دستیابی به نیازهای بعدی نمود. یکی از این نیازها مالکیت وبالطبع دست یافتن به هر چیز ممکن در حد فراوان بود لذا برای حفظ آن باید می‌دانست چه واحدی از هر چیز را در اختیار دارد تا در نگهداری وتزاید آن بکوشد. در این مقطع ابتدا بکمک انگشتان دست و سپس با انداختن سنگ در محلی به شمارش روزها و شبها و یا ماه‌ملك خود پرداخت در این رابطه برای راحتی بدنبال روش آسانتری بود که به مرور زمان کشیدن خط بسر روی دیوار را آموخت که هم‌اینك در آثار بجا مانده بر روی دیوار غارهای قدیمی مشهود است. این امر همچنان ادامه داشت و او پیوسته در پی راههای ساده‌تری برای محاسبه بود که با گذشت زمان و پیشرفتهای حاصله "چرتکه" را ساخت تا در شمارش ارقام بزرگ و اقلام بیشتر دچار مشکل نشود. بدیهی است زمانی فرا رسید که نقش "وقت" در محاسبه او را به فکر ساختن وسیله‌ای برای استفاده در محاسبه اعداد وارقام بزرگتر وجلوگیری از اتلاف وقت بیشتر نمود. این زمانی بود که ماشین حساب دستی به میان آمد. اما با پیشرفت جوامع بشری و وجود پدیده‌های مختلف این وسیله نیز هرگز پاسخگوی جمع‌آوری اطلاعات مورد نیاز نبود. با اختراع برق، ماشینهای برقی مکانیکی وارد بازار شد که تا دیرزمانی در محاسبات مورد استفاده واقع میشد. ناگفته نماند هر چه زمان پیش میرفت و بشر با این پیشروی در زمینه‌های مختلف پیشرفت میکرد جایگاه زمان وارزش واقعی آن بیشتر خودرا نشان میداد. با اختراع ترانزیستور تحول عظیمی در دانش انسان به وقوع پیوست که چون موجی فراگیر غالب ابزار و وسائل مورد استفاده و در دسترس بشر را در بر گرفته و موجبات تغییر و دگرگونی اساسی در آنها گردید که در این میان سهم محاسبات، در اختیار داشتن ماشینهای الکترونیکی بود که با امکانات و تسهیلات بیشتری واحدهای مورد نظر بزرگتری را با آن شمارش یا اندازه‌گیری کرد که اینك نیز مورد استفاده میباشد. این اختراع نیز پس از مدتی بدلیل فقدان حافظه مورد نیاز در مقابل داده‌های در حجم کلان

ناتوان مانده و طولی نکشید که کامپیوتر یعنی ماشین دارای حافظه زیاد جایگزین ماشینهای حسابگر قبلی شد که در مدت زمان کمتری برای پردازش حجم زیادی از اطلاعات نقش بسزائی را بعهده گرفت و طی چهار نسل پیشرفتهای شگرفی در این ماشینها حاصل شد که امروزه نسل چهارم آن مورد استفاده میباشد و شکی نیست که انسان در مسیر پیشرفتهای علمی با سیری ناپذیریکه در اندوختن علم دارد انتظار میرود همچنان در راستای تفکری نو برای نوآوری در تلاش ایجاد ساختاری بدیع باشد و تردیدی باقی نمیماند که روزی برای این ماشینها نیز جایگزینی ارزنده‌تر اختراع شود.

موضوع

همانگونه که در پیشگفتار بیان شد برای شناخت پدیده‌های گوناگون اجتماعی ـ اقتصادی پس از جداسازی یا بعبارتی طبقه‌بندی نیاز به اندازه‌گیری آنها احساس گردید لذا در همینجا نقش شمارش ظاهر شد و بطوریکه مقدمتا" گفته شد ابتداء بصورت دستی این هدف دنبال شد و بتدریج تنوع و فزونی هر روز روش جدیدی در این امر پدیدار گشت ، یکروز بوسیله انگشتان دست و روزدیگر بوسیله پرتاب سنگ و زمانی بسد با کشیدن خط بر روی دیوار ، این روند بدلیل نیاز و داشتن اطلاعات بیشتر در هر برهه زندگی جدید بخود گرفت تا زمانیکه اطلاعات بصورت نوشتار بر روی پوست درختان و حیوانات و سرانجام کاغذ منتقل شد و با بالا رفتن سطح آگاهیها و در اختیار داشتن داده‌های بیشتر از پیش لزوما" شمارش عناوین مختلف و موضوعات گوناگون به‌سختی صورت گرفته و یا امکان‌پذیر نبود بدین جهت بشر به فکر چاره افتاد و قطعا" پس از بحث و بررسی و ارائه طرق مختلف "کد" را انتخاب ودر مقابل عناوین و موضوعات مورد نیاز که با روشهای قبلی به سهولت شمارش نشده و یا اصلا" غیر قابل شمارش بود قرار داد .

"کد" در فرهنگ به معنای رمز یا بعبارت دیگر شماره میباشد وآن عددی است یا رقمی وبسته بمورد بیشترکه پس ازقرارگرفتن در مقابل هر عنوان قابل شمارش است . برای روشن شدن مطلب گفته میشود که اگر بخواهیم دریک جامعه بزرگ تعداد مردها و زنها را شمارش کرده ارائه دهیم نیاز به جایگزین برای کلمه مرد و زن داریم چون شمارش کلمه مرد و زن با وجود سادگی کلمه ، مشکل است چه رسد به بعضی عناوین طولانی که بعدا" توضیح داده خواهد شد .بدینمنظور بعنوان مثال دریک جدول در مقابل ستون کلمه مرد و زن یک ستون تحت عنوان ستون کد شکل میگیرد و سپس بجای هر کلمه مرد مثلا" عدد ۱ و برای زن عدد ۲ منظور میشود بنابر شمارش اعداد ۱ و ۲ با کمک جداول واسط بسهولت امکان پذیربوده و در نهایت حاصل جمع نماینده تعداد مرد و زن خواهد بود.تجربه استفاده از "کد" راهگشای حل بسیاری از مشکلات و عملا" جمع‌آوری داده‌ها و پردازش آنها و نهایتا" انتشار نتایج میباشد بدین سبب که در آمار مباحثی است که بجز پرداختن به "کد" راه دیگری

برای ضبط آنها وجود ندارد یعنی آنجا که سخن از مبحث آمار بازرگانی خارجی است مشاهده میشود که حدود ۲۵۰۰ تا ۳۰۰۰ نوع کالای وارداتی وجود دارد که اگر بخواهیم تمام آنها را با تمام مشخصه سازیم امکان جمع و جور کردن آنها وجود ندارد چون آنقدر کالاهای متنوع با مشخصات خاص در بخشهای صنعتی، غذائی، دارویی و غیره میباشد که فهرست کردن نام آنها خود، کاری بس عظیم است لذا مرتبا " با تدوینکاری امکان تفکیک و طبقه‌بندی آنها میسر است "

مقایسه داده‌های آماری کشورها در سطح منطقه، قاره و جهان از دیدگاههای مختلف امکان‌پذیر نیست مگر اینکه، برای جمع‌آوری و تجزیه و تحلیل داده‌ها از تعاریف و مفاهیم و طبقه‌بندیهای یکسان در کشورهای مختلف استفاده شود. به همین دلیل است که سازمانهای بین‌المللی آماری برای پدیده‌های مختلف اجتماعی - اقتصادی و فرهنگی طبقه‌بندیهای مشخصی را تهیه و استفاده از آن را به کشورهای عضو توصیه کرده‌اند. طبعا " چون این طبقه‌بندیها با توجه به شرایط عمومی اجتماعی- اقتصادی و فرهنگی کشورهای مختلف تهیه میشود، ممکن است در بعضی از موارد با شرایط خاص هر کشور دقیقا " منطبق نباشد. بدین جهت توصیه شده است که سازمانهای آماری هر یک از کشورها با توجه به شرایط موجود و با حفظ چارچوب اصلی، تغییراتی را که لازم میدانند در طبقه‌بندیهای بین‌المللی بعمل آورند.

در حال حاضر چهار نوع طبقه‌بندی بین‌المللی در این ارتباط وجود دارد که هر کدام گروهی از عناوین را شامل میشود برای طبقه‌بندی فعالیتهای اقتصادی بنا به توصیه کمیسیون آماری، شورای اقتصادی و اجتماعی سازمان ملل در سال ۱۹۴۸ متن اولیه طبقه‌بندی استاندارد بین‌المللی کلیه رشته فعالیتهای اقتصادی (INTERNATION STANDARD INDUSTRIAL CLASSIFICATION OF ALL ECONOMIC ACTIVITIES) را که به اختصار (ISIC) میباشد پذیرفت و استفاده از آن را به یکی از دو طریق زیر به دولتهای عضو توصیه کرد:

الف - پذیرفتن این نظام طبقه‌بندی بعنوان یک استاندارد ملی "

ب - تنظیم مجدد اطلاعات آماری بر طبق این نظام بشرطی که قابلیت مقایسه بین‌المللی به علت

استفاده وسیع و گسترده از (ISIC) در طبقه‌بندی اطلاعات بر حسب نوع فعالیت اقتصـــادی در زمینه‌های جمعیت، تولید، اشتغال، درآمد ملی و سایر آمارهای اقتصادی، تاکنون سه بار این طبقه‌بندی مورد تجدیدنظر قرار گرفته که بار اول طی سالهای ۱۹۵۸ ـ ۱۹۵۶ و بار دوم طی سالهای ۱۹۶۸ ـ ۱۹۶۵ و سرانجام تجدید نظر سوم در سال ۱۹۹۰ طی نشریه‌ای برای استفاده به کشورهای عضو توصیه شده‌است.

در ایران نیز بمنظور طبقه‌بندی فعالیتهای اقتصادی، با اقتباس از توصیه‌های سازمـــان ملل طبقه‌بندیهای متعددی انجام گرفته است که عبارتند از:

الف: "طبقه‌بندی فعالیتهای دسته‌جمعی اقتصاد و کد مشاغل" تهیه شده از سوی وزارت کشور، اداره کل آمار عمومی، اداره محاسبات ملی.

ب: "طبقه‌بندی جدید فعالیتهای اقتصادی ایران" در سال ۱۳۴۸ از طرف وزارت اقتصاد، دفتر آمار تهیه گردید.

ج: "فهرست کدهای گروه بندی فعالیتهای اقتصادی" که توسط مرکز آمار ایران در بهمن ماه سال ۱۳۵۵ تهیه شده است.

د: "طبقه‌بندی بین‌المللی فعالیتهای اقتصادی (فهرست موضوعی و الفبائی)" که توسط مرکز آمار ایران در سال ۱۳۶۵ تدوین گردید. این طبقه‌بندی بمنظور استفاده در کدگذاری فعالیتهـــای اقتصادی کارگاههای فهرست شده و فعالیت اصلی محل کار اعضاء شاغل خانوار در سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ تهیه شده که در کدگذاری بیش از دو میلیون و پانصد هزار فعالیت اقتصـــادی کارگاههای فهرست شده و فعالیت اقتصادی خانوار طرح آمارگیـــری جاری جمعیت ۱۳۷۰ مورد بهره‌برداری قرار گرفت.

طبقه‌بندی دیگر در ارتباط با کالاهای وارداتی و صادراتی بنام "طبقه‌بندی هماهنگ کالاهای بازرگانی خارجی" (HARMONISED SYSTEM) که مختصرا" (HS) تعریف شده مـــی‌باشد و همانگونه که گفته شد بیش از ۲۵۰۰ تا ۳۰۰۰ نوع کالا را که بعضی از نظر ظاهر مشابه یکدیگر لیکن از نظر پاره‌ای مختصات از جمله خصوصیت شیمیایی (تنوع آلیاژ) کاملا" با هم متفاوت بوده و کاربردی جداگانه

دارندکه در این طبقه‌بندی با کمک کدگذاری شناسائی میشوند . سومین طبقه‌بندی تدوین شده بنام " طبقه‌بندی مرکزی کالاها وخدمات داخل کشور"(CENTRAL PRODUCT CLASSIFICATION) است که اختصارا " (CPC) گفته میشود که خود بخش اعظمی از کالاها و خدمات داخل کشور را شامل میشود . شایان ذکر است که با عنایت به پیشرفتهای روزافزون علمی پیوسته شاهد شکوفائی و ایجاد رشته‌های گوناگون در این زمینه هستیم یک روز فقط سخن از خواندن بود و روز دیگر خواندن و نوشتن و این روند با سرعت ، از ابتدائی‌ترین دانستنیها تا علوم پیچیده ریاضی ، فضائی ، پزشکی ، انسانی ، فلسفی ، عرفانی و ... را پیموده و همچنان در گستره علم پیش میرود و در هر زمان مولودی تازه تقدیم جوامع بشری میکند حال برای ضبط و بررسی و اندازه‌گیری این تنوع در مطالعات اجتماعی ، اقتصادی و ... چه باید کرد ؟ اینجا است که لزوم ایجاد یک طبقه‌بندی برای این امر لازم‌الاجراء است . چرا که در بخش فعالیتهای اقتصادی مشاهده میشود که یک رشته فعالیت به تنهائی تعداد بسیاری از مشاغل را دربر دارد ولذا در مقایسه با مشاغل موجود در هر فعالیت چنانچه بخواهیم مشاغل آن فعالیت را فهرست کرده و بهمراه فعالیت مربوطه کدگذاری کنیم برای رشته فعالیت یاد شده صرفا" یک کُد استفاده میشود لیکن برای مشاغل بسته به تنوع لزوما" از چندین کد باید استفاده شود که در اینجا نیز اهمیت نقش کد کاملا " هویدا میباشد . پس با این مقایسه چهارمین طبقه‌بندی بنام " طبقه‌بندی بین المللی مشاغل"(INTERNATION OCCUPATIONAL CLASSIFICATION) که به اختصار (IOC) میباشد ، معرفی میگردد . البته لازم به ذکر است که در ارتباط با رشته و پایه تحصیلی نیز یک طبقه‌بندی وجود دارد که برای کدگذاری در رشته و پایه تحصیلی مورد استفاده قرار میگیرد . با ایجاد طبقه‌بندی ، پدیده‌ها و عناوین و موضوعات متناظر در یک گروه قرار گرفتند . حال با توجه به کدهای انتخابی که بصورت قراردادی تعیین شد کار کدگذاری آغاز میشود . بدیهی است دیگر صحبت از یک کلمه کوچک نیست بلکه در هر طبقه عناوین و موضوعاتی مطرح است که غالبا" بیش از یک کلمه را شامل میشود مثل شماره تلفن منزل ... ، شماره خیابان ... شماره اتومبیل ... ، شماره یک قطعه از ماشین در یک سیستم پیچیده و بسیاری از این قبیل که تنها با جایگزینی یک یا چند رقم قابل شناسائی و شمارش میگردد . وقتی صحبت از

شماره تلفن منزل آقای X و یا شماره اتومبیل پیکان دو لوکس سفید رنگ مدل سال ۷۰ میشود یک عدد چند رقمی بجای یک جمله طولانی قرار گرفته و قابلیت شمارش و شناسائی پیدا میکند . مضافا" بجهت حجم کلان داده‌ها امکان استخراج بوسیله جداول مقدماتی، واسط و سرانجام نهائی ضعیف و یا غیرممکن است بهمین دلیل باید از کامپیوتر کمک گرفت که تاریخچه استفاده از کامپیوتر برمیگردد به سال ۱۸۸۰ میلادی که نشریه‌ای که توسط سازمان امکاپ منتشر شده و مبانی داده‌پردازی ماشینی برای آمارشناسان را مطرح کرده، آمده است که سال ۱۸۸۰ اداره آمار ایالات متحده دکتر هرمن هلیریت را برای ساختن دستگاه مکانیکی پردازش اطلاعات سرشماری نفوس استخدام نمود . ماشینی‌که‌او اختراع کرد بر مبنای وسایل کارت منگنه بود که بعدها" درکنترل‌ماشینهای‌بافندگی‌صنایع‌نساجی‌مورداستفاده قرار گرفت . در این سیستم اطلاعات کُدبندی میشد و کدها بوسیله سوراخهای موجود درکارتهایشان داده میشد و ماشین کارتها را به ترتیبی که از قبل معین شده بود، ردیف میکرد تا تولید مجموعه آماری را امکان‌پذیر سازد . این سیستم جدول برای اولین بار در سرشماری ۱۸۹۰ ایالات‌متحده‌امریکا مورد استفاده قرار گرفت و بنابراین اولین کاربرد وسیع پردازش خودکار اطلاعات بود. بدین‌ترتیب بجای کلمات اعداد مورد بررسی و محاسبه قرار گرفتند البته باید گفت که درهنگام جمع‌آوری اطلاعات‌بعضی از عناوین میتوانند بصورت SELF CODE باشدکه حجم فراوانی از کار مامور آمارگیر را تقلیل میدهد ولی بهر حال با توجه به تنوع بسیار زیاد داده‌های قابل کدگذاری روش استخراج دستی برای جمع‌آوری نتایج عناوین کدگذاری شده مستلزم صرف وقت زیاد و نیروی انسانی فراوان آموزش دیده میباشد که هرگز نیز خالی از خطا نخواهد بود و لذا برای پوشش خطاهای کدگذاری، بازبینی‌یا بازنگری مجدد را می‌طلبد که نزوما" صرف وقت مجدد و نیروی انسانی کار آمد مورد نیازاست اکنون برای اجتناب از صرف وقت زیاد و استفاده ازنیروی انسانی بسیار با توجه‌به پیشرفتهای ماشینی میتوان بکمک کامپیوتر (CAC) اطلاعات را از طریق (ONLINE) کدگذاری نمود .

حال قبل از پرداختن به شیوه کدگذاری با کمک کامپیوترلازم است گفته شود که امروزه "کد" در شناسائی تقسیمات کشوری نقش بسزائی دارد . بدین‌معنی که برای معرفی استانها، شهرستانها،

شهرها، دهستانها و آبادیها که هرکدام با اسامی مختلف و بعضاً " تشابه اسمی نامگذاری شده‌اند تنها با استفاده از "کد" قابل بررسی و شمارش هستند. بعنوان مثال یک آبادی از یک دهستان و بخش و شهرستان و استان مربوط به آن بوسیله یک عدد چند رقمی منحصر بفرد شناسائی میشود حتی اگر مشابهت اسمی وجود داشته باشد که هر قسمت از آن عدد چند رقمی نماینده واحدی از تقسیمات کشوری است. بدیهی است برای بررسی اطلاعات مربوط به بلوکها یا آبادیها و یا حوزه‌ها و ... از یک شهرستان یا کل کشور امکان شمارش و احتمالاً " جدول‌گیریهای خاص برای منظورهای ویژه وجود ندارد ولی با کدگذاری بکمک کامپیوتر حجم عظیمی از تمام یا قسمتی از یک کار بسرعت جمع‌آوری و قابل بررسی و تحلیل میگردد و اصولاً " در مقوله کد و کدگذاری چیزی بدون شماره معرفی نمیشود. در باب فعالیت باید گفته شود فعالیتهای متنوع موجود در ایران به ۹ بخش عمده کشاورزی، معدن، صنعت، آب و برق و گاز، ساختمان، عمده‌فروشی و خرده فروشی، حمل و نقل، خدمات مالی و خدمات عمومی طبقه‌بندی شده که هر طبقه در داخل خود به رشتهٔ فعالیت‌های دیگر تقسیم میگردد تا جائی که برای هر فعالیت یک "کد" چهار رقمی منظور میشود. این امر در مشاغل و رشته و پایه تحصیلی و ... نوعاً " صادق است.

در زمینه کدگذاری به کمک کامپیوتر COMPUTER ASSISTED CODING که به صورت خلاصه (CAC) نامیده میشود ابتدا، تاریخچه آن، سپس تعریف و بعد نحوه عمل کدگذاری بیان میگردد.

تاریخچه CAC

اکثر سازمانهای آماری در جهان برای اجتناب از مخارج سنگین عملیات کدگذاری و استخراج به‌روش دستی که نیروی انسانی فراوان و لوازم مورد نیاز و از همه مهمتر صرف وقت زیاد را میطلبد، مضافاً " مشکلات عدیده استخدامی و کمبود محدودیت در آموزش نیروها، شرایط مناسبی برای معرفی تکنولوژی طراحی کامپیوتر اتلاف وقت و نیروی انسانی و غیره را بنام CAC فراهم نمود.

در اداره سرشماری آمریکا یک سیستم کدگذاری خودکار را پیش از سرشماری اقتصادی ۱۹۶۲ طراحی کردند تا کدهای صنعتی را به جوابهای داده شده در فرم سرشماری تخصیص دهند. در این

سیستم یک فایل از کدهای صنعتی براساس جوابهای دوره‌های قبل سرشماری ساخته که عمل تطابق کد با عبارات صورت میگیرد ولی از همین فایل برای اختصاص کدها به جوابهای جدید نیز استفاده شده است. در سوئد نیز سیستمهای کدگذاری خودکار برای اصناف گوناگون آماری وغیر آماری طراحی شده است. از این سیستم در آمار سال ۱۹۷۸ محاسبات مخارج خانگی برای اجناس خریداری شده بهره‌برداری شد همچنین فیلد شغل در سرشماری نفوس ۱۹۸۰ در این کشور بشکل خودکار کدگذاری شده. ایسن روش کدگذاری در اوایل خرداد ماه سال ۱۳۶۹ در ارتباط با اسامی شهرستانهای فایلهای استانی یکی از مدارک سرشماری عمومی کشاورزی سال ۱۳۶۷ و سپس در اوایل شهریورماه سال ۱۳۶۹ برای عبارات فعالیتهای اقتصادی در فایل فرم یک پیامد زلزله استانهای گیلان و زنجان با موفقیت و در مدت زمان کوتاهی انجام گردید و اکنون پیشنهاد شده برای سرعت بخشیدن کار در امر کدگذاری ۲ فیلد کاربر شغل و فعالیت اقتصادی که در سرشماریها بخصوص بخش اعظم کار کدگذاری است، بوسیله این سیستم کدگذاری شود.

تعریف CAC

قسمتی از عملیات کدگذاری در تطابق عنوان یا عبارت با جدول یا فهرست کد و اختصاص کد مناسب که بصورت خودکار صورت میگیرد را CAC میگویند. خاطر نشان میگردد که این سیستم یک جایگزین در امر کدگذاری نیست بلکه صرفاً کمک کننده میباشد چرا که هنوز هم تصمیم نهایی کدگذار در تشخیص کد مناسب برای یک عبارت نقش عمده را دارد.

نحوه عمل کدگذاری

برای روشن شدن نحوه عمل کدگذاری در سیستم CAC ابتدا، کدگذاری به روش دستی توضیحاً بیان میشود: در روش کدگذاری دستی کدگذار برای یافتن کد مورد نظر جهت یک عنوان یا عبارت با مراجعه به یک جدول یا فهرست مجموعه کد به جستجو پرداخته، پس از پیدا کردن کد مناسب آنرا در مقابل عبارت مندرج در هر مدرک ثبت میکند که در این حالت چنانچه حجم کار کدگذاری کم باشد امکان اشتباه نیز کم خواهد بود ولی چنانچه این حجم زیاد باشد امکان اشتباه فراوان میگردد بخصوص

اگر جداول کد پیچیده و حجیم باشد، اما در سیستم CAD با استفاده از برنامه‌های داده شده عمل تطابق با عبارت یا عباراتی که باید کدگذاری شود از طریق جداول کد مربوطه انجام میگیرد. بدین ترتیب که اپراتور توسط یک دستگاه ورودی مثل KEYBOARD عبارت را به کامپیوتر داده سپس کامپیوتر با جستجو در جدول کد مربوطه (که قبلا" به کامپیوتر داده شده) تمامی عباراتی را که نزدیکترین تطابق را با عبارت ورودی داشته به همراه کد مربوط به آن لیست میکند. سپس کدگذار میتواند مناسبترین عبارت را انتخاب نموده تا عمل تطابق مجددا" انجام شود. کوشش فراوان در راستای تمام اتوماتیک کردن سیستم CAD با شبیه سازیها و مراجع گوناگون کمکی دیگر شده لیکن این کار عملیات محاسباتی بسیار را خواهد داشت.

والسلام

روش تهیهٔ طرحهای فنی-خاصه طرح فنّی هزینه و درآمد خانوار

علی‌اکبر مظهری

مرکز آمار ایران

معمولا " چند روش برای اجرای آمارگیری به کار می‌رود که یا به صورت سرشماری یا به صورت نمونه گیری ویا ترکیبی از آنهاست . دراین مقاله بحثی اجمالی راجع به نمونه گیری را ارائه می‌دهم . البته نه بحثی راجع به خود نمونه گیری وروشهای آن بلکه راجع به طرحهایی که به صورت نمونه گیری اجراء می‌شود وچگونگی کیفیت آنها واصولی که درآنها باید رعایت شود .

درتهیه هرطرح نمونه گیری که اصطلاحا " درمرکز آمار طرح فنی گفته می‌شود اصولی را باید مد نظر داشت . هیچ طرح فنی خارج از این اصول نیست واگر عمدا " یاغیر عمد بعضی از اصول رعایت نشود درهر مرحله امکان بروزاشکال وجود دارد . ناگفته نماند درهر طرح فنی علاوه بر رعایت اصول عمومی به بعضی نکات فنی مخصوص همان طرح نیز باید توجه نمود .

خلاصه این اصول عمومی به صورت زیراست :

۱ ـ مقدمه

۲ ـ اهمیت وسابقه طرح

۳ ـ هدف یا اهداف

۴ ـ شناخت عوامل محدود کننده

۵ ـ تعاریف ومفاهیم

۶ ـ اقلام آماری

۷ ـ جامعه آماری

۸ ـ واحد آماری

۹ ـ قالب آماری

۱۰ ـ زمان آماری

۱۱ ـ زمان آمارگیری

۱۲ـ روش جمع آوری اطلاعات

۱۳ـ روش تعیین تعداد نمونه

۱۴ـ صفت یا وضعیت مشخصه

۱۵ـ روش انتخاب واحدهای نمونه

۱۶ـ فرمولهای برآورد

اصولی را که ذکر شد مسلما " در تمام طرحهای فنی کم وبیش باید رعایت نمود. البته این اصول جدا از اصول تهیه خود طرح موضوعی نیست . لازم میدانم قبل از شرح اصول ذکر شده مقداری راجع به طرح هزینه ودرآمد خانوار صحبت نمایم .

دراین طرح خانوار به منزله واحد اصلی جامعه بهترین محمل برای مطالعه وضع اقتصادی، تغذیه ، اجتماعی وفرهنگی هرجامعه است . طرح هزینه ودرآمد خانوار یکی از ابزار برنامه ریزی اقتصادی واجتماعی است که برمبنای خانوار تدوین شده است . این طرح نه به صورت امروزی ، بیشتر به منظور جوابگویی به نیازهای فوق ازسال ۱۳۴۲ توسط مرکزآمارایران اجراء شده است . هرسال کوشش براین بوده که طرح با تمامات اجرایی وفنی آن نسبت به سال قبل کمتر شود واز جهت علمی وفنی تکمیل و تصحیح گردد . البته ازسال ۱۳۴۲ درنقاط روستایی وازسال ۱۳۴۷ درنقاط شهری اجراء میشود ونتایج آن نیز هرساله منتشر شده است .

قابل ذکر است که طرح شهری آن ازسال ۱۳۱۷ به بعد توسط بانک ملی وبانک مرکزی اداره آمار عمومی کاها " به صورت پراکنده دربعضی نقاط کشور ویابعضا " درکل کشور به اجراء درآمده است .

به علت ضرورت واهمیتی که نتایج این طرح برای سازمانهای مختلف دارد درسال ۱۳۶۲ـ ۱۳۶۱ به دنبال جلسات متعدد با کارشناسان ودست اندرکاران درمرکز وسایر سازمانها مخصوصا " بانک مرکزی انجام شده تغییرات زیادی روی خود طرح انجام گرفت . بنابراین طرحی که درحال حاضر به اجراء درمی‌آید دارای تغییرات جزیی نسبت به طرح سال ۱۳۶۳ـ۱۳۶۶ است .

واما اصولی که درمقدمه ذکرشد

۱ـ مقدمه : مقدمه باید حاوی انگیزه وپیشنهاد طرح باشد که شامل علتهای مختلف که ضرورت

اجرای آنرا طرح می‌نماید باشد این ضرورت وانگیزه اجرا، طرح در ابتـــداء مقاله ذکر شد .

۲ ـ اهمیت وسابقه طرح : اگـــر طرح یا طرحهای مشابهی قبـــلا " اجراء شده است : دانستـــن زمان اجراء ، روشهای فنی ، روشهای اجرایی ونتیجه بدست آمده و تجزیه وتحلیل نتایج برای راهنمایـــی بیشتر ضروری است . یکی دیگر از مسائل کـــه دانستن آن تا حدودی برای تهیه طـــرح کمک می‌کنـــد پوشش جغرافیایـــی طرحهای اجراء شده وروشهای طبقه بندی اگر بکار رفته باشد وخصوصا استخراج آنهـا است بادانستن این مسائل میتوانیم طرح را به تصویـــب واجراء درآوریم . اگر بتوانیم علت یا علتهایـــی که بـــاعث به حذف ترمیـــم طرحهـــای اجراشده را پیدا کنیم برای رهنمود راه گشایـــی بسیار مفید خواهد بود .

۳ ـ هدف یـا اهداف : معمولا " هرطرح موضوعی دارای هدف یا اهـداف کلی مشخص وتعریف شـــده است . این هدفها باید جامع ، کمی ، قابل لمس وقابل دسترسی باشد به گونه ای کـــه بعداز اجـــراء بتوان هریك از آنها را به صورت کمی وبارقـام واعداد بیان نمود . ضمنا " باید مشخص باشد کـــه اطلاعات درچه سطحی از نظر پوشش جغرافیایی قابل ارائه وانتشار است . باتوجه به این مطالب ضـــروری میدانیم برهدف کلی واهداف تفصیلی طرح هزینه ودرآمد خانوار :

هدف کلی طرح هزینه ودرآمد خانوار بررسی درآمدها ، هزینه های مصرفی یك خانوار شهری یـــا روستایی ، مقدار مصرف مواد خوراکی وتعیین سهم هزینه های مختلف درهزینه کل وتشخیص انواع درآمدهای خانوار ومطالعه تغییـــرات حاصل از آنهـــا درطول زمان میباشد واهداف تفصیلی آن بشرح زیرمیباشد :

الف ـ برآورد کلیه هزینه های مصرفی خانوارها به عنوان مواد اولیه برای حسابهای ملی واحیانـــا " منطقه‌ای

ب ـ تهیه ومحاسبه ضرایب اهمیت کالاهـــا وخدمات مصرفی خانوار برای محاسبه شاخص بهـــای کالاها وخدمات مصرفی .

ج ـ تهیه آمارهای لازم به منظور ارزیابی اثـرات اقتصاد برنامه‌ای روی سطح زندگی ویافتن معیارهای جدید برای رفاه از نقطه نظر بررسی نحوه ترکیب هزینه های مصرفی

د ـ بررسی اطلاعات آماری بدست آمده از درآمد وهزینه خانوار وآگاهی از نحوه توزیع آن درمیـــان

طبقات از نظر درآمد برای تصمیم گیری های اقتصادی رفاهی بمنظور تامین حداقل معیشت برای خانوارها وفراهم آوردن رفاه نسبی برای آنها

م - برآورد غیر مستقیم تولید بعضی از مواد مصرفی خانوار که به سهولت ومستقیما" قابل اندازه‌گیری نیست .

ن - ارزیابی ویژگیهای جمعیتی جامعه واثرات متقابل این ویژگیها وپدیده‌های اقتصادی‌واجتماعی بریکدیگر

و - یافتن یک سری روندهای مشخص درمورد عادات مصرفی خانوار وتعیین چگونگی تغییرایستن عادات وجابجایی گروههای اجتماعی دررابطه باگروههای اقتصادی واجتماعی .

هـ - آینده نگری وساختن الگوهایی برای پیش بینی ومقدار ونحوه تقاضای کالاها وخدمات

البته موارد فوق ویااصولا" هدف کلی واهداف تفصیلی درقالب جداول نهایی از قبیل یعنی درزمان تهیه طرح موضوعی حتما" باید آمده باشد یعنی هیچ یک از اهداف بدون داشتن جداول مربوطه با جداول نهایی قابل دسترسی نیست .

شاید جالب توجه باشد که‌از نتایج طرح علاوه برکارشناسان ومحققان سازمانها ونهادهای زیادی استفاده می‌نمایند ومهمترین آنها را میتوان به صورت زیرنام برد واین سازمانها عموما" هرساله خواهان نتایج طرح میباشند .

الف - ستاد بسیج اقتصادی برای تدوین مقدار، بعضی ازکالاهای سهمیه بندی

ب - بانک مرکزی جمهوری اسلامی برای تهیه حسابهای ملی وشاخص قیمت کالاها وخدمات وبررسیهای اقتصادی مورد نظر

ج - وزارت کشاورزی برای پیش بینی تولیدات و واردات بعضی از مواد غذایی .

د - استیتو خواروبار و تغذیه برای بررسی وضعیت تغذیه خانوارها درمناطق مختلف کشور .

م - وزارت مسکن وشهرسازی جهت بررسی ومطالعه رابطه بین درآمد ووضعیت محل سکونت ومسکن .

ن - وزارت ارشاد اسلامی برای بررسی وضع تسهیلات عمده محل زندگی وبخصوص وسایل ارتباط جمعی

و - وزارت بازرگانی برای بررسی مصرف بعضی ازاقلام خوراکی مانند گندم ، نان ، گوشت وغیره

هـ - مرکزآمارایران برای محاسبه شاخص قیمت کالاها وخدمات وتهیه حسابهای ملی واحیانا " منطقه‌ای

۲ - شناخت عوامل محدودکننده

اصولا " درهرطرح عوامل محدودکننده ای وجود داردکه رسیدن به بعضی اهداف را دشوار واحیانا " غیر ممکن میسازد این عوامل باید شناسایی وحدود تاثیرات آنها کاملا " مشخص باشد . معمولا " این عوامل محدود کننده را اشتباهات غیر نمونه‌ای می‌گوئیم .

درطرح هزینه ودرآمد خانوار این عوامل شناخته شده به صورت زیر وبطور خلاصه بیان می‌شود .

الف - میدانیم که روش اجراء طرح یعنی نحوه تکمیل پرسشنامه طی مدت طولانی تغییر نکرده است . البته کوششهایی دراین‌زمینه شده ولی درنهایت به نتیجه مطلوب نرسیده . روش تکمیل پرسشنامه ها مراجعه ومصاحبه مستقیم با خانوار است که دراین مصاحبه باتوجه به حجم زیاد سئوالات خانوار ممکن است برای پاسخ بعضی سئوالات حضور ذهن وآمادگی نداشته باشد واحیانا " پاسخ را فراموش نماید . این حجم زیاد سئوالات وعدم حضور ذهن پاسخگو را به سرعت خسته می‌کند واین خستگی عموما " به مامور آمارگیر منتقل میشود ورفته رفته مامور ازپرسش بعضی سئوالات یااقلام آماری خودداری می‌کند . مامورآمارگیر ممکن است براثر خستگی با تجربیات گذشته تصورات ذهنی خودرا درمورد بعضی‌از سئوالات بکار برده واقدام به عدد سازی نماید . دراینجا قصد براین نیست که کار ماموران آمارگیر را که واقعا " از صدیق ترین اقشار جامعه هستند خدای نکرده لوث کرده باشیم وبآنها را متهم نماییم . این یك واقعیتی است که درهرطرح که به صورت طولانی مدت وهرساله به اجراء درآید ممکن است اتفاق بیافتد .

ب - عدم شناسایی خانوارهای نمونه با اهداف طرح پیش فرضهایی را برای آنها تداعی می‌کند و سبب میشود که پاسخ بعضی سئوالات غیرواقعی باشد .

ج - معمولا " چارچوب مورد استفاده طرح مربوط به طرحهای سرشماری است وچون انجام سرشماریها

معمولا " توام با دشواری وهزینه بسیار است درفواصل بین سرشماریها معمولا " چارچوب تفکیکی را نشان میدهد که ممکن است دربعضی مواقع تعدادی از خانوارها به طبقات مختلف شناسایی نشوند وآمارگیری از آنها میسر نباشد یا احیانا " بلوک یا آبادی نمونه کلا " قابل دسترسی نباشد . باتوجه به گسترش نقاط شهری وایجاد بلوکهای جدید درحاشیه این نقاط معمولا " این نقاط درچارچوب وجود ندارد ونتایج برای آنها قابل تعمیم وگسترش نیست به عبارتی چارچوب بهنگامی خودرا از دست داده باشد .

د - دربعضی از بلوکها یاآبادیها که تعداد خانوار آنها کم باشد احتمال اینکه یک یاچند خانوار طی دوسال متوالی بعنوان نمونه انتخاب شود وجود دارد وتجربه نشان داده است که مراجعه مکرر به یک خانوار اثر نامطلوبی به روی خانوار نمونه میگذارد .

هـ - دربعضی از سالها یافصلها یابه علتی بعضی از استانها امکان آمارگیری از تمام خانوارهای نمونه را ندارند لذا واریانس محاسبه شده از دقت کافی برخوردار نیست .

۵-تعاریف ومفاهیم بکار رفته

کلیه تعاریف ومفاهیم که درطرح از آن استفاده میشود بایدکاملا " تعریف شده باشد . به نحوی که خواننده وآمارگیر یک برداشت یکنواخت ازآن داشته باشند . اگر تعاریف به گونه ای باشد که برداشتها ومنظور ازآن یکسان نباشد آنوقت اطلاعات بدست آمده قابل تعمیم نیست .

درطرح هزینه ودرآمد خانوار تعاریف بکاررفته اصلی بشرح زیراست :

الف - هزینه : عبارتست از ارزش پولی یا ریالی کالا یا خدمت تهیه شده توسط خانوار میباشد که شامل خرید ، تولید خانگی ، دربرابرخدمت ، از محل کسب یا پایاپای وسجانی است که به صورت پولی برآورد شده است . البته نکاتی نیز دراین جهت وجود دارد که درطرح موضوعی کاملا " تعریف شده است .

ب - مصرف : استفاده نهایی از مواد خوراکی مصرف نامیده میشود . نکات ضروری درطرح موضوعی آمده است .

ج - درآمد : کلیه وجوه کالاهایی که دربرابر کار انجام شده یا سرمایه به کار افتاده یا

از طرق ومنابع دیگر مانند حقوق ودستمزد ، حقوق بازنشستگی ، درآمدهای اتفاقی وغیره آن به صورت نقدی یا غیر نقدی که درزمان آماری به خانوار تعلق گرفته باشد درآمد محسوب میشود . نکات اضافی وموارد قابل ذکردرطرح موضوعی وجوددارد .

د - ارزش افزوده : تفاوت بین ارزش کالاها وخدمات تولید شده وارزش کالاها وخدمات بکاررفته درجریان تولید راارزش افزوده می‌گویند .

م - کار : هرنوع فعالیت فکری یابدنی است که قانونا " مجاز بوده وبمنظور کسب وایجاد درآمد صورت بگیرد . نکات قابل ذکر درطرح موضوعی آمده است .

ن - کالاهای بادوام : منظور آن دسته کالاهایی است که عمر مورد انتظار آنها بیش از یکسال بوده وارزش پولی زیادی ندارد .

و - کالاهای کم دوام : منظور آن دسته کالاهایی است که عمر مورد انتظارآنهاعموما " حدود یکسال است وارزش کمتری نسبت به کالای بادوام دارند .

ه - کالاهای بی دوام : منظور آن دسته کالاهایی است که عمر مورد انتظارآنها عموما " کمتراز یکسال است .

۶ - اقلام آماری

کلیه اقلام آماری مورد نیاز باید به نحوی مشخص شود که جمع آوری و ارائه آنها درقالب جداول نهایی و هدفهای کلی وموضوعی را به صورت کامل تامین نماید .

درطرح هزینه ودرآمد خانوار اقلام آماری درقالب جداول نهایی آمده است . این جداول بایستی به گونه ای باشد که بهترین شکل ممکن بیانگر کمی هدفهای طرح ومنطبق با اهداف تفصیلی باشند بطوریکه مراجعه کننده یا خواننده با مطالعه آنها به نتایج طرح آگاهی یابد .

۷ - جامعه آماری

جامعه آماری باید دقیقا " تعریف ومشخص شود بطوریکه هیچ گونه ابهامی درشناخت آن وجود نداشته باشد . معمولا " جامعه آماری را به صورت زیرتعریف می‌نمایند : به مجموعه افراد واشیاء که حداقل دارای یک خاصیت مشترک باشند جامعه آماری گفته میشود . درطرح هزینه ودرآمد خانوار

جامعه آماری عبارتست از :کلیه خانوارهای معمولی ساکن درنقاط شهری وروستایی ۰منظور از نقـــــــاط شهری نقاطی است که درزمان سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ دارای شهرداری بوده اند ۰نقـــاط روستایی نقاطی است که درزمان سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ فاقد شهرداری بوده‌اند ۰

منظور از خانوار معمولی ساکن : خانواری است که ازیك یا چنـد نفر تشکیل میشود که باهـــــــم دریك اقامتگاه ثابت زندگی می‌کنند وبایکدیگر هم خرج هستند ومعمولا " (نه همیشه) باهم غـــــــذا می‌خورند ۰ نکات بسیار زیادی راجع به تعریف خانوار وجود دارد که درطرح موضوعی بطور مفصل آمـــده است ۰

۸ ـ واحد آمـــاری

واحد آماری معمولا " باتوجه به تعریف جامعه آماری بطور واضح وروشن تعریف میشود بطوریکـــــه شناخت آن درزمان اجراء به آسانی توسط آمارگیر یاپرسشگر میسر باشد ۰ درطرح هزینه ودرآمـــــــــد خانوار واحد آماری عبارت است از یك خانوار معمولی ساکن درنقاط شهری یاروستایی است ۰

درتعریف جامعه آماری وواحد آماری توجه بفرمایید که خانوارهای معموله غیرساکن وخانوارهـــــــای دستجمعی جزء جامعه آماری طرح هزینه ودرآمد نیستند ۰

۹ ـ قالب آمـــاری

آن واحدهای کار یست کـــه درچارچوب بطور منظم ومشخص وجود دارد ۰ ودارای مشخصات کلـــــــی و آدرس کامل است ۰

درحال حاضر درطرح هزینه ودرآمد خانوار قالب آماری فهرست کلیه حوزه‌های سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ وفرم فهرست برداری هریك ازآن حوزه هـــاست فرم فهرست برداری آنهاشامل مشخصات جغرافیایی،تعداد خانوار ،تعداد بلوك برای حوزه های شهری وتعداد آبادی ســـــــــــــرای حوزه‌های روستایی وتعداد خانوار هریك از بلوکها یاآبادیهاست ۰

توجه شود که حوزه یك محـــــدود مشخص است که توسط مسئولین تهیه طرح موضوعی تعیین میگردد ۰

۱۰ ـ زمان آماری

دوره پرسش سئوالات درپرسشنامه هـــا است وجوابها نیز باید درهمان زمـــان نوشته شوند ۰ درطرح

هزینه درآمد خانوار زمان آماری برای اطلاعات مختلف متفاوت است و بسته به مورد روز گذشته دو روز گذشته ماه گذشته و ۱۲ ماه گذشته است .

۱۱-زمان آمارگیری

همان زمان اجرا، آمارگیری است که باتوجه به امکانات وشرایط اقلیمی ومشکلات اجرایی این زمان را برای کسب اطلاعات وتنظیم پرسشنامه‌ها درنظرمیگیرند .

درطرح هزینه ودرآمد خانوار زمان آمارگیری از اول اردیبهشت ماه هرسال شروع وتاپایان فروردین ماه سال بعد ادامه خواهد داشت .

۱۲-روش جمع آوری اطلاعات

معمولا " روش جمع آوری اطلاعات همان روش تنظیم پرسشنامه‌ها وچگونگی پرسش سئوالات است که درایران عموما " به صورت مصاحبه مستقیم با پاسخگو ست البته روشهایی دیگر نیز درست مطالعه وبررسی است که تابحال جواب خوبی نداده است مگر در مواردی که طرح به صورت موردی ودریک منطقه خواص به اجرا، درآمده است .

در طرح هزینه ودرآمد خانوار همانطور که قبلا " اشاره شد روش جمع آوری اطلاعات به صورت مصاحبه مستقیم است یعنی مامورآمارگیر مستقیما " به خانوار نمونه مراجعه می‌نماید ونسبت به تنظیم پرسشنامه ها اقدام می‌نماید .

۱۳-روش تعیین تعداد نمونه

درهر طرح فنی که تهیه میشود بایدروش تعیین تعداد نمونه کاملا " مشخص باشد معلوم باشد که برای تعیین حجم نمونه لازم از چه فرمولهایی/وباچه مقدار خطا وباچه سطح باصفاتی میزان وتعدادنمونه مورد نیاز درسطح جغرافیایی معین تعیین شده است . روش تعیین تعداد نمونه باید منطقی بریکی از روشهای علمی آماری باشد .

درطرح هزینه ودرآمد خانوار روش تعیین تعداد خانوار نمونه فعلا " بشرح زیراست . نمونه های مورد نیاز هراستان به صورت شهری یا روستایی وبطور جداگانه محاسبه میگردد . برای محاسبه آن از نتایج طرح از سالهای گذشته استفاده میشود .

درحال حاضر بااستفاده از فرمول زیر

$$n = \frac{t^2 \cdot S^2}{d^2} \quad , \quad d = \alpha \bar{y}$$

بدست می آید که درفرمولهای فوق n تعداد نمونه مورد نیاز استان ، t حدود اعتماد ۹۵٪ که بااستفاده از جدول استودنت برای این سطح عدد ۱/۹۶ استفاده میشود ، S^2 انحراف معیار صفت مورد نظر است که ازآن برای محاسبه تعداد نمونه استفاده کرده ایم ، α میزان اشتباه پیش بینی شده میانگین است که معمولا" بین ۵٪ تا ۱۰٪ درنظر گرفته میشود ، البته ناگفته نماند میزان خطا برای کلیه صفات دیگر که درزیر آمده است معمولا" درسطح استان محاسبه میگردد .

۱ ـ آرد ورشته وغلات

۲ ـ گوشت

۳ ـ سبزی وتخم پرندگان

۴ ـ چربیها وروغن مایع

۵ ـ میوه وسبزی‌ها

۶ ـ قندوشکر

۷ ـ پوشاك وکفش

۸ ـ بهداشت ودرمان

۹ ـ تحصیل وآموزش

۱۰ ـ خدمات شخصی

۱۱ ـ دخانیات

۱۲ ـ اجاره بهای مسکن

۱۳ ـ آب وسوخت وروستایی

۱۴ ـ حمل ونقل وارتباطات

۱۵ ـ کالای مصرفی بی‌دوام

۱۶ ـ هزینه خوراکی

۱۷ ـ هزینه غیرخوراکی

۱۸ ـ هزینه کل

۱۹ ـ درآدکل

۱۴ ـ صفت یا صفات مشخصه

آن صفاتی است که می‌خواهیم برآورد برای آنها ارائه نمائیم بابدست آوریم که معمولا" شامل میانگین ، واریانس ، حدود اعتماد برای استان ، کل کشور وبرای صفات مختلف میباشد البته هر پارامتر دیگری را نیز میتوانیم بافرمولهای برآورد بدست آوریم . درطرح مصرف صفت مشخصه شامل هزینه کل است که براساس اطلاعات بدست آمده ازآن تعداد نمونه را نیز برای سال بعد محاسبه می‌نمائیم ولی

همانطورکه گفته شد می‌توانیم اطلاعات مربوط به صفات دیگر را نیز محاسبه نمائیم .

توضیح یک نکته این جا ضروری است که درطرح هزینه ودرآمد خانوار بعدازاینکه برآورد برای دیگر صفات که درقسمتهای قبل گفته شد محاسبه نمودیم واز روی آنها تعداد نمونه مورد نیاز براساس آنها بدست آوریم ملاحظه شد که تعداد نمونه بدست آمده بسیار زیاد است وامکانات اجرایی هیچ یک واحدهای اجرایی درحد وتوان انجام آن تعداد نمونه نبود

درهرحال با محاسبه تعداد نمونه از روی صفت هزینه کل وباتوجه به امکانات اجرایی واحدها (استانها) مقدار α معمولا " بین ۵٪تا ۱۰٪بیانگین‌همان صفت درنظر گرفته شد .

۱۵- روش انتخاب واحدهای نمونه

منظور از روش انتخاب نمونه ها است که باید منطبق برروشی از روشهای علمی آماری باشد . در طرح هزینه ودرآمد خانوار روش انتخاب برای آبادیهای نمونه درنقاط روستایی وبلوکهای نمونه درنقاط شهری به صورت زیر بوده است . روش انتخاب دومرحله‌ای‌است مرحله اول مربوط به انتخاب بلوک یا آبادی نمونه است ومرحله دوم مربوط به انتخاب خانوار نمونه درداخل واحد نمونه مرحله اول است .

برای انتخاب بلوک یا آبادی نمونه تعداد خانوار نمونه هراستان را به عدد ۵ تقسیم می‌کنیم تا تعداد حوزه نمونه بدست آید وبعد باروش $P.P.S$ وسیستماتیک خطی آبادی یا بلوک‌نمونه‌راباستفاده از جدول اعداد تصادفی درداخل‌هرحوزه‌مشخص گردد . برای این منظور از فهرست حوزه های سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ استفاده شده است .

اطلاعات این حوزه معمولا " شامل تعداد خانوار ، تعداد آبادی یا بلوک وتعداد خانوار مربوط به آنها وجمع تجمعی خانوارهای حوزه ها میباشد .

بعد حوزه ها انتخاب شده را بطور مساوی بین چهارفصل سال تقسیم می‌کنیم

درمرحله دوم برای تعیین خانوارهای نمونه درداخل بلوکها یا آبادیها : باتوجه به فرم فهرست آنها تعداد خانوار آنها را به ۵ تقسیم می‌کنیم با فاصله تصادف بدست آید وبا استفاده از جدول اعداد تصادفی ۵ خانوار نمونه مورد نیاز از داخل آنها انتخاب می‌نمائیم . این روش انتخاب سیستماتیک دایره ای است .

توجه شود انتخاب ۵ خانوار نمونه درداخل بلوك یا آبادی نمونه یك نظریه كارشناسی است وعلاوه بر آن براساس محاسبات بسیار زیاد تقریبا" مشخص شده است كه ۵ خانوار بین خود دارای كمترین واریانس هستند، وبه همین علت است كه درفرمولهای برآورد از واریانس بین نمونه‌ها معمولا" صرف نظر میشود.

درمواردا ستثنایی كه بلوك یا آبادی انتخاب شده حداقل لازم را ندارد با اضافه كردن بلوك یا آبادی مجاور حداقل لازم را تامین می نمائیم.

۱۶- فرمولهای برآورد

ازدیگر مراحل مهم تهیه یك طرح آماری اگر به صورت نمونه گیری باشد تهیه فرمولهای برآورد است كه باید با توضیحات كامل درباره‌تعریف از علامات وسمبلهای بكاررفته وطریق محاسبه هرفرمول بكار رفته كاملا" توضیح داده شود تا استفاده كننده از نتایج وواحدهایی كه‌كار استخراج نتایج را بعهده دارند هیچگونه اشكال وابهامی برای انجام محاسبات مربوط وبرآورد هرصفت خواسته‌شده یا مورد نظر درطرح را نداشته باشند.

در طرح هزینه ودرآمد خانوار از فرمولهای زیر برای برآورد صفات مورد نظر استفاده شده است

۱ - برآورد میانگین صفت دراستان i ام

$$\bar{\bar{y}}_{i} = \frac{1}{n_i} \sum_{k=1}^{n_i} \bar{y}_{ik}$$

۲ - برآورد میانگین كل كشور صفت موردنظر

$$\bar{\bar{y}} = \sum_{i=1}^{L} W_i \bar{\bar{y}}_i \qquad W_i = \frac{N_i}{N}$$

۳ - برآورد واریانس صفت مورد نظر دراستان i ام

$$V(\bar{\bar{y}}_i) = \frac{1}{n_i(n_i-1)} \sum_{k=1}^{n_i} (\bar{y}_{ik} - \bar{\bar{y}}_i)^2$$

۴ ـ برآورد واریانس صفت موردنظر در کل کشور

$$V(\bar{y}) = \sum_{h} W_h^2 V(\bar{y}_h)$$

که در آنها :

m_h : تعداد خانوار نمونه در استان h ام است

i : اندیس آبادی یا بلوک نمونه است

h : اندیس استان است که از ۱ تا ۲۴ تغییر می‌کند و h شماره استان است

$\bar{y}_h$: میانگین صفت موردنظر در آبادی یا بلوک i ام از استان h ام که از فرمول زیر بدست می‌آید

$$\bar{y}_h = \frac{\sum_{i} y_{hi}}{m}$$

m : است تعداد خانوار نمونه در آبادی یا بلوک i ام است

y_{hi} : اندازه صفت موردنظر در استان h ام از آبادی یا بلوک i ام از خانوار j ام است

M_h : تعداد کل خانوار شهری یا روستایی در استان h ام است

M : تعداد کل خانوار شهری یا روستایی کل کشور است یعنی :

البته برای برآورد بعضی از صفات در داخل هر گروه مانند (هزینه سمت در کل و غیره) نیز محاسباتی انجام میشود که ذکر آنها به درازا خواهد کشید . علاقمندان میتوانند در مرکز اطلاعات مربوط را تهیه نمایند و فرمولهای آنرا نیز ملاحظه فرمایند .

شناخت ویژگیهای جمعیت معلول با استفاده از روشهای آماری

سیدرضا معینی

مرکز آمار ایران

تاریخچه

مطالعات مربوط به معلولیتها از دیدگاه پزشکی، روانشناسی، روانپزشکی و توانبخشی به سبك نوین درکشور قدمت چندانی ندارد. فعالیتهای آماری درزمینه معلولیتها وتولید عدد و رقم با استفاده از روشهای مدرن آماری حتی سابقه کمتری ازمطالعات فوق دارد.

این ویژگی تنها به ایران اختصاص ندارد بلکه هنوز در بسیاری دیگر ازکشورهای جهان نیز معلولان مورد شناسایی دقیق آماری قرار نگرفته‌اندوآمارجهانی یامنطقه‌ای مورد اطمینان در زمینه معلولان وجود ندارد. برای دانستن تعداد معلولان درجهان به برآوردهایی قناعت می‌شود. به عنوان مثال سازمان جهانی بهداشت نسبت معلولان جهان را بین ۷ تا ۱۰ درصد کل جمعیت می‌داند.

ازعمده‌ترین مشکلات که موجب بروز ابهام آماری درمورد معلولان شده است وجود کمبودها درتعاریف مربوط به تعاریف ومفاهیم وطبقه‌بندیهای لازم دراین زمینه است.

سازمانهای بین‌المللی درارائه تعاریف ومفاهیم وطبقه‌بندیهای استاندارد بین المللی درمورد برخی ازاقلام آماری ازسالها قبل اقدام کرده‌اند. به عنوان مثال قدمت طبقه بندی استاندارد بین‌المللی فعالیتها ومشاغل که توسط سازمان بین‌المللی کار (ILO) انجام شده بالنسبه زیاداست. ضمنا" آنکه طبقه‌بندی مزبور درچندگاه یکبار مورد تجدید نظر قرار گرفته و به روز در می‌آید. سازمان آموزشی، علمی وفرهنگی ملل متحد (یونسکو) نیز دست داشتن طبقه بندیهای تهیه، آخرین توصیه‌های خود را نیز درمورد طبقه بندی استاندارد بین المللی آموزش درسال ۱۹۷۵ ارائه کرده است. اخیرا" نیز سازمان جهانی بهداشت (WHO) برای طبقه بندی نقص عضو، ناتوانیها و معلولیتها اقدام انجام شده

زیرا تنها درسال ۱۹۸۰ است که یک طبقه بندی استاندارد بین المللی توسط سازمان مزبور ارائه شده است. (۱)

علیرغم وجود تاخیر نسبی در برپاداشتن تعاریف ومفاهیم وطبقه‌بندی درزمینه فوق ، اقدامات سازمانهای بین‌المللی دردهه ۸۰ برای شناخت آماری نقص عضو ناتوانی ومعلولیت وسعت گرفته است . طلیعه این کوششهای جهانی توسط سازمانهای بین‌المللی برای شناخت بهتر آماری افراد دارای نقص عضو ، ناتوانی ومعلولیت مربوط به سال ۱۹۸۱ است که باانتخاب آن سال به عنوان سال بین‌المللی معلولان شروع شد . همزمان بااین امر سازمان ملل متحد دربرنامه جهانی کار ، گسترش روشهای مناسبی برای جمع‌آوری و استخراج منظم داده‌های مربوط به معلولان راخواستار شد .

دربرنامه‌جهانی کار مربوط به سال ۱۹۸۲ نیز مجمع عمومی سازمان ملل متحد ، اداره آمار آن سازمان وچند مرجع دیگر جهانی ازجمله شورای اقتصادی اجتماعی ، دبیرخانه سازمان ملل متحد وسازمانهای تخصصی بین‌المللی ومنطقه‌ای را به همکاری باکشورهای درحال رشد دربرپاداشتن یا تقویت سیستم جمع‌آوری آمارعمومی یاانتخاب نمونه‌های قابل‌تعمیم درزمینه نقص عضوناتوانی‌یا معلولیتها ترغیب کرده است . علاوه برآن برتهیه دستورالعملها ومدارک فنی برای چگونگی استفاده ازخانه‌وارها درآمارگیریهای ویژه ناتوانی ومعلولیت به منظور بهبود شرایط معلولان تاکید شده است .

دربرنامه جهانی کار ، سازمان ملل متحد کوششهای درسطح ملی‌وادرزمینه‌هایی که به بهبود شرایط اجتماعی ، توانبخشی وشناخت آماری افراد دارای نقص عضو ناتوانی ومعلولیت می‌انجامد خواستار شده است وبه این منظور توصیه‌های زیر راارائه کرده است :

شرکت دادن بیشترافراد دارای نقص عضو درتصمیم‌گیری

پیشگیری‌ازنقص عضو وناتوانی‌ومعلولیت

توسعه بخشیدن به برنامه‌های توانبخشی

به تساوی درآوردن امتیازات وموقعیتهای افراد دارای نقص عضو باسایرافراد

افزایش فعالیتهای گروهی افراد دارای نقص عضو

آموزش کارکنان برای برنامه‌های ویژه ناتوانان معلولان وافراد دارای نقص عضو

فراهم آوردن موقعیتهای آموزش وپرورش عمومی برای افراد دارای نقص عضو وتولید اطلاعات

---

(1) International Classification of Impairments Disabilities and Handicaps; World Health Organization , Geneva , 1980?

درسال ۱۹۸۸ کوششهای اداره آمار سازمان ملل متحد دربرپاداشتن یک بانک اطلاعاتی به نام " بانک اطلاعاتی سازمان ملل متحد درمورد معلولان " که به طور خلاصه DISTAT خوانده می‌شود به ثمر رسید. بانک اطلاعاتی مزبور داده‌هایی درزمینه آمارهای معلولان حاصل ازآمارگیریهای عمومی ازخانوارها، سرشماریهای جمعیتی وآمارهای ثبتی مربوط به ۵۵ کشور جهان رادر شروع فعالیت خود دارا بود.

دربانک اطلاعاتی دفتر آمار سازمان ملل متحد درابتدای فعال شدن، داده‌های مربوط به نقص عضو وناتوانی ومعلولیت حاصل ازسرشماریها وآمارگیریهای نمونه‌ای ضبط شده است. داده‌های مربوط به سرشماریهای ایران در زمینه نقص عضو تازمان طرح این مقاله به بانک اطلاعاتی مزبور نرسیده است. تعداد کشورهایی که داده‌های خودرابه بانک اطلاعاتی مزبور فرستاده‌اند به شرح زیربوده است.

	تعداد کشور	زمان آماری
سرشماریها	۳۲	۱۹۷۶-۸۴
آمارگیریهای نمونه‌ای	۲۴	۱۹۷۶-۸۳
آمارهای ثبتی	۴	۱۹۷۸-۱۹۸۵
جمع	۶۳	

× × ×

هرچندکه فعالیتهای درزمینه تولیدآمارهای نقص عضو، ناتوانی ومعلولیت یک فعالیت جدید شمرده می‌شود اماقبل ازهراقدام دراین زمینه وجود تعاریف ومفاهیم یکسان وواستانداردارد اهمیت بسیار دارد.

تعریف ومفهوم معلولیت

درمورد تعریف ومفهوم معلولیت، دولتها، متخصصان، قانونگزاران، ناتوانان ونمایندگان آنها اتفاق نظر ندارند. اما تعاریف سازمان جهانی بهداشت به شرح زیراست:

- نقص عضو Impairment شامل هرنوع فقدان عضویاوضع غیرطبیعی روانی یاجسمی یاساختمانی بدن یاوظایف آن است. نقص عضو به تنهایی در سطح اندام اطلاق می‌شودکه شامل نقص یافقدان اندام، عضو یاقسمتی ازساختمان بدن است. نقص وضایعات روانی نیز دراین زمره است.

مثال دراین مورد درسرشماریها وآمارگیریهای نمونه‌ای شامل نابینایی، ناشنوایی، نابینایی یک چشم، فلج اندام، قطع اندام، عقب افتادگی ذهنی، ضعف بینایی، ازدست دادن قدرت تکلم ولالی است.

ـ ناتوانی Disability عبارت است ازعدم توانایی انجام فعالیت ناشی از نقص عضو درحدی که برای یک انسان عادی، طبیعی به شمارمی‌آید. ناتوانی شامل محدودیت انجام وظایف متاثر از نقص عضو است. ناتوانی به تعبیری نارسایی وظایف الاعضای فرداست. نمونه پرسشهای ناتوانی درسرشماریها وآمارگیریهای نمونه‌ای شامل: اشکال در بینایی، صحبت کردن، شنیدن، حرکت کردن، ازپله بالارفتن، گرفتن، درازکردن دست، حمام کردن و خوردن است.

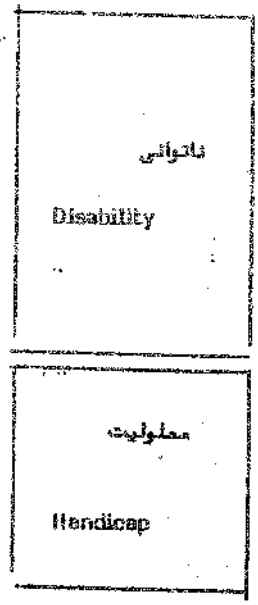
ـ معلولیت Handicap عبارت از وضع نامناسبی است که برای فردی دراثر یک نقیصه یاناتوانی پیش آمده واورا ازانجام نقش طبیعی خود دراقتضای باوضع سنی جنسی اجتماعی فرهنگی‌اش بازمی‌دارد. معلولیت به تعبیری یک نارسایی درایفای نقش اجتماعی اقتصادی فرد دارای نقص عضو یاناتوان است که اورازسایر افراد متمایز می‌سازد. این نارسائی‌ها در رویارویی فرد بامحیط‌ها یا فرهنگهای ویژه حاصل می‌شود. مثال برای معلولیت که درسرشماریها وآمارگیریهای نمونه‌ای یاآمارهای ثبتی مورد پرسش قرار می‌گیرد شامل افراد فاقد "بستری، زمین‌گیر، ناتوان برای استفاده از وسایط نقلیه عمومی، قادر نبودن برای اجرای شغل، کمکاربودن ونیز منزوی‌بودن است. (نمودار ۱)

بنابر تعاریف فوق درسرشماریهاوآمارگیریهای نمونه‌ای وآمارهای ثبتی به تشخیص افراد دارای نقص عضویا افراد ناتوان پرداخته می‌شود درحالی که معلولیت وعلیل بودن به تعبیر سازمان جهانی بهداشت درسرشماریها وآمارگیریهاوآمارهای ثبتی به عنوان مفهومی برای تمیز افرادداری نقص عضو به کار نمی‌رود بلکه صرفا" به عنوان ابزار تحلیلی برای تشخیص موقعیت اجتماعی اقتصادی افراد دارای نقص عضو وناتوان است.

نمودار ۱ـ نقص عضو، ناتوانی ومعلولیت وانواع آن

نقص عضو
Impairment

- نابینایی دوچشم
- نابینایی یک چشم
- ضعف بینایی
- ناشنوایی
- ازدست دادن قدرت تکلم
- لالی
- فلج اندام
- قطع اندام
- عقب‌ماندگی ذهنی



روشهای جمع‌آوری داده‌های مربوط به نقص عضو، ناتوانی و معلولیت

کشورهای جهان به کمک سه روش عمده داده‌های مربوط به نقص عضو، ناتوانی و معلولیت‌ها را جمع آوری می‌کنند. این روشها عبارتند از:

سرشماریهای جمعیت و خانوار

آمارگیریهای نمونه‌ای از خانوار

ثبت آمارهای نقص عضو، ناتوانی و معلولیت

الف- سرشماریهای جمعیت و خانوار

ازمنابع عمده سرشماریهای جمعیتی پوشش کامل جمعیت کشور است. اگر از خانوار در مورد نقص عضو، ناتوانی یا معلولیت در سرشماری پرسیده شود، قصد بر آن است که به وجود یا عدم اعضایی از خانوار که دارای این صفات هستند، پی برده شود. گرچه این امر مستلزم حجم زیادی از اطلاعات است اما ۸۵ درصد از کشورهایی که در بانک اطلاعاتی اداره آمار سازمان ملل متحد دارای اطلاع هستند از سرشماری جمعیت برای جمع‌آوری داده‌های نقص عضو، ناتوانی و معلولیت استفاده کرده‌اند مثل کویت در سال ۱۹۸۰.

در سرشماریهای جمعیتی پرسشهای در این زمینه را که مربوط به هریک از اعضای خانوار است می‌توان از ۱۰۰ درصد از خانوارها یا در صورت تمایل از ۱۰ یا ۲۰ درصد از خانوارها پرسید که حالت دوم وضعی است که از ترکیب سرشماری و نمونه‌گیری حاصل می‌شود.

دربرخی‌از کشورهاپرسشهای مربوط به نقص عضو ناتوانی ومعلولیت تنها بخش کوچکی ازپرسشهای سرشماری را به خود اختصاص می‌دهد. دراین حالتها کل جامعه آماری دارای نقص عضو یاناتوان ویامعلول آمارگیری نمی‌شود. به عنوان مثال دربرخی ازکشورها در پاسخ به علل غیرفعال بودن ازنظر اقتصادی درکنار افراد خانه‌دار ومحصل وبازنشسته، معلولان رانیز گنجانده‌اند مثل برمه درسال ۱۹۸۳، ونزوئلا درسال ۱۹۸۱. دربرخی دیگر ازکشورها، معلولیت وناتوانی در شمار علل به مدرسه نرفتن آمده است مثل مکزیک درسال ۱۹۸۰ وکوبا درسال ۱۹۸۱. بدیهی است دراین دوحالت از کل افراد ناتوان یا معلول آماری به دست نخواهد آمد.

محاسن گنجاندن پرسشهای نقص عضو ناتوانی ومعلولیت درسرشماریها

۱ـ ارزش سرشماریها داده‌های مربوط به نقص عضوناتوانیها ومعلولیت در حوزه‌های کوچک جغرافیایی قابل حصول است.

۲ـ نسبتها درحوزه‌های کوچک جغرافیایی قابل محاسبه است زیرا اطلاع کل جمعیت ازآن حوزه‌ها درقسمت است.

۳ـ جداول تفصیلی باخطای نمونه مواجه نیست. امکان مطالعه بندی موارد نادرو تاکید برآنها وجود دارد.

۴ـ امکان مقایسه نسبتها در سریهای زمانی درصورت تغییرنکردن پرسشها وجود دارد.

۵ـ تعداد ناتوانان آمارگیری شده زیاداست وجداول‌تفصیلی با اطلاعات جزئی درمورد آنها به دست می‌آید.

۶ـ چارچوب مفیدی برای آمارگیریهای نمونه‌ای ازنقص عضو، ناتوانی ومعلولیت را به دست می‌دهد.

معایب گنجاندن پرسشهای مربوط به نقص عضو ناتوانی ومعلولیت در سرشماریها

۱ـ فواصل سرشماریها ازیکدیگر زیادومعمولاً ۱۰ سال است وداده‌ها دیرهنگام می‌شود.

۲ـ سرشماریها به داده‌های پایه‌ای اقتصادی اجتماعی می‌پردازد و پرسشهای زرفانگر درمورد ناتوانی نمی‌تواند مطرح شود.

۳ـ در سرشماریها نسبت بدون پاسخ برای پرسشهای نقص عضو وناتوانی بالااست وکم شماری هم درسطح بالایی است زیرا برای پرسشهای نقص عضو ناتوانی ومعلولیت، مشکلات وحساسیت برای تحقیق ابزار محدود وجود دارد.

۴ـ سرشماری فعالیتی پرهزینه است و نیاز به صرف وقت زیادی برای طراحی دارد. برای پیداکردن نسبت کمی ازجامعه که دارای صفت موردنظر هستند ازبیش از ۹۰ درصد از جمعیت پرسشی مطرح می‌شودکه به

پاسخ منفی روبرواست وازآن پرسش قصد برآن است که بخش کوچکی ازجمعیت که ناتوان است مشخص شود.

ب- باتوجه به انبوه اطلاعاتی که طی آموزش باید به ماموران سرشماری منتقل شود احتمال دارد مقدار آموزش کسب شده درمورد نقص عضو، ناتوانی ومسئولیت محدود باشد درحالی که این تحقیق نیاز به آموزش تفصیلی دارد.

ب- آمارگیریهای نمونه‌ای ازخانوار

درآمارگیریهای نمونه‌ای برخلاف سرشماریها قصد برآمارگیری ازکلیه خانوارها وافراد نیست بلکه هدف عمده آن است که نمونه‌ها نماینده واقعی جامعه باشند. حتی اگر نسبت نمونه به یک صدم، یک‌هزارم یاکمترازآن برسد، انتظار ازکاربردتکنیکهای نمونه‌گیری این است که پاسخهای خانوار یا فرد انتخاب شده ازنظرآماری به منزله پاسخ گرفتن ازیکایک خانوارها یافرد فرد جمعیت باشد.

دربرخی‌ازکشورها پرسشهای مربوط به نقص عضو، ناتوانی ومعلولیت را درآمارگیریهای نمونه‌ای ویژه نیروی کار، بهداشت ودرمان وآمارگیریهایی درزمینه شرایط‌زندگی مطرح می‌کنند. دراین حالت ازطریق طرحهای فوق دریچه‌ای برای پی‌بردن به داده‌های نقص عضو ناتوانی و معلولیت ایجاد می‌شود. به عنوان مثال درآمارگیری نیروی کار کانادا درسال ۱۹۸۱ ودرآمارگیری بهداشت و درمان مصر درسال ۱۹۷۹ ونیز درسال ۱۹۸۱ درهمان کشور پرسشهای نقص عضو ناتوانی ومعلولیت گنجانده شده است. دراین نوع آمارگیریهای نمونه‌ای بابافته شدن فرد دارای نقص عضو، ناتوان یامعلول، مطلب به صورت ژرفانگر مورد بررسی قرار می‌گیرد برخلاف پرسشهای مختصری که دراین باره درسرشماریها وجود دارد. تعداد زیادی پرسش از خانوار یافرد مطرح می‌شود. دراین گونه طرحها این امتیازوجود داردکه به برخی ازاعضای گروه پرسشگر آموزش ویژه‌ای در مورد مصاحبه بانمونه‌های دارای نقص عضو، ناتوانی ومسئولیت داده می‌شود.

دربرخی ازآمارگیریهای نمونه‌ای مصاحبه گر برای پرسشهای مربوط به نقص عضو، ناتوانی ومسئولیت یک آمارگیر ساده است. دربرخی دیگر یک متخصص اجتماعی یا پیراپزشک است که برای آمارگیری آموزش خاصی دیده است. دربارهای ازحالات یک عضو گروه پزشکی مثلا" یک پرستار یا حتی پزشک آموزش‌دیده‌مصاحبه را انجام می‌دهد. درحالت اخیر که آمارگیران ازتخصص بالا برخوردار هستند هزینه‌های عملیات اجرایی سنگین تر می‌شود.

محاسن وکارگیری روش آمارگیری نمونه‌ای در بررسیهای ویژه نقص عضو، ناتوانی ومسئولیت

امکان‌دردست داشتن یک چارچوب آماری به‌اجرا درآوردن یک آمارگیری نمونه‌ای درمقایسه با

سرشماری ساده تر است .

۲- درآمارگیری نمونه‌ای می‌توان با انضباط بیشتری به بررسی ژرفانگر در عناوین مورد مطالعه پرداخت .

۳- درآمارگیریهای نمونه‌ای درمقایسه باسرشماریها، پوشش جغرافیایی وتعداد مصاحبه محدودتر است، بنابراین امکان کنترل بیشتری بر شرایط مشاهده آماری ومصاحبه‌ها وجود دارد .

۴- برای نظارت برعملیات اجرایی وآموزش خاص عملیات اجرایی وبازبینی مقدماتی پرسشنامه‌ها فرصت بیشتری وجود دارد .

۵- طراحی ممکن است به نحوی انجام شود که تاکیدزیادی برپافتن افراد ناتوان وجود داشته‌باشد . به‌عنوان مثال ازکل جمعیت درخوشه‌های نمونه فهرست برداری شـــود وبااستفاده ازپرسشهـــــــای متعدد، افراد ناتوان درخوشه نمونه پیداشوند .

۶- باتوجه به آنکه اطلاع ازکل جمعیت در قسمتهای نمونه پرسش می‌شود ، محاسبه نسبتهـــــا ودرصدها امکان پذیر است .

۷- اگر در طرح تحقیق موارد قابل مقایسه گنجانده شده باشد، برای مقایسه در سریهای زمانـــی ونیز مقایسه با داده‌های سرشماری مفید است .

معایب بکارگیری روش آمارگیری نمونه‌ای در بررسیهای ویژه نقص عضو،ناتوانی ومعلولیت

۱- بررسیهای ژرفانگر ازاین قبیل نیاز به آموزش زیاد برای عملیات اجرایی، مجریان وآمارگیـــران دوزمینه ناتوانی ومعلولیت دارد .

۲- باتوجه به اینکه تعداد ناتوانان در جامعه احتمالا " حداکثر از ۲۰ درصد جمعیت تجاوز نمی‌کند لذا اندازه نمونه محدوداست . به این علت باید به بررسی گسترده‌گی زیادی داد زیراکسانی که ناتوان‌هستـــــد تعداد کمی ازجمعیت را تشکیل می‌دهند .

۳- احتمال پوشش ناقص برخی ازخانوارهای دسته‌جمعی مثل مراکز نگهداری معلولین، خانه‌المتدان وجمعیت بی‌خانمان کم است .

۴- محاسبه نسبتها ودرصدها برای تقسیمات جغرافیایی کوچک بامحدودیت روبروست . تعـــــداد محدود نمونه وبالارفتن اشتباه نمونه گیری درتقسیمات کوچک ازموجبات این امر است .

۵- مقایسه سریهای زمانی بالمونه‌گیریهای کم دقت اطمینانی ندارد .

۶- روش ثبت آمارهای نقص عضو ،ناتوانی ومعلولیت

آمارهای‌ثبتی جمعیتی مثل ثبت ولادت ، آمارهای ثبتی تامین اجتماعی،آمارهای ثبتی بهداشتـــی ،

آمارهای ثبتی صنعتی مربوط به مجروحان ناشی از شغل هرکدام بالقوه می‌توانند به موازات سرشمار یها وآمارگیریهای نمونه‌ای برای جمع‌آوری آمارمربوط به نقص عضو، ناتوانی ومعلولیت به کار روند. بررسی کردن موسسه‌ای که در سطح ملی به ثبت داده‌های مربوط به معلولیتها بپردازد نیز راه دیگری برای ایجاد داده‌های آماری است.

برای ثبت داده‌های مربوط به ناتوانی، به صورت کلی به دوگونه اقدام می‌توان دست زد. حالت اول فراخوانی درسطح ملی ازافراد دارای نقص عضو، ناتوان یامعلول یاخانوارآنها است که طی آن برخی از داده‌ها به ثبت درآیدودرمراحل بعدی، خانوار برای کسب اطلاعات بیشتر مورد مصاحبه قرارگیرد. این نوع فراخوانی به صورت مراجعه اعضای خانوار دارای فرد ناتوان به مراکز ثبت،درکشور اردن درسال ۱۹۷۹ ودرلبنان درسال ۱۹۸۰ انجام شده است. سپس افراد مزبور بااستفاده ازپرسشنامه ای مورد آمارگیری قرار گرفتند.

درحالت دوم ازمعتمدان محل خواسته می‌شود که افراد دارای نقص عضو، ناتوان یا افراد معلولی راکه می‌شناسند گزارش کنند. سپس همان معتمدان محلی، والدین یا نگهدارندگان ناتوانان مورد مصاحبه قرار می‌گیرند. مثل آنچه در اتیوپی درسال ۱۹۸۱ انجام شد.

درفراخوانی‌های عمومی نبایدازتاریخ وساعتی که برای مراجعه مشخص شده تخطی‌کرد.

حالت دیگری هم وجود داردکه ازتلفیق آمارهای ثبتی وآمارگیری حاصل می‌شود. به این ترتیب کلیه افرادی‌که داده‌های آنها به صورت عمومی ثبت شده است بایک نمونه‌گیری مورد پرسش قرار می‌گیرند. در اطریش آلمان فدرال وسنگاپور ازاین روش استفاده شده است.

محاسن استفاده از روش ثبت داده‌های مربوط به نقص‌عضو، ناتوانی ومعلولیت

۱ـ برخلاف سرشماریها وآمارگیریهای نمونه‌ای به تعداد زیادی پرسنل که شناختن همه آنها نیز دشوار است،نیازندارد.

۲ـ جدول‌گیری براساس آمارهای ثبتی برای تقسیمات کوچک جغرافیایی امکان دارد.

۳ـ جمع‌آوری داده‌هامی‌تواند به آسانی برای ارائه خدمات ویژه به کار رود.

۴ـ درمحدوده آمارهای ثبتی سریهای زمانی کوتاه مدت وبلند مدت به آسانی به دست می‌آید.

۵ـ برای محاسبه میزان وقوع وسیزان‌فراوانی صورت کسر را به دست می‌دهد.

۶ـ چارچوب نمونه‌گیری برای مطالعات ژرفانگر درمورد افراد ناتوان ویژه مثلاً" دارای یکنوع ناتوانی خاص مثل ناشنوایی رافراهم می‌کند.

معایب کاربرد روش ثبتی درجمع‌آوری آمارهای نقص عضو، ناتوانی ومعلولیت

۱ـ سازماندهی ثبت آمار نقص عضو ناتوانی ومعلولیت نیاز به پرسنل آموزش گرفته آماری وبهداشتی

پزشکی دارد وتهیه اطلاعات در طول زمان مشکل است .

۲ـ ثبت ناتوانی درهنگامی است که تشخیص طبی داده شده ونه درهنگامی که ناتوانی رخ داده‌است .

۳ـ میزان پوشش مشخص نیست . به عبارت دیگر نمی‌دانیم تعداد ثبت شده، نماینده چه‌تعـــداد ازجامعه ناتوانان هستند . خطای نمونه هم قابل برآورد نیست .

۴ـ امکان برآورد کل جمعیت ناتوان با اطلاعات دردست، سوال برانگیز است .

۵ـ باتوجه به اینکه ممکن است اطلاع دهندگان متعدد باشند وفهرستهای مختلفی ازناتوانـــان تهیه می‌شود، امکان دوباره شماری وچندباره شماری وجود داردو تشخیص دادن وپیداکردن دوبـــار ثبت شده‌ها مشکل است .

۶ـ کل جمعیت مورد بررسی که می‌تواند درخطر معلولیت قرار گیرد معلوم نیست، بنابرایـــن نمی‌توان میزانها ونسبتها را محاسبه کرد .

۷ـ سیستم‌های ثبتی درمحتوی وروش انعطاف کمتری دارد .

ارزشیابی سه روشی که درفوق مورد بررسی قرارگرفت وتحلیل مزایا ومعایب هرکدام از روشهـــا به این نتیجه می‌انجامدکه روش آمارگیریهای نمونه‌ای می‌تواند داده‌های باقابلیت اعتماد بیشتری در مـــورد نقص عضو، ناتوانی ومعلولیت را به دست دهد . بنابراین توصیه می‌شودکه علاوه برگنجاندن پرسشهـــای مربوطبه نقص عضو، ناتوانی ومعلولیت درسرشماریهای جمعیتی، آمارگیریهای نمونه‌ای ویژه ایـــن موضوع درسطح کشور در فاصله سرشماریها نیز به‌کوشش سازمانهای مربوطه به عمل آید تا به شناخـــت ویژگیهای جمعیت معلول ازراه روشهای آماری کمک کند .

سرشماریهای نفوس درکشور و داده‌های نقص عضو

باتوجه به مزایاومعایب هریک ازسه روش سرشماری، نمونه‌گیری وآمارهای ثبتی وامکانات موجـود، بسیاری ازکشورهای جهان به جمع‌آوری داده‌های آماری درمورد نقص عضو، ناتوانی ومعلولیـــــت پرداخته‌اند . درکشور مانیز سابقه جمع آوری آمارهای مربوط به معلولیت که به تدریج تعاریف ومفاهیم سازمان جهانی بهداشت زیرعنوان نقص یا قطع عضو قرار می‌گیرد به سال ۱۳۵۵ برمی‌گردد کـــــه در سرشماری جمعیت درآن سال برای اولین بار پرسشهایی دراین زمینه مطرح شده است . درسرشماری سال ۱۳۶۵ نیز این اطلاعات جمع‌آوری شده است . درهردوسرشماری مورد اشاره نقص عضو یا قطع عضـو بارز مورد نظر بوده است . زیرا جمع آوری اطلاعات درمورد انواع ناتوانیهایی که نیاز به دقـــت در تشخیص فنی به کمک وسایل پزشکی دارد ازعهده عملیات اجرایی سرشماری عمومی جمعیت ساختـــه نیست .

در سرشماری نفوس ومسکن مهر ۱۳۶۵ افرادی معلول شناخته شده‌اند که نابینا، ناشنوایعنی کرولال یا کرباشند، ونیز افرادی که دارای قطع دست یا پا یاهردو بوده وهمچنین افراد دارای نقص دست چپ، نقص پای راست، نقص پای چپ و فلج کامل معلول به حساب آمده‌اند، باتوجه به این تعریف کمتر از یک درصد ازکل جمعیت کشور یکی ازشاخص جسمی فوق راداشته‌اند، این نسبت باآمار به دست آمده ازسایرکشورها قابل مقایسه است، (جدول ۱)،

جدول ۱ـ نسبت افراد دارای ناتوانی به کل جمعیت درکشورهای منتخب

نام کشور	روش آماری	سال	درصد ناتوانان: مرد	زن	جمعیت مورد بررسی
ایران	سرشماری	۱۹۸۶	۱/۱	۰/۲	کلیه سنین
		۱۹۷۶	۰/۹	۰/۸	" "
بحرین	"	۱۹۸۱	۱/۱	۰/۹	" "
مصر	"	۱۹۷۶	۰/۴	۰/۲	" "
هنگ کنگ	"	۱۹۸۱			" "
اندونزی	"	۱۹۸۰			" "
کویت	"	۱۹۸۰	۰/۵	۰/۴	" "
مالی	"	۱۹۷۶	۳/۱	۳/۰	" "
پاکستان	"	۱۹۸۱	۰/۴	۰/۵	" "
پاناما	"	۱۹۸۰	۰/۸	۰/۶	۰/۴۹ ۷۹ سال
پرو	"	۱۹۸۱			کلیه سنین
لهستان	"	۱۹۷۸			" "
سریلانکا	"	۱۹۸۱	۰/۶	۰/۴	" "
تونس	"	۱۹۸۴	۱/۱	۰/۷	" "
ترکیه	"	۱۹۷۵	۱/۲	۱/۴	" "
ایالات متحده	"	۱۹۸۰	۹/۰	۸/۰	۱۶ـ۶۴
					جمعیت غیرفعال ازنظر اقتصادی ۱
مکزیک	"	۱۹۸۰	۲/۵۰	۲/۳	۱۵+
یوگسلاوی (سابق)	"	۱۹۸۲	۰/۲	۰/۳	۱۰+
جمهوری فدرال آلمان	"	۱۹۷۵	۱/۲	۱/۰	۱۰+
کانادا	"	۱۹۸۱	۳/۲	۰/۸	۱۵+
فرانسه	"	۱۹۸۱	۴/۲	۲/۶	۱۵+
مکزیک	"	۱۹۸۰	۲/۹	۲/۷	۱۵+
فیلیپین	"	۱۹۸۰			۱۵+
اسپانیا	"	۱۹۸۱	۶/۳	۴/۸	کلیه سنین
ونزوئلا	"	۱۹۸۱	۸/۴	۴/۰	۱۲+

ضمیمه جدول ۱ـ

فرصت ناتوانان

نام کشور	روش آماری	سال	مردان	مرد	زن	جمعیت مورد بررسی
استرالیا	آمارگیری نمونه‌ای بازماندگان	۱۹۸۱	۱۳/۲			کلیه سنین
اتریش	"	۱۹۷۶	۲۰/۹	۱۹/۹	۲۱/۹	" "
کانادا	"	۱۹۸۶	۱۳/۲	۱۲/۲	۱۴/۸	" "
چین	"	۱۹۸۲	۳/۹			" "
آلمان فدرال	"	۱۹۸۳		۱۱/۸	۹/۸	" "
ژاپن	"	۱۹۸۰	۲/۴			۱۸۰
نپال	"	۱۹۸۰	۳/۰			کلیه سنین
فیلیپین	"	۱۹۸۰	۴/۴	۵/۱	۳/۷	" "
اسپانیا	"	۱۹۸۶	۱۵/۰	۱۴/۸	۱۵/۷	" "
یوگسلاوی	"	۱۹۸۵-۸۶	۲/۱	۳/۷	۲/۶	۱۵ـ۰۰ ساله
یوگسلاوی	"	۱۹۸۵-۸۶	۱۴/۲	۱۲/۱	۱۶/۱	۱۶+

۱ - سرشماری عمومی نفوس و مسکن سهر ماه ۱۳۶۵ ، نتایج تفصیلی معلولیت کل کشور ۶۵ - ۱۳
تهران ، مرکز آمار ایران ، اردیبهشت ۱۳۷۱ .

- "International Classification of Impairments, Disabilities and Handicaps", World Health Organization, Geneva 1980.
- "Development of Statistics of Disabled Persons," Case Studies Statistical Papers, series Y, No 2 United Nations Publications, Sales no 86 XVII, 17. New York 1986
- "Disability Statistics Compendium," Statistics on Special Population Groups Series Y No4 United Nations Publications, Sales: No E 90 XVII 17 New York 1990

مسائل آموزش آمار برای دانشجویان علوم کشاورزی

مسعود منصوری

دانشگاه ارومیه

علم آمار چراغی است فرا راه حقیقت‌پژوهان و دانشجویان علوم. این علم ایشان را یاری می‌کند تا مجموعه‌ای از گفتارها را که بر مبنای قواعد منطقی با یکدیگر در ارتباط بوده، انتظام داده که بخشی از واقعیتها را تبیین کنند. در این راستا با بررسی و آزمایش تجربی متین (EXPLANDUM) در عمل (تطبیق تئوری با واقعیت)، علم آمار رسالت خود را نشان می‌دهد. تجربیات تدریس آمار در مقاطع آموزش عالی در علوم تجربی به‌طور اعمّ و علوم کشاورزی به‌طور اخص حاکی است:

- دانشجوی نوعی: آمار را به‌طور مجرد فرا می‌گیرد، ماهیّت آن را بدان گونه که باید باشد نمی‌شناسد و در نتیجه در کاربرد آن چه در آمار توصیفی و چه در آمار استنباطی با مشکل مواجه می‌شود.
- مدرّس نوعی: یا در قالبهای ریاضی (آمار ریاضی) است که در این حالت خویشتن را با تفکر علمی-تخصصی دانشجو نمی‌آمیزد و در نتیجه تصویری یک بعدی از مسئله را به دانشجو ارائه می‌دهد. مدرّس نوعی در قالب متخصص شاخه‌ای از علوم کشاورزی معمولاً خود را با آموزش آمار مشغول نمی‌کند و اگر غیر از این باشد بعضاً ناقد بینش همه‌سونگر از قلمرو دانش آمار در متنظمات علوم کشاورزی به‌طور عام می‌باشد. در حالت خاص مدرّس متخصص علوم کشاورزی با بصیرت و معرفت کامل از دانش آمار در تطابق اهداف آموزشی دانش آمار بعنوان راهگشای علم منظور، مورد توجیه است.
- کتاب و منبع مورد استفاده که دانشجوی علوم کشاورزی را از آغاز تا انتها و جزء به جزء شاخه‌های آن را هدف‌دار تغذیه کند، حکم کیمیا دارد.

در این مقاله آموزش آمار برای علوم کشاورزی از دیدگاه دانشجو، مدرّس و مآخذ، مورد ارزیابی قرار می‌گیرد و در بخش پایانی مقاله خطوط اساسی رسیدن به وضع مطلوب ترسیم و ارائه می‌شود.

نظارت و کنترل کیفی در طرحهای بزرگ آماری

حسین منطقی
مرکز آمار ایران

مقدمه

معمولا" درکنار تشکیلات اجرایی هرسازمان ، موسسه نهاد و ... ، تشکیلات دیگری نیز به منظور سازماندهی نظارت طراحی می‌شود . این سازماندهی مسئولان ومدیران سازمانها را برای انجام مطلوب وظایف سازمان ، درقالب زمانهای تعیین شده برنامه‌های نظارت وبااستفاده از ابزارهای نظارتی یاری می‌دهد . اعمال نظارت درحین اجرای پروژه‌ها تصاویرواقعی از چگونگی روند پیشرفت کمی وکیفی امور ، عقب افتادگیهای احتمالی از برنامه ، مسائل بازدارنده و ... باخبر ساخته و بدینوسیله ابزار تصمیم‌گیری برای مدیریت سازمان را درهرموردوبرحسب مقتضیات فراهم می‌سازد .

تشکیلات و سازمان اجرایی طرحهای بزرگ آماری باتوجه به حجم و تنوع کاروبحث زمان محدود وتعدادنیروی انسانی زیاد ، مستلزم طراحی سازمان کارآمد بوده ونتیجتا" برای کنترل کمی وکیفی کار نیازبه سازماندهی نظارت ویژه دارد .

نظارت دراینگونه سازمانهای موقت بانظارت درسازمانهای دائمی تفاوت فاحش دارد . درسازمانهای دائمی بعلت داشتن فرصت کافی نظارت به صورت بازرسی وبا زدید ازقسمتهای مختلف سازمان انجام می‌گیرد ومعمولا" گزارشها به صورت مکتوب به مدیران ذیصلاح ارائه می‌شود . نظربه اینکه پروژه‌ها بلندمدت و وظایف مستمربوده وکارکنان دائمی می‌باشند با فرصت کافی می‌توان نسبت به ارزشیابی کارآنها اقدام نمودو برای بهبود کیفیت کار ، از طریق ارتقاء دانش آنان برنامه‌های

آموزشی کوتاه‌مدت ویلندمدت تهیه‌وتنظیم نمود . درسازمانهای موقت آمارگیریها امکان هرگونه مانور و اصلاح ، تغییر و تصحیح کاملا " محدود بوده و زمان اجرا کوتاه تر ازآن است که بتوان به روش سازمانهای دائمی درجهت کارآیی سازمان وانجام وظایف تعمق وچاره اندیشی کرد .

بدین جهت درسازمانهای موقت ، ازجمله درآمارگیریهای بزرگ ، سازمان نظارت به گونه‌ای تنظیم می‌شود که حتی الامکان درآن وظایف وحیطه اختیارات نظارتی هر رده سازمان اجرایی کاملا " تعریف و توجیه شده است . درنتیجه ، اعمال نظارت از طریق سلسله مراتب سازمانی انجام می‌گیرد و هر رده سرپرستی بلافصل ماموراًمارگیری ، علاوه بروظایف مترتب برپست سازمانی خود ، وظایف نظارت بر رده پایین تر ازخود را نیزبه عهده داشته باشد . بدین ترتیب غالبا " نظارت بانوعی کنترل کیفی ، به‌منظور ارتقاء کیفیت اطلاعات جمع‌آوری شده وکنترل گردش کار درسازمان اجرایی آمارگیری همراه است .

بدین لحاظ ، نظارت را نمی‌توان صرفا " ازدیدگاه بازرسی مطرح نمود ، بلکه لازم است نظارت دراینجا ، جامع‌تر از بازرسی تعریف شود . به‌عبارت دیگر دراین سازماندهی ، نظارت ، به صورت مجموعه فعالیتهای مربوط به‌کنترل کیفی ، بازرسی ، بازبینی ، مراقبت ، نقد ، اصلاح ، تهیه گزارش و بالاخره تایید نحوه کار ، تعریف می‌شود . از طرف دیگر سازماندهی نظارت مستلزم دراختیارداشتن عوامل وابزارهای زیراست :

- نیروی انسانی مجرب وآموزش دیده
- بودجه کافی
- وسایل حمل ونقل

که اگر بخواهد جدا از سازمان اجرایی طراحی شود ، درصد قابل توجهی از عوامل فوق را به طرح کلی آمارگیریهای بزرگ تحمیل می‌کند ، که تامین آنها نه تنها مشکل ، بلکه دربرخی موارد غیرممکن می‌نماید .

دراین سخنرانی ، طرح نظارت درآمارگیری جاری جمعیت ۷۰ که نسبت به چهار سرشماری جمعیتی اجراشده درایران ویژگیهای متفاوت وگسترده بیشتری داشته ودارای طرح مستقلی بوده است ، مورد بررسی قرارمی‌گیرد .

اهداف نظارت :

- ارتقاء کیفیت اطلاعات جمع آوری شده

ـ برقراری یک جریان نظارت آگاهانه برای کسب اطلاعات وانتقال به موقع آن به سطوح مدیریتی به منظور فراهم آوردن امکان استفاده بهینه از امکانات و منابع موجود .

ـ نظارت آگاهانه مورد نظر برکلیه مراحل کار، اعم از تهیه واجرای طرح ، پردازش اطلاعات ، ذخیره سازی ، بازیابی و بالاخره استخراج وانتشار نتایج .

ـ انتقال اطلاعات لازم درزمینه طرح به سازمانهای آماری منطقه‌ای و بین‌المللی .

ـ استفاده ازمشاورحقوقی برای حل و فصل مسائل حقوقی احتمالی به منظور جلوگیری ازتاثیر نامطلوب مسائل ضمن کار برکیفیت اجرای آمارگیری .

سازمان اجرای طرح

قبل از شروع بحث ، ابتدا ، ویژگیهای سازمان اجرایی این قبیل طرحهای بزرگ به اختصار معرفی می‌شود :

ـ این قبیل سازمانها دارای ستاد مرکزی می‌باشند .

ـ ستاد مرکزی ازگروههای تخصصی مختلفی تشکیل شده است ، که ستاد را درانجام وظایف خود یاری می‌دهند .

ـ سازمان اجرایی دارای بودجه مستقل و مصوب می‌باشد .

ـ معمولا " این سازمانها موقتی هستند، پرسنل هرکدام ازرده‌های سازمان برای مدت معینی ، تامین شده و آموزش لازم را می‌بینند و پس از شرکت درآزمون و کسب حدنصاب امتیاز به کار گمارده می‌شوند و درپایان عملیات میدانی به کارآنها درسازمان اجرایی طرح خاتمه داده می‌شود .

ـ ماموران آمارگیری معمولا " از بین دانشجویان یا کارکنان سازمانهای دولتی انتخاب می‌شوند .

ـ رده‌های سرپرستی بلافصل ماموران و متصدیان رده‌های بالاتر ، ازکارکنان سازمانهای دولتی انتخاب می‌شوند تا کارآیی سازمان ازدیدگاه مدیریت افزایش یابد .

ـ راهنماهای فنی واجرایی و شرح وظایف مستقل برای کلیه رده‌های سازمانی تدوین می‌شود و درکلاسهای آموزشی ، مورد آموزش و توجیه قرارمی‌گیرد .

ـ کلاسهای آموزشی معمولا " ۶۵ ساعت درطول ۱۱ تا ۱۲ روز برگزارمی‌شود ، دراین کلاسها شبیه سازی فرمها و پرسشنامه‌ها و همچنین عملیات میدانی نیز جزء برنامه می‌باشد و در

پایان آزمون برای انتخاب ماموران و منصبیان رده‌های مختلف انجام می‌گیرد .

ـ سازمان اجرایی دارای نقشه‌های مخصوص آمارگیری می‌باشد که قبلا " توسط واحد ذی ربط بهنگام شده به‌عنوان ابزار اصلی کار در اختیار کارکنان سازمان قرار می‌گیرد .

ـ نحوه انجام وظایف توسط هریک از رده‌های سازمان ، تابع برنامه‌زمانی کلی طرح بوده و در سرتاسر کشور هریک از رده‌های سازمانی در یک روز معین کار خود را شروع و در تاریخ معینی وظایف خود را به پایان می‌رسانند و مرحله بعدی کار بلافاصله شروع می‌شود . این هماهنگی و همزمانی یکی از خصوصیات بارز این قبیل سازمانهای اجرایی است . بنابراین لحاظ کنترل کمی و کیفی کار توانا " انجام می‌گیرد تا ضمن جلوگیری از عقب افتادگیهای احتمالی از برنامه زمانی ، کار با کیفیت مورد نظر و مطلوب انجام بگیرد .

به‌منظور آشنایی بیشتر ، نمودار سازمان اجرایی طرح آمارگیری جاری جمعیت ضمیمه این گزارش می‌باشد (پیوست شماره ۱) . همانطورکه ملاحظه می‌شود درکنار ستاد طرح ، ۹ گروه برحسب ضرورت ستاد را در برنامه‌ریزی ، تهیه طرح ، اجرا ، استخراج و ... یاری می‌دهند .

یکی از گروههای مزبور گروه نظارت و پیگیری می‌باشد . این گروه وظیفه داشت ضمن سازماندهی نظارت ، برنامه زمانبندی اجرای طرح (CPM) را نیز درمقاطع قبل ازاجرا و بعداز اجرا طراحی نموده و اجرای فعالیتهای مربوط به کلیه واحدهای ستادی واجرایی را پیگیری نماید .

سازماندهی نظارت

روش نظارت درآمارگیری جاری جمعیت سال ۱۳۷۰

برای این آمارگیری ، طرح نظارت به دو صورت ، طراحی و اعمال شد :

ـ نظارت درون سازمانی

ـ نظارت برون سازمانی

نظارت برون سازمانی

با استفاده از تجارب اعضاء و بهره گیری ازمدارک مربوط به طرحهای بزرگ آماری و سرشماری کشورها و با هماهنگی گروه اجرایی ، نظارت برون سازمانی مورد نظر ، در سازمان اجرایی آمارگیری جاری جمعیت به عهده رده سرپرستی بلافصل مامور آمارگیری محول و عنوان ناظر فنی به آن اطلاق شد . عنوان

رده بالای ناظر فنی ، معاون فنی نامیده شد . متصدیان این دوره سازمان اجرایی طرح ، علاوه بر شرکت در کلاسهای آموزشی ، در کلاس توجیهی سرپرستی نیز شرکت داده شدند تا بتوانند ضمن مدیریت و سرپرستی بر افراد تحت نظر خود ، نظارت و مراقبت کمی و کیفی لازم را نیز در طول مدت اجرا اعمال نمایند .

نظارت ناظر فنی و معاون فنی در این آمارگیری ثبت مستمر نتایج بازدید از نحوه مصاحبه آمارگیر در حین آمارگیری و نحوه تکمیل فرمهای ۱ ، ۲ و ۳ بود . (در این مرحله از کار فرمهای ۲۲ نظارت مورد استفاده قرار گرفت) در این رابطه ، ناظر با برنامه خاصی به ماموران آمارگیری مراجعه کرده ضمن بازبینی فرمهای تکمیل شده چگونگی کار را مورد نظارت و ارزیابی قرار داده ، با نقاط قوت و ضعف ماموران در رابطه با تکمیل هریک از فرمها آشنا شده و اشتباهات و نقائص کار آنها را به موقع برطرف می‌کرد . ناظر برای درج مشاهدات و ارزیابی کار ماموران از فرم ۲۲ نظارت استفاده می‌کرد . در هر مراجعه با تکمیل یک ستون از فرم مزبور کیفیت کار ماموران ثبت می‌شد و در مراجعه بعدی این امکان برای ناظر فنی وجود داشت تا نحوه پیشرفت کیفی کار مامور را در مراجعات مکرر مورد مقایسه و ارزیابی قرار دهد .

سوالات مطرح شده در این فرم با نظر گروه تهیه طرح و آموزش از بین سوالات کلیدی انتخاب و در فرم مزبور ملحوظ گردید .

این فرم عمده‌ترین ابزار نظارت درون سازمانی بود . معاون فنی نیز در مراجعات خود موظف بود به اجرای دقیق وظایف نظارتی ناظر فنی نظارت کرده و مواردی از نظارت اعمال شده توسط وی را کنترل کند .

فرم ۲۲ با توجه به اهداف مربوطه تهیه و برای نظارت برنامه‌ریزی شده ، طراحی گردید و در راهنمای آمارگیری قرار گرفت . در نظارت درون سازمانی کلیه رده‌های سازمان آمارگیری ، از جمله مدیر اجرای آمارگیری شهرستان ، قائم مقام و مدیر اجرای آمارگیری استان موظف به اعمال نظارت و کنترل کمی و کیفی کار آمارگیری بودند و یکی از وظایف آنها کنترل سیستم گردش کار آمارگیری ، نظارت و شناسایی " ارتقاء کیفیت اطلاعات کسب شده بود .

فرم ۲۲ (پیوست شماره ۲) همراه با نمودارهای حاصل از استخراج آن در جلد شهرستان ، تحت عنوان " نمودار افزایش دقت برحسب مراحل نظارت " با استفاده از Harvard Graphics ترکیب رنگ نمایش داده خواهد شد .

نظارت برون‌سازمانی

درنظارت برون‌سازمانی ، دوروش پیش بینی شد ، نظارت برون سازمانی به روش مستقیم‌ونظارت برون‌سازمانی‌به روش غیرمستقیم .

الف ـ نظارت برون سازمانی به روش مستقیم :

نظارت برون سازمانی توسط گروه نظار اعمال گردید واعضای آن با پیشنهادگروه نظارت‌تعیین واعزام شدند . این‌گروه برگردش کار سازمان اجرایی ازستاد استان تاسطح معاون‌فنی شهرستان نظارت می‌کردند . نظارت برون‌سازمانی شامل مواردی است که درذیل عنوان‌می‌شود :

۱ ـ نظارت قبل از اجرای آمارگیری

۲ ـ نظارت حین‌اجرای آمارگیری

۳ ـ نظارت بعدازاجرای آمارگیری

۱ ـ نظارت قبل‌ازاجرای‌آمارگیری

۱ـ۱ـ جمع‌آوری برنامه زمانبندی فعالیت‌گروههای ستاد آمارگیری و تهیه CPM مربوط به فعالیتهای مختلف گروهها اعم‌ازمراحل آموزش ، تشکیل ستادهای اجرایی استان‌وشهرستان،اجرا ، جمع‌آوری فرمها‌و پرسشنامه ها ، استخراج نتایج (زودرس تفصیلی) وانتشارنتایج .

۱ـ۲ـ ایجادارتباط باگروهها و پیگیری‌انجام فعالیتهای مربوط به آنها توسط رابطین‌گروه نظارت و پیگیری وارائه نتایج‌به ستاد وانجام پیگیری تصمیمات‌متخذه ستاد .

۱ـ۳ـ ایجاد ارتباط‌بین‌گروهها به منظور حصول اطمینان‌ازانجام به موقع فعالیتهای آنها و کسب اطمینان ازانجام فعالیتهایی‌که به تعبیری پیش نیاز فعالیتهای سایرمراحل به شمار می‌آید .

۱ـ۴ـ نظارت برنحوه پیگیری تامین نیروی انسانی لازم برای ردههای سازمان‌اجرایی و مسئولین آموزش و ...

۱ـ۵ـ نظارت‌بر نحوه برگزاری کلاسهای آموزشی مطابق با برنامه آموزش (شامل چگونگی آموزش ، تشکیل به موقع کلاسها ، هماهنگی و پیشرفت آموزش کلاسها ، انجام عملیات میدانی ، آزمونها و ...) و تشکیل کلاسهای بازآموزی وامور مترتب‌برآنها وارائه گزارشهای لازم به ستاد .

۱-۶- نظارت برنحوه تامین وسایل نقلیه جهت اجرای طرح .

۱-۷- نظارت برتامین ملزومات مورد نیاز و بررسی کیفیت آنها .

۱-۸- نظارت برنحوه توزیع ملزومات وارسال به موقع آنها به استانها .

۱-۹- پیگیری درتشکیل به موقع ستادهای اجرایی طرح دراستان و شهرستانها طبق ضوابط سازمان اجرایی .

۱-۱۰- تهیه فرم ۲۱، ساعات شرکت اعضاء گروهها وجمع بندی آنها به منظور تعیین حجم کارهای انجام شده درگروهها .

۱-۱۱- استفاده از مشاور حقوقی به منظور پیش بینی موارد حقوقی لازم درزمینه نحوه احقاق حق قانونی مرکز وعوامل اجرایی ، طرح و پیگیری حقوقی درمراحل مختلف اجرا تا حصول نتیجه .

۲- نظارت حین اجرای آمارگیری

وظایف نظارت درون سازمانی حین اجرا ، دروظایف اجرای معاونین وناظران فنی (مسئولین بلافصل مامورآمارگیری) گنجانیده شده است . نظر به اهمیت وظایف نظارتی ، درسازمان اجرایی این طرح فقط سه مامور تحت سرپرستی یک ناظر فنی قرارداده شد تا به ناظر فنی فرصت رسیدگی وکنترل بیشتر کار داده شود . دراین بخش فقط وظایف نظارتی برون سازمانی حین اجرا، که بطور عمده گروه نظار عامل آن بوده اند، توضیح داده می شود . مهم این وظایف شامل نظارت برموارد زیر است :

۲-۱- تشکیل گروههای آمارگیری .

۲-۲- استقرار به موقع ماموران آمارگیری درحوزه های مربوط .

۲-۳- شروع همزمان کارآمارگیری .

۲-۴- انجام وظایف تعیین شده معاونان فنی شهرستان وناظران فنی .

۲-۵- جمع آوری وانتقال مدارک .

۲-۶- تهیه گزارشهای دوره ای و گزارشهای موردی .

۳- نظارت بعدازاجرای آمارگیری

این مرحله ازنظارت بعدازپایان عملیات میدانی انجام می شود و شامل نظارت و پیگیری مراحل

زیر سی باشد :

۱-۲- دریافت فرمهای کامل و کنترل شده حوزه‌ها از ناظر فنی .

۲-۲- بسته بندی مدارک مربوط به شهرستانها و تحویل آن به استان .

۳-۲- انتقال فرمها و مدارک از استان به مرکز طبق برنامه تنظیم شده .

۴-۲- تحویل و تحول مدارک آمارگیری بین استانها و مدیریت استخراج دستی .

۵-۲- تفکیک مدارک و پیشرفت کار .

۶-۲- پیشرفت کار منگنه اطلاعات مندرج در فرمهای ۱ ، ۲ ، ۳ و فرمهای گزارش‌گیری .

۷-۲- ساخت فایلهای اطلاعاتی فرمهای ۱ ، ۲ ، ۳ و فرمهای گزارش گیری .

۸-۲- ادیت و اصلاح فایلهای اطلاعاتی .

۹-۲- استخراج نتایج داده‌های آزمایشی .

۱۰-۲- استخراج نتایج فرمهای گزارش‌گیری .

۱۱-۲- استخراج نتایج فرمهای ۱ ، ۲ ، ۳ (فرم ۱ - فهرست برداری، فرم ۲ - پرسشنامه عمومی خانوار و فرم ۳ - پرسشنامه تفصیلی خانوار)

۱۲-۲- استخراج نتایج مربوط به چارچوب سازی جمعیت ، فعالیتهای اقتصادی در خانوار و کارگاهها .

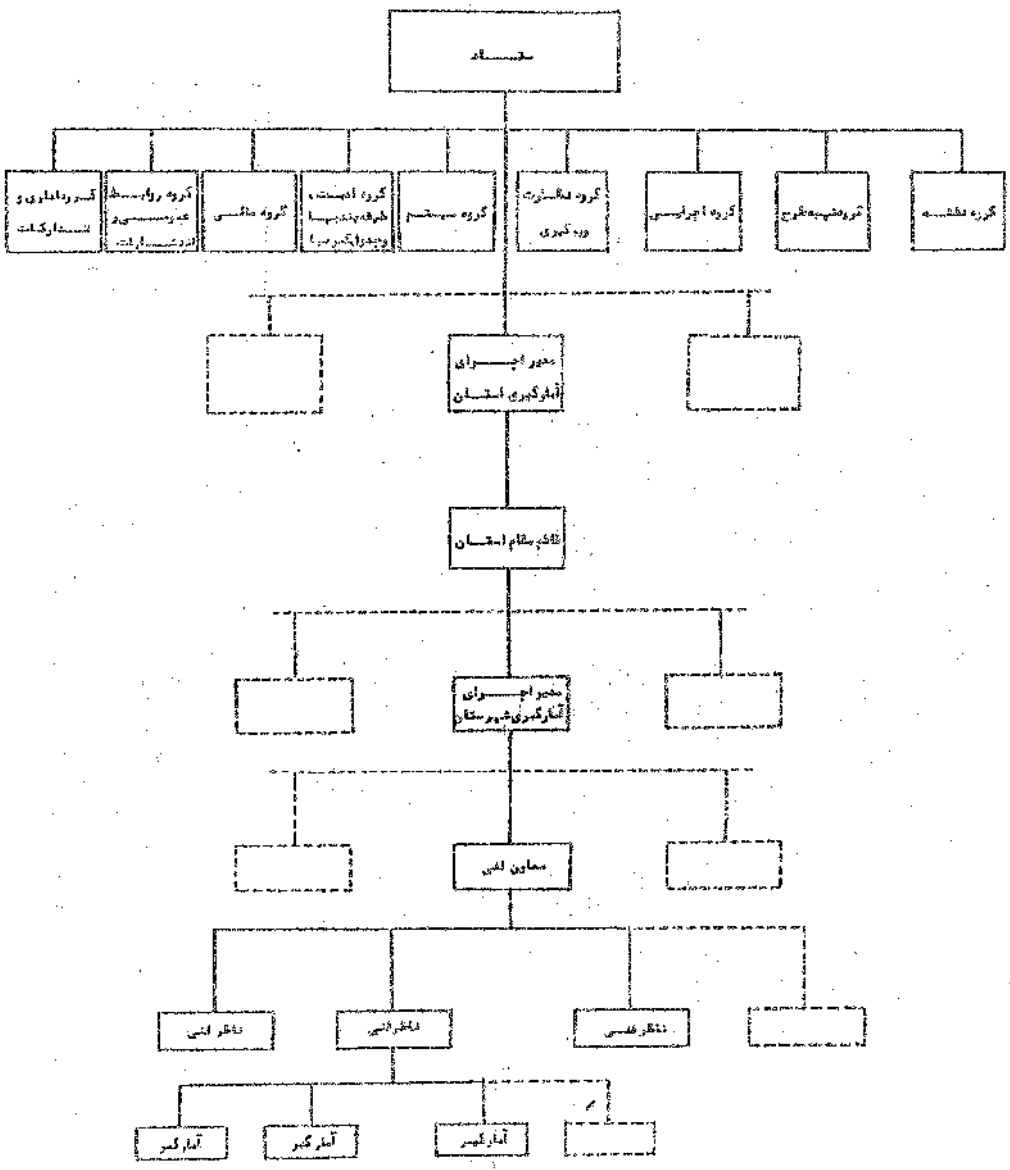
ب - نظارت برون سازمانی به روش غیرمستقیم

باتوجه به اینکه در اجرای کلیه طرحها و فعالیتها ، مردم یکی از ارکان اصلی بشمار آمده و همواره مسئولان از حمایتها و کمکهای آنان برخوردارند . دراین طرح نیز برای حصول اطمینان از پوشش کامل آمارگیری ، بر همکاری خانوارها تکیه شده و تعدادی از آنان به عنوان عوامل نظارت برون سازمانی ، در حین اجرا ، در نظر گرفته شدند . برای این منظور فرم ۴ طراحی گردید و همراه با نامه‌ای برای تعدادی از خانوارها در سطح کشور ارسال و از آنها درخواست گردید تا پس از مراجعه به مامور آمارگیری و تکمیل پرسشنامه آن خانوار (درصورتی که آمارگیری شده باشد) یا ، شا پس از پیمتم سرماه (درصورتی که آمارگیری نشده باشد) بنا به مورد با تکمیل یکی از دو بخش فرم مزبور ، ستاد آمارگیری را یاری دهند . فرم مزبور و نتایج آن نیز دراین سرشماری با استفاده از Overhead نشان داده می‌شود .

نتیجه گیری :

علیرغم تجربه اندك کشور درمقایسه با کشورهایی که تاکنون بیش از یك قرن دراجـــــرای سرشماریهای جمعیتی تجربه دارند ، بنظر می‌رسد ، اعمال نظارت و کنترل کیفی اطلاعـــــات جمع‌آوری شده درسرشماریها ، با وجود نوسانهایی ، روند صعودی داشته و منجر به‌ارتقا کیفیـــت اطلاعات شده‌است . مقدار شاخص ویپل محاسبه‌شده توسط‌مرکزآمارایران درسرشماریهای ۴۵ تـــا ۶۵ وهمچنین محاسبه شاخص‌تلفیقی درسرشماریهای جمعیتی ایران که روند روبه بهبودی رانشـــان می‌دهد، می‌تواند عمدتا " ناشی ازاعمال روشهای دقیق تر ومناسبتر کنترل کیفیت طرحهای بـــزرگ آماری باشد، هرچند که بهرحال تاثیرافزایش میزان باسوادی وارتقاء دانش آماری نیزدراین مـــورد نباید نادیده بماند .

" سازمان اجرایی طرح آمارگیری بازرگانی جمعیت ۱۳۷۰ "



پیوست شماره ۱

بسمه تعالی

جمهوری اسلامی ایران

وزارت بهداشت، درمان و آموزش پزشکی

مرکز بهداشت شهرستان

فرم ۱۲ ـ نظارت (مخصوص مناطق روستایی)

شماره ..... / ..... / .....

شماره خانه‌بهداشتی ..... / ..... / .....

در هر بار نظارت، شماره خانه و یا ملیت مشخص شده، مورد نظر قرار گرفته و به ترتیب در ستون‌های مربوطه علامت (×) گذاشته شود. ستون‌های مربوط به تکمیل فرم، حتی‌المقدور سالی یکبار که در هر بار نظارت، گزارش نظارت شامل نظارت تکمیل فرم ۲ هم باشد.

کد جغرافیائی

شماره خانه / تاریخ نظارت	مرحله اول	مرحله دوم	مرحله سوم	مرحله چهارم	مرحله پنجم	مرحله ششم	مرحله هفتم	مرحله هشتم
۱	۲	۳	۴	۵	۶	۷	۸	۹
نحوه نظارت و بررسی نظارتی:	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز	در منزل / در مرکز
فرم ۱								
درج نام مادر								
مشخص بودن محل سکونت در نقشه								
رعایت مربوط به شماره و ردیف خانوار								
صحیح بودن اعداد جابجا شده در ستون								
مشخص بودن وضعیت اشتغال و بیکاری								
درج نام کارگاه								
مشخص بودن محل کارگاه								
مشخص بودن نوع کارگاه								
فرم ۲								
صحیح بودن اعضای خانوار در صفحه مربوطه								
مشخص بودن محل تولد								
مشخص بودن وضع تاهل								
مشخص بودن افراد باسواد								
مشخص بودن وضع فعالیت								
مشخص بودن رشته تحصیلی								
فرم ۳								
استعمال صحیح اطلاعات در فرم ۳								
مشخص بودن استعمال								
مشخص بودن نوع زایمان								
مشخص بودن محل تولد مرتبط								
مشخص بودن اطلاعات مراقبت								

ایرادات توسط ناظر فنی تکمیل شود.

نام و نام خانوادگی ناظر فنی .....

تاریخ نظارت ..... محل نظارت .....

شرح نظارت :

.....

.....

تاریخ نظارت ..... محل نظارت .....

شرح نظارت :

.....

.....

نام و نام خانوادگی مناطق روستایی .....

پیوست شماره ۲

سیستمهای خبرهٔ استقرایی و کاربرد آن در آمار

سید باقر میراشرفی* - حسن صادقی

دانشگاه فردوسی مشهد

خلاصه

سیستمهای خبره بر اساس هوش مصنوعی (AI) وبا استفاده از زبانهای LISP , PROLOG و یا زبانهای برنامه سازی پیشرفته دیگر ساخته میشوند .

آمار وهوش مصنوعی بطور متقابل به توسعه یکدیگر کمک میکنند .از یک طرف روشهای آماری برای توصیف مفاهیم مبهم در سیستمهای خبره مورد استفاده قرار میگیرد .از طرف دیگر سیستمهای خبره بعنوان ابزار فراگیری مفاهیم میتواند برای حل تعدادزیادی از مسائل طبقه بندی کردن آماری مورد استفاده قرار گیرد .بخصوص در روشهای تحلیل رگرسیون، تحلیل ممیزی، تحلیل خوشه ای و تحلیل سریهای زمانی کاربرد فراوانی دارد .

در این مقاله یکی از الگوریتمهایی را که کاربردزیادی دارد مورد بررسی قرار میدهیم . این الگوریتم توسط Quinlan در سال ۱۹۷۹ تحت عنوان ID3 ، برای حالتی که متغیر کلاس کیفی باشند ، ارائه وتاسال ۱۹۸۶ اصلاح وتکمیل گردید .نتیجه این الگوریتم قواعدی است که بر اساس آن به مسائل طبقه بندی جواب داده میشود .در کنار این الگوریتم برای حالتی که متغیرها پیوسته باشند الگوریتم Regression Tres توسط Breiman (۱۹۸۴) مطرح میشود .

۱- مقدمه

شناسی سیستمهای خبره استقرائی بر اساس قاعده استقراء از داده‌های آماری بوده و متعلق به دسته‌ای از روشها تحت عنوان "یادگیری از روی مثالها" می‌باشد. یادگیری از روی مثالها را می‌توان جزئی از تکنیک یادگیری ماشینی (ML) دانست. مقصد بر اینکه یادگیری ماشینی بخشی از هوش مصنوعی را تشکیل می‌دهد.

در کنار سیستم‌های خبره استقرائی، تعدادی از الگوریتم‌های شبکه عصبی و نیز روشهای آماری کلاسیک، نظیر کوسیون، سریهای زمانی و تحلیل تمیزی، از طریق یادگیری از روی مثالها تحلیل می‌شوند.

سیستم‌های خبره استقرائی پیشرفت اساسی در جهت تهیه فرآیندهای اتوماتیک و فراگیری دانسته‌ها داشته است. با استفاده از تجربیات گذشته که بصورت نمونه‌هائی در دسترس هستند، می‌توانیم قوانینی بر اساس دانسته‌ها تولید کنیم. بیشتر سیستم‌های خبره استقرائی که در حال حاضر مورد استفاده قرار می‌گیرند و کاربرد تجاری نیز پیدا کرده است، بر اساس مقاله مشترک (۱۹۸۴) Quinlan و (۱۹۸۴) Breiman می‌باشد.

که در بخش بعدی به معرفی این الگوریتمها و کاربرد آن می‌پردازیم.

۲- تحلیل تمیزی و طبقه بندی

در تحلیل تمیزی و طبقه بندی به روش کلاسیک، معمولاً بر اساس روش کمترین مربعات خطا و یا ماکزیمم کردن تابع زیان، و با استفاده از نمونه‌ها، یک تابع تمیز تشخیص جهت تحلیل دسترسی پیدا می‌کنیم.

اما در روشهای سیستم‌های خبره استقرائی بر اساس الگوریتم‌های مشخص، که تشریح یکی از این الگوریتمها بعداً خواهد آمد، به یک درخت تصمیم گیری رسیده و بر

اساس درخت، قواعد تصمیم‌گیری مشخص خواهد شد. و با استفاده از قواعد بدست آمده، به تحلیل مسائل مورد نظر می‌پردازیم.

۳- الگوریتم ID3 (۱۹۸۶)

اولین بار در سال ۱۹۷۹ این الگوریتم توسط Quinlan مطرح که پس از اصلاح در سال ۱۹۸۶ بصورت کاملتری تحت عنوان "Intractive Dachotomizer 3" ارائه گردید. در این روش با استفاده از اطلاعات نمونه‌ها و بر اساس دستورالعمل های زیر به نتایج مورد نظر خواهیم رسید. فرم کلی داده‌ها در جدول شماره ۱ نشان داده می‌شود. این الگوریتم زمانی مورد استفاده قرار می‌گیرد، که مقادیر مربوط به طبقه‌ها (کلاسها) مقادیر کیفی (گسسته) باشند.

دستور العمل ID3

۱- ابتدا کل نمونه‌ها را بعنوان یک مجموعه S در نظر می‌گیریم، که اصطلاحاً به آن گره نامیم.

۲- اگر تمام نمونه‌ها در گره S دارای کلاس مشترک C باشند، آنگاه دسته بندی مربوط به این گره C خواهد بود و به دستورالعمل بعدی می‌پردازیم.

در غیر اینصورت خصیصه A را که دارای بیشترین اطلاعات می‌باشد انتخاب می‌کنیم. سپس S را به مجموعه های جدا از هم بر اساس مقادیر A افراز می‌کنیم و هر یک از این مجموعه‌های افراز شده را بعنوان گره S در نظر گرفته و دستور العمل مرحله ۲ را برای تمام این گره ها اجرا می‌کنیم.

۳- در این مرحله تمامی گره‌های نهایی کلاس یکسانی داشته و الگوریتم خاتمه پیدا می‌کند.

جدول شماره ۱ ـ فرم کلی داده‌ها در یادگیری از روی مثالها

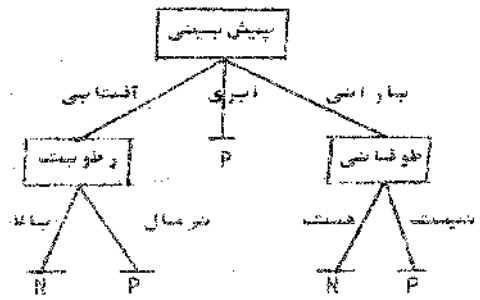
شماره‌ها	خصیصه‌ها (متغیرها)				طبقه‌ها (کلاسها)
۱	a_{11}	a_{12}	. . .	a_{1m}	C_1
۲	a_{21}	a_{22}	. . .	a_{2m}	C_2
۳					
۴					
۵					
n	a_{n1}	a_{n2}	. . .	a_{nm}	C_n

مثال ۱ ـ در یک تحقیق وضعیت هوای صبح چهارده روز مورد مطالعه قرار گرفت. هدف از این تحقیق این است که بر اساس مقادیر متغیرهای مورد بررسی، از مثبت بودن شرایط هوا برای یک فعالیت ورزشی، مطلع شویم. متغیرها عبارتند از پیش بینی هوا بر حسب آفتابی، ابری و بارانی و متغیر دما بر حسب درجه سانتی‌گراد و نوع متغیر رطوبت به مقادیر بالا و نرمال، متغیر فوق‌الذکر با مقادیر درست و غلط. اطلاعات مربوط به این تحقیق در جدول شماره ۲ آمده است.

جدول شماره ۲ ـ یک مجموعه اجرائی کوچک

نمونه ها	خصیصه ها (متغییرها)			کلاس ها
	طوفانی	رطوبت	دما	پیش بینی هوا	
۱	نمیدم	بالا	گرم	آفتابی	N
۲	هست	"	"	"	N
۳	نمیدم	"	"	ابری	P
۴	"	"	معتدل	بارانی	P
۵	"	نرمال	سرد	"	P
۶	هست	"	"	"	N
۷	"	"	"	ابری	P
۸	نمیدم	بالا	معتدل	آفتابی	N
۹	"	نرمال	سرد	"	P
۱۰	"	"	معتدل	بارانی	P
۱۱	هست	"	"	آفتابی	P
۱۲	"	بالا	"	ابری	P
۱۳	نمیدم	نرمال	گرم	"	P
۱۴	هست	بالا	معتدل	بارانی	N

با استفاده از اطلاعات جدول فوق (مجموعه اجرائی) و بر اساس الگوریتم ID3 درخت تصمیم گیری نمودار ۱ حاصل می‌شود. که بر اساس درخت حاصل قواعد تصمیم گیری بصورت ذیل خواهد شد .



نمودار ۱ ـ درخت تصمیم گیری ساده

قواعد متصمیم شیری

۱- اگر هوای صبح روز شنبه آفتابی و رطوبت بالا باشد آنروز برای فعالیت بطور نامساعد است .

۲- اگر هوای صبح روز شنبه آفتابی و رطوبت نرمال باشد آنروز برای فعالیت بطور مساعد است .

۳- اگر هوای صبح روز شنبه ابری باشد آنگاه آنروز برای فعالیت بطور مساعد است .

۴- اگر هوای صبح روز شنبه بارانی و طوفانی هم باشد آنگاه آنروز برای فعالیت بطور نامساعد است .

۵- اگر هوای صبح روز شنبه بارانی و طوفانی نباشد آنگاه آنروز برای فعالیت بطور مساعد است .

در دستور الصملهای ID3 ، بعضی مواقع در مرحله دوم لازم است که خصیصه دارای بیشترین اطلاعات را متعیین کنیم برای این منظور از تکنیک زیر استفاده می‌کنیم .

در این روش ابتدا اطلاعات مربوط به کل نمونه‌ها را بدست آورده وسپس بر اساس آن ، میزان موثر بودن تمامی متغیرها را محاسبه می‌کنیم و درمیان این مقادیر ، بزرگترین مقدار را مشخص کرده ودر نتیجه متغیر مربوطه ، دارای بیشترین اطلاعات می‌باشد .

فرض کنید متغیر کلاس مربوط به مجموعه اجرائی وجدول داده‌ها دارای K رده باشد .همچنین فرض کنید خصیصه (متغیر) A دارای مقادیر A1 و A2 و ... و Av و متغیر نمونه‌های مجموعه (کره) S را بر اساس n باشد برای محاسبه میزان عدم اطلاع غیر شرطی کلاس (آنتروپی کل کلاس) از فرمول زیر استفاده می‌کنیم .

$$I(C_1,\dots,C_K)=-\sum_{i=1}^{K} P_i \log_2 P_i \quad , \quad P_i=\frac{\#C_i}{n}=\frac{\text{تعداد نمونه‌های دارای کلاس } C_i}{\text{تعداد کل نمونه‌های در } S}$$

اکنون برای محاسبه میزان موثر بودن متغیر دلخواه A ، ابتدا متوسط آنتروپی کلاس با فرض متغیر را بدست آورده و از آنتروپی کل متغیر کلاس کم کرده و مقدار حاصل میزان موثر بودن A برای پیش بینی کلاس خواهد بود. برای بدست آوردن متوسط آنتروپی کلاس با فرض متغیر A ، براساس فرمول میانگین موزون زیر عمل می‌کنیم .

$$E(A)=\sum_{i=1}^{V}\frac{\#A_i}{n} I(C_1,C_2,\dots,C_K/A_i)$$

که در آن

$$I(C_1,C_2,\dots,C_K/A_i)=-\sum_{x=1}^{K} P_x \log_2 P_x$$

و در آن

$$P_x=\frac{\#(C_x/A_i)}{\#A_i}=\frac{\text{تعداد نمونه‌هایی که براساس افراز } A_i \text{ دارای کلاس } C_x \text{ هستند}}{\text{تعداد نمونه‌هایی که در افراز } A_i \text{ قرار دارند}}$$

و در نتیجه :

$$\text{میزان موثر بودن } A = I(C_1,C_2,\dots,C_K)-E(A)$$

این میزان برای تمامی متغیرها محاسبه و ماکزیمم میزان را مشخص می‌کنیم که متغیر مربوطه ، همان متغیر دارای بیشترین اطلاعات خواهد بود. چون در اینجا I (C1,C2,...,Cn) برای تمامی متغیرها یکسان است بنابر این کافی است، متوسط آنتروپی کلاس با فرض متغیرها را محاسبه و از بین متوسط ها ، مینیمم را مشخص کرده و متغیر مربوطه همان متغیر دارای بیشترین اطلاعات خواهد بود .

این محاسبات را برای مثال ذکر انجام میدهیم. در اینجا تعداد نمونه‌ها $n=14$ و تعداد کلاسها $K=2$ و بعنوان مثال برای متغیر پیش بینی $V=3$ است، ابتدا برای متغیر پیش بینی که دارای سه مقدار آفتابی، ابری و بارانی می‌باشد متوسط آنتروپی را بدست می‌آوریم .

برای مقدار آفتابی داریم : $\#N_1 = 3$ و $\#P_1 = 3$ ... وبنابراین

$$I(N,P/A_1 = \text{آفتابی}) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = /۹۴ \text{ بیت}$$

وبطور مشابه برای مقادیر ابری وبارانی داریم :

$$I(N,P/A_1 = \text{ابری}) = 0 \quad \text{و} \quad \#P_2 = 4 \text{ و } \#N_2 = 0$$

$$I(N,P/A_3 = \text{بارانی}) = /۹۷۱ \quad \text{و} \quad \#P_3 = 3 \text{ و } \#N_3 = 2$$

ودر نتیجه :

$$E(\text{پیش بینی}) = \frac{5}{14} I(N,P/A_1 = \text{آفتابی}) +$$
$$\frac{4}{14} I(N,P/A_2 = \text{ابری}) + \frac{5}{14} I(N,P/A_3 = \text{بارانی}) = /۶۹۴$$

وبنابراین

بابیت ۰/۲۴۶ = (پیش‌بینی) E − ۰/۹۴ = میزان موثر بودن متغیر پیش‌بینی .

وبطور مشابه برای دیگر متغیرها داریم .

بابیت ۰/۰۲۹ = موثر بودن دما

بابیت ۰/۱۵۱ = موثر بودن رطوبت

بابیت ۰/۰۴۸ = موثر بودن طوفانی

و از روی این مقادیر واضح است که متغیر پیش بینی بعنوان متغیر دارای بیشترین اطلاعات تشخیص می‌شود .ونمونه‌ها بر اساس مقادیر این متغیر تقسیم می‌شوند وهر افرازی از این متغیر بعنوان یک گره (مجموعه S) در نظر گرفته وبا توجه به متغیرهای باقیمانده ، مرحله دوم دستورالعملهای ID3 را روی آن پیاده می‌کنیم .یکی از حالتهای خاص وقتی اتفاق می‌افتد که یکی از مجموعه‌های افراز شده (گره‌ها) خالی از نمونه باشد .در اینصورت ID3 برای آن گره مقداری تهی (NULL) را در نظر گرفته ودرپایان اجرای الگوریتم ، به گره‌هایی تهی ، یکی از کلاسهایی را که بیشترین فراوانی را در مجموعه نمونه‌ها داشته

باشد . انتخاب می‌شود .

حالت خاص دیگر وقتی اتفاق می‌افتد که یکی از متغیرها دقیقا " منطبق بر طبقه بندی باشد .فرض کنید متغیر A دارای مقادیر A1 و ... و AV و مجموعه نمونه‌های S را بر اساس مقادیر متغیر A افراز می‌کنیم .چنانچه برای هریک از مجموعه‌های افراز شده ، نسبت‌نمونه‌هایی را که دارای کلاس C (K ، ... ،۲ ،۱) هستند برابر نسبت نمونه‌ها در کل مجموعه S دارای آن کلاس باشند . آنگاه این متغیر را دقیقا " منطبق بر طبقه بندی گوئیم .بنابراین وجود این متغیر در درخت تصمیم گیری ضرورتی نخواهد داشت .

الگوریتم ID3 برای حل این مسئله بر اساس میزان موثر بودن متغیرها ، از متغیرهای غیرموثر صرف نظر می‌کند .و درخت تصمیم گیری را بر اساس متغیرهای موثر تشکیل می‌دهد . بدینصورت که چنانچه برای هر متغیر ، میزان موثر بودن آن از قدر مطلق یا درصدی از یک معیار مشخص کوچکتر باشد ، متغیر را غیر موثر شناخته و در محاسبات شرکت نمی‌دهد .در غیراینصورت موثر شناخته و بر اساس روشی که گفته شد ، بدنبال موثرترین متغیر و سپس ادامه الگوریتم می‌باشد .

این الگوریتم برای مقادیر گمشده و اطلاعات دارای نوسان (NOISE) نیز راه حلهای مناسبی ارائه می‌دهد .جهت اطلاع بیشتر در این زمینه وهمچنین در رابطه با الگوریتم Regressio Tree می‌توانید به پایان نامه تحصیلی اینجانب تحت عنوان باقر میراشرفی دانشکده علوم دانشگاه فردوسی مشهد رجوع نمائید .

نرم افزارهایی نیز در این رابطه از سوی شرکت‌های مختلف تهیه ومورد استفاده قرار می‌گیرند .

دو نمونه از این نرم افزارها تحت عناوین NEWID , KET بوده که بصورت راهنمای مربوطه در اختیار قرار دارد .درکنار این دو نرم‌افزار ، نرم‌افزار

دیگری نیز اینجانب با همکاری گروه کامپیوتر دانشگاه فردوسی مشهد تهیه شموده ومورد استفاده قرار دادیم .

۵- نتیجه گیری

همانگونه که در این مقاله آمده است ، ما بدنبال سیستمی هستیم که با استفاده از آن بتوانیم براحتی وبا دقت بیشتر ونیز سرعت زیاد ، تجزیه وتحلیلهای مختلف آماری را که در زمینه های مختلف کاربردی مورد استفاده قرار میگیرد ، به انجام برسانیم .

برای نیل به این هدف از سیستمهای خبره ای که بر اساس روند فکر کردن افراد خبره تشکیل شده باشند ، استفاده میکنیم .و انتظار داریم که در آینده ای نه چندان دور برای تمامی تجزیه وتحلیلهای آماری ، سیستمهای خبره ای تهیه و مورد استفاده علوم ودر خدمت بشر قرار گیرد .

در پایان لازم میدانم از جناب آقای ارقامی استادیار دانشگاه فردوسی مشهد به خاطر زحماتی که در رابطه با این مقاله متقبل شدند تشکر وقدردانی نمایم .

مراجع :

۱- پایان نامه کارشناسی ارشد آمار - سیدباقر میرآخوری - دانشکده علوم - دانشگاه فردوسی " مشهد "

2) Breiman,L.et al(1984),Classification and Regression Tree.Wandsworth Belmont.

3) Nakhaeizadeh, G.(1991 a),Inductive Expert Systems and their Application in Statistics. Erscheint in Proceedings of Statsoft 91.

4) Nakhaeizadeh, G.(1991b),Application Of Machine Learning to solving industrial problems. Eyscheint in Proceedings of SOR 91 Trier, Physica Verlag.

5) Quinlan,J.R.(1986), Induction of Decision Trees.Machine Learning. VOL. 1,81-106.

الگوی پیشنهادی در مورد تصحیح یافته‌های سرشماریهای جمعیتی در زمینهٔ موالید

محمدحسین رجاتیان

مرکز آمار ایران

مقدمه وطرح مسئله

گردآوری اطلاعات آماری درزمینه جمعیت و مشخصات آن سابقه‌ای دیرینه دارد وشاید جامعه بشری ازاین در به قلمرو آمار وارد شده باشد. ازاین رو بحث درمورد کیفیت داده‌های جمعیتی نیز ازقدمتی دیرینا برخوردار است وازگذشته‌های دور کوششهایی برای ارتقاء یا حداقل اندازه‌گیری کیفیت داده‌های جمعیتی به عمل آمده است که بعضاً به ارائه شیوه‌های مناسبی نیز منجر شده است ولی ازآنجا که این شیوه‌ها بسته به روش گردآوری اطلاعات متفاوت می‌باشد در زیر به تقسیم‌بندی این شیوه‌ها بسته به روش گردآوری اطلاعات اشاره و نقاط ضعف و قوت هر یک بحث می‌شود.

گردآوری اطلاعات مربوط به موالید ومرگومیر از طریق سیستم ثبت وقایع حیاتی اولین و مهمترین شکل گردآوری این اطلاعات است که در برنامه ازطریق مشاهده و پرسشهای اطلاعات ضبط شده ونیز روش تهیه آنها مستقیماً مورد بررسی قرار می‌گیرد ولی اساساً گفتنی است که درآغاز تشکیل نگرفته است. پیشنی ثبت هر واقعه نوعی باور و اعتقاد مذهبی و نوع دعوا با صدور یا ابطال مدرک رسمی ... این سیستم در جوامعی که ... مناسبی به کار گرفته و اطلاعات جامع و دقیقی در ... تهیه و میزان‌های موالید ومرگومیر به دست داده است. ولی در اکثر کشورها نیز با مشکلاتی مواجه است ...

بوده است . این مشکلات که ناشی از بازتابهای اجتماعی و فردی درمقابل ارزش قانونی ثبت یا عدم ثبت واقعه مورد نظر است ، ابدا " ازارزش فعالیتهایی که دراین زمینه به عمل میآید نخواهد کاست .

از جمله مشکلاتی که دراین زمینه وجود دارد، میتوان ازعدم مراجعه به موقع برای ثبت واقعه مورد نظر ، عدم انطباق محدوده مورد عمل ادارات ثبت یا سیستم تقسیم بندی کشور ونیز امکان ثبت وقایع مربوط به یک منطقه در منطقه‌ای دیگرازکشور، نام برد .

علاوه براین با پیشرفت علوم ودستیابی به روشهای پیچیده تر تحلیل اطلاعات ضرورت کسب اطلاعات دیگر دموگرافیک همراه با آمار وقایع حیاتی نیز افزایش یافته است که سیستم ثبت وقایع ماهیتا " نمیتواند چنین اطلاعاتی راد راختیار بگذارد . بدین لحاظ برای نیل به این داده‌ها، سایر روشهای گردآوری اطلاعات ،ازجمله آمارگیریهای نمونه‌ای وسرشماریهای جمعیتی نیز به خدمت گرفته شده است .

طبیعی است ، اطلاعاتی که ازطریق آمارگیریهای نمونه‌ای یاسرشماریهای جمعیتی نیز بد ست میآید نمیتواند عاری ازخطا، به ویژه خطاهای غیرنمونه‌ای باشد . تجربه نشان داده است که در مورد وقایع ولادت ، عمده ترین منبع خطا ، اشتباهات حافظه‌ای است . چرا که دراین مورد عموما " راجع به وقایعی که اززمان وقوع آنها مدت درازی (مثلا " ۱۲ ماه یا یکسال) گذشته است استفسارمیشود بدین لحاظ احتمال زیاد وجود دارد که یاواقعه ولادت فراموش شود یاوقوع آن دردوره زمانی مورد نظر Reference Time محل تردید قرار گیرد . به ویژه اگرنوزاد بافاصله کوتاهی پس ازتولد، فوت شده وپاسخگونیز شخص دیگری به جزمادر وی باشد این احتمال افزایش خواهد یافت . داده‌های خام آمارگیریهای نمونه‌ای وسرشماریهای جمعیتی نیز بگونه بودن این خطاها وبه عبارت دیگر وجود نوعی اریب درآنهارا نشان میدهد .

۲ـ روشهای تصحیح وتعدیل داده‌ها

طبیعی است که وجود خطاهای مذکور نباید مانع ازکاربرد این اطلاعات شود . بلکه باید باشناخت نوع وسیزان خطاها کوشش شود اثرات آنهارا ختی کرد. کوششهایی که دراین زمینه به عمل آمده عمدتا " از دوجهت مختلف به مسئله پرداخته است .

اول : طرح سوالهای مناسبتر یاسوالهای کمکی که به طورحال پاسخگو را دریادآوری واقعه مورد نظر

وارائه پاسخ صحیح یاری دهد.

دوم: پیشنهاد روشهای مناسبی برای تصحیح وتعدیل یافته‌ها متناسب با سوالهای مطرح شده.

درهرمورد نظریه‌های نسبتاً متفاوتی ارائه شده وبرآن اساس مدلهایی برای تصحیح یافته شده ارائه شده است. بارها ازاین نظریات چه به لحاظ کارآیی آنها وچه به لحاظ سهولت کاربرد نزد جمعیت شناسان مقبولیت ودرنتیجه رواج بیشتری یافته است. ازآن جمله می‌توان ازروش پیشنهادی Willion Brass و J.T. Rele یادکرد.

ویلیام براس بامسوالیدرمورد تعداد کل فرزندانی که هرزن ازدواج کرده درطول دوران باروری خود به دنیاآورده ونیز تعدادی ازآنها که درزمان آمارگیری زنده‌اند وپذیرش این فرض که اریبی پاسخ بدین سوالها برحسب سن مادر تغییر می‌کند وبه آن وابسته است، شیوه‌ای برای تعدیل یافته‌های سرشماریها وآمارگیریهای نمونه‌ای، چه درمورد باروری وچه درمورد مرگومیر، پیشنهاد کرده است.

وی مطالعات اولیه خودراهنگامی که دریکی ازکشورهای افریقایی ماموریت داشت انجام دادوبعد همین دلیل نیز روش پیشنهادی را ابتدا درچند کشور افریقایی به آزمایش گذاشت که نتایج قابل قبول و بعضاً مطلوبی به دست داد وباتکیه براین تجربیات به تکمیل روش پیشنهادی پرداخت. درحال حاضر این روش در بسیاری ازکشورها ازجمله کشورما ایران نیز به کارگرفته می‌شود.

شیوه‌نگرش J.T. Rele براساس فرض وجود همبستگی قوی بین نسبت کودک بزن (Child- Woman Ratio) ازیکطرف ومیزان تجدید نسل ازطرف دیگر مبتنی است وبه این ترتیب باتوجه به تعداد کودکان ۵- ساله (یا ۹-۵ ساله) وزنان ۴۹-۱۵ ساله (یا ۵۴-۲۰ ساله) بله برآورد رگرسیونی ازمیزان تجدید نسل به دست می‌دهد.

۱-۲- سابقه امر در ایران

درکشور مانیزان علاوه بر کاربرد روشهای پیش گفته وسایر روشهایی که دراین زمینه ارائه شده است، پژوهشهای دیگری نیز به عمل آمده وروش ویژه‌ای نیز پیشنهاد شده است. این تحقیقات که توسط استاد گرامی جناب عباسقلی خواجه نوری که درآن زمان استاد آمار دانشگاه تهران ومؤسسه آموزش عالی آمار بودند همراه با دکتر عباس جامعی معاون وقت وزارت علوم وآموزش عالی واستاد مؤسسه آموزش عالی

آمار انجام شده است . دراین پژوهش ابتدا کوشش شده است ازروشی که در هندوستان مورد استفاده قرارمی‌گرفته ، سود برده شود . روش مزبور با استفاده پرسش " تعداد وقایع ولادت در ۱۲ ماه گذشته " وبرپایه رسم یک منحنی برای جواب مربوط به هریک ازماههای ۱۲ ماه گذشته بوجودآمده بود که با استفاده ازمنحنی مزبور تعداد وقایع گزارش شده برحسب ماه وقوع آن طی ۱۲ ماه گذشته تعدیل می‌شد . براساس مدارک موجودآزمایش روش مزبور درایران " نتیجه خوبی " نداشته است . به همین دلیل روش جدیدی پیشنهاد ودر دوآبادی به آزمایش گذاشته شده که نتایج حاصله " امیدوارکننده " توصیف شده است . دراین روش سوال، " تعداد وقایع در ۱۲ ماه گذشته " از پرسشنامه حذف وبجای‌آن" تاریخ آخرین دفعه‌ای که یک واقعه حیاتی درخانوار رخ داده است " مورد پرسش واقع شده است .

این روش همراه با نظریه ریاضی وروش استنتاج ازداده‌های حاصل ازآن ، در"سمپوزیوم سنتو درباره آمارهای جمعیت شناسی" که درتابستان ۱۳۴۷(درفتا ۲ تا ۱۲ اوامبر ۱۹۶۸) درکراچی تشکیل شده بود ارائه شد ومورد توجه قرار گرفت . ازآن تاریخ تا سال ۱۳۵۵ وانجام سرشماری عمومی نفوس وممکن در سال مزبور ، همین شیوه در پاره‌ای آمارگیریهای جمعیتی مرکزآمارایران مورد استفاده قرار گرفته است .

۲ـ روش پیشنهادی

قبل‌ازهرگونه توضیحی درمورد روش پیشنهادی ، لازم می‌داندکه ازبرادر گرامی جناب آقای دکتر جمشید ی معاون محترم سازمان برنامه وبودجه کشور و رئیس مرکزآمارایران تشکر کند که در واقع پیشنهاد اولیه برای انجام این پژوهش درسال ۱۳۶۷ (پیدازآن بدست آمدن نتایج سرشماری عمومی نفوس وممکن سال ۱۳۶۵) ازطرف ایشان مطرح شدو با حمایتهای همه جانبه مشارالیه به ویژه درامر تهیه اطلاعات مورد نظر ، امکان بانجام رسیدن آن فراهم آمد . علاوه براین جادارد از زحمات همکاران گرامی برادران طه نورالفهیی، مجید نوروزی ، محمداسماعیل ناصحی کارشناسان مدیریت آمارهای اجتماعی، اقتصادی خانوار درمرکزآمارایران به خاطر کمک درامر تهیه برنامه‌های ماشینی و انجام محاسبات مربوط سپاسگزارانه یادشود .

روش پیشنهادی برپایه مقایسه بین " تعداد افراد سفرسالانه " و " متوفیین ۱۲ ماه گذشته "

بنا شده است . هریك از این اطلاعات در اثر آمارگیریهای جمعیتی قابل حصول است ولی اساسا دو چیز که وجود دارد آنست که :

۱ـ چون افراد صفر ساله در واقع بازماندگان متولدین ۱۲ ماه گذشته هستند، لازم است تعداد وقایع فوت مربوط به متولدین ۱۲ ماه گذشته نیز در دست باشد .

۲ـ این اطلاعات باید مربوط به یك جامعه واحد باشد و برای انجام مقایسه‌های کنترلی ضرورتا" باید از یك آمارگیری واحد به دست آمده باشد .

خوشبختانه اطلاعات حاصل از سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ کلیه شرایط مورد نظر را دارا بود و در این بررسی مورد استفاده قرار گرفته است .

اساس تفکر در این روش اینست که اگر در خانواری P_0 نفر فرد صفرساله ، D_0 واقعه فوت مربوط به نوزادان صفرساله گزارش شده است ، باید به تعداد $B = P_0 + D_0$ واقعه ولادت در ۱۲ ماه گذشته گزارش شده باشد ، به طوری که همواره رابطه $B = P_0 + D_0$ (۱) برقرار باشد .

هرچند مطلب به ظاهر ساده می‌نماید ، ولی با توجه به نوع اطلاعاتی که از سرشماری در دست بود، دو مشکل عمده راه را برای ادامه کار تنگ کرده بود .

ـ اول آنکه در سرشماری سال ۱۳۶۵ سوالی مربوط به تعداد افراد، از جمله تعداد افراد صفرساله از سرپرست یا یکی از اعضای خانوار پرسیده می‌شد که می‌توانست فردی غیر از مادر آن طفل باشد . درحالیکه سوال مربوط به تعداد نوزادان در ۱۲ ماه گذشته (یا ۳۶۵ روز) گذشته از زنانی که نوزادان مزبور را به دنیا آورده بودند، پرسیده می‌شد . بنابراین ، چنانچه در فاصله زمانی بین تولد نوزاد و انجام مصاحبه با خانوار ، مادر نوزاد مزبور فوت یا به هر دلیلی از جمعیت خانوار مورد آمارگیری خارج می‌شد، واقعه ولادت در این خانوار به ثبت نمی‌رسید ولی وجود فرد صفرساله گزارش می‌شد .

ـ دوم اینکه ، شیوه پرسش در مورد سن افراد فوت شده ، برای استفاده مورد نظر در این روش تسجیلی، طراحی نشده بود و با آن مناسبت نداشت ، به این ترتیب که چون فقط سن فرد متوفی پرسیده شده بود، امکان داشت فردی که در ۱۲ ماه گذشته فوت شده و هنگام فوت نیز کمتر از یکسال (صفرساله) بوده است

نوزما" متولد ۱۲ ماه گذشته نبوده باشد. طبیعی است که وجود چنین حالتی در برقرای رابطه

$P_{ij} + P_{ij} = 1$ اخلال ایجاد خواهد کرد.

درمورد مشکل اول، یعنی امکان عدم حضور مادر نوزادی که به عنوان فرد مفرساله گزارش شده است،درنتیجه عدم ثبت واقعه تولد، اقدامی به عمل نیامد. چراکه انتظار می‌رفت حالت بسیار نادری شد وعلاوه برآن هیچگونه اطلاعی برای تعدیل دردست نبود. البته انتظار می‌رود دریکی از آمارگیریهای نمونه‌ای سوالات به گونه‌ای طراحی شوندکه فراوانی چنین حالتی نیز به دست آید. در واقع چنین سوالی درطرح آمارگیری جاری جمعیت سال ۱۳۷۰ پیش بینی شده بود ولی به لحاظ تراکم سوالات ازآن صرفنظر شد. ولی درمورد اشکال دوم یعنی احتمال اینکه طی ۱۲ ماه قبل ازآمارگیری فوت مربوط به طفل صفر ساله‌ای گزارش شده باشد بدون آنکه آن طفل متولد ۱۲ ماه گذشته بوده باشد، بررسیهایی انجام شد وتعدیل لازم به عمل آمدکه برای احتراز از اطاله مطلب ازذکر آن خودداری می‌شود.

۱-۲-مراحل اجرایی روش پیشنهادی

به طوری که گفته شد در روش پیشنهادی ۳ عامل دخالت داشته است:

۱- وجود فرد صفرساله در بین اعضای خانوار

۲- گزارش ولادت طی ۱۲ ماه گذشته

۳- گزارش فوت برای افراد صفرساله طی ۱۲ ماه گذشته.

آنچه که ازنتایج سرشماری عمومی نفوس ومسکن قابل حصول بود، ترکیب یک به یک، دوبه دو سه به سه این اقلام بود. به همین دلیل هریک از ۱۰ میلیون خانوار کشور جداگانه مورد بررسی قرارگرفته وازاین دیدگاه یکی از ۸ حالت ذیل درآن مشخص شد تا فراوانی هریک ازاین حالات به دست آید. برای سهولت استفاده از ارقام تصمیم گرفته شد که چنانچه درخانواری بیش ازیک فرد صفرساله وجود داشته یا بیش از یک واقعه فوت یا ولادت گزارش شده بوده است بسته به مورد آن رابه یکی ازهشت حالت تجزیه کرده و درمجموع نتیجه کاررا به صورت وجود یاعدم وجود یکی ازحالات موردنظر قابل ارائه نمائیم. به این ترتیب جدول زیر به دست آمد:

شماره ردیف	شرح حالت: وجود فرد صفر سالگی در بین اعضای خانوار	گزارش ولادت طی ۱۲ ماه گذشته	گزارش واقعه فوت مرتبط با فرد صفر سالگی طی ۱۲ ماه گذشته	فراوانی	ارزیابی
۱	+	+	-		قابل قبول
۲	-	+	+		قابل قبول
۳	-	-	-		قابل قبول
۴	+	-	-		قابل بررسی
۵	+	-	+		قابل بررسی
۶	-	-	+		قابل بررسی
۷	-	+	-		
۸	+	+	+		قابل بررسی

برپایه یافته‌های جدول شماره ۱ که درواقع مغیرغم تراکم فعالیتهای استخراجی دردفتر سیستم و برنامه نویسی واستخراج مرکزآمار به همت برادر گرامی جناب‌سرکاروندی مدیرکل دفتر مزبور به دست‌آمد، امکان اجرای تعدیل مورد نظر فراهم گردید. براساس نظریه‌ای که ابراز شده، جز درمورد سه حالت قابل قبول ردیفهای ۱، ۲ و ۳ جدول شماره ۱، اطلاعات زیر درمورد بقیه حالتها به عمل آمد.

جدول شماره ۱ـ روش پیشنهادی برای اصلاح حالات قابل بررسی

شماره ترتیب	شـــرح حالـــت			روش اصـــلاح
	وجودفرد صفرساله ویابین اعضاء خانوار	گزارش ولادت طی ۱۲ ماه گذشته	گزارش واقعه فوت فرد صفر ساله طی ۱۲ ماه گذشته	
(۱)	(۲)	(۳)	(۴)	(۵)
۱	+	-	-	به تعداد فراوانی به تعداد موالید اضافه شود
۲	-	+	-	به تعداد فراوانی ضربدر (۱+$\frac{۱۲}{۲۳}$) به تعداد موالید اضافه شود
۳	-	-	+	$\frac{۱۲}{۲۳}$ فراوانی به تعداد موالید اضافه شود
۴	+	-	+	به تعداد فراوانی ضربدر ($\frac{۱۲}{۲۳}$ +۱) به تعداد موالید اضافه شود
۵	+	+	+	به تعداد فراوانی ضربدر ($\frac{۱۲}{۲۳}$ +۱) به تعداد موالید اضافه شود

* $\frac{۱۲}{۲۳}$ میزان موالید نوزادان طی یکسال اول زندگی

خلاصه آنکه اگر تعاریف زیر را در نظر بگیریم:

P_1BD فراوانی حالتی که درآن وجود فرد صفرساله و واقعه تولد در ۱۲ ماه گذشته گزارش شده‌باشد. (متناظر با سطراول جدول شماره ۱).

DDD_0 فراوانی حالتی که درآن فقط واقعه ولادت در ۱۲ ماه گذشته و " واقعه فوت مربوط فـــــرد صفرساله " در ۱۲ ماه گذشته ثبت شده است (متناظر با سطردوم جدول شماره ۱)

P_0DD فراوانی حالتی که درآن فقط وجود فرد صفر ساله در پرسشنامه درج شده است (متناظر بـا سطر چهارم جدول شماره ۱)

DDD فراوانی حالتی که درآن فقط واقعه تولد در ۱۲ ماه گذشته گزارش شده باشد (متناظر با سطر هفتم جدول شماره ۱)

DDD_0 فراوانی حالتی که درآن فقط فوت فرد صفرساله در پرسشنامه آمده باشد (متناظر با سطرششم جدول شماره ۱)

$P_0D\ D_0$ فراوانی حالتی که درآن وجودفرد صفر ساله ونیز وقوع فوت مربوط به فرد صفرساله در ۱۲ ماه گذشته در پرسشنامه تهیه شده باشد (متناظر با سطر پنجم جدول شماره ۱)

$P_1B\ D_0$ فراوانی حالتی که درآن هم وجود فرد صفر ساله وهم وقوع تولد وهم وقوع فوت فرد صفر سالـه در ۱۲ ماه گذشته در پرسشنامه به ثبت رسیـــده باشد. (متناظر باسطر هشتم جدول شماره ۱)

درآنصورت مدل پیشنهادی به شرح زیر خلاصه خواهد شد:

$$0 = P_1BD + DDD_0 + P_0DD + \left(1 - \frac{1^{m_0}}{2}\right) DDD + \left(\frac{12}{12}\right) DDD_0 + \left(1 + \frac{12}{24}\right)(P_0DD_0 + P_1BD_0)$$

پس از انجام اصلاحات لازم به شرح متون ذیجدول ۲، تعداد موالید اعلام و بر آن اساس میزان خـــــام موالید برای نقاط شهری و روستایی کشور و نیز هریک از استانها محاسبه شده و نتیجـــه در جدول شماره ۲ آمده است.

جدول شماره ۲ـ میزانهای خام موالید محاسبه شده برمبنای اطلاعات تعدیل شده به روش پیشنهادی

نام استـــــــان	میزان خـــــام موالیـــد (در هزار)	
	مناطق شهری	مناطق روستائی
جمع کل	۳۹/۴۲	۴۲/۸۲
تهران	۳۶/۳۸	۴۲/۷۴
مرکزی	۴۰/۵۰	۴۸/۰۲
گیلان	۳۱/۹۱	۴۲/۸۲
مازندران	۳۵/۲۰	۴۷/۹۷
آذربایجان شرقی	۳۷/۱۰	۴۱/۵۰
آذربایجان غربی	۳۷/۰۱	۴۸/۷۴
باختران	۴۱/۹۴	۴۴/۰۱
خوزستان	۴۲/۵۳	۵۰/۹۷
فارس	۳۹/۴۶	۴۶/۲۲
کرمان	۴۱/۹۳	۴۳/۵۱
خراسان	۴۴/۱۱	۴۲/۲۱
اصفهان	۳۷/۹۵	۴۲/۰۵
سیستان وبلوچستان	۵۴/۸۵	۴۷/۹۶
کردستان	۴۲/۳۸	۴۴/۱۱
همدان	۳۸/۸۰	۴۷/۱۰
چهارمحال وبختیاری	۴۲/۲۵	۴۹/۴۵
لرستان	۴۳/۱۸	۴۵/۹۷
ایلام	۴۲/۸۹	۴۵/۸۹
کهگیلویه وبویراحمد	۵۴/۹۶	۵۶/۱۰
بوشهر	۴۶/۶۶	۴۷/۷۰
زنجان	۴۲/۶۶	۴۲/۵۷
سمنان	۳۸/۶۴	۴۵/۵۳
یزد	۳۸/۲۹	۴۷/۸۶
هرمزگان	۴۶/۲۶	۵۰/۳۶

۴ـ آزمون نتایج

هرچندباتوجه به شناختی که ازمیزان موالید در کشور وهریک ازاستانها وجود دارد ، ارزیابی نتایج براساس شناخت قبلی میسر بود ، ولی به لحاظ قطعیت بخشیدن به این ارزیابی میزان خام موالید با استفاده از ۵ منبع دیگر نیز محاسبه باکردآوری شده وبرای ارزیابی نتایج حاصل ازروش پیشنهادی با آنها مورد مقایسه قرار گرفته است . این چهار منبع عبارتست از :

۱ـ نتایج حاصل ازکاربرد روش براس Brass

۲ـ نتایج حاصل ازکاربرد روش رله J. R . Rele

۳ـ نتایج حاصل براساس آمار منتشره ازسوی سازمان ثبت احوال کشور

۴ـ ارقام حاصل ازمیانگین ۳ روش فوق (موردنظر بخش جمعیت دفتر مطالعات وتحقیقات شهرسازی وسعماری ـ وزارت مسکن وشهرسازی)

۵ـ نتایج حاصل ازکاربرد روش ثبت مضاعف اطلاعات Dual Record System برمبنای یافته‌های سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ ، (که چنانچه فرصت باقی بماند در اینمورد ونیز نتایج جنبی جالبی که این روش به دست داده است توضیح مختصری ارائه خواهد شد .

این ارقام در جدول شماره ۴ آمده است .

نتیجه

مقایسه بین ارقام موجود ونیز باتوجه به شناختی که صاحب نظران وکارشناسان موضوع در مورد ویژگیهای جمعیت کشور دارند حاصل این بررسی را می‌توان به شرح زیر خلاصه کرد.

۱- روش ویلیستام‌براس Grass ، ارقام را بیش ازحد مورد انتظار به دست‌می دهد و در واقع Overestimate می‌کند.

۲- نتایج حاصل از روش رله Rele نیز هرچند نه به اندازه روش براس ولی به هرحال ارقام بالایی را به دست می‌دهد.

۳- نتایج حاصل از اطلاعات مربوط به سازمان ثبت احوال کشور به دلایل پیش گفته ارقامی به مراتب پائین تراز حد معقول وسازگار با سایر منابع به دست می‌دهد.

۴- میانگین ارقام حاصل از ۳ روش فوق، هرچند ازنظر نگارنده روش توجیه پذیری نیست ولی ارقام قابل قبولی به دست داده است.

۵- ارقام حاصل از کاربرد روش ثبت مضاعف اطلاعات نیز کم‌وبیش ارقام بالایی به دست داده است. هرچند این پدیده دراین روش به شدت روشهای براس ورله نیست ولی می‌تواند معلول عوامل شناخته شده‌ای باشد که درمدیریت آمارهای اجتماعی-اقتصادی خانوار مرکزآمارایران کاربر روی آن ادامه دارد.

۶- ارقام حاصل ازکاربرد روش پیشنهادی ، علاوه بر آنکه دربسیاری موارد با ارقام حاصل ازمیانگین ۳ روش براس، رله و ثبت احوال سازگاری دارد ودرموارد کاملا" برآن منطبق است ، در پاره‌ای موارد اشعامات موجود درآن روش رانیز ندارد. به عنوان مثال پائین بودن میزان خام موالید در استان آذربایجان شرقی به ویژه مناطق روستایی استان مزبور که درروش یادشده رقمی در حدود میزان خام موالید درمناطق شهری استان تهران است، در روش پیشنهادی وجود ندارد. هرچند که خودنیزعاری از نقاط ضعف نیست. مثلا" در استانهای سمنان، زنجان و یزد که میزان خام موالید در نقاط شهری بیشتر ازنقاط روستایی به دست آمده است قابل تامل است . ولی به هرحال در اولین گام خود، نتایج قابل قبولی را به دست داده وچنین می‌نماید که چنانچه مطالعه دراین زمینه ادامه یابد نتایج مطلوبتری به دست خواهد داد.

جدول ۲ـ تخمین سه میزان شاخص مواليد، میزانهای شهری و روستایی کشور براساس روشهای مختلف (به هزار)

نام استان	روش بهشتیادی		روش بهرامی		ثبت سجل احوال		میانگین ارقام ثبت احوال و روش بهشتیادی		ثبت احوال		روش برآورد	
	شهری	روستایی	شهری	روستایی	شهری	روستایی	شهری	روستایی	شهری	روستایی	شهری	روستایی
کل جمع	۳۹/۴۲	۴۲/۸۲	۴۲/۱۲	۵۱/۸۹	۴۱/۴۶	۴۷/۱۲	۴۱/۴	۴۲/۸	۴۱/۴	۵۰/۴۲	۳۸/۹	۵۰/۶
تهران	۳۶/۳۸	۴۴/۷۲	۳۷/۴۳	۴۸/۴۹	۳۵/۴	۴۸/۱۰	۳۶/۵	۴۳/۴	۳۶/۴	۴۳/۴	۳۲/۸	۴۵/۰
مرکزی	۴۰/۵۰	۳۸/۰۳	۵۱/۱۵	۵۴/۵۲	۴۰/۹	۴۷/۰	۳۸/۴	۴۸/۴۰	۳۸/۴	۴۳/۴۱	۳۹/۱	۴۱/۸
گیلان	۳۱/۹۱	۳۲/۸۲	۳۸/۴۸	۴۳/۸۹	۳۱/۲	۳۲/۲	۳۱/۴	۳۲/۸	۳۵/۸	۳۴/۱	۳۴/۲	۳۷/۲
مازندران	۳۵/۷۰	۲۷/۹۳	۲۲/۱۶	۲۵/۶۹	۲۵/۰	۲۸/۰	۲۷/۲	۴۰/۴۲	۴۵/۲	۲۴/۱	۲۸/۴	۴۲/۲
آذربایجان شرقی	۲۷/۱۰	۲۱/۵۰	۲۲/۲۲	۵۲/۱۸	۲۶/۰	۴۰/۱۸	۲۶/۱	۲۵/۲	۲۲/۲	۴۵/۲	۴۰/۲	۵۰/۵
آذربایجان غربی	۲۷/۰۱	۲۸/۲۲	۲۲/۸۸	۵۲/۲۸	۴۰/۱	۴۱/۸	۲۷/۹	۲۴/۸	۲۳/۲	۲۹/۵	۲۸/۲	
باختران	۲۱/۹۲	۲۴/۰۱	۲۲/۲۲	۵۲/۹۲	۵۰/۲	۵۲/۲	۲۲/۰	۴۴/۱	۲۱/۲	۲۹/۲	۲۴/۷	۵۰/۲
خوزستان	۴۴/۵۲	۵۰/۹۲	۲۸/۲۲	۵۵/۸۲	۵۲/۶	۶۹	۲۲/۵	۶۱/۲	۲۲/۵	۵۲/۵	۶۲/۲	۵۲/۶
فارس	۲۹/۹۶	۲۶/۲۷	۲۵/۶۶	۵۲/۲۲	۲۷/۲	۵۴/۹	۲۹/۱	۲۷/۹	۲۸/۲	۲۲/۲	۲۰/۸	۵۲/۲
کرمان	۲۱/۲۶	۲۲/۵۱	۲۸/۶۰	۵۱/۰۹	۲۷/۲	۴۷/۵	۲۲/۲	۲۶/۴	۲۸/۲	۴۵/۲	۲۲/۴	۵۱/۲
خراسان	۲۴/۱۱	۲۲/۲۱	۲۸/۲۱	۵۶/۵۱	۲۷/۲	۴۲/۰	۲۳/۶	۲۰/۱	۲۲/۱	۲۰/۲	۲۴/۲	۵۱/۸
اصفهان	۲۷/۹۵	۲۴/۰۵	۲۱/۲۰	۵۸/۲۱	۲۰/۲	۲۷/۲	۲۵/۲	۲۶/۸	۲۲/۲	۵۰/۸	۲۰/۲	۴۸/۲
سیستان و بلوچستان	۵۲/۸۵	۲۷/۹۲	۵۲/۹۸	۵۲/۱۵	۵۰/۲	۲۷/۶	۵۲/۲	۸۷/۱	۵۲/۸	۶۶/۲	۵۸/۵	۵۰/۲
کردستان	۲۲/۲۸	۲۲/۱۱	۲۵/۰۹	۵۵/۲۱	۲۸/۷	۲۸/۵	۲۲/۱	۲۸/۰	۵۱/۵	۴۷/۰	۲۴/۲	۵۲/۶
کهگیلویه	۲۸/۸۰	۲۷/۱۰	۲۲/۵۱	۵۴/۰۰	۲۲/۲	۵۵/۹	۲۸/۲	۲۲/۸	۲۳/۴	۴۳/۵	۲۱/۵	۵۲/۹
چهارمحال و بختیاری	۲۲/۲۵	۲۹/۲۵	۲۷/۹۵	۵۲/۲۰	۲۲/۱	۵۲/۲	۲۲/۹	۸۴/۰	۲۰/۲	۲۵/۲	۵۸/۲۰	
لرستان	۲۲/۱۸	۲۵/۶۲	۲۷/۲۱	۲۵/۸۱	۲۵/۱	۵۱/۰	۲۴/۲	۲۸/۲	۲۴/۲	۲۷/۲	۲۴/۰	۵۲/۶
ایلام	۲۲/۸۱	۲۵/۸۹	۵۱/۲۱	۵۵/۲۹	۵۱/۱	۵۲/۸	۲۷/۲	۸۲/۱	۲۲/۲	۵۲/۹	۵۱/۲	۵۷۹۶
کهگیلویه و بویراحمد	۲۲/۹۲	۵۶/۱۰	۵۲/۲۲	۶۱/۲۶	۵۹/۲	۵۴/۰	۸۲/۲	۸۲/۲	۵۸/۲	۹۱/۵	۵۰/۰	۲۵۹/۲
بوشهر	۴۴/۲۶	۴۷/۲۰	۵۲/۰۰	۵۴/۲۲	۵۱/۲۷	۵۰/۲	۵۰/۲	۲۵/۰	۵۵/۰	۲۵/۳	۲۸/۸	۵۲/۰
زنجان	۴۲/۲۶	۴۴/۵۲	۲۴/۲۲	۵۶/۲۲	۵۱/۲	۴۹/۶	۲۰/۵	۲۴/۵۵	۲۴/۲	۴۲/۸	۲۰/۰	۵۰/۰
سمنان	۲۸/۲۲	۲۵/۵۲	۲۱/۲۱	۴۲/۲۹	۲۹/۰	۲۵/۵	۲۵/۲	۲۸/۸	۲۰/۲	۲۱/۰	۲۸/۲	۲۱/۰
یزد	۲۸/۲۹	۲۷/۸۲	۲۲/۸۱	۲۷/۰۵	۲۲/۲	۲۲/۸	۲۲/۰	۲۰/۱	۲۵/۲	۲۷/۲	۲۷/۲	۲۱/۸
هرمزگان	۴۶/۲۲	۵۰/۲۲	۵۲/۲۲	۵۲/۰۰	۴۵/۲	۲۲/۶	۲۸/۲	۵۰/۲	۲۱/۲	۲۲/۲	۵۱/۲۲	۶۵/۸

عمده‌ترین مزایایی که برای این روش می‌توان برشمرد عبارتست از :

۱ـ سهولت کاربرد، هرچند در اجرای فعلی به علت عدم پیش بینی اقلام مورد نظر به هنگام طراحی سیستم استخراج ، ذیل به اطلاعات مورد نظر به طول انجامید وزمان وهزینه قابل ملاحظه‌ای صرف آن شد ولی چنانچه این اقلام ازقبل وبه عنوان بخشی ازنتایج سرشماریها وآمارگیریهای نمونه منظور شود به سهولت وبدون هیچگونه هزینه اضافی قابل حصول خواهد بود .

۲ـ برخورداری از منطق قوی وقابل پذیرش

۳ـ به دور بودن از نظرباتی که ممکن است در پاره‌ای جوامع صادق باشد ودرپاره‌ای دیگرازجوامع صادق نباشد .

۴ـ سازگاری داخلی نتایج به دست آمده ونیز سازگاری این نتایج با نظرات کارشناسی و ترکیب سنی جنسی جمعیت کشور .

ازجمله نکاتی که ذکر آن دراینجا لازم می‌نماید، آنکه درکاربرد روش ثبت مضاعف اطلاعات Dual Record System به طورضمنی، میزان پوشش هریک از دوصفت مورد بررسی یعنی "گزارش فرد صفرساله " و " گزارش واقعه ولادت در ۱۲ ماه گذشته " نیز به دست آمده است . این ارقام نشان می‌دهد که در آمارگیریهای جمعیتی در مقیاس سرشماری فقط ۷۸/۲ افراد صفرساله و ۲۰/۵٪ نوزادان کمتراز یکسال گزارش میشوند . البته همانگونه که انتظار می‌رود گزارش افراد صفرساله در نقاط شهری وضع مطلوبتری نسبت به نقاط روستایی دارد . (۹۰ درصد در مقابل ۷۸ درصد) .

ضمن آنکه باید اذعان کرد این ارقام در روزهای پایانی تهیه این گزارش محاسبه شده وهنوز قابلیت بررسی بیشتری دارند ولی از این فرصت استفاده کرده وپیشنهاد می‌کنم که چنین محاسباتی در مورد سیستم ثبت وقایع کشور می‌تواند راهگشای بسیاری ازمسایل مرتبط با این پدیده باشد .

منابع پرداخت مورد استفاده :

۱ـ مرکز آمار ایران، تهران ۱۳۶۶ـ۱۳۶۷ : نتایج تفصیلی سرشماری عمومی نفوس ومسکن مهرماه ۱۳۶۵ـ

(شامل مربوط به کل کشور و هریک از ۲۴ استان)

۱۳۴۸

۲ـ عباسقلی خواجه نوری ، عباس جعفری یک روش جمع‌آوری آمارهای نفوس ، تهران موسسه آموزش عالی آمار

۳ـ

Rele. j. R
Fertility analysis Through extension of stable Population Concepts.

University of California , Berkeley - 1976

۴ـ محمدحسین نجاتیان ، نگاهی تازه به سیستم ثبت مضاعف اطلاعات ، تهران مرکز آمار ایران خرداد ۱۳۶۴

۵ـ حبیب‌الله زنجانی، جمعیت وشهرنشینی درایران، مرکز تحقیقات ومطالعات شهرسازی ومعماری ـ

وزارت مسکن وشهرسازی ، تهران ۱۳۶۸ .

تهیهٔ جدول عمر جمعیّت کشور در سال ۱۳۶۵

طه نوراللهی

مرکز آمار ایران

مقدمه:

باافزایش کاربرد آمار به ویژه توسعه علم احتمالات واستفاده ازآن درآمار، دامنه این علم به آینده نیز گسترش یافته واین امکان رافراهم ساخته است ، که ازطریق آمار نمونه‌ای به آینده گشوده شود ، برای گشودن همین دریچه کوششها وپژوهشهای بسیاری انجام شده و الگوهای مختلفی ساخته شده است ، که ازبین آنها پاره‌ای به دلیل عدم سازگاری، شرایط واقعی جامعه مورد بررسی، به تدریج منسوخ شد ، و پاره‌ای دیگر برعکس هرروز مسیر تکامل راطی‌کرده و برکاربردهای‌آن افزوده شده است ، جدول عمرکه درواقع تحول کمی جمعیت یک جامعه را تحت قانون مرگ‌ومیر مبنایی برای آینده ترسیم می‌کند ازجمله الگوهایی است که علاوه بر مطلوبیتی که درجمعیت شناسی به دست‌آورده و به صورت ابزاری مهم در تحلیلهای جمعیتی درآمده ، درسایرعلوم نیز کاربرد یافته است ، یکی از موضوعات علم جمعیت شناسی، که اغلب برنامه‌ریزی‌های توسعه اجتماعی ـ اقتصادی بدان نیاز دارند، پیش بینی جمعیت برای دوره‌های کوتاه مدت ، میان مدت وبلند مدت است ، دربین روشهای پیش بینی جمعیت ، پیش بینی به روش ترکیبی به لحاظ چند وجهی بودن نوع استفاده ، کاربرد بیشتری دارد ، استفاده ازاین روش بدون آگاهی از سطح مرگ‌ومیرناشی‌از زندگی و ضرایب احتمال بقاء ممکن نیست واینها شاخصهایی هستند که ماتریس جدول عمر به دست می‌آید ،

دلیل عدم کاربردهای فراوان و مهمی که جدول عمر در تحلیلهای جمعیتی دارد، متاسفانه به لحاظ مشکلاتی که در تهیه آن رخ می‌نماید، در کشور ما به ندرت به آن پرداخته شده است. تا آنجا که اطلاع نشان می‌دهد، تاکنون جزء جداولی که براساس یافته‌های طرح اندازه‌گیری رشد جمعیت، که توسط مرکز آمار ایران تهیه شده و جداول عمر خلاصه‌ای که توسط یکی از کارشناسان مدیریت آمارهای اجتماعی و نفوس، در مرکز آمار ایران فراهم آمده و جدول عمری که توسط دانشکده بهداشت دانشگاه تهران و وزارت بهداشت، درمان و آموزش پزشکی تهیه شده است، جدول عمر قابل استفاده‌ای در کشور ارائه نشده است.

بدنبال انتشار نتایج سرشماری سال ۱۳۶۵ و تحولاتی که در سالهای اخیر در کشور ما به وجود آمده، ضرورت تهیه جدول عمر جدید بر پایه آمارهای مرکوط به جمع‌آوری شده در سرشماری سال ۱۳۶۵، احساس می‌شد زیرا دست اندرکاران پیش بینی جمعیتی در محاسبات خود، ناچار به استفاده از همان جداول عمر تهیه شده سالهای قبل بوده‌اند که عمدتا" عدم انطباق آنها با زمان مورد پیش بینی نیازمند اعمال نظرهای کارشناسی، اغلب متفاوت بوده و از دقت پیش بینی‌ها می‌کاهد.

ضرورت دستیابی به این جداول موجب شده، حتی کشورهائیکه فاقد جداول عمر هستند در چنین مواردی ناچار به استفاده از مدلهای جداول عمر پیشنهادی سازمان ملل (تهیه شده در سال ۱۹۵۵) و جداول عمر تنظیمی "کول رمزی" (تهیه شده در سال ۱۹۶۶) هستند، که نحوه انتخاب مناسبترین جدول از بین این مدلها نیز خود نیازمند تجربه کافی در زمینه استفاده از این جداول است ولی در اینجا فقط اشاره می‌کند که این انتخاب، متناسب با خصوصیات جامعه مورد نظر به ویژه سطح مرگومیر آن می‌باشد.

مهمترین عناوین مورد بررسی در این مقاله عبارتند از:

۱- جدول عمر چیست؟

۲- انواع جدول عمر

۳- روشهای تهیه جدول عمر

۴- تحلیل به بررسی نتایج تهیه جدول عمر کشور در سال ۱۳۶۵

بدین منظور ابتدا می‌رود کوششی که در این زمینه به عمل آمده است مورد استفاده اهل تحقیق واقع و با هدایت آنان کامل شود.

۱ـ جدول عمر چیست ؟

جدول عمر ، تاریخ مرگ‌وزندگی مجموعه‌ای ازجمعیت است که بتدریج بامرگ از دورزندگی بیرون خارج می‌شوند،که این راباحتمال بقاء (${}_nP_x$) ازسنی‌به سن دیگر تعریف می‌کنند . به عبارت دیگر جدول‌عمر ، تاریخچه زندگی یک گروه وبایک نسل رابه طور نظری (تئوری) ارائه می‌نماید، زیرابراساس مرگومیر هایی که درطی یکسال خاص واقع شده‌اند جمع‌آوری ومحاسبه شده است . بنابراین جدول عمری‌که براساس اطلاعات یکسال تقویمی ساخته می‌شود معرف مرگومیرکلی جامعه درهمان سال مورد بررسی است .

نقش اصلی‌واساسی جداول‌عمر نشان دادن چگونگی تغییرات جمعیت تحت تاثیر سطح مرگومیر موجود می‌باشد . بنابراین مناسب‌نرخواهدبود که این تغییرات جمعیت به جای آنکه برای کلیه افراد جامعه اعمال‌گردد، باهمان قانون مرگومیر موجود درجامعه، برای عده مشخصی مثلا " ۱۰۰۰ نفر یا ۱۰،۰۰۰ نفر یا ۱۰۰،۰۰۰ نفرعرضه شود . بااین تصهید (مفروض داشتن عده مشخص) می‌توان جداول عمر رامستقل از تغییر تعداد افراد جامعه، تهیه نمود . بنابراین جدول عمر برمبنای این فرض تهیه می‌شود که اگرتعداد مشخصی از افراد دریک لحظه معین باهم متولد شوند، تحت قانون مرگومیر حاکم برجامعه، چگونه خواهند زیست واحتمال زندگی‌ویامرگومیرو... آنها درطول زندگی چه خواهد بود . این تعداد افراد فرضی راکه پایه یامبنای جدول زندگی نامیده‌اندچنانکه قبلا " نیز گفته شد، معمولا " ۱۰۰۰ یا ۱۰،۰۰۰ ویا ۱۰۰،۰۰۰ نفرویاهر توان دیگر از ۱۰ فرض می‌کنند . (درتهیه جدول عمرکشور، استانها ومناطق شهری وروستایی برای سال ۱۳۶۵ تعداد تولد معین اولیه ۱۰۰،۰۰۰ نفر . فرض شده است) . مسلما " این گروه فرضی، برای رسیدن ازسنی به سن دیگر مرتبا " کاهش یافته واین کاهش تحت قانون مرگومیر جامعه است که این جدول برای‌آن جامعه ساخته‌می‌شود .

عده‌ای راکه درشروع سن بعدی هنوز زنده هستند، بازماندگان گروه دراین سن می‌نامند . به عنوان مثال اگر از ۱۰۰،۰۰۰ نفرکه باهم درلحظه معینی متولد شده اند، ۶۰۰۰ نفر درطول اولین سال زندگی فوت نمایند، بازماندگان این گروه درپایان یکسالگی وشروع سال دوم زندگی‌عبارت خواهد بود از ۹۴۰۰۰ نفر بنابراین در دومین سال زندگی به جای ۱۰۰۰۰۰ نفر فقط ۹۴،۰۰۰ نفر تحت مطالعه قرار خواهندبود . مشاهده می‌شود که برای ساختن جدول عمر فقط اطلاع ازنسبت مرگ ومیر ویژه سنی (${}_nm_x$) وتعداد افرادی که گروه اولیه را تشکیل داده‌اند،کافی خواهد بودوسایرارقام متن جدول عمر (ارقام سایرتوابع) یعنی احتمال مرگ (${}_nq_x$) برای همه سنین، احتمال بقاء (${}_nP_x$)،تعداد مرگ میزلا (${}_nd_x$) ، تعداد زندگان (${}_nl_x$)، تعداد نفر-سالهای زندگی درطی یکسال (${}_nL_x$)، تعداد کل‌نفر سالهای زندگی (T_x) برای همه سنین بعدی،متوسط سالهای زندگی (e_x^0) برای هرفرد رابراساس همین اطلاع می‌توان به دست‌آورد .

۲ـ انواع جدول عمر

۱ـ۲ـ جدول عمر گروهی یا نسلی

این جداول برمبنای تعقیب زندگی یک گروه مشخص ازافراد بنا می‌شود. به این ترتیب که یک گروه مثلا "۱۰۰،۰۰۰ نفری که در ۷۵ یا ۸۵ سال پیش باهم زاده شده وبدون مهاجرت تاکنون زندگی کرده باشند، مورد مطالعه قرار می‌گیرند. باتمام مزایا ومحاسنی که این جداول دارند به دلایل مختلف ازجمله نداشتن اطلاعات وارقام مربوط به مرگ ومیر در ۷۵ یا ۸۵ سال قبل، رواج چندانی نداشته وبیشتر جداول عمر نوع دوم (جدول عمر جاری) مرسوم می‌باشد.

۲ـ۲ـ جدول عمر جاری یاجدول عمر در مقطع زمانی

این جداول رایج ترین ومتداولترین نوع جدول عمر است ،که براساس نرخهای مرگ‌ومیر جاری جمعیت بنا میشوند. اصل کلی که در تشکیل جداول عمر جاری مورد استفاده قرار می‌گیرد این است که میزانهای مرگ‌ومیر ویژه سنی برحسب جنس‌برای کلیه افرادی که درآن گروه سنی‌جنسی زندگی کنند، ثابت فرض می‌شود.

جداول عمرناظر ازجاری یاگروهی در دوشکل کامل وخلاصه شده تهیه می‌شوند. جدول عمر کامل، کلیه توابع جدول عمر رابرحسب سنین منفرد شامل می‌شود. درحالیکه جدول عمر خلاصه شده ، معمولا " توابع رابرحسب گروههای سنی پنجساله وباده ساله نشان می‌دهد. البته جداول عمردیگری‌نیز وجود دارند که شاید بتوان آنهارا تلفیقی ازاین دونوع دانست ومبارتند از جداولی که پارامترهای جدول عمر رابرحسب گروههای سنی وبرای سنین منفرد ولی فقط برای سنین معینی (نه کلیه سنین) به دست می‌دهد.بطورخلاصه‌جدول‌عمرراهی‌توان برحسب جنس ، نژاد ، مناطق مختلف یک کشور ، حرفه‌های مختلف وغیره تهیه نمود.

۲- روشهای تهیه جدول عمر

برتهیه جدول عمر روشهای مختلفی به کار گرفته می‌شود که برحسب منبع اطلاعات اولیه می‌توان آنها را به دوگروه عمده، روشهای مستقیم وروشهای غیر مستقیم تقسیم نمود.

درروش مستقیم، آمار مرگ ومیر برحسب جنس وسن مورد نیاز است. تنها در سه مورد روش مستقیم درایران برای محاسبه جدول عمر کشور استفاده شده است.(۱)

روش غیرمستقیم خود برحسب اطلاعات موجود ویا بکارگرفته شده می‌تواند تقسیماتی داشته باشدکه ازجمله اطلاعات اولیه آن مبنی برتوزیع سنی وجنسی جمعیت حاصل ازسرشماریهای عمومی نفوس ومسکن می‌باشد. درایران یک مورد بااستفاده ازاین نوع روش غیر مستقیم برای محاسبه جدول عمر کشور استفاده شده است.(۲)

روش غیرمستقیم دیگر، استفاده ازرابطه سطح مرگومیر با پاره‌ای ازشاخص های اجتماعی، اقتصادی وبهداشتی می‌باشد کهآنها که سوابق نشان می‌دهد، ازاین روش نیز یک مورد برای محاسبه جدول عمرکشور استفاده شده است.(۳)

درصورتیکه ثبت وجمع‌آوری آمارمرگومیر وموالید ازدقت وپوشش کامل برخوردار باشد، ساده‌ترین و کم‌خرج‌ترین روش، روش مستقیم محاسبه جدول عمربااستفاده‌ازآمارهای‌ثبتی‌است. متاسفانه به علت عدم پوشش ودقت کافی آمارهای ثبتی درایران همانند اغلب کشورهای درحال توسعه یاتوسعه نیافته، این آمار نمی‌تواند مستقیما" به عنوان اطلاعات اولیه محاسبات به کاربرده شود.

درعین حال این گونه آمارها به علت عدم تغییر فاحش عوامل موثر در ثبت وجمع‌آوری، برای نشان دادن روند مرگومیر جامعه مفید وبه کارگرفته می‌شود.

---

(۱)- درسالهای ۱۳۵۵، ۱۳۵۶، ۱۳۶۲ و ۱۳۶۴ به ترتیب ازنتایج طرحهای " اندازه‌گیری رشدجمعیت ایران سالهای ۵۵-۱۳۴۲ مرکزآمارایران " ، " طرح میزانهای حیاتی ایران سالهای ۵۲-۱۳۵۲ دانشکده بهداشت دانشگاه تهران " و "طرح آمارگیری نمونه‌ای سال ۱۳۶۳، وزارت بهداشت، درمان وآموزش پزشکی " بااستفاده‌ازروش مستقیم جدول عمرکشورمحاسبه وارائه شده است.

(۲)- درسال ۱۳۶۱ ،ازنتایج سرشماری‌های عمومی‌نفوس‌ومسکن سالهای ۵۵-۱۳۴۵ بااستفاده ازروش غیرمستقیم (ساخت سنی‌وجنسی جمعیت کشور توسط سیدحسن خسرو پاکارتماس مدیریت آمارهای اجتماعی‌ونفوس مرکزآمار ایران جدول عمرکشور محاسبه‌وارائه‌شده‌است.

(۳) درسال ۱۳۶۷، ازرابطه مرگومیربپاره‌ای‌ازشاخصهای اجتماعی، اقتصادی‌وبهداشتی ازروش غیرمستقیم ازتوسط دکتره‌حمیده‌برزابی جدول عمرکشورتهیه‌وارائه‌شده است.

۴- روش مطالعه و بررسی نتایج

درگزارش حاضر، جدول عمر از روش مستقیم با استفاده از آمار مرگ و میر جمع‌آوری شده در سرشماری سال ۱۳۶۵ ، محاسبه و ارائه شده است .

هرگاه آمارهای مرگ‌ومیر، حسب سن و جنس، درست باشد و از درستی و جامعیت کامل برخوردار باشند براساس فرمولهای مربوطه به راحتی می‌توان میزان مرگ‌ومیر ویژه سنی و جنسی را محاسبه و جدول عمر تهیه نمود . درمحاسبه میزان مرگ‌ومیر ویژه سنی و جنسی براساس نتایج سرشماری سال ۱۳۶۵ ، کم‌شماری‌های آشکار تعداد مرگ‌ومیر کاملا " خودداشان می‌دهد . مثلا " میزان مرگ‌ومیر در گروه سنی " ۰ - ۴ ساله براساس نتایج سرشماری درکشور ، ۴/۲ نفر درهزار به دست می‌آید . درحالیکه واقعیت‌های جامعه نشان می‌دهد که میزان واقعی مرگ‌ومیر در این گروه سنی درسال مزبور بالاتر از رقم فوق‌الذکر است .

به دلیل اختلاف بسیار فاحش آمار مرگ‌ومیر سرشماری با واقعیت‌های جامعه برای به دست آوردن تعداد مرگ و میر، از آمارهای مربوط به تعداد فرزندان زنده به دنیا آمده وتعداد فرزندان درحال حاضر (زمان سرشماری) زنده، استفاده به عمل آمده است . اختلاف بین تعداد فرزندان زنده به دنیا آمده وتعداد فرزندان درحال حاضر زنده، مبنای مناسبی برای محاسبه میزان مرگ‌ومیر درجامعه بوده و دراین رابطه به منظور تصحیح میزانهای مرگ‌ومیر به دست آمده از نتایج سرشماری سال ۱۳۶۵، از ضرایب پیشنهادی " ویلیام براس[1] " وجداول تنظیمی " کول ودمنی[2] " استفاده به عمل آمده است .

با استفاده از میزان مرگ‌ومیر ویژه سنی-جنسی (تصحیح شده)، جداول عمر برای کل کشور و مناطق شهری و روستایی محاسبه گردیده‌که درجداول شماره ۱-۶ تا ۹-۶ ارائه شده است . ازبین ستونهای $_nL_x$ و T_x (جمعیت ایستان جدول عمر) که برای پیش بینی‌های جمعیتی مورد استفاده قرار می‌گیرند ، ستون (e_0^0) یا امید زندگی درلحظه تولد یا درابتدای هرگروه سنی نیز شاخص خوبی برای نشان دادن میزان مرگ و میر درتمام سنین می‌باشد چه سطح مرگ‌ومیر درتمام گروههای سنی دومیزان (e_x^0) تاثیر دارد .

بررسی امید زندگی درلحظه تولد برای کل کشور نشان می‌دهد که درایران هرنوزاد دختر متولد به طور متوسط ۵۹/۰ سال امید زندگی دارد . امید زندگی نوزادان دختر درایران نیز همانند سایر نقاط دنیا بیشتر از پسرها و برابر با ۶۰/۰ سال بوده است . درحالیکه نوزادان پسر درهنگام تولد امید زندگی برابر با ۵۸/۲ سال دارند .

---

(1) Willam Brass Method.

(2) Coale and Demeny "Model life Tables" 1966

درمناطق شهری امید زندگی درهنگام تولد ۸/۴ سال بیشتر ازمناطق روستایی است .
به عبارت دیگر، امیدزندگی مناطق شهری ۶۴/۷ ودرمناطق روستایی کشور ۵۴/۲ سال برده است .

بالاترین میزان امید زندگی درلحظه تولد متعلق به زنان شهری است . نوزادان دختر در شهرها امید می‌رود که به طور متوسط ۶۴/۳ سال زندگی نمایند درحالیکه نوزادان دختر درمناطق روستایی به طور متوسط ۵۴/۶ سال انتظار حیات دارند . نوزادان پسر در روستاها نیز با ۵۴/۰ سال کمترین امید زندگی رادارا می‌باشند . درحالیکه درشهرها انتظار می‌رود نوزادان پسر به طور متوسط ۶۲/۰ سال عمر نمایند .

باگذشت یکسال از زندگی هرنوزاد پواردشدن به گروه سنی ۴-۱ ساله درمناطق شهری به طور متوسط درحدود ۳/۰ سال ودرمناطق روستایی نزدیک به ۵/۰ سال‌به امید زندگی افزوده می‌شود . این امر به دلیل بالابودن میزان مرگومیر اطفال کمتر از یکسال درمناطق روستایی کشور است .

بررسی میزانهای مرگومیر درسال ۱۳۶۵ نشان می‌دهدکه مناطق روستایی ایران هنوز ازمیزان مرگومیر بالایی به خصوص درمیان کودکان کمترازیکسال برخوردار بوده است .

جداول ۱، ۲و۳ امید زندگی دربدوتولد، میزانهای مرگ‌ومیر عمومی واطفال کمترازیکسال را قبل وبعد ازکاربرد روش " ویلیام براس " نشان می‌دهد . همچنین جدول شماره ۴ ، امید زندگی (e_0^0) کشور، مناطق شهری وروستایی رابرحسب سن وجنس وجدول شماره ۵ امید زندگی دربدو تولد دراستانهای کشور رابرحسب جنس به تفکیک مناطق شهری وروستایی ارائه می‌دهد . به دلیل ضیق وقت از بررسی این جداول دراین گزارش اجتناب می‌شود .

جدول ۱ـ امید زندگی در بدو تولد (e_0) بمعدل کاربرد روش ویلیام براس ـ سال ۱۳۶۵

(سال)

شـــــرح	دوجنس	مرد	زن
کل کشور	۵۸/۹۹	۵۸/۲۸	۵۹/۷۰
مناطق شهری	۶۲/۶۸	۶۲/۰۶	۶۳/۳۰
مناطق روستایی	۵۴/۲۲	۵۳/۹۹	۵۴/۶۲

جدول ۲ـ مقایسه میزان مرگ ومیر عمومی قبل وبعدازکاربرد روش ویلیام براس ـ سال ۱۳۶۵

(درهزار) .

شـــــرح	دوجنس		پسر		دختر	
	اصلاح نشده	اصلاح شده	اصلاح نشده	اصلاح شده	اصلاح نشده	اصلاح شده
کل کشور	۲/۸۲	۱۰/۶۸	۳/۶۳	۱۰/۹۴	۲/۰۰	۱۰/۴۰
مناطق شهری	۲/۹۳	۸/۵۹	۳/۸۲	۸/۶۳	۲/۲۸	۸/۵۴
مناطق روستایی	۲/۶۷	۱۳/۸۱	۲/۰۷	۱۴/۳۸	۱/۹۱	۱۳/۲۲

جدول ۳ـ مقایسه میزان موگومیر اطفال قبل وبعدازکاربرد روش ویلیام براس ـ سال ۱۳۶۵ (درهزار)

شـــــرح	دوجنس		پسر		دختر	
	اصلاح نشده	اصلاح شده	اصلاح نشده	اصلاح شده	اصلاح نشده	اصلاح شده
کل کشور	۱۸/۱۰	۲۸/۳۰	۲۰/۳۹	۸۰/۶۳	۱۵/۶۹	۲۵/۸۹
مناطق شهری	۱۸/۱۵	۶۰/۲۱	۲۱/۰۵	۶۱/۵۸	۱۵/۱۵	۵۸/۸۰
مناطق روستایی	۱۸/۱۲	۱۰۲/۴۲	۱۹/۸۰	۱۰۵/۷۴	۱۶/۴۲	۱۰۰/۹۸

جدول ۴ـ امید زندگی (بدو تولد) کشور، مناطق شهری و روستایی برحسب سن و جنس سال ۱۳۶۵

گروه سنی	کشور			مناطق شهری			مناطق روستایی		
	دوجنس	مرد	زن	دوجنس	مرد	زن	دوجنس	مرد	زن
۰ سالگی	۵۸/۹۹	۵۸/۲۸	۵۹/۷۰	۶۲/۶۸	۶۲/۰۶	۶۳/۳۰	۵۴/۲۴	۵۳/۹۹	۵۴/۶۳
۱ـ۴ سالگی	۶۲/۷۲	۶۲/۲۶	۶۳/۳۹	۶۵/۵۵	۶۴/۹۸	۶۶/۱۳	۵۹/۱۳	۵۸/۹۹	۵۹/۴۰
۵ـ۹ سالگی	۶۰/۷۷	۶۰/۰۷	۶۱/۵۹	۶۲/۹۳	۶۲/۱۹	۶۳/۶۶	۵۷/۹۹	۵۷/۵۸	۵۸/۵۵
۱۰ـ۱۴ سالگی	۵۶/۲۰	۵۵/۶۲	۵۷/۱۶	۵۸/۴۰	۵۷/۶۴	۵۹/۱۴	۵۳/۸۶	۵۳/۳۵	۵۴/۴۶
۱۵ـ۱۹ سالگی	۵۱/۸۵	۵۱/۰۸	۵۲/۷۴	۵۳/۷۴	۵۲/۹۶	۵۴/۵۲	۴۹/۲۱	۴۸/۸۸	۵۰/۱۱
۲۰ـ۲۴ سالگی	۴۷/۴۷	۴۶/۶۸	۴۸/۲۹	۴۹/۲۱	۴۸/۴۴	۵۰/۰۲	۴۵/۲۱	۴۴/۶۲	۴۵/۹۷
۲۵ـ۲۹ سالگی	۴۳/۱۹	۴۲/۲۵	۴۴/۱۲	۴۴/۸۴	۴۴/۰۷	۴۵/۶۲	۴۱/۲۰	۴۰/۵۸	۴۱/۹۹
۳۰ـ۳۴ سالگی	۳۸/۹۸	۳۸/۲۰	۳۹/۰۰	۴۰/۲۶	۳۹/۶۶	۴۱/۲۸	۳۷/۱۹	۳۶/۵۱	۳۸/۰۶
۳۵ـ۳۹ سالگی	۳۴/۷۸	۳۳/۹۶	۳۵/۸۵	۳۶/۱۰	۳۵/۲۶	۳۶/۹۶	۳۳/۳۲	۳۳/۴۷	۳۴/۱۵
۴۰ـ۴۴ سالگی	۳۰/۴۲	۲۹/۷۹	۳۱/۰۴	۳۱/۷۸	۳۰/۹۲	۳۲/۶۸	۲۹/۲۸	۲۸/۴۰	۳۰/۲۶
۴۵ـ۴۹ سالگی	۲۶/۵۰	۲۵/۲۳	۲۷/۶۷	۲۷/۵۶	۲۶/۷۰	۲۸/۲۶	۲۵/۲۰	۲۴/۶۱	۲۶/۳۲
۵۰ـ۵۴ سالگی	۲۲/۶۳	۲۱/۸۵	۲۳/۴۷	۲۳/۴۹	۲۲/۶۶	۲۴/۲۲	۲۱/۶۳	۲۰/۹۳	۲۲/۵۲
۵۵ـ۵۹ سالگی	۱۸/۹۰	۱۸/۲۱	۱۹/۸۵	۱۹/۶۳	۱۸/۸۲	۲۰/۲۸	۱۸/۰۶	۱۷/۲۶	۱۸/۸۶
۶۰ـ۶۴ سالگی	۱۵/۲۱	۱۴/۸۳	۱۶/۲۲	۱۶/۰۳	۱۵/۲۸	۱۶/۲۲	۱۴/۲۲	۱۴/۲۲	۱۵/۲۱
۶۵ـ۶۹ سالگی	۱۲/۲۷	۱۱/۸۰	۱۲/۹۲	۱۲/۲۵	۱۲/۲۲	۱۳/۲۸	۱۱/۷۱	۱۱/۳۲	۱۲/۱۱
۷۰ـ۷۴ سالگی	۹/۲۶	۹/۱۱	۹/۹۷	۹/۸۲	۹/۲۲	۱۰/۲۰	۹/۰۲	۸/۷۹	۹/۲۲
۷۵ـ۷۱ سالگی	۷/۰۶	۶/۸۱	۷/۴۹	۷/۳۲	۷/۰۶	۷/۵۲	۶/۷۲	۶/۵۲	۶/۹۲
۸۰ سالگی و بیشتر	۵/۰۰	۴/۸۵	۵/۱۲	۵/۲۲	۵/۰۶	۵/۲۹	۴/۷۲	۴/۶۲	۴/۸۲

جدول شاخص امید زندگی در بدو تولد (e0) در استانهای کشور به تفکیک مناطق شهری و روستایی و جنس
سال ۱۳۶۵

نام استان	مناطق شهری		مناطق روستایی	
	مرد	زن	مرد	زن
کشور	۶۲/۰۶	۶۳/۲۰	۵۲/۹۹	۵۴/۶۲
آذربایجان	۶۵/۵۰	۶۷/۲۲	۶۰/۶۵	۶۱/۳۰
مرکزی	۶۰/۸۷	۶۲/۰۸	۵۰/۷۶	۵۱/۰۸
گیلان	۶۵/۸۶	۶۷/۶۵	۵۹/۹۲	۶۱/۰۵
مازندران	۶۳/۶۱	۶۵/۰۰	۵۲/۷۲	۵۱/۲۳
آذربایجان شرقی	۵۸/۲۲	۵۹/۲۸	۵۱/۲۴	۵۱/۶۵
آذربایجان غربی	۵۷/۵۶	۵۸/۴۵	۴۹/۷۲	۴۹/۸۸
باختران	۶۱/۶۲	۶۲/۹۰	۵۲/۷۷	۵۴/۲۵
خوزستان	۶۵/۱۲	۶۷/۰۰	۵۷/۴۹	۵۸/۲۲
فارس	۶۵/۱۲	۶۶/۲۲	۵۹/۰۵	۶۰/۱۲
کرمان	۶۱/۱۱	۶۲/۲۲	۵۶/۷۱	۵۷/۵۵
خراسان	۵۶/۵۱	۵۷/۶۵	۴۶/۸۸	۴۶/۹۰
اصفهان	۶۱/۹۵	۶۳/۲۵	۵۵/۹۸	۵۶/۸۰
سیستان و بلوچستان	۵۹/۲۲	۶۰/۲۸	۵۵/۹۶	۵۶/۲۳
کردستان	۵۲/۷۲	۵۵/۵۵	۴۴/۶۵	۴۴/۶۳
همدان	۶۱/۰۱	۶۲/۲۳	۵۲/۷۲	۵۲/۲۵
چهارمحال و بختیاری	۶۲/۰۱	۶۴/۲۰	۵۷/۸۵	۵۸/۱۰
لرستان	۶۲/۰۲	۶۳/۲۰	۵۵/۲۱	۵۶/۸۰
ایلام	۶۰/۵۱	۶۱/۷۰	۵۲/۶۶	۵۸/۵۵
کهگیلویه و بویراحمد	۶۲/۴۰	۶۴/۸۳	۵۵/۲۵	۵۶/۰۸
بوشهر	۶۴/۰۲	۶۵/۵۰	۵۸/۲۷	۵۹/۲۲
زنجان	۵۹/۲۲	۶۰/۲۹	۵۰/۶۲	۵۰/۵۱
سمنان	۶۲/۰۲	۶۴/۲۲	۵۲/۲۵	۵۵/۰۵
یزد	۶۲/۶۰	۶۳/۹۸	۵۶/۱۰	۵۶/۸۵
هرمزگان	۶۲/۲۰	۶۴/۸۱	۵۷/۵۲	۵۸/۲۸

توضیح ستونهای جدول عمر مختصر (Abridged Life Table)

شماره ستون در جدول عمر	عنوان ستون	شرح ستون	فرمول محاسبه
۱	گروههای سنی (X_{X+n})	گروههای سنی، عبارت از فاصله زمانی بین دو سن بیان شده در ابتدای دو گروه متوالی می‌باشند، که در آن (X) حد پائین و $(X+n)$ حد بالای یک گروه سنی است	ـــــ
۲	میزان مرگ و میر ویژه سنی $({}_nm_X)$	این ستون میزان مرگ و میر ویژه سنی است که در آن (X) آغاز گروه سنی و $(X+n)$ تعداد سالهای گروه سنی است.	${}_nm_X = \frac{{}_nd_X}{{}_nP_X}$ که در آن: ${}_nd_X$ = تعداد مرگ و میر در هر گروه سنی ${}_nP_X$ = تعداد جمعیت گروه سنی در وسط سال
۳	احتمال مرگ $({}_nq_X)$	این ستون، نسبت افرادی که در یک نسل که در شروع هر فاصله سنی (X) زنده بوده و قبل از رسیدن به آخر آن فاصله سنی $(X+n)$ می‌میرند.	در این بررسی از فرمول زیر استفاده گردیده است: $${}_nq_X = \frac{2n({}_nm_X)}{2+n({}_nm_X)}$$ که در آن: n = فاصله گروه سنی ${}_nm_X$ = میزان مرگ و میر ویژه سنی
۴	تعداد زندگان $({}_nl_X)$	این ستون، تعداد افراد بازمانده از ۱۰۰۰۰۰ نفر اولیه را در شروع هر دوره سنی نشان می‌دهد، مقادیر $({}_nl_X)$ بر مبنای اندازه‌های $({}_nq_X)$ متناظر محاسبه می‌شوند.	در این بررسی از فرمول زیر محاسبه گردیده است: $${}_nl_{X+n} = {}_nl_X - {}_nd_X$$ که در آن: ${}_nd_X$ = تعداد مرگ و میر در هر گروه سنی ${}_nl_X$ = تعداد افرادی که در شروع سن (X) سالگی زنده مانده‌اند. ${}_nl_{X+n}$ = تعداد افرادی که در شروع سن $(X+n)$ سالگی زنده مانده‌اند.

توضیح ستونهای جدول عمر خلاصه (Abridged life table)

شماره ستون جدول عمر	عنوان ستون	شـــــرح ستون	فرمول محاسبه
۵	تعداد درگذشتگان (${}_nd_x$)	این ستون، معرف تعداد فوت شدگان بین سنین (x) و ($x+n$) می‌باشد.	در این بررسی از فرمول زیر محاسبه گردیده است : $${}_nd_x = {}_nl_x \times {}_nq_x$$
۶	نسبت بازماندگی (${}_nP_x$)	این ستون، معرف احتمال زنده ماندن یا نسبت کسانی که در هر فاصله سنی زنده می‌مانند، می‌باشد.	در این بررسی از فرمول زیر محاسبه گردیده است : $${}_nP_x = \frac{{}_nL_{x+n}}{{}_nL_x}$$ ${}_nL_{x+n}$ = تعداد نفر سالهای زندگی شده بین سن (x) و ($x+n$) ${}_nL_{x+n}$ = تعداد نفر سالهای زندگی شده بین سن (x) و ($x+n$) ${}_nP_x$ = نسبت بازماندگی (احتمال بقاء) فرمول محاسبه فوق برای کلیه سنین به استثنای ۵۰ ساله و ۸۰+ ساله می‌باشد. فرمول محاسبه برای ۵۰ ساله : $$P_{۵۰} = \frac{۵L_{۵۰}}{۵ \times l_{۵۰}}$$ برای ۸۰ ساله و بیشتر : $$P_{۷۷۵} = \frac{T_{۸۰}}{T_{۷۵}}$$ می‌باشد.

توضیح ستونهای جدول عمر خلاصه (Abridged life Table)

شماره ستون در جدول عمر	عنوان ستون	شـــــــــرح ستون	فرمول محاسبه
۷	تعداد سالهای زندگی باقیمانده بین سن (X) و ($X+n$)	این ستون، تعداد سالهایی که یک گروه به طور متوسط در سن (X) تا ($X+n$) امرمی نمایند به ($_nL_x$) نشان می دهند. مفهوم دیگر که از این متغیر استنباط می شود عبارت است از تعداد جمعیت یک گروه سنی در جمعیت متوقف (Stationary Population) که در این جمعیت تعداد موالید و مرگ و میر یکسان بوده و در آن مهاجرت صورت نمی گیرد، به عبارت دیگر رشد جمعیت برابر با صفر است.	فرمول محاسبه : $_nL_x = \frac{n(l_x + l_{x+n})}{2}$ که در آن : $_nL_x$ = تعداد افراد باز مانده در سن درست l_x = تعداد افراد باز مانده در سن درست n = فاصله گروه سنی $_nL_x$ = تعداد سالهای زندگی را که یک گروه به طور متوسط از یک سن به سن دیگر طی می کند. فرمول محاسبه فوق برای کلیه سنین به استثنای صفر ساله و ۱-۴ ساله و ۸۰ ساله و بیشتر می باشد. برای تعیین صفر ساله از رابطه : $_1L_0 = 0.27 l_0 + 0.73 l_1$ برای سنین ۱-۴ ساله از رابطه : $_4L_1 = 0.33 l_1 + 1.34 l_5 + 1.33 l_{10}$ و برای گروه سنی ۸۰ ساله و بیشتر : $_\infty L_{80} = \frac{T_{80}}{T_{80}}$ که در آن T_{80} نسبت سالهای زندگی مرد و میر ویژه سنی جامعه می باشد.

توضیح ستونهای جدول عمر خلاصه (Abridged life Table)

شماره ستون در جدول عمر	عنوان ستون	شـــــرح ستون	فرمول محاسبه
۸	تعداد کل نفر سال زندگی شده از سن (X) ببعد (T_x)	این ستون، جدول عمر بعضی (T_A) نشان دهنده تعداد جمعیت متوقف از سن (X) به بعد می‌باشد و به عبارت دیگر (T_x) تعداد کل نفر سالهای زندگی شده ازسن (X) به بعد را نشان میدهد .	$T_x = \sum_{x=0}^{80} l_x + T_{80}$ که : $T_{80} = l_{80}$
۹	امید زندگی یا متوسط عمر باقیمانده (e_x°)	امید زندگی یا متوسط عمر باقیمانده ، عبارت است از تعداد سالهایی که احتمال می‌رود هریک ازافراد جامعه تحت قانون مرگ و میر خاص همان جامعه ، درآینده زنده خواهند ماند .	$e_x^\circ = \frac{T_x}{l_x}$ امید زندگی در موقع تولد که در آن : $X = o$ $e_o = \frac{T_o}{l_o}$

جدول ا-ک الجدول عمـــــــــر جمعیت کشور (مردان) سرشماری ۱۳۳۵

گروه سنی	$_nq_x$	q_x	l_x	d_x	P_x	L_x	T_x	e_x
۰ سالگی	۰/۰۷۸۳۰	۰/۰۷۵۲۵	۱۰۰۰۰۰۰	۷۵۲۵	۰/۹۲۱۲۱۰	۹۴۵۳۵	۵۸۹۸۷۵۳	۵۸/۹۰
۱_۴ "	۰/۰۰۸۰۱	۰/۰۳۱۵۳	۹۲۴۶۵	۲۹۱۵	۰/۹۷۵۲۸	۳۶۲۰۰۷	۵۸۰۴۲۰۸	۶۲/۷۲
۵_۹ "	۰/۰۰۲۱۵	۰/۰۱۰۶۹	۸۹۵۵۰	۹۵۷	۰/۹۹۰۵۶	۴۴۵۲۵۸	۵۲۴۲۲۰۱	۵۰/۷۷
۱۰_۱۴ "	۰/۰۰۱۶۶	۰/۰۰۸۱۲	۸۸۵۹۳	۷۲۴	۰/۹۸۹۶۹	۴۴۱۱۵۵	۳۹۹۶۸۴۳	۵۶/۴۰
۱۵_۱۹ "	۰/۰۰۲۵۱	۰/۰۱۲۴۷	۸۷۸۶۹	۱۰۹۶	۰/۹۸۵۸۶	۴۳۶۶۰۵	۳۵۵۵۶۸۸	۵۱/۸۵
۲۰_۲۴ "	۰/۰۰۳۴۴	۰/۰۱۵۸۲	۸۶۷۷۳	۱۳۷۴	۰/۹۸۲۶۴	۴۳۰۴۳۰	۴۱۱۹۰۸۲	۴۷/۴۷
۲۵_۲۹ "	۰/۰۰۳۸۱	۰/۰۱۸۸۷	۸۵۳۹۹	۱۶۱۱	۰/۹۷۹۸۰	۴۲۲۹۴۸	۳۶۸۸۶۵۲	۴۳/۱۶
۳۰_۳۴ "	۰/۰۰۴۲۶	۰/۰۲۱۵۶	۸۳۷۸۸	۱۸۰۶	۰/۹۷۶۳۸	۴۱۳۴۲۵	۳۲۶۵۶۸۵	۳۸/۴۸
۳۵_۳۹ "	۰/۰۰۵۲۱	۰/۰۲۵۷۲	۸۱۹۸۲	۲۱۰۵	۰/۹۷۱۱۱	۴۰۴۶۳۸	۲۸۵۱۴۶۰	۳۴/۷۸
۴۰_۴۴ "	۰/۰۰۶۵۲	۰/۰۳۲۱۳	۷۹۸۷۳	۲۵۶۶	۰/۹۶۴۹۰	۳۹۴۹۵۰	۳۰۴۴۶۲۲	۳۰/۶۳
۴۵_۴۹ "	۰/۰۰۸۶۲	۰/۰۴۲۶	۷۷۲۰۷	۳۲۹۵	۰/۹۴۹۲۶	۳۷۸۲۷۳	۲۰۵۳۶۷۳	۲۶/۵۷
۵۰_۵۴ "	۰/۰۱۳۶۹	۰/۰۵۹۶۱	۷۶۰۴۲	۴۶۱۴	۰/۹۲۸۴۶	۵۶۹۱۷۵	۱۵۷۲۶۹۹	۲۰/۶۳
۵۵_۵۹ "	۰/۰۱۷۶۳	۰/۰۸۲۴۳	۶۹۵۴۸	۵۸۷۹	۰/۸۹۵۰۲	۳۴۴۳۹۲	۱۳۱۶۰۲۴	۱۸/۹۰
۶۰_۶۴ "	۰/۰۲۶۷۹	۰/۱۲۰۳۳	۵۳۷۶۹	۷۹۸۱	۰/۸۴۷۶۴	۲۹۸۷۷۳	۹۸۲۶۰۱	۱۸/۳۱
۶۵_۶۹ "	۰/۰۴۰۳۵	۰/۱۸۳۲۶	۵۵۷۶۰	۱۰۲۱۹	۰/۷۷۷۱۷	۲۰۴۳۶۵	۶۸۳۲۰۸	۱۲/۲۷
۷۰_۷۴ "	۰/۰۶۲۷۲	۰/۲۷۱۲۸	۴۵۵۴۱	۱۲۳۶۴	۰/۷۷۲۱۲	۱۴۶۸۲۰	۴۴۰۶۵۵	۹/۶۶
۷۵_۷۹ "	۰/۰۹۸۰۲	۰/۳۶۲۷۰	۳۲۱۸۷	۱۳۰۶۶	۰/۶۳۰۰۷	۱۲۳۲۷۰	۲۲۲۸۲۵	۷/۰۵
۸۰ سالبه‌بعد	۰/۲۰۰۰۸	۱/۰۰۰۰۰	۱۸۶۷۴	۱۸۶۷۴	۰/۰۰۰۰۰	۱۰۰۵۶۵	۱۰۰۵۶۵	۵/۰۰

جدول عمر جمعیت مردان شهر در سال ۱۳۶۵"

گروه	n_x	n_x	l_x	d_x	p_x	q_x	T_x	e_x
۰ سال	۰/۰۸۰۶۳	۰/۰۲۷۵۱	۱۰۰۰۰۰	۲۷۵۱	۰/۹۹۱۴۶۹	۰/۰۰۲۷۵۱	۵۸۳۷۵۹۰	۵۸/۳۸
۱ ۴	۰/۰۰۷۲۴	۰/۰۰۲۸۹۲	۹۹۷۲۴۹	۲۸۷۰	۰/۹۹۷۰۷۲	۰/۰۰۲۸۹۲	۵۷۲۳۱۵۲	۵۷/۴۶
۵ ۹	۰/۰۰۰۲۰۵	۰/۰۰۱۰۲۰	۹۹۴۷۹	۹۱۴	۰/۹۹۹۰۰۶	۰/۰۰۱۰۲۰	۵۲۸۱۳۴۰	۵۰/۰۲
۱۰ ۱۴	۰/۰۰۱۵۴	۰/۰۰۰۷۶۷	۸۸۶۶۵	۶۸۰	۰/۹۹۹۰۰۷	۰/۰۰۰۷۶۷	۴۸۱۶۲۵	۵۵/۰۴۷
۱۵ ۱۹	۰/۰۰۰۳۶۶	۰/۰۰۱۲۲۴	۸۷۹۸۵	۱۰۷۵	۰/۹۹۸۵۲۸	۰/۰۰۱۲۲۴	۴۳۲۳۲۸	۵۱/۰۸
۲۰ ۲۴	۰/۰۰۰۲۴۸	۰/۰۰۱۲۲۵	۸۶۹۱۰	۱۰۶۹	۰/۹۹۸۲۲۱	۰/۰۰۱۲۲۵	۴۳۰۸۰۲	۴۹/۰۲۸
۲۵ ۲۹	۰/۰۰۰۲۷۰	۰/۰۰۱۳۷۳	۸۵۴۰۱	۱۱۷۶	۰/۹۹۸۰۴۰	۰/۰۰۱۳۷۳	۴۲۲۶۰۴۳	۴۹/۴۵
۳۰ ۳۴	۰/۰۰۰۴۲۲	۰/۰۰۲۰۸۸	۸۴۸۲۵	۱۷۵۱	۰/۹۹۴۸۲	۰/۰۰۲۰۸۸	۴۱۷۴۲۸	۴۸/۲۰
۳۵ ۳۹	۰/۰۰۰۵۱۲	۰/۰۰۲۵۵۲	۸۲۰۹۲	۲۰۹۵	۰/۹۹۰۵۱	۰/۰۰۲۵۵۲	۴۰۵۲۳۳	۴۹/۹۶
۴۰ ۴۴	۰/۰۰۰۸۴۷	۰/۰۰۲۳۵۴	۷۹۹۹۹	۲۶۸۲	۰/۹۹۰۲۸	۰/۰۰۲۳۵۴	۳۹۳۲۸۳۴	۴۹/۷۹
۴۵ ۴۹	۰/۰۰۰۹۲۹	۰/۰۰۲۸۸۲	۷۷۲۱۲	۲۰۴۷	۰/۹۹۳۰۳۷	۰/۰۰۲۸۸۲	۳۷۷۷۰۲	۴۸/۹۳۷
۵۰ ۵۴	۰/۰۰۱۳۵۲	۰/۰۰۲۶۴۲	۷۴۲۶۲	۴۸۲۷	۰/۹۹۲۰۷۸	۰/۰۰۲۶۴۲	۳۶۵۲۷۸	۴۹/۱۸۵
۵۵ ۵۹	۰/۰۰۱۹۷۲	۰/۰۰۲۳۱۷	۶۸۲۶۰	۴۲۶۸	۰/۹۸۸۳۰۰	۰/۰۰۲۳۱۷	۳۳۵۰۶۹	۴۸/۲۱
۶۰ ۶۴	۰/۰۰۲۹۵۵	۰/۰۰۳۷۵۲	۶۲۶۶۲	۸۵۲۴	۰/۹۸۲۲۹۱	۰/۰۰۳۷۵۲	۲۹۰۸۲۵	۴۶/۸۳
۶۵ ۶۹	۰/۰۰۴۴۲۵	۰/۰۰۵۴۱۳	۵۳۸۶۸	۲۰۴۲۷	۰/۹۷۶۰۷۳	۰/۰۰۵۴۱۳	۲۰۳۷۳۹	۳۸/۰۰
۷۰ ۷۴	۰/۰۰۶۷۶۷	۰/۰۰۸۲۴۲	۴۳۱۶۱	۱۲۲۸۵	۰/۹۶۸۷۹۰	۰/۰۰۸۲۴۲	۱۸۲۳۹۳	۴۲۳۱۱
۷۵ ۷۹	۰/۰۱۰۶۲۶	۰/۰۱۲۸۴۳	۳۰۶۵۶	۱۳۶۸۷	۰/۹۳۱۷۵۹	۰/۰۱۲۸۴۳	۱۲۱۵۶۳	۳۹۸۱
۸۰ و بیشتر	۰/۰۲۰۶۱۶	۱/۰۰۰۰۰۰	۱۷۳۶۹	۱۷۳۶۹	۰/۰۰۰۰۰۰	۱/۰۰۰۰۰۰	۸۷۱۶۰	۵/۰۲

جدول ۶-۳ "جدول عمر جمعیت زنان کشور در سال ۱۳۶۵"

گروه سنی	m_x	q_x	l_x	d_x	L_x	P_x	T_x	e_x
۰ سالگی	۰/۰۷۳۸۹	۰/۰۷۳۱۲	۱۰۰۰۰۰	۷۳۱۲	۹۳۷۰۶	۰/۶۱۳۷۸	۶۹۷۰۰۶۳	۶۹/۷۰
۱_۴ "	۰/۰۰۸۹۰	۰/۰۳۴۲۰	۹۲۶۸۸	۳۱۷۰	۳۶۲۱۸۲	۰/۹۷۲۱۴	۶۰۷۲۰۲۷	۶۳/۳۹
۵_۹ "	۰/۰۰۲۴۶	۰/۰۱۱۲۲	۸۹۵۱۸	۱۰۰۶	۴۴۵۰۷۵	۰/۹۹۰۰۱	۵۵۱۴۱۷۵	۶۱/۵۹
۱۰_۱۴ "	۰/۰۰۱۷۵	۰/۰۰۸۷۲	۸۸۵۱۲	۷۷۲	۴۴۰۶۳۰	۰/۸۸۹۲۹	۵۰۶۸۱۰۰	۵۷/۲۶
۱۵_۱۹ "	۰/۰۰۲۵۶	۰/۰۱۲۷۱	۸۷۷۴۰	۱۱۱۶	۴۳۵۹۱۰	۰/۹۸۵۲۳	۴۶۲۷۹۷۰	۵۲/۷۶
۲۰_۲۴ "	۰/۰۰۳۴۰	۰/۰۱۶۸۶	۸۶۶۲۴	۱۴۶۰	۴۲۹۶۷۰	۰/۹۸۱۸۸	۴۱۹۱۱۵۶۰	۴۸/۳۹
۲۵_۲۹ "	۰/۰۰۳۹۲	۰/۰۱۹۴۱	۸۵۱۶۴	۱۶۵۰	۴۲۱۶۸۸	۰/۹۷۹۱۶	۳۷۲۲۰۴۰	۴۳/۱۷
۳۰_۳۴ "	۰/۰۰۴۵۱	۰/۰۲۲۳۰	۸۳۵۱۱	۱۸۶۲	۴۱۲۹۰۰	۰/۹۲۵۳۱	۳۲۷۰۴۰۲	۳۰/۰۰
۳۵_۳۹ "	۰/۰۰۵۲۵	۰/۰۲۵۹۱	۸۱۶۴۹	۲۱۱۶	۴۰۲۹۵۵	۰/۹۷۱۷۲	۲۶۷۷۳۰۲	۳۲/۸۵
۴۰_۴۴ "	۰/۰۰۶۲۳	۰/۰۳۰۶۷	۷۹۵۳۳	۲۴۳۹	۳۹۱۵۰۸	۰/۹۶۵۵۷	۲۵۲۲۵۵۷	۳۱/۷۲
۴۵_۴۹ "	۰/۰۰۷۸۱	۰/۰۳۸۴۰	۷۷۰۹۴	۲۹۵۲	۳۷۸۰۸۸	۰/۹۵۲۵۶	۲۱۲۶۶۷۲	۲۷/۶۷
۵۰_۵۴ "	۰/۰۱۰۸۶	۰/۰۵۲۸۶	۷۴۱۴۱	۳۹۱۹	۳۶۰۴۰۸	۰/۹۳۷۱۹	۱۷۵۴۸۹۱	۲۳/۶۷
۵۵_۵۹ "	۰/۰۱۵۴۲	۰/۰۷۳۶۱	۷۰۲۲۲	۵۱۴۸	۳۴۸۲۰۰	۰/۹۰۸۹۸	۱۳۶۳۸۸۳	۱۹/۳۵
۶۰_۶۴ "	۰/۰۲۳۶۱	۰/۱۰۰۰۴	۶۵۰۷۴	۶۵۱۶	۳۰۳۹۰۳	۰/۸۶۳۰۶	۰۰۶۵۰۲۲	۱۶/۰۰۰
۶۵_۶۹ "	۰/۰۳۵۸۰	۰/۱۴۷۲۰	۵۷۹۰۷	۸۵۱۷	۲۵۷۵۰۰	۰/۷۶۵۷۹	۰۰۸۸۰۶۰	۱۰/۱۹۰
۷۰_۷۴ "	۰/۰۵۷۴۶	۰/۲۵۱۹۸	۴۸۳۹۳	۱۲۱۸۹	۲۱۱۶۸۰	۰/۷۰۷۱۸	۶۰۲۳۶۰	۶/۶۷
۷۵_۷۹ "	۰/۰۹۴۰۳	۰/۳۴۷۴۱	۳۶۱۹۲	۱۲۵۷۹	۱۴۹۵۵۵	۰/۳۴۸۲۶	۴۷۱۰۶۰	۴/۴۹
۸۰ سالگی و بیشتر	۰/۱۹۳۲۲	۱/۰۰۰۰۰	۲۳۶۲۲	۲۳۶۲۲	۱۲۱۵۰۵	۰/۰۰۰۰۰	۱۱۱۵۰۵	۵/۱۲

جدول زیر "جدول عمر جمعیت مناطق شهری کشور در سال ۱۳۷۸"

گروه سنی	m_x	q_x	l_x	d_x	P_x	L_x	T_x	$\mathring{e}_x$
۰ سالگی	۰/۰۶۰۲۱	۰/۰۵۸۴۵	۱۰۰۰۰۰	۵۸۴۵	۰/۹۳۳۲۰	۹۵۷۴۸	۶۳۶۸۴۲۸	۶۳/۶۸
۱-۴ "	۰/۰۱۵۴۹	۰/۰۲۰۹۲	۹۴۱۵۵	۱۹۷۲	۰/۹۸۴۶۲	۳۷۱۳۳۳	۶۱۷۲۰۶۰	۶۵/۵۵
۵-۹ "	۰/۰۰۱۵۶	۰/۰۰۷۷۷	۹۲۱۸۳	۷۱۶	۰/۹۹۳۱۲	۴۵۹۱۲۵	۵۸۰۰۷۲۷	۶۲/۹۳
۱۰-۱۴ "	۰/۰۰۱۲۰	۰/۰۰۵۹۸	۹۱۴۴۷	۵۴۷	۰/۹۹۳۲۶	۴۵۶۹۶۸	۵۲۲۱۹۰۲	۵۸/۲۰
۱۵-۱۹ "	۰/۰۰۱۹۰	۰/۰۰۹۴۶	۹۰۹۲۰	۸۶۰	۰/۹۸۸۷۳	۴۵۲۴۵۰	۴۸۸۵۴۳۴	۵۳/۷۲
۲۰-۲۴ "	۰/۰۰۲۶۴	۰/۰۱۳۰۶	۹۰۰۶۰	۱۱۷۶	۰/۹۸۴۲۸	۴۴۷۲۶۰	۴۴۲۳۱۸۴	۴۹/۲۲
۲۵-۲۹ "	۰/۰۰۲۹۰	۰/۰۱۴۹۰	۸۸۸۸۴	۱۳۸۰	۰/۹۸۴۵۵	۴۴۱۴۲۰	۳۹۸۵۸۲۴	۴۴/۸۴
۳۰-۳۴ "	۰/۰۰۳۳۲	۰/۰۱۶۵۱	۸۷۴۰۲	۱۴۴۶	۰/۹۸۱۷۶	۴۳۴۰۵	۳۵۴۴۰۰۲۴	۴۰/۴۶
۳۵-۳۹ "	۰/۰۰۴۰۶	۰/۰۲۰۵۱	۸۶۱۵۸	۱۷۴۶	۰/۹۷۷۰۷	۴۳۴۲۸۰	۳۱۱۰۱۹۹	۳۶/۱۰
۴۰-۴۴ "	۰/۰۰۵۲۵	۰/۰۲۵۹۱	۸۴۴۳۴	۲۱۸۸	۰/۹۶۹۲۸	۴۲۴۷۰۰	۲۶۸۲۷۱۹	۳۱/۷۸
۴۵-۴۹ "	۰/۰۰۷۲۶	۰/۰۳۵۶۵	۸۲۲۴۶	۲۹۲۲	۰/۹۵۶۵۱	۴۰۳۹۰۰	۲۲۴۷۰۱۶	۲۷/۵۶
۵۰-۵۴ "	۰/۰۱۰۶۰	۰/۰۵۱۶۲	۷۹۳۱۶	۴۰۹۵	۰/۹۳۶۹۶	۳۸۴۲۳۳	۱۸۴۳۳۱۶	۲۳/۲۹
۵۵-۵۹ "	۰/۰۱۵۶۰	۰/۰۷۵۰۲	۷۵۲۱۹	۵۶۴۷	۰/۹۰۶۴۸	۳۶۱۴۲۸	۱۴۷۴۷۸۶	۱۹/۶۳
۶۰-۶۴ "	۰/۰۲۳۲۶	۰/۱۱۳۰۳	۶۹۶۷۲	۷۸۶۴	۰/۸۶۰۶۲	۳۲۸۴۰۰	۱۱۰۲۸۰۸	۱۶/۰۲
۶۵-۶۹ "	۰/۰۳۸۶۹	۰/۱۷۰۰۹	۶۱۷۰۸	۱۰۴۲۲	۰/۶۲۳۱۵	۲۸۲۷۵۵	۷۸۶۴۰۸	۱۲/۷۳
۷۰-۷۴ "	۰/۰۵۸۴۲	۰/۲۷۴۳۱	۵۱۶۷۲	۱۴۰۴۹	۰/۶۹۴۰۸	۲۲۰۷۲۸	۵۰۲۱۵۲	۹/۸۲
۷۵-۷۹ "	۰/۰۹۲۲۷	۰/۳۷۴۸۸	۳۸۲۲۵	۱۴۳۲۰	۰/۴۹۶۱۶	۱۵۵۴۰۰	۲۸۰۲۰۵	۷/۳۲
۸۰ سال و بیشتر	۰/۱۹۱۰۰	۱/۰۰۰۰۰	۲۴۸۰۵	۲۴۸۰۵	۰/۰۰۰۰۰	۱۲۵۱۰۵	۱۲۵۱۰۵	۵/۰۴

جدول شماره "جدول عمر ... مردان مناطق شهری در سال ۱۳۶۵"

گروه سنی	m_x	q_x	l_x	d_x	P_x	L_x	T_x	e_x
۰ سال	۰/۰۰۶۱۵۸	۰/۰۰۵۹۷۴	۱۰۰۰۰۰	۵۹۷۴	۰/۹۳۴۲۰	۹۵۴۷۵	۶۲۰۵۸۰۱	۶۲/۰۶
۱-۴ "	۰/۰۰۲۷۱	۰/۰۰۱۸۶۶	۹۴۰۱۲۶	۱۷۵۵	۰/۹۸۴۰۴	۳۷۱۲۲۵	۶۱۱۰۱۲۶	۶۴/۹۸
۵-۹ "	۰/۰۰۱۴۹	۰/۰۰۷۴۴	۹۲۲۷۱	۶۸۵	۰/۹۹۳۴۲	۴۵۹۴۴۳	۵۷۳۸۷۰۱	۶۲/۱۹
۱۰-۱۴ "	۰/۰۰۱۱۵	۰/۰۰۵۷۴	۹۱۵۸۶	۵۲۵	۰/۹۹۲۲۸	۴۵۶۶۱۸۰	۵۲۷۹۰۵۸	۵۷/۶۴
۱۵-۱۹ "	۰/۰۰۱۹۱	۰/۰۰۹۵۱	۹۱۰۶۱	۸۶۶	۰/۹۸۸۵۵	۴۵۴۱۴۰	۴۸۳۳۶۰۰	۵۲/۹۶
۲۰-۲۴ "	۰/۰۰۲۷۰	۰/۰۱۳۴۱	۹۰۱۹۵	۱۲۱۰	۰/۹۸۴۲۹	۴۴۹۹۵۰	۴۳۶۴۲۰۰	۴۸/۳۶
۲۵-۲۹ "	۰/۰۰۲۸۲	۰/۰۱۴۰۰	۸۸۹۸۵	۱۲۴۴	۰/۹۸۵۰۲	۴۴۱۸۱۰	۴۲۷۱۳۵۰	۴۸/۰۷
۳۰-۳۴ "	۰/۰۰۳۲۰	۰/۰۱۵۸۷	۸۷۷۴۱	۱۳۹۲	۰/۹۸۲۲۱	۴۳۵۱۶۵	۴۲۷۹۵۴۰	۴۸/۰۵
۳۵-۳۹ "	۰/۰۰۳۹۹	۰/۰۱۹۷۵	۸۶۳۴۲	۱۷۰۵	۰/۹۷۶۸۲	۴۲۷۴۷۳	۴۰۴۳۳۵	۴۵/۲۶
۴۰-۴۴ "	۰/۰۰۵۲۱	۰/۰۲۶۶۹	۸۴۶۴۲	۲۲۵۹	۰/۹۶۷۵۶	۴۱۲۵۸۳	۳۶۱۶۸۵۳	۴۰/۹۲
۴۵-۴۹ "	۰/۰۰۷۸۲	۰/۰۳۸۲۵	۸۲۲۸۳	۳۱۵۹	۰/۹۵۳۷۳	۴۰۴۰۱۸	۳۱۹۹۲۸۹	۳۸/۷۰
۵۰-۵۴ "	۰/۰۱۱۶۲	۰/۰۵۴۵۵	۷۹۲۲۴	۴۳۸۰	۰/۹۴۰۰۱	۳۸۴۹۴۰	۱۷۶۵۲۷۱	۲۲/۶۶
۵۵-۵۹ "	۰/۰۱۷۵۹	۰/۰۸۰۳۵	۷۴۷۶۹	۶۲۹۲	۰/۸۹۵۹۵	۳۵۷۴۹۸	۱۳۱۰۳۵۱	۱۸/۸۷
۶۰-۶۴ "	۰/۰۲۴۸۲	۰/۱۳۵۰۷	۶۸۰۳۷	۸۶۰۲	۰/۸۴۶۴۲	۳۲۰۷۳۰	۱۰۵۳۲۷۳	۱۵/۴۸
۶۵-۶۹ "	۰/۰۲۰۸۹	۰/۱۸۵۲۹	۵۹۸۰۵	۱۱۱۰۱	۰/۷۷۴۳۵	۲۷۱۷۷۳	۷۳۱۶۰۳	۱۲/۲۲
۷۰-۷۴ "	۰/۰۴۲۵۱	۰/۲۷۴۰۶	۴۸۷۹۴	۱۳۳۵۸	۰/۶۷۲۲۰	۲۱۰۳۲۵	۴۵۰۱۷۰	۹/۲۲
۷۵+ "	۰/۰۹۹۰۲	۰/۲۹۶۸۶	۲۵۲۸۶	۱۲۰۴۳	۰/۴۲۲۲۶	۱۴۱۸۲۳	۲۲۹۸۴۵	۹/۰۶
نامعلوم و پیوسته	۰/۱۹۷۵۸	۱/۰۰۰۰۰	۲۱۳۹۳	۲۱۳۹۳	۰/۰۰۰۰۰	۱۰۸۰۲۳	۱۰۸۰۲۳	۵/۰۶

جدول شماره ... جدول عمر ... زنان مناطق شهری کشور در سال ۱۳۷۰

گروه سنی	m_x	q_x	l_x	c_x	P_x	L_x	T_x	e_x
۰ سالگی	۰/۰۰۵۸۸۰	۰/۰۰۵۷۱۲	۱۰۰۰۰۰	۵۷۱۲	۰/۹۳۴۱۹	۹۵۸۶۵	۶۲۲۶۸۲۶	۶۲/۲۰
۱_۴ "	۰/۰۰۰۵۹۰	۰/۰۰۲۳۴۱	۹۴۲۸۸	۲۱۹۹	۰/۹۸۱۲۷۷	۳۷۱۲۲۹	۶۲۲۷۲۷۱	۶۶/۱۲
۵_۹ "	۰/۰۰۰۱۶۲	۰/۰۰۰۸۱۲	۹۲۰۸۹	۷۴۸	۰/۹۹۴۲۸۰	۳۶۸۵۷۸	۵۸۶۲۷۶۶	۶۳/۶۶
۱۰_۱۴ "	۰/۰۰۰۱۲۶	۰/۰۰۰۶۲۸	۹۱۳۰۳	۵۷۲	۰/۹۹۶۲۱۲	۳۶۵۲۷۵	۵۲۰۲۱۶۴	۵۶/۱۶
۱۵_۱۹ "	۰/۰۰۰۱۹۰	۰/۰۰۰۹۴۶	۹۰۷۶۸	۸۵۹	۰/۹۸۸۲۹۱	۴۵۱۶۹۳	۴۶۴۸۸۸۶	۵۱/۵۲
۲۰_۲۴ "	۰/۰۰۰۲۵۰	۰/۰۰۱۲۷۲	۸۹۹۰۹	۱۱۳۴	۰/۹۸۶۲۲۳	۴۴۶۲۸۵	۴۶۹۷۱۱۶	۵۰/۰۹
۲۵_۲۹ "	۰/۰۰۰۲۶۹	۰/۰۰۱۳۸۴	۸۸۷۶۵	۱۲۱۷	۰/۹۸۴۲۹۶	۴۴۰۵۳۳	۴۰۵۰۵۱۱	۴۵/۶۳
۳۰_۳۴ "	۰/۰۰۰۳۳۷	۰/۰۰۱۶۳۰	۸۷۰۳۸	۱۵۰۲	۰/۹۸۱۱۹	۴۳۳۲۸۰	۴۶۰۴۹۷۸	۵۱/۲۸
۳۵_۳۹ "	۰/۰۰۰۴۱۲	۰/۰۰۲۰۳۶	۸۵۹۴۴	۱۷۵۷	۰/۹۷۷۷۶۶	۴۲۵۲۲۸	۴۱۷۲۶۹۸	۴۸/۹۶
۴۰_۴۴ "	۰/۰۰۰۵۰۸	۰/۰۰۲۵۰۸	۸۴۱۸۷	۲۱۱۱	۰/۹۷۱۲۱	۴۱۵۶۵۸	۳۷۲۱۱۲۰	۴۴/۶۸
۴۵_۴۹ "	۰/۰۰۰۶۶۳	۰/۰۰۳۲۶۱	۸۲۰۷۶	۲۶۷۶	۰/۹۶۰۸۷	۴۰۲۶۹۰	۳۳۳۵۵۱۲	۴۰/۴۶
۵۰_۵۴ "	۰/۰۰۰۹۲۹	۰/۰۰۴۵۸۷	۷۹۴۰۰	۳۶۴۲	۰/۹۴۲۵۰	۳۷۷۸۹۵	۲۹۴۱۸۴۲	۳۷/۳۳
۵۵_۵۹ "	۰/۰۱۳۴۳	۰/۰۶۴۳۷	۷۵۷۵۸	۵۲۸۰	۰/۹۱۱۴۲۹	۳۶۵۵۹۰	۲۵۴۳۱۲۷	۳۰/۲۸
۶۰_۶۴ "	۰/۰۲۰۷۸	۰/۰۹۸۷۷	۷۰۰۷۸	۶۶۹۱	۰/۸۷۶۳۳	۳۳۶۹۸۸	۱۸۷۸۰۳۷	۱۶/۷۲
۶۵_۶۹ "	۰/۰۳۶۸	۰/۱۶۱۰۶	۶۳۵۱۷	۹۹۵۵	۰/۸۰۹۹۱	۲۹۵۶۹۸	۸۴۴۲۲۶	۱۳/۳۸
۷۰_۷۴ "	۰/۰۵۳۵۳	۰/۲۳۶۰۶	۵۴۶۲۲	۱۲۷۲۶	۰/۷۱۱۵۲	۲۲۷۷۸۸	۵۲۶۲۷۵۱	۱۰/۴۰
۷۵_۷۹ "	۰/۰۸۶۱۴	۰/۳۵۰۹	۲۱۱۹۴	۱۴۷۱۰	۰/۴۵۷۶۶	۱۴۹۱۹۰	۲۱۱۹۴۴	۷/۵۷
۸۰ سالگی و بیشتر	۰/۱۸۵۱۹	۱/۰۰۰۰۰	۲۶۴۸۲	۲۶۴۸۲	۰/۰۰۰۰۰	۱۴۲۷۷۲	۱۹۲۷۷۲	۷/۲۹

جدول ۶-۷ "جدول عمر" جمعیت مناطق روستایی کشور در سال ۱۳۶۵"

گروه سنی	q_x	m_x	l_x	d_x	P_x	L_x	T_x	e_x
۰ سالگی	۰/۰۲۰۴۲۲	۰/۰۲۰۹۸۲۲	۱۰۰۰۰۰	۲۰۴۲	۰/۸۸۶۰۶	۹۲۸۸۰	۵۲۳۶۰۱۵	۵۲/۳۶
۱_۴ "	۰/۰۱۲۰۸	۰/۰۲۷۱۸	۹۷۹۵۸	۱۱۸۶	۰/۹۶۲۵۲	۳۸۱۱۲۶	۵۲۲۱۱۲۵	۵۳/۱۲
۵_۹ "	۰/۰۰۳۰۰	۰/۰۱۴۸۹	۹۸۶۷۲	۱۴۷۶	۰/۶۸۶۹۲	۴۹۳۶۶۲	۴۶۸۱۹۷۱	۴۷/۴۹
۱۰_۱۴ "	۰/۰۰۲۲۶	۰/۰۱۱۲۲	۹۸۰۳۳	۹۵۱	۰/۹۸۶۰۷	۴۹۰۷۸۸	۴۶۲۵۶۰۸	۴۷/۸۲
۱۵_۱۹ "	۰/۰۰۳۳۶	۰/۰۱۶۶۶	۹۷۶۸۲	۱۶۲۲	۰/۹۸۰۳۸	۴۸۶۴۲۵	۴۱۲۴۲۵	۴۶/۳۱
۲۰_۲۴ "	۰/۰۰۳۵۸	۰/۰۱۷۸۴	۹۷۳۸۸	۱۸۰۴	۰/۶۷۹۱۱	۴۸۰۶۸۲	۳۷۱۶۸۲۵	۴۵/۲۱
۲۵_۲۹ "	۰/۰۰۵۱۰	۰/۰۲۵۱۸	۹۷۰۳۵	۲۰۲۵	۰/۹۷۳۱۲	۴۸۶۰۶۳	۳۲۱۳۱۱۲	۳۱/۴۰
۳۰_۳۴ "	۰/۰۰۵۸۱	۰/۰۲۸۶۲	۹۷۸۰۰	۲۲۴۵	۰/۶۲۸۹۷	۴۸۶۳۸۸	۲۶۱۶۰۶۹	۲۷/۱۹
۳۵_۳۹ "	۰/۰۰۶۸۱	۰/۰۳۳۷۸	۹۶۱۵۵	۲۵۵۰	۰/۹۶۲۱۲	۴۷۳۰۷۰	۲۱۲۶۶۶۱	۲۲/۲۲
۴۰_۴۴ "	۰/۰۰۸۲۲	۰/۰۴۰۲۷	۹۷۶۰۵	۳۹۷۱	۰/۲۹۵۴۵۶	۴۶۰۵۹۸۰	۲۱۵۵۲۶۱	۲۹/۲۸
۴۵_۴۹ "	۰/۰۱۰۴۱	۰/۰۵۰۷۳	۹۷۰۶۳۴	۴۵۸۲	۰/۹۴۰۱۵	۴۶۲۲۱۳	۱۷۶۶۰۶۲	۲۵/۴۰
۵۰_۵۴ "	۰/۰۱۴۲۹	۰/۰۶۹۲۵	۹۶۰۵۱	۶۶۵۷	۰/۹۱۷۶۲	۴۵۲۶۱۲	۱۶۵۰۰۵۰	۲۱/۰۶
۵۵_۵۹ "	۰/۰۲۰۲۲	۰/۰۹۶۲۴	۹۳۶۹۶	۹۰۰۵	۰/۸۸۴۷۹	۴۶۹۲۵۸	۱۱۲۵۰۳۷	۱۸/۰۶
۶۰_۶۴ "	۰/۰۳۰۱۰۱	۰/۱۴۰۲۰	۹۲۸۶۹	۱۳۰۱۷	۰/۸۳۱۷۰	۴۶۰۱۵۳	۸۶۶۸۷۹	۱۴/۷۳
۶۵_۶۹ "	۰/۰۰۵۵۳	۰/۰۲۰۰۷۵	۹۰۸۶۷۳	۱۷۳۱	۰/۷۵۸۷۹	۴۱۸۰۶۳	۴۱۶۷۲۶	۱۱/۴۱
۷۰_۷۴ "	۰/۰۶۸۷۵	۰/۲۴۱۸۸	۱۳۰۸	۳۸۷۲۱	۰/۶۳۵۴۲۸	۱۶۵۴۲۵	۳۶۹۵۶۲	۴/۰۳
۷۵_۷۹ "	۰/۱۰۵۲۷	۰/۳۱۰۶۶	۱۱۴۳۱	۳۷۶۳۳	۰/۴۱۰۰۶۷	۱۰۸۵۸۸	۱۸۳۶۵۸	۶/۷۲
۸۰ سالگی و بیشتر	۰/۲۳۱۱۲۷	۱/۰۰۰۰۰	۱۴۰۰۲	۶۰۰۲	۰/۰۰۰۰۰	۷۵۶۷۰	۷۵۶۷۰	۴/۷۳

جدول شماره [illegible] جدول عمر مردان مناطق روستایی ایران سال ۱۳۴۵

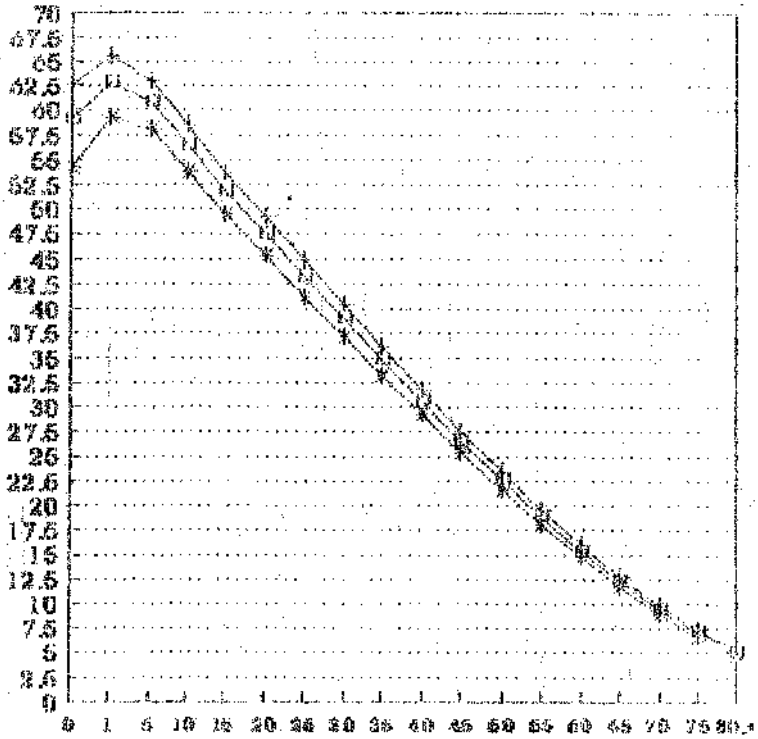
گروه سنی		q_x	S_x	l_x	d_x	P_x	L_x	T_x	e_x
۰	سال	۰/۱۰۰۷۲	۰/۱۰۰۹۲	۱۰۰۰۰۰	۱۰۰۲۲	۰/۸۸۴۲۲	۹۴۷۱۰	۶۷۶۸۹۴۱	۶۷/۶۹
۱-۴	"	۰/۰۱۱۰۱	۰/۰۴۳۰۹	۸۹۹۷۸	۳۸۷۴	۰/۹۹۶۸۰	۳۶۴۳۸۶	۶۳۰۶۲۱۲	۵۸/۹۹
۵-۹	"	۰/۰۰۲۷۸	۰/۰۱۴۸۰	۸۶۰۸۱	۱۱۸۸	۰/۹۸۷۹۹	۴۲۷۴۲۵	۵۹۶۶۸۴۶	۶۷/۶۸
۱۰-۱۴	"	۰/۰۰۲۰۵	-۰/۰۱۰۲۰	۸۴۸۲۳	۸۶۶	۰/۹۸۷۰۷	۴۲۳۰۰	۶۲۶۹۴۹۱	۶۳/۲۵
۱۵-۱۹	"	۰/۰۰۳۱۶	۰/۰۱۵۴۸	۸۴۰۲۷	۱۳۱۸	۰/۹۸۱۰۶	۴۱۶۸۴۰	۵۱۰۷۰۹۱	۶۰/۸۸
۲۰-۲۴	"	۰/۰۰۳۵۰	۰/۰۲۲۲۵	۸۲۷۰۹	۲۸۹۰	۰/۹۷۴۸۸	۴۰۸۹۲۵	۴۶۹۰۲۵۱	۵۶/۶۲
۲۵-۲۹	"	۰/۰۰۴۸۶	۰/۰۲۴۰۱	۸۰۸۶۹	۱۹۴۲	۰/۹۷۳۳۳	۴۰۶۹۴۹۰	۳۶۸۱۳۰۶	۴۰/۵۸
۳۰-۳۴	"	۰/۰۰۵۵۰	۰/۰۲۷۲۷	۷۸۹۲۷	۴۳۴۰	۰/۹۶۴۸۴	۳۸۴۲۳۵	۲۸۸۱۸۱۴	۳۶/۵۱
۳۵-۳۹	"	۰/۰۰۶۷۲	۰/۰۳۳۰۴	۷۶۷۴۷	۴۵۴۶	۰/۹۶۲۳۲	۳۷۷۶۹۱	۲۶۹۲۵۸۱	۳۲/۴۷
۴۰-۴۴	"	۰/۰۰۸۴۷	۰/۰۴۲۲۹	۷۲۲۲۱	۴۱۲۹	۰/۹۵۱۳۶	۳۶۳۳۰۸	۲۶۱۵۰۸۴	۳۸/۴۹
۴۵-۴۹	"	۰/۰۱۱۳۲	۰/۰۵۵۲۸	۶۱۰۹۲	۴۹۲۰	۰/۹۴۲۴۸	۳۴۵۰۳۵	۱۷۵۱۲۷۸	۲۹/۰۶
۵۰-۵۴	"	۰/۰۱۵۸۸	۰/۰۷۶۲۷	۶۰۱۶۲	۵۱۲۹	۰/۹۰۹۵۲	۵۴۲۹۸۸	۱۴۰۶۱۴۳	۲۰/۸۹
۵۵-۵۹	"	۰/۰۲۲۲۲	۰/۱۰۵۷۵	۵۴۰۳۳	۶۵۴۰	۰/۸۷۲۲۲	۲۹۲۷۶۵	۱۰۸۲۱۵۵	۱۷/۳۵
۶۰-۶۴	"	۰/۰۳۲۹۰	۰/۱۵۲۰۰	۵۵۲۷۲	۸۲۲۲	۰/۸۱۸۹۲	۲۵۶۲۸۵	۷۸۹۲۹۰	۱۴/۲۳
۶۵-۶۹	"	۰/۰۴۸۲۷	۰/۲۲۵۶۰	۴۷۰۴۱	۱۰۲۳۱	۰/۷۳۴۰۷	۲۰۶۸۷۸	۵۳۳۱۰۵	۱۱/۳۳
۷۰-۷۴	"	۰/۰۶۲۹۱	۰/۲۰۷۲۲	۳۶۹۱۰	۱۱۲۵۵	۰/۶۴۰۷۳	۱۵۶۱۶۳	۲۲۲۲۲۷	۸/۷۶
۷۵-۷۹	"	۰/۱۱۰۸۰	۰/۴۲۲۸۲	۲۵۵۵۵	۱۱۰۸۷	۰/۴۰۱۰۸	۱۰۰۰۵۸	۱۰۶۰۰۶۵	۶/۵۲
۸۰+	سال و بیشتر	۰/۲۱۵۲۳	۱/۰۰۰۰۰	۱۴۴۶۸	۱۴۴۶۸	۰/۰۰۰۰۰	۶۷۰۰۶	۶۷۰۰۶	۴/۶۳

جدول ۶-۹ "جدول عمر ... زنان مناطق روستایی کشور سال ۱۳۶۵"

گروه سنی	m_x	q_x	l_x	d_x	L_x	P_x	T_x	e_x
۰ سال	۰/۱۰۰۹۸	۰/۰۹۶۱۲	۱۰۰۰۰۰	۹۶۱۲	۹۴۰۲۰	۰/۸۸۲۹۸	۵۲۶۱۶۴۶	۵۲/۶۲
۱-۴ "	۰/۰۱۴۱۹	۰/۰۵۱۴۰	۹۰۳۸۷	۴۶۴۶	۳۸۹۵۰	۰/۹۶۲۲۰	۵۳۶۸۹۰۶	۵۹/۴۰
۵-۹ "	۰/۰۰۲۲۲	۰/۰۱۵۹۷	۸۵۲۴۱	۱۳۶۸	۴۲۵۲۸۳	۰/۹۸۵۸۱	۵۰۱۹۹۵۶	۵۸/۵۵
۱۰-۱۴ "	۰/۰۰۲۴۹	۰/۰۱۲۲۷	۸۳۳۷۳	۱۰۲۳	۴۱۹۲۵۰	۰/۹۸۴۹۹	۴۵۲۳۶۴۳	۵۴/۲۴
۱۵-۱۹ "	۰/۰۰۳۵۴	۰/۰۱۷۴۹	۸۲۳۲۸	۱۴۳۲	۴۱۲۹۵۵	۰/۹۷۹۶۷	۴۱۷۵۲۲۳	۵۰/۱۱
۲۰-۲۴ "	۰/۰۰۴۶۶	۰/۰۲۳۰۲	۸۱۸۵۲	۱۸۸۵	۴۰۴۵۵۸	۰/۹۷۵۴۶	۴۱۳۳۲۶۸	۵۰/۹۷
۲۵-۲۹ "	۰/۰۰۵۲۲	۰/۰۲۶۳۰	۷۹۹۶۹	۲۱۰۳	۳۹۳۵۸۸	۰/۹۷۱۹۸	۳۲۵۷۹۱۰	۴۱/۲۹
۳۰-۳۴ "	۰/۰۰۶۰۵	۰/۰۲۹۸۰	۷۷۸۶۶	۲۳۲۰	۳۸۲۵۴۰	۰/۹۶۸۲۲	۲۹۶۳۲۲۴	۳۸/۰۶
۳۵-۳۹ "	۰/۰۰۶۸۸	۰/۰۳۴۸۲	۷۵۵۴۶	۲۵۵۵	۳۷۱۲۲۲	۰/۹۶۲۸۱	۲۵۷۹۷۹۲	۳۴/۱۵
۴۰-۴۴ "	۰/۰۰۷۸۸	۰/۰۳۸۴۲	۷۲۹۹۱	۲۸۲۰	۳۵۷۹۰۵	۰/۹۵۷۷۰	۲۲۰۸۴۴۹	۳۰/۲۶
۴۵-۴۹ "	۰/۰۰۹۲۴	۰/۰۴۶۱۱	۷۰۱۷۱	۴۲۳۶	۳۴۲۷۶۵	۰/۹۲۵۶۴	۱۸۵۰۵۲۴	۲۶/۲۲
۵۰-۵۴ "	۰/۰۱۲۸۸	۰/۰۶۲۲۹	۶۶۴۳۵	۴۱۷۶	۳۲۳۲۳۵	۰/۹۲۶۹۰	۱۵۰۲۷۷۹	۲۲/۵۳
۵۵-۵۹ "	۰/۰۱۷۶۵	۰/۰۸۴۵۲	۶۲۷۶۵	۵۳۰۶	۳۰۰۵۲۵	۰/۸۶۵۹۶	۱۱۸۳۶۴۲	۱۸/۸۶
۶۰-۶۴ "	۰/۰۲۶۷۵	۰/۱۲۵۳۷	۵۷۴۵۹	۷۲۰۳	۲۶۱۲۶۸	۰/۸۶۳۰۶	۸۸۳۰۰۶	۱۵/۳۷
۶۵-۶۹ "	۰/۰۳۹۹۷	۰/۱۸۱۶۹	۵۰۴۵۲	۹۱۳۰	۲۲۸۴۳۵	۰/۸۷۵۰۹	۶۱۳۷۷۰	۱۲/۱۱
۷۰ سال و بیشتر	۰/۰۹۳۱۳	۰/۲۷۴۶۲	۴۱۱۲۲	۱۱۲۶۱	۱۷۷۵۸۲	۰/۶۷۵۴۸	۳۸۵۲۰۶	۹/۳۷
۷۵-۷۹ "	۰/۰۹۸۷۱	۰/۲۹۵۸۴	۲۹۹۰۱	۱۱۸۴۱	۱۱۹۹۵۳	۰/۴۲۲۵۳	۲۰۷۲۷۲	۶/۹۲
۸۰ ساله و بیشتر	۰/۲۰۵۸۸	۱/۰۰۰۰۰	۱۸۰۷۰	۱۸۰۷۰	۸۷۷۰	۰/۰۰۰۰۰	۸۷۷۰	۴/۸۰

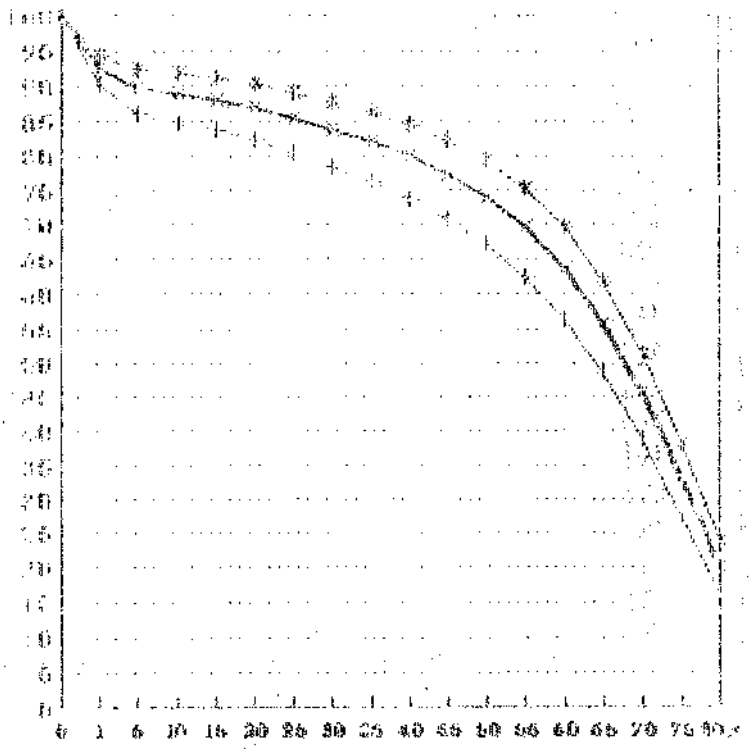
نمودار امیدزندگی (e_0^0) کشور، مناطق شهری و روستایی برحسب سن

سال ۱۳۶۵



نمودار ۲ـ بازماندگان (l_x) از ۱۰۰۰۰۰ نفر اولیه در شروع هر دوره سنی

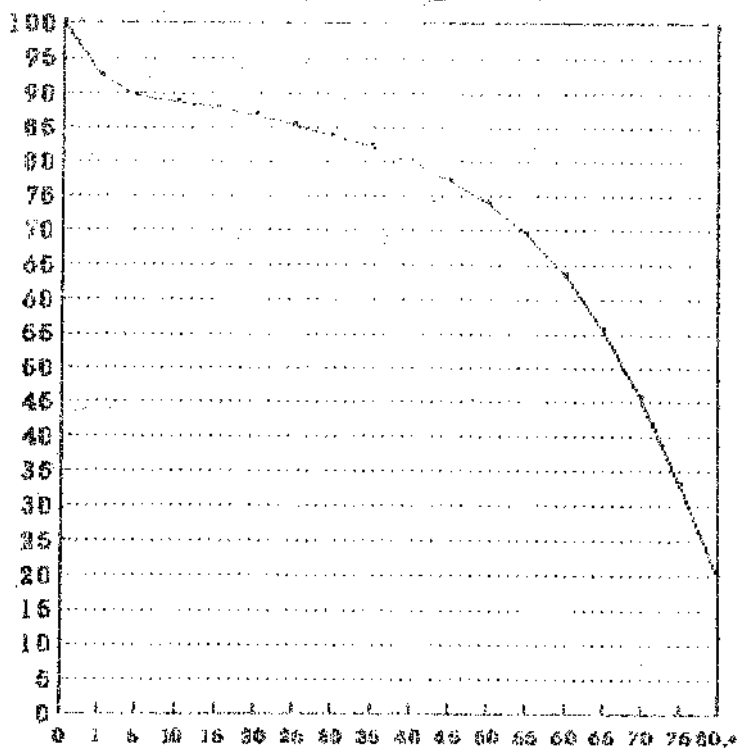
کشور، مناطق شهری و روستایی سال ۱۳۶۵



کشور ـــ شهری ـ ـ روستایی ـ ـ

نمودار ۲-بازماندگان (l_x) از ۱۰۰،۰۰۰ نفر اولیه در شروع هر دوره سنی

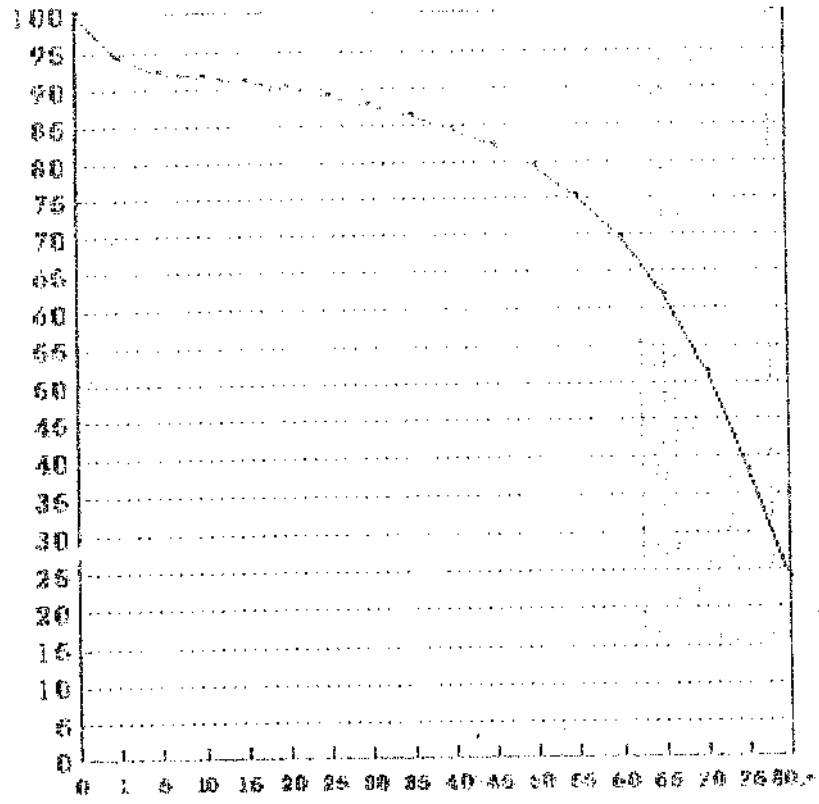
سال ۱۳۶۵



کشور ...

نمودار ۴ـ بازماندگان (l_x) از ۱۰۰۰۰۰ نفر اولیه در شروع هر دوره سنی

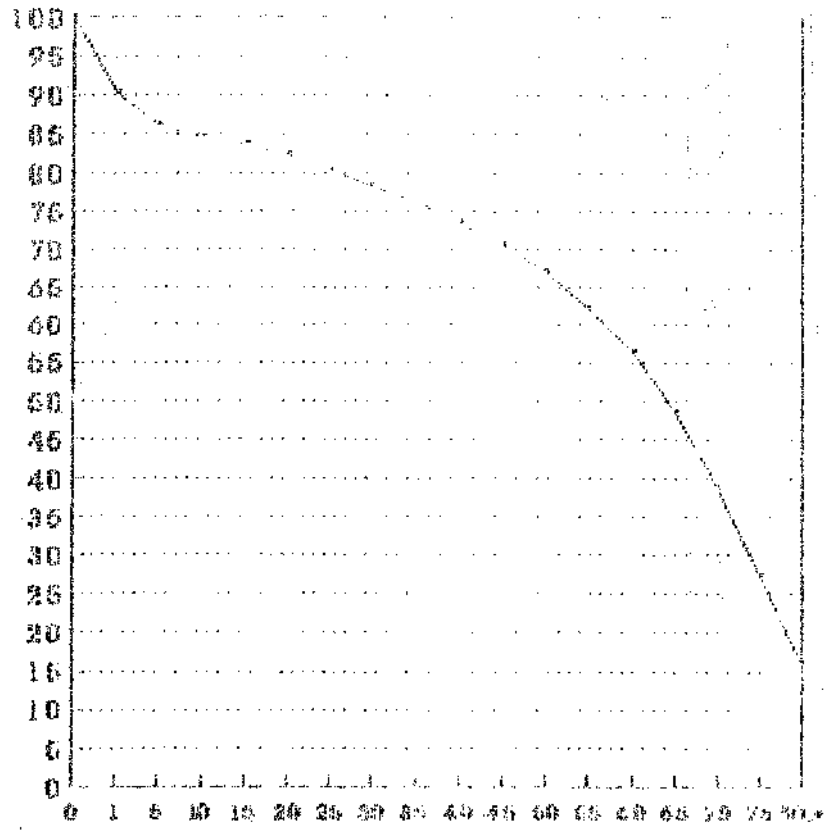
سال ۱۳۶۵



شهری ـــ

نمودار ۵- بازماندگان (l_x) از ۱۰۰۰۰۰ نفر اولیه در شروع هر دوره سنی

سال ۱۳۶۵



روستایی

مآخذ ومنابع:

۱ـ نشریه نتایج تفصیلی سرشماری عمومی نفوس ومسکن مهرماه ۱۳۶۵، جلد ۶ کل کشور ، سازمان برنامه وبودجه مرکز آمار ایران، اردیبهشت ۱۳۶۷، جداول شماره ۳۳،۳۴،۳۵.

۲ـ میزانهای حیاتی ایران مرگومیر، باروری رشد جمعیت وجدول عمر ، دکتر واتکس نهاوندیان، حبیب خزانه، انتشارات دانشکده بهداشت دانشگاه تهران سال ۱۳۵۶.

۳ـ باروری وجمعیت، ترجمه زهرا پورانصاری، علی ناصری، مرکز نشر دانشگاهی ، تهران سال ۱۳۶۸.

۴ـ محاسبه آمارهای حیاتی استان آذربایجان شرقی، نشریه شماره ۶۹/۶ ، سازمان برنامه وبودجه استان آذربایجان شرقی ـ دیماه ۱۳۶۹.

۵ـ لغت نامه جمعیت شناسی ، تالیف وتدوین مهدی امانی ودیگران ، انتشارات دانشگاه تهران ، تهران ۱۳۵۲.

۶ـ روشهای مقدماتی جمعیت شناسی، محمدحسین نجاتیان، مرکز آمار ایران، تهران ۱۳۵۶.

۷ـ جدول امید به زندگی در ایران ، حسن شمس، مرکز آمار ایران ، تهران، ۱۳۶۴.

۸ـ اندازه گیری رشد جمعیت ایران، در سالهای ۱۳۵۲ تا ۱۳۵۵ ، مرکز آمار ایران، تهران، ۱۳۶۳.

1- Coal & Demeny , Model life tables, 1966.

2- Henryk . Moyrock Siegel. The Methods and Materials of Demography. 1976

3- Abridged life Tables for Iran and Its RURAL - URBAN AReas, 1973-76
by K.J.Kouli

برآورد در فرآیندهای میانگین متحرک با استفاده از روش کمترین مربعات

حسینعلی نیرومند

دانشگاه فردوسی مشهد

۱.۱ مقدمه

هدف ما دراین مقاله بسط راهی مستقیم و ساده جهت محاسبهٔ برآوردکنندهٔ کمترین مربعات در فرآیندهای میانگین متحرک است . می خواهیم پارامترهای الگوهای میانگین متحرک را با استفاده از الگوریتمی که بدست خواهیم آورد برآورد کنیم . بطوری که مطلع هستیم برای برآورد پارامترهای یک فرآیند میانگین متحرک چند روش میسر است ولی چون اجرای این روشها پیچیده به نظر می رسد لذا جستجوی روشی ساده طبیعی می باشد. ما دراین مقاله دو الگوریتم را برای محاسبه برآوردکننده کمترین مربعات فرآیند هایی MA(1) و MA(2) معرفی می کنیم .

۲.۱ برآورد

فرض کنید $Y_t = e_t + \beta e_{t-1}$ $t = 1, 2, 3, \ldots$ یک فرآیند میانگین متحرک مرتبه اول باشد که در آن $e_1, e_2, e_3, \ldots$ دنباله ای از متغیرهای تصادفی مستقل هم توزیع با میانگین صفر و واریانس σ^2 است . برآورد کننده کمترین مربعات β

مقداریست مانند $\theta \in (-1,1)$ که رابطه زیر را می نیمم کند.

$$S_n(\theta) = \sum_{t=1}^{n} \{e_t(Y;\theta)\}^2 \qquad (1.2.1)$$

که در آن

$$e_t(Y;\theta) = Y_t - \theta\, e_{t-1}(Y;\theta) \quad , \quad e_0(Y;\theta) = 0$$

و می توان نوشت :

$$\begin{aligned} e_t(Y;\theta) &= Y_t - \theta\, e_{t-1}(Y;\theta) \\ &= Y_t - \theta(Y_{t-1} - \theta\, e_{t-2}(Y;\theta)) \\ &= Y_t - \theta Y_{t-1} + \theta^2 e_{t-2}(Y;\theta) \\ &= Y_t - \theta Y_{t-1} + \theta^2(Y_{t-2} - \theta\, e_{t-3}(Y;\theta)) \\ &= Y_t - \theta Y_{t-1} + \theta^2 Y_{t-2} - \theta^3 e_{t-3}(Y;\theta) \\ &\;\;\vdots \end{aligned}$$

بنابراین

$$e_k(Y;\theta) = Y_k - \theta Y_{k-1} + \theta^2 Y_{k-2} - \cdots + (-1)^{k-1}\theta^{k-1} Y_1 \quad , \quad k = 1,2,\ldots \qquad (1.2.1)$$

که می توان نوشت :

$$\begin{bmatrix} e_1(Y;\theta) \\ e_2(Y;\theta) \\ \vdots \\ e_n(Y;\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\theta & 1 & & 0 \\ (-\theta)^2 & -\theta & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ (-\theta)^{n-1} & (-\theta)^{n-2} & \cdots & 1 \end{bmatrix} \times \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{n-1} \\ Y_n \end{bmatrix}$$

$$\underline{Y}^T = (Y_1, Y_2, \cdots, Y_n)$$

که

و

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\theta & 1 & \cdots & 0 \\ (-\theta)^2 & -\theta & & 0 \\ \vdots & \vdots & & \vdots \\ (-\theta)^{n-1} & (-\theta)^{n-2} & \cdots & 1 \end{bmatrix}$$

بنابراین $Q_n(\theta)$ یک چندجمله ای از درجهٔ کـ $2n$ نسبت به θ برده و بسرای n زوج داریم :

$$Q_n(\theta) = \sum_{i=1}^{n} Y_i^2 + \left(\sum_{i=1}^{n-1} Y_i^2 + 2 \sum_{i=1}^{n-2} Y_i Y_{i+2} \right) \theta^2 +$$

$$\left(\sum_{i=1}^{n-2} Y_i^2 + 2 \sum_{i=1}^{n-3} Y_i Y_{i+2} + 2 \sum_{i=1}^{n-4} Y_i Y_{i+4} \right) \theta^4 + \cdots +$$

$$\left(\sum_{i=1}^{\frac{n}{2}+1} Y_i^2 + 2 \sum_{i=1}^{\frac{n}{2}} Y_i Y_{i+2} + \cdots + 2 \sum_{i=1}^{2} Y_i Y_{i+n-2} \right) \theta^{n-2} +$$

$$\left(\sum_{i=1}^{\frac{n}{2}} Y_i^2 + 2 \sum_{i=1}^{\frac{n}{2}-1} Y_i Y_{i+2} + \cdots + 2 \sum_{i=1}^{1} Y_i Y_{i+n-2} \right) \theta^n +$$

$$\left(\sum_{i=1}^{\frac{n}{2}-1} Y_i^2 + 2 \sum_{i=1}^{\frac{n}{2}-2} Y_i Y_{i+2} + \cdots + 2 \sum_{i=1}^{1} Y_i Y_{i+n-4} \right) \theta^{n+2} + \cdots$$

$$+ Y_1^2 \theta^{2n-2} - 2 \left(\sum_{i=1}^{n-1} Y_i Y_{i+1} \right) \theta - 2 \left(\sum_{i=1}^{n-2} Y_i Y_{i+1} + \sum_{i=1}^{n-3} Y_i Y_{i+3} \right) \theta^3$$

$$- \cdots - 2 \left(\sum_{i=1}^{\frac{n}{2}} Y_i Y_{i+1} + \sum_{i=1}^{\frac{n}{2}-1} Y_i Y_{i+3} + \cdots + \sum_{i=1}^{1} Y_i Y_{i+n-1} \right) \theta^{n-1}$$

$$- 2 \left(\sum_{i=1}^{\frac{n}{2}-1} Y_i Y_{i+1} + \sum_{i=1}^{\frac{n}{2}-2} Y_i Y_{i+3} + \cdots + \sum_{i=1}^{1} Y_i Y_{i+n-3} \right) \theta^{n+1} - \cdots - 2 Y_1 Y_n \theta^{2n-2}$$

این عبارت نسبتاً پیچیده است و آنرا بصورت زیر می توان خلاصه نمود. فرض کنیـــــد B یک ماتریس $n \times n$ باشدکه بصورت زیر تعریف می شود.

$$B=\begin{bmatrix} Y_1 & 0 & 0 & \cdots & 0 & 0 \\ Y_2 & -Y_1 & 0 & \cdots & 0 & 0 \\ Y_3 & -Y_2 & Y_1 & \cdots & 0 & 0 \\ \vdots & & & & & \\ Y_n & -Y_{n-1} & -Y_{n-2} & \cdots & (-1)^{n-2}Y_2 & (-1)^{n-1}Y_1 \end{bmatrix}$$

فرض کنید $\underline{\theta}$ بردار ستونی $n \times 1$ باشد که بصورت زیر تعریف می شود

$$\underline{\theta}^T = (1 \quad \theta \quad \theta^2 \quad \cdots \quad \theta^{n-1})$$

می توان نشان داد که

$$B^T B = \begin{bmatrix} \sum_{i=1}^{n} Y_i^2 & \sum_{i}^{n-1} Y_i Y_{i+1} & \cdots & -\sum Y_i Y_{i+n-1} \\ -\sum_{i=1}^{n-1} Y_i Y_{i+1} & \sum_{i=1}^{n-1} Y_i^2 & & \sum Y_i Y_{i+n-2} \\ \vdots & \vdots & \cdots & \vdots \\ -\sum_{i}^{1} Y_i Y_{i+n-1} & \sum_{i}^{1} Y_i Y_{i+n-2} & \cdots & \sum_{i=1}^{1} Y_i^2 \end{bmatrix}$$

لذا می توان نوشت

$$Q_n(\theta) = \underline{\theta}^T (B^T B)\, \underline{\theta}$$

ازاین شکل ساده در یک برنامهٔ کامپیوتری استفاده کرده و θ را برآورد می کنیم .

یک مطالعهٔ شبیه سازی

برای $\sigma^2 = 1$ و برخی از مقادیر β واقع در ناحیه وارون پذیری ، صد مشاهده را از الگوی زیر تولید نموده و با استفاده ازروش کمترین مربعات و با صد بار تکرار این طرز عمل، β را برآورد می کنیم .

$$Y_t = e_t + \beta e_{t-1} \quad , \quad t = 1, 2, 3, \ldots$$

$\overline{\hat{\beta}}$ متوسط برآوردهای β در صد مرتبه تکرار طرز عمل فوق را نشان می دهد . جدول ۱ متوسط و انحراف معیار این متوسط ها را در صد مرتبه تکرار می دهد .

جدول ۱ متوسط $\overline{\hat{\beta}}$ و انحراف معیار آن با استفاده از روش کمترین مربعات

β	$\overline{\hat{\beta}}$	$S.d(\overline{\hat{\beta}})$
— ۰/۹	— ۰/۸۸۰۲	۰/۰۰۵۷
— ۰/۷	— ۰/۶۹۷۵	۰/۰۰۸۲
— ۰/۵	— ۰/۵۰۰۰	۰/۰۰۸۸
— ۰/۳	— ۰/۲۹۹۰	۰/۰۱۱۰
— ۰/۱	— ۰/۰۹۶۶	۰/۰۱
۰/۱	۰/۱۰۶۹	۰/۰۱
۰/۳	۰/۳۰۶۳	۰/۰۱۰۱
۰/۵	۰/۵۰۲۵	۰/۰۰۹۲
۰/۷	۰/۷۰۶۸	۰/۰۰۸۰
۰/۹	۰/۸۸۰۶	۰/۰۰۶۰

براساس نتایج شبیه سازی جدول ۱ ملاحظه می کنیم که به جز برای $|\beta| = 0/9$ اریب آشکاری برای سایر مقادیر β در فاصلهٔ وارون پذیری مشاهده نمی شود. این مطلب با آزمون $\alpha(\bar{\beta}) = \frac{(\bar{\beta} - \beta)}{s.d.(\bar{\beta})}$ نیز تأیید می شود. فرمول مجانبی برآورد کنندهٔ کمترین مربعات که همان فرمول مجانبی برآورد کنندهٔ درستنمایی ماکزیمم بوده و برابر $\left(\frac{1-\beta^2}{N}\right)^{1/2}$ است با مقادیر بدست آمده از شبیه سازی مطابقت می نماید. اگر علامت β را عوض کنیم فقط علامت $\bar{\beta}$ تغییر می کند و انحراف معیار تغییر نمی کند.

با استفاده از روش کمترین مربعات یک الگوی مرتبه اول را به تفاضلهای مرتبه اول داده های IBM STOCK PRICES برازش نموده و الگوی $\nabla Y_t = e_t + .89\, e_{t-1}$ نتیجه می شود.

۲ - برآورد در فرآیندهای میانگین متحرک مرتبهٔ دوم

حال نشان می دهیم که می توانیم ایدهٔ بالا را به یک الگوی میانگین متحرک مرتبه دوم تعمیم داده و برآورد کنندهٔ کمترین مربعات پارامترهای یک فرآیند MA (۲) را به طریق زیر محاسبه نمود.

۲ - ۱ برآورد

فرض کنید $Y_t = e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2}$ ، $t = 1, 2, \ldots$ یک الگوی میانگین متحرک مرتبه دو باشد که در آن $\ldots, e_0, e_1, e_2$ دنباله ای از متغیرهای تصادفی مستقل هم توزیع با میانگین صفر و واریانس σ^2 است.

برآوردکنندهٔ کمترین مربعات (β_1, β_2) آن مقدار (θ_1, θ_2) در ناحیـــــه

وارون پذیری است که عبارت زیر را می مینیمم کند،

$$Q_n(\theta_1, \theta_2) = \sum_{t=1}^{n} \left[e_t(Y; \theta_1, \theta_2) \right]^2 \qquad (1.1.2)$$

که در آن

$$e_t(Y;\theta_1,\theta_2) = Y_t - \theta_1 e_{t-1}(Y;\theta_1,\theta_2) - \theta_2 e_{t-2}(Y;\theta_1,\theta_2), \quad t=1,2,\dots,n$$

و

$$e_0(Y;\theta_1,\theta_2) = e_{-1}(Y;\theta_1,\theta_2) = 0$$

می نویسیـــم ،

$$Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}$$
$$= (1 + \theta_1 B + \theta_2 B^2) e_t$$

$$e_t = \frac{Y_t}{1+\theta_1 B + \theta_2 B^2} = \frac{Y_t}{(1+\phi_1 B)(1+\phi_2 B)}$$

$$= \left[\frac{\phi_1/(\phi_1-\phi_2)}{1+\phi_1 B} - \frac{\phi_2/(\phi_1-\phi_2)}{1+\phi_2 B} \right] Y_t$$

که درآن B عملگر پسرو و $\theta_1 = \phi_1 + \phi_2$ و $\theta_2 = \phi_1 \phi_2$ است ، دراین صورت

می توان نوشت

$$e_t = \frac{\phi_1}{(\phi_1-\phi_2)} \left(1 - \phi_1 B + \phi_1^2 B^2 - \cdots\right) Y_t - \frac{\phi_2}{(\phi_1-\phi_2)} \left(1 - \phi_2 B + \phi_2^2 B^2 - \cdots\right) Y_t$$

$$e_t = \frac{1}{(\phi_1-\phi_2)}\left[(\phi_1-\phi_2)-(\phi_1^2-\phi_2^2)B+(\phi_1^3-\phi_2^3)B^2-\dots\right]Y_t$$

$$e_t = Y_t-(\phi_1+\phi_2)Y_{t-1}+(\phi_1^2+\phi_1\phi_2+\phi_2^2)Y_{t-2}-\dots+(-1)^{K-1}(\phi_1^{K-1}+\phi_1^{K-2}\phi_2+\phi_1^{K-3}\phi_2^2+\dots+\phi_2^{K-1})Y_{t-K+1}$$

$$e_t = Y_K-(\phi_1+\phi_2)Y_{K-1}+(\phi_1^2+\phi_1\phi_2+\phi_2^2)Y_{K-2}+\dots+(-1)^{K-1}(\phi_1^{K-1}+\phi_1^{K-2}\phi_2+\dots+\phi_2^{K-1})Y_1$$

لذا می توان نوشت

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{n-1} \\ e_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -(\phi_1+\phi_2) & 1 & \dots & 0 \\ (\phi_1^2+\phi_1\phi_2+\phi_2^2) & -(\phi_1+\phi_2) & \dots & 0 \\ & & & \\ (-1)^{n-1}(\phi_1^{n-1}+\phi_1^{n-2}\phi_2+\dots+\phi_2^{n-1}) & & \dots -(\phi_1+\phi_2) & 1 \end{bmatrix} \times \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_{n-1} \\ Y_n \end{bmatrix}$$

آنگاه

$$\hat{Q}_2(\phi_1,\phi_2) = \sum_{t=1}^{n}\left[e_t(Y;\phi_1,\phi_2)\right]^2 = \underline{Y}^T A^T(\phi_1,\phi_2)A(\phi_1,\phi_2)\underline{Y}$$

که

$$\underline{Y}^T = (Y_1, Y_2, \dots, Y_n)$$

$$A=\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -(\phi_1+\phi_2) & 1 & & 0 & 0 \\ (\phi_1^2+\phi_1\phi_2+\phi_2^2) & -(\phi_1+\phi_2) & & 0 & 0 \\ \vdots & & \ddots & & \\ (-1)^{n-1}(\phi_1^{n-1}+\phi_1^{n-2}\phi_2+\cdots+\phi_2^{n-1}) & & \cdots & -(\phi_1+\phi_2) & 1 \end{bmatrix}$$

بنابراین $Q_n(\phi_1,\phi_2)$ یک چند جمله ای از درجه $2n-2$ نسبت به ϕ_1 و ϕ_2 است . مجددا" ملاحظه می کنیم که شکل اولیه پیچیده است ولی آنرا می توان بصورت زیر خلاصه نمود.

فرض کنید B یک ماتریس $n\times n$ باشد

$$B=\begin{bmatrix} \gamma_1 & 0 & \cdots & 0 & 0 \\ \gamma_2 & -\gamma_1 & & 0 & 0 \\ \gamma_3 & -\gamma_2 & & 0 & 0 \\ \vdots & & \ddots & & \\ \gamma_n & -\gamma_{n-1} & \cdots & (-1)^{n-2}\gamma_2 & (-1)^{n-1}\gamma_1 \end{bmatrix}$$

و همچنین فرض کنید $\underline{u}$ یک بردار ستونی $n \times 1$ باشد

$$\underline{u}^T = (u_0, u_1, u_2, \ldots, u_{n-1})$$

که در آن u_i ، $i = 1, 2, 3, \ldots, n-1$ در روابط برگشتی زیر صدق می‌کند

$$u_0 = 1, \quad u_i = \phi_1 u_{i-1} + \phi_2^i \quad , \quad i = 1, 2, \ldots, n-1$$

لذا می توان نوشت :

$$Q_n(\phi_1, \phi_2) = \underline{u}^T (B^T B) \underline{u}$$

ما از این شکل ساده در یک برنامهٔ کامپیوتری استفاده کرده ابتدا ϕ_1 و ϕ_2 و سپس با توجه به روابط زیر θ_1 و θ_2 را برآورد می کنیم .

$$\hat{\theta}_1 = \hat{\phi}_1 + \hat{\phi}_2$$
$$\hat{\theta}_2 = \hat{\phi}_1 \hat{\phi}_2$$

یک مطالعه شبیه سازی

برای $\sigma^2 = 1$ و مقادیر ثابت β_1 و β_2 در ناحیه وارون پذیری یک صد مشاهده را از الگوی زیر تولید می کنیم

$$Y_t = e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} \quad , \quad t = 1, 2, 3, \ldots$$

با استفاده از روش فوق β_1 و β_2 را برآورد می نمائیم، $\overline{\hat{\beta}_1}$ و $\overline{\hat{\beta}_2}$ بترتیب متوسطهای برآوردهای β_1 و β_2 را در صد مرتبه تکرار روش فوق نشان می دهند.

جدول۲ متوسطهای $\overline{\hat{\beta}_1}$ و $\overline{\hat{\beta}_2}$ وانحراف معیار این متوسطها با استفاده ازروش کمترین مربعات

β_1	β_2	$\overline{\hat{\beta}_1}$	$\sigma(\overline{\hat{\beta}_1})$	$\overline{\hat{\beta}_2}$	$\sigma(\overline{\hat{\beta}_2})$
۰	-۰/۹	۰/۰۱۶۴	۰/۰۰۶۵	-۰/۸۵۰۳	۰/۰۰۸۴
۰	-۰/۵	۰/۰۱۲۰	۰/۰۱۱۰	-۰/۲۸۲۵	۰/۰۱۰۶
۰/۲	-۰/۵	۰/۲۸۹۷	۰/۰۰۹۷	-۰/۴۹۳۵	۰/۰۱۰۲
۰	-۰/۱	-۰/۰۱۰۵	۰/۰۱۱۲	-۰/۱۰۰۵	۰/۰۱۱۲
۰/۲	-۰/۱	۰/۲۹۹۸	۰/۰۱۰۹	-۰/۱۰۶۶	۰/۰۱۱۰
-۰/۲	-۰/۱	-۰/۳۰۸۱	۰/۰۱۰۱	-۰/۹۶۲۰	۰/۰۱۰۹
۰/۶	-۰/۱	۰/۵۸۶۳	۰/۰۱۱۳	-۰/۱۰۳۸	۰/۰۱۲۰
-۰/۶	-۰/۱	-۰/۵۸۷۷	۰/۰۱۲۰	-۰/۱۱۰۰	۰/۰۱۱۵
۰	۰/۱	۰/۰۱۰۸	۰/۰۱۱۴	۰/۱۰۷۵	۰/۰۱۰۷
۰/۲	۰/۱	۰/۳۰۴۶	۰/۰۱۰۴	۰/۱۰۶۴	۰/۰۱۰۱
-۰/۲	۰/۱	-۰/۲۹۹۰	۰/۰۱۰۹	۰/۰۹۶۸	۰/۰۱۱۲
۰/۶	۰/۱	۰/۵۸۹۵	۰/۰۱۲۰	۰/۰۱۵۷	۰/۰۱۱۴
-۰/۶	۰/۱	-۰/۵۸۴۱	۰/۰۱۴۱	۰/۰۹۷۶	۰/۰۱۰۶
۱/۲	۰/۵	۱/۱۲۸۰	۰/۰۱۰۸	۰/۴۸۰۱	۰/۰۱۱۲
-۱/۲	۰/۵	-۱/۱۵۰۴	۰/۰۱۲۵	۰/۴۸۱۱	۰/۰۱۰۷
-۰/۳	۰/۹	-۰/۲۹۱۲	۰/۰۰۸۷	۰/۸۴۸۶	۰/۰۰۷۸
۰/۶	۰/۹	۰/۵۸۴۵	۰/۰۱۰۰	۰/۸۴۴۷	۰/۰۰۸۱
-۰/۶	۰/۹	-۰/۵۸۹۹	۰/۰۰۸۴	۰/۸۴۹۲	۰/۰۱۱۰
۰/۹	۰/۹	۰/۸۲۷۷	۰/۰۰۷۶	۰/۸۴۳۵	۰/۰۰۶۴
۱/۲	۰/۹	۱/۱۵۶۲	۰/۰۰۸۷	۰/۸۴۹۰	۰/۰۱۰۲
-۱/۲	۰/۹	-۱/۱۶۰۰	۰/۰۰۸۴	۰/۸۴۶۵	۰/۰۰۶۲
۱/۵	۰/۹	۱/۲۴۶۱	۰/۰۰۶۲	۰/۸۵۱۲	۰/۰۰۹۷
-۱/۵	۰/۹	-۱/۲۵۴۲	۰/۰۰۷۶	۰/۸۴۹۷	۰/۰۰۸۱

بر مبنای نتایج شبیه سازی شدهٔ جدول ۲ ملاحظه می کنیم که برای $|\rho_1| < 0.7$ و $|\rho_2| < 0.5$ اریبی آشکار دیده نمی شود. ولی در $|\rho_1| > 0.7$ و $|\rho_2| = 0.9$ اریبی آشکاری وجود دارد. این موضوع با توجه به آمارهٔ آزمون $d(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{S.d(\hat{\beta}_1)}$ و $d(\hat{\beta}_2) = \frac{\hat{\beta}_2 - \beta_2}{S.d(\hat{\beta}_2)}$ مورد تأیید قرار می گیرد. فرمول مجانبی انحراف معیار برآورد کنندهٔ درستنمایی ماکزیمم پارامترهای β_1 و β_2 عبارت است از $\left(\frac{1-\rho_2^2}{N}\right)^{1/2}$ که با مقادیر جدول که از شبیه سازی بدست آمده اند مطابقت دارد.

بررسی اثرات عوامل اجتماعی و اقتصادی در باروری
ارائه یک مدل رگرسیونی

فاطمه هرندی
مرکز آمار ایران

خلاصه

این مطالعه به منظور بررسی عوامل اجتماعی ـ اقتصادی موثر بر باروری در ایران انجام گرفته است .

بدین منظور از اطلاعات بدست آمده از اجرای مرحله اول طرح آمارگیری از خصوصیات اجتماعی ـ اقتصادی خانوار استفاده شده است . این اطلاعات از ۴۷۶۲ خانوار نمونه که از کلیه نقاط شهری و روستایی کشور با استفاده از روشهای نمونه گیری انتخاب شده اند ، جمع آوری شده است .

جمعا" ۴۱۴۳ زن ازدواج کرده مورد بررسی قرار گرفتند و با استفاده از برنامه از پیش نوشته شده SPSS تجزیه وتحلیلهای آماری صورت گرفت .

نتایج بدست آمده عبارتنداز :

۱ ـ متوسط تعداد فرزند زنده بدنیا آمده هرزن ازدواج کرده ۵/۱ است

۲ ـ سن زن هنگام ازدواج در باروری موثر است .

۳ ـ اختلاف معنی دار بین متوسط تعداد فرزند زنان شهری وزنان روستایی وجود دارد .

۴ ـ درآزمون همبستگی رابطه شدید بین باروری ووضع سواد دیده شده است .

۵ ـ عوامل وضع مهاجرت مادر ووضع فعالیت مستقل بر باروری موثر نیستند .

۱ ـ مقدمه

موضوع باروری درکشورهای درحال توسعه از جمله کشور ما از اهمیت خاص برخوردار است . میزان کنونی باروری زنان ایران فقط از تعداد معدودی از کشورهای بسیارعقب مانده آسیایی و افریقایی کمتر است وعنوان سیزدهمین کشور پرزاد و ولد جهان را یدک می‌کشد .

رشد جمعیت حدود ۳/۲ درصد درسال که بمنزله تنوان دوبرابر شدن جمعیت طی یك دوره کوتاه ۲۲ ساله است ، جوانی جمعیت ، بار تکفل اقتصادی شدیدی که به خانوارها تحمیل شده است ، گسترده شدن ابعاد خانوار و ... علاوه بر گرفتاریها ومشکلات زیادی که بر سر راه حرکت های تکاملی ورشد سازندگی کشور ایجاد کرده ، نیل سریع به آهنگ متناسب رشد جمعیت را نیز مشکل کرده است ، چرا که جوانی بسیار شدید جمعیت فعلی و پتانسیل وتوان زایشی بالقوه زیادی که معمولا " در جمعیتی اینسان جوان نهفته است وطی ۳۰ سال آینده از قوه به فعل درمی‌آید ، یکی از موانع عمده کاهش سریع میزان باروری ودرنتیجه رشد طبیعی جمعیت کشوراست .

به این ترتیب ضرورت کاهش نرخ رشد جمعیت اجتناب ناپذیر است .

عملکردهایی صرفا " مبتنی بر روشهای مختلف جلوگیری از بارداری های ناخواسته ، گرچه درکاهش موالید موثراست ، اما به تنهایی کافی نیست وبه نتایج مورد انتظار منتهی نمی‌شود ؛ چون رفتارهای باروری خانوارها عموما " تحت تاثیر وضعیت اجتماعی اقتصادی آنها است .

قشرهایی از جامعه که دچار انفجار جمعیتند غالبا " همانهایی هستند که بطور نسبی در مجموعه‌ای از شرایط نامناسب اجتماعی ـ اقتصادی وآکنده از محرومیت و کاستی به سر می‌برند . دراین جوامع ازدواج زودرس دختران که عموما " به معنی ترك آموزشهای دوره راهنمایی ودبیرستانی نیز هست ، نرخ رشد جمعیت را بالا می‌برد .

اصولا " کاهش موالید وکنترل جمعیت دراین اقشار کاهش بار تکفل اقتصادی خانوار ، افزایش توان اقتصادی خانوار وبرخوردارشدن کودکان از مواهب توسعه نظیر آموزش را به دنبال دارد .

۲ ـ بررسی مطالعات انجام شده

دراین قسمت نتایج بررسی های انجام شده قبلی درمورد باروری وارتباط آن با فاکتورهای

اجتماعی اقتصادی ارائه می‌شود.

۱-۲- درطرح نمونه گیری باروری که درسال ۱۳۶۸ توسط سازمان ثبت احوال فر۵ استان همدان، ایلام، آذربایجان غربی، همدان وکردستان انجام گردید، نتایج زیر بدست آمده است:

- سطح سواد مادر با تعداد فرزندان او همبستگی منفی دارد.

میانگین تعداد فرزندان برحسب سطح سواد مادر

استان	سطح سواد	
	کمتراز راهنمایی	راهنمایی وبالاتر
همدان	۴/۳	۲/۳
کردستان	۴/۲	۲/۲

استان	بی‌سواد	کم سواد	پایان راهنمایی وبالاتر
ایلام	۴/۶	۳/۲	۲/۵
آذربایجان غربی	۴/۲	۲/۵	۲/-

استان	بی سواد	تا پایان راهنمایی	دیپلم وبیشتر
سمنان	۵	۲/۹	۲/۱

- متوسط تعداد فرزند مادران شهری کمتراز متوسط تعداد فرزند مادران روستایی است.

میانگین تعداد فرزندان برحسب محل سکونت مادر

استان	شهری	روستایی
همدان	۳/۷	۴/۴
کردستان	۴/-	۴/۲
ایلام	۲/۱	۴/۳
آذربایجان غربی	۳/۲	۴/۱
سمنان	۲/۸	۴/۱

- متوسط تعداد فرزند زنان شاغل کمتر از متوسط تعداد فرزند زنان غیر شاغل است .

متوسط تعداد فرزندان برحسب وضعیت اشتغال مادر

استان	وضع فعالیت	
	غیر شاغل	شاغل
همدان	۴/۱	۲/۱
کردستان	۴/۲	۱/۷
ایلام	۴/۳	۳/۷
آذربایجان غربی	۴/۷	۲/۴
سمنان	۴/۴	۲/۳

۲-۳- دربررسی باروری زنان ایران که درسال ۱۳۵۶ توسط مرکز آمار ایران در رابطه با طرح W.F.S (WORLD FERTILITY SURVEY) انجام شد اطلاعاتی از ۴۸۹۰ زن ازدواج کرده ۲۹-۱۵ ساله درمناطق شهری وروستایی کشور جمع آوری گردید که نتایج آن به شرح زیر می‌باشد:

- متوسط تعداد فرزندان زنان شهری ۲/۲ و زنان روستایی ۲/۹ می‌باشد که حاکی از اختلاف درمیزان بهره‌وری زنان شهری وروستایی نسبت به یکدیگر است، به‌طوری که تقریبا " سهم باروری زنان روستایی ۱۵ تا ۲۰ درصد بیش از زنان شهری گزارش شده است .

- بین سن در اولین ازدواج وسطح سواد زن همبستگی مثبتی وجود دارد .

۲-۴- دربررسی رابطه باروری وعوامل اقتصادی اجتماعی که توسط مرکز جمعیت شناسی دانشگاه شیراز درسال ۱۳۵۲ در ۶ ده استان فارس انجام گردید نتایج زیر بدست آمده است :

- رابطه مستقیم بین شغل شوهر ومیزان باروری وجود دارد . به طوری که میانگین تعداد فرزندان زنده بدنیا آمده خانوارهای کشاورز ۵/۱ و خانوارهای کارگر (بخش کشاورزی یا غیر کشاورزی) ۳/۹ است .

- رابطه مستقیم مثبتی بین زمین داری وباروری وجود دارد . چه میانگین تعداد فرزند زنده بدنیا آمده درمیان زنان خانوارهایی که ۱۰ هکتار ویا بیشتر زمین دارند ۵/۸ است وبرای خانوارهایی که کمتر از ۱۰ هکتار زمین دارند ۴/۹ وبرای خانوارهای بدون زمین ۳/۸ است .

۴-۳- در مطالعه‌ای که در سال ۱۳۴۷ در دفتر جمعیت ونیروی انسانی سازمان برنامه وبودجه انجام گرفت نتایج زیر ارائه گردید : ابعاد بالای باروری در ایران از کاهش نرخ مرگ ومیر کودکان ، پیش‌رسی ازدواج ، عمومیت ازدواج ، سطح آموزش وآگاهیهای اجتماعی وشرایط زیست مردم جامعه تاثیر می‌پذیرد .

- میزان باروری درجوامع شهری کمتر از جوامع روستایی است .

- نقاط محروم وعقب مانده ، قشرهای کم بینه وفقیر وشاغلان بخشهای سنتی در قیاس با نقاط پیشرفته ، قشرهای پردرآمد وشاغلان بخشهای مدرن باروری بیشتری دارند . استانهای کهگیلویه وبویراحمد، چهارمحال وبختیاری ، سیستان وبلوچستان ، ایلام ، لرستان ، کردستان ، هرمزگان ، بوشهر وقسمتهایی از کرمان وجنوب خراسان بیش از سایر نواحی کشور از باروری بالا آسیب دیده‌اند (بخصوص جوامع روستایی آنها) .

- در آزمون همبستگی رابطه شدید بین باروری باباسوادی ، وضع اشتغال وسن ازدواج زن دیده شده است .

در بین شهرنشینان کارکنان ساختمانی ، کارکنان خرده‌پای خدماتی ، مزد وحقوق بگیران سطوح پایین خدمات دولتی وبطور کلی حاشیه نشینان شهری که غالبا" مهاجران روستایی اند باروری بالاتری دارند .

۴-۴- مطالعه‌ای درشهر تنکابن درمورد باروری زنان روستایی ورابطه آن با موقعیت اجتماعی-اقتصادی خانوار انجام گرفت . براساس نتایج این مطالعه درآمد ونسبت باسوادی درخانوار بابارووری بستگی منفی دارد .

۴-۵- بررسی درسال ۵۲ درمورد باروری درشهر تهران توسط مهدی امانی انجام گرفت . براساس نتایج این بررسی دررابطه با میزان باروری زنان باشغل شوهران آنها تفاوت بارزی وجود دارد . بطوریکه درشغل های بادرآمد بیشتر میزان باروری کمتر است . سواد زن وشوهر هردو دربارووری موثرند (در اکثر موارد وقتی زن باسواد است شوهر هم باسواد است) .

- سن زن هنگام ازدواج عاملی موثر دربارووری است .

- باروری باوضع فعالیت زن (شاغل یاغیر شاغل بودن او) ارتباط دارد .

۳ـ انگیزه بررسی

انگیزه این بررسی تایید فرضیه های موجود (درصورت صحت)درمورداثر عوامل اجتماعی اقتصـــادی بر باروری ورسیدن به حقایق جدید است .

۴ـ هـــدف

هدف این مطالعه تعیین نوع وشدت ارتباط فاکتورهای اجتماعی واقتصادی برباروری درکشـــــور با استفاده از رابطه رگرسیونی است .

۵ـ روش پـــژوهش وجمعیت تحت مطالعه

برای انجام این بررسی از اطلاعات فایلهای طرح آمارگیری از خصوصیات اجتماعی ـ اقتصادی خانـــوار (مرحله اول) استفاده شده است و جهت تجزیه وتحلیل وانجام محاسبات مورد نظـــر بسته نرم افزاری SPSS بکار گرفته شده است .

علل انتخاب طرح فوق عبارتنداز : تنوع اطلاعات جمع آوری شده درآن ، پوشش طرح که کلیه نقـاط شهری ، روستایی کشور را دربر میگیرد ، و تعاقبی بودن طرح که درمطالعات آینده امکان بررســـــی روند تغییرات درطول زمان را فراهم می آورد .

به منظور آشنایی بیشتر بافایل های مورد استفاده ، خلاصه ای از طرح آمارگیری از خصوصیـــات اجتماعی ـ اقتصادی خانوار دوذیل ارائه می شود .

۵ـ۱ـ کلیات طرح آمارگیری از خصوصیات اجتماعی ـ اقتصادی خانوار

ـ هـدف : طرح آمارگیری از خصوصیات اجتماعی ـ اقتصادی خانوار ، به منظور اطلاع از ویژگیهـــای اجتماعی ـ اقتصادی جامعه وروند تغییرات آنها درمناطق شهری وروستایی تهیه شده است . بـــــــــرای ثبت روند مذکور تغییر ویژگیهای خانوار درطول زمان مورد پیگیری قرارمی گیرد . به همین دلیل ایـــــن روش را روش پیگیری وهریک از مراجعات را یک مرحله از اجرای طرح می نامند . این طرح از ۱۲ مرحلـــه تشکیل شده وفاصله زمانی بین هردومرحله ۳ ماه است .

ـ جامعه آماری : کلیه خانوارهای معمولی ثابت ایرانی مقیم درنقاط شهری وروستایی کل کشور

ـ واحد آماری : خانوار معمولی ثابت ایرانی

- چارچوب : درنقاط روستایی از فهرست آبادیها درسرشماری ۱۳۵۵ استفاده شده است . درنقاط شهری نیز از فهرست بلوکها که درفاصله سالهای ۶۳-۶۰ تهیه شده ، استفاده گردیده است .

- زمان آمارگیری : دهه دوم ماه اول هرفصل است . اولین مرحله طرح ۱۰ تا ۲۰ فروردین ۱۳۶۶ اجرا شد .

- روش جمع آوری اطلاعات : تکمیل پرسشنامه های این طرح از طریق مصاحبه مستقیم با واحد آماری (خانوار) صورت پذیرفته .

- روش نمونه گیری : روش نمونه گیری این طرح خوشه‌ای سومرحله ای بود . واحد مرحله اول یک خوشه ۳۰ خانواری است (منظور از خوشه یک یا چند بلوک یا قسمتی از یک بلوک ونیز یک یا چند آبادی ویا قسمتی از یک آبادی است ، که حدود ۳۰ خانوار را دربرگرفته باشد . واحد مرحله دوم خانوار است . ابتدا استانها براساس ضریب محرومیت (منتج از طرح مقدماتی شناسایی مناطق محروم کشور - نتایج اجرایی مرحله سوم) به ۶ طبقه تقسیم شد . سپس با استفاده از روش PPS* (متناسب با تعداد خانوار) تعداد ۱۶۰ خوشه نمونه بین طبقات توزیع شد . تعداد خوشه های نمونه شهری ۹۸ خوشه وتعداد خوشه های نمونه روستایی ۶۲ خوشه است . برای انتخاب محدوده ای که خوشه نمونه درآن قرار دارد (بلوک یا آبادی) نیز از روش PPS استفاده شد (متناسب با تعداد خانوار) درصورتی که بلوک یا آبادی انتخاب شده کمتر از ۳۰ خانوار داشت با بلوکها ویا آبادیهای مجاورش ادغام گردید .

- فرمهای آمارگیری :

الف - فرم ۱ (فهرست برداری) : منظور از تکمیل این فرم فهرست کردن کلیه خانوارها و مکانهای واقع در محدوده ای است که خوشه نمونه درآن واقع شده است .

ب - فرم ۲ (پرسشنامه خانوار مرحله اول) : منظور از تکمیل این فرم ثبت خصوصیات اعضای خانوار، درآمد ، هزینه وپاره ای از اطلاعات مربوط به خانوارهای نمونه درمرحله اول است .

ج - فرم ۳ (کاربرگ مخصوص گزارشگر) : منظور از تکمیل این فرم درهر مرحله ثبت خصوصیات اعضاء ، ومشخصه‌های مسکن وپاره‌ای اطلاعات دیگر مربوط به خانوارهای نمونه است .

د - فرم ۴ (فرم واسط) : منظور از تکمیل این فرم پس از پایان آمارگیری درهر مرحله ، انتقال

* PROBABILITY PROPERTIONAL TO SIZE

اطلاعات مندرج درفرم ۲ است .

در مرحله اول آمارگیری فرمهای ۱ و۲ تکمیل شد ودر مراحل بعدی ابتدا کار برگ مخصوص گزارشگر توسط آمارگیر تکمیل شده است ، سپس اطلاعات مندرج در کار برگ به فرم واسط منتقل وبه مرکز ارسال شد . کار برگ مخصوص گزارشگر تاپایان آخرین مرحله آمارگیری نزد سرپرست آمارگیری باقی ماند .

- روش تعیین خوشه نمونه : پس از تکمیل فرم فهرست برداری درمحدوده ای که خوشه نمونه درآن واقع است،با انتخاب یک مبدا، تصادفی ، اولین خانوار واقع درخوشه نمونه تعیین شد و شماره مکان این خانوار به عنوان نقطه شروع خوشه منظور شد . به دنبال آن تا ۲۰ خانوار از روی فرم فهرست بطور مسلسل انتخاب شد . فاصله مکان اولین خانوار تا آخرین خانوار نمونه ، خوشه نمونه را تشکیل می‌دهد .

- حجم نمونه : تعداد خانوار نمونه درمرحله اول طرح ۴۲۶۲ خانوار وجمعیت این خانوارها ۲۲۲۹۱ نفر است . ازاین تعداد ۱۲۲۲۶ نفر مرد و۱۱۹۵۵ نفر زن هستند .

- عملیات میدانی : دراین طرح از وجود پرسشگران محلی استفاده شد . این پرسشگران از بین آموزگاران محلی انتخاب شدند .

۲-۴- تعریف متغیرها وواژه‌ها :

خانوار معمولی ثابت - عبارت از چند نفر که باهم دریک اقامتگاه زندگی می‌کنند ، بایکدیگر هم خرج هستند ، معمولا " باهم غذا می خورند و در اقامتگاه ثابتی واقع درمحدوده یکی از شهرها یا آبادیهای کشور زندگی می‌کنند و اقامتگاه خود را تغییر نمی دهند و یا درصورت تغییر این تغییر به طور سالانه ، منظم و مستمر نیست .

شهر - هریک از نقاطی که تاپایان سال ۶۲ دارای شهرداری بوده است .

سن - تعداد سالهای کاملی که از تولد فرد گذشته است .

مهاجران وارد شده - درهرشهر یا آبادی ، آن دسته از اعضای خانوارهای معمولی ثابت که متولد آنجا نباشند مهاجر شناخته می‌شوند .

محل تولد - منظور از محل تولد ، آبادی یا شهری است که درزمان تولد فرد اقامتگاه معمولی صاحب در

اوبوده است .

باسواد ـ کسانی که می‌توانند به فارسی یا هر زبان دیگری متن ساده ای را بخوانند وبنویسند ، بنابراین سواد تلقی می‌شوند .

کار ـ کار ، هرفعالیت فکری یا بدنی است که قانوناً مجاز باشد وبه منظور کسب درآمد (نقدی ، غیر نقدی) صورت گیرد .

شاغل ـ افراد زیر شاغل به حساب می‌آیند :

الف ـ کسانی که دارای شغلی هستند ، اعم از اینکه در ۷ روز پیش از مراجعه مامور آمارگیری کار کرده ویا به عللی از قبیل مرخصی ، بیماری ، تعطیلی موقت محل کار و ... کار نکرده‌اند و پس از رفع علت به کار خود ادامه خواهند داد .

ب ـ کسانی که شغل مستمر نداشته ولی در ۷ روز پیش از مراجعه مامور آمارگیری ، حداقل یکروز کار کرده‌اند .

ج ـ کسانی که تحت عنوان کارآموز درهفت روز پیش از مراجعه مامور آمارگیری به کار اشتغال داشته اند .

د ـ کسانی که دارای شغلی هستند ، ولی درهفت روز پیش از مراجعه مامور آمارگیری به اقتضای فصل وماهیت فصلی کار خود ، کار نکرده اند (بیکاران فصلی) مشروط براینکه در جستجوی کار دیگری نباشند .

هـ ـ افرادی که به انجام خدمت وظیفه عمومی مشغول هستند .

و ـ کسانی که درحال گذراندن دوره های آموزشی درداشکده ها وآموزشگاههای مربوطه نیروهای نظامی وانتظامی هستند .

درآمد اعضای خانوار ـ منظور از درآمد ، کلیه وجوه یا کالاها یا ترکیبی از آنها است که دربرابر کار انجام شده یا سرمایه بکار افتاده ویا از طریق دیگر نظیر حقوق بازنشستگی ، دریافتی‌های اتفاقی و ... به صورت نقدی یا غیرنقدی درزمان معینی به فرد تعلق گیرد .

وضع زناشویی ـ برای هریک از اعضای ۹ ساله وبالاتر خانوار وضعیت زناشویی درزمان آمارگیری

مد نظر قرار می گیرد ، که بسته به مورد دارای همسر ، بی همسر براثر طلاق ، بی‌همسربراثرفوت همسر ، وهرگزازدواج نکرده ، محسوب می‌شوند .

تعدادکل فرزندان زنده به دنیا آورده -منظوراز " تعدادکل فرزندان زنده بدنیا آورده " برای هرزن ازدواج کرده تعداد اطفالی که وی تازمان آمارگیری زنده بدنیا آورده است ، نوزادی که بعداز تولد حداقل یکی از آثار حیات دراو مشاهده شود ، حتی اگر بلافاصله فوت کند ، زنده متولد شده به حساب می‌آید .

۶ ـ یافته ها *

۶ـ۱ـ جداول توزیع فراوانی نسبی

توزیع فراوانی نسبی ومتوسط تعداد فرزند زنده بدنیا آورده‌زنان ازدواج‌کرده‌برحسب سن‌مادر

سن مادر	کمتراز ۱۶ سال	۱۶ـ۲۵	۲۶ـ۳۵	۲۶ـ۴۵	۴۶ـ۵۵	۵۶وبیشتر	کل
فراوانی‌نسبی	۰/۶	۲۲/۵	۳۱/۱	۲۱/۷	۱۳/۷	۹/۴	۱۰۰
متوسط تعداد فرزندزنده بدنیاآورده	۰/۲	۱/۹	۴/۴	۶/۷	۷/۷	۷/۷	۵/۱

توزیع فراوانی‌نسبی‌ومتوسط تعدادفرزند زنده‌بدنیا آورده‌زنان ازدواج‌کرده برحسب وضع سوادمادر

وضع سواد مادر	باسواد	بی‌سواد	کل
فراوانی‌نسبی	۳۷/۹	۶۲/۱	۱۰۰
متوسط تعداد فرزندزنده بدنیاآورده	۳/۰	۶/۳	۵/۱

* یافته های ارائه شده دراین بررسی مربوطبه زنان ازدواج کرده ای است که همسر سرپرست خانوارمیباشند .

توزیع فراوانی نسبی و متوسط تعداد فرزند بدنیا آورده زنان ازدواج کرده برحسب محل سکونت مادر

محل سکونت مادر	شهر	روستا	کل
فراوانی نسبی	۶۰/۴	۳۹/۶	۱۰۰
متوسط تعداد فرزند به زنده بدنیا آورده	۴/۲	۶/۱	۵/۱

توزیع فراوانی نسبی و متوسط تعداد فرزند زنده بدنیا آورده زنان ازدواج کرده برحسب وضع فعالیت مادر

وضع فعالیت مادر	شاغل	غیرشاغل	کل
فراوانی نسبی	۷/۸	۹۲/۲	۱۰۰
متوسط تعداد فرزند زنده بدنیا آورده	۴/۵	۵/۱	۵/۱

توزیع فراوانی نسبی و متوسط تعداد فرزند زنده بدنیا آورده زنان ازدواج کرده برحسب وضع مهاجرت مادر

وضع مهاجرت مادر	مهاجر	غیرمهاجر	کل
فراوانی نسبی	۵۴/۰	۴۶/۰	۱۰۰
متوسط تعداد فرزند زنده بدنیا آورده	۵/۱	۵/۰	۵/۱

توزیع فراوانی نسبی و متوسط تعداد فرزند زنده بدنیا آورده زنان ازدواج کرده برحسب سن مادر هنگام ازدواج

سن مادر هنگام ازدواج	کمتر از ۱۶ سال	۱۶-۲۵	۲۶-۳۵	۳۶ و بیشتر	کل
فراوانی نسبی	۳۱/۷	۶۴/۹	۳/۲	۰/۲	۱۰۰
متوسط تعداد فرزند زنده بدنیا آورده	۵/۸	۴/۸	۳/۶	۳/۱	۵/۱

۶-۲- مدل رگرسیون

به منظور تعیین اینکه تعداد فرزند زنده بدنیا آورده ،هر زن ازدواج کرده با چه عواملی مرتبط است رگرسیون ۶متغیره ای به شرح زیر محاسبه گردید (حجم نمونه مورد بررسی ۱۲۱۲ نفر زن ازدواج کرده است) .

متغیر وابسته : تعداد فرزند زنده بدنیا آورده هرزن ازدواج کرده

متغیرهای مستقل :

۱ ـ سن مادر

۲ ـ وضع سواد مادر (باسواد = ۱ بی سواد = ۰)

۳ ـ سن مادر هنگام ازدواج

۴ ـ محل سکونت مادر (شهر = ۱ روستا = ۰)

۵ ـ وضع مهاجرت مادر (مهاجر = ۱ غیر مهاجر = ۰)

۶ ـ وضع فعالیت مادر (شاغل = ۱ غیر شاغل = ۰)

به منظور دستیابی به مدل مناسب از روش محاسبه " گام به گام* " استفاده گردید . بدین ترتیب ابتدا متغیری که بیشترین میزان همبستگی را به متغیر وابسته داشت تعیین ومعادله رگرسیون براساس این متغیر محاسبه شد . درمرحله بعد با ثابت نگه داشتن این متغیر ، از بین باقیمانده متغیرها ، متغیری که بیشترین همبستگی را با متغیر وابسته دارا بود انتخاب وبراساس این دومتغیر معادله رگرسیون محاسبه گردید . ودر مراحل بعدی نیز به همین ترتیب عمل شد . از بین مدلهای بدست آمده مدلی که مناسب ترین شرایط را دارا بود انتخاب وارائه گردید .

---

* STEP WISE

ـ ماتریس ضرایب همبستگی ساده

نام متغیر	تعداد کل فرزندان زنده بدنیا آورده	سن مادر	وضع سواد مادر	سن مادر هنگام ازدواج	محل سکونت مادر	وضع مهاجرت مادر	وضع فعالیت مادر
تعداد کل فرزندان زنده به دنیا آورده	۱						
سن مادر	۰/۶۱	۱					
وضع سواد مادر	۰/۴۸-	۰/۳۸-	۱				
سن مادر هنگام ازدواج	۰/۱۷-	۰/۱۴	٪۹	۱			
محل سکونت مادر	۰/۲۷-	۰/۱-	۰/۳۲	٪۶	۱		
وضع مهاجرت مادر	۰/۰۱	۰/۰۱	۰/۰۲	۰/۰۲	۰/۲۲-	۱	
وضع فعالیت مادر	۰/۰۵-	۰/۰۱	٪۵	۰/۱۴	۰/۰۷-	۰/۰۲	۱

.. خلاصه محاسبات مراحـــل مختلف

مرحله	متغیرهای مستقل داخل معادله				سایر متغیرهای مستقل	
	نام متغیر	ضریب همبستگی چندگانه *	ضریب تعیین ** رگرسیون	ثابت معادله	نام متغیر	ضریب همبستگی جزئی ***
۱	سن مادر	۰/۶۱	۰/۳۷	۲۴۵۲/۹	وضع سواد مادر سن مادر هنگام ازدواج محل سکونت مادر وضع مهاجرت مادر وضع فعالیت مادر	۰/۳۴ — ۰/۳۲ — ۰/۲۶ — ۰/۰۰۲ — ۰/۰۷ —
۲	سن مادر وضع سواد مادر	۰/۶۷	۰/۴۵	۱۶۷۰	سن مادر هنگام ازدواج محل سکونت مادر وضع مهاجرت مادر وضع فعالیت مادر	۰/۲۹ — ۰/۱۶ — ۰/۰۱ ۰/۰۵ —
۳	سن مادر وضع سواد مادر سن مادر هنگام ازدواج	۰/۷۰	۰/۴۹	۱۳۵۰/۷	محل سکونت مادر وضع مهاجرت مادر وضع فعالیت مادر	۰/۱۶ — ۰/۰۲ ۰/۰۱ —
۴	سن مادر وضع سواد مادر سن مادر هنگام ازدواج محل سکونت مادر	۰/۷۱	۰/۵۱	۱۰۶۵	وضع مهاجرت مادر وضع فعالیت مادر	۰/۰۲ — ۰/۰۲ —

*MULTIPLE CORRELATION COFFICIENT
**COFFICIENT OF DETERMINATION
***PARTIAL CORRELATION COFFICIENT

با ورود متغیرهای فوق در معادله رگرسیون ضمنا چون دیگر متغیرهای مستقل مورد مطالعه با متغیر وابسته ضعیف بوده ودرنتیجه F آنها درسطح $\alpha = 0/05$ معنی دارنیست ، باین ترتیب مدلی که مناسب تشخیص داده شد مدل زیر میباشد :

$$\hat{y} = 4/53 + 0/14 x_1 - 1/28 x_2 - 0/20 x_3 - 0/83 x_4$$

که درآن :

$\hat{y}$: تعداد فرزند زنده بدنیا آورده　　x_1 : سن مادر　　x_2 : وضع سوادمادر

x_3 : سن مادر هنگام ازدواج　　x_4 : محل سکونت مادر

- جدول آنالیز واریانس :

منبع تغییرات	SS	df	MS	F
رگرسیون	۱۲۳۹۸/۲	۴	۵۹۷۲/۵	۱۰۶۵/۰۱
باقیمانده	۲۲۲۱۲/۵	۲۱۳۸	۵/۶	

باتوجه به مقدار F محاسبه شده درجدول فرض صفر $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ بسطح ۹۹٪ اطمینان رد می‌شود ، بنابراین حداقل یکی از β_i ها مخالف صفر است .

- حدود اعتماد ۹۵ درصد ضرایب معادله رگرسیون عبارتنداز :

$$-1/25 < \beta_2 < -1/31 \qquad 0/134 < \beta_1 < 0/147$$

$$-0/99 < \beta_4 < -0/67 \qquad -0/22 < \beta_3 < -0/18$$

این فواصل اعتماد نشان می‌دهند که فرض صفر بودن هریک از ضرایب β_1 ، β_2 ، β_3 و β_4 با ۹۵ درصد اطمینان رد می‌شود .

- مقادیر استاندارد ضرایب معادله رگرسیون

$$b'_1 = 0/55 \qquad b'_2 = -0/21$$

$$b'_3 = -0/22 \qquad b'_4 = -0/13$$

- ضریب همبستگی چند گانه

$$R = 0/71$$

- ضریب تعیین رگرسیون

$$R^2 = 0/51$$

- ضریب تعیین رگرسیون تعدیل شده*

$$\bar{R}^2 = ۰/۵۱$$

مقدار R^2 نشان می‌دهد که ۵۱ درصد تغییرات Y توسط مدل به دست آمده توجیه می‌شود.

- تست خود همبستگی** با استفاده از آزمون دوربین-واتسون***

آزمون دوداملـــــه

H_0 : خود همبستگی وجود ندارد (مثبت یامنفی)

H_1 : خود همبستگی وجود دارد (مثبت یامنفی)

$$d_L = ۱/۴۲ \quad , \quad d_U = ۱/۶۳$$

$$\text{محاسبه شده } d = ۱/۹$$

خود همبستگی مثبت وجود ندارد $\Longleftarrow d > d_U$

خودهمبستگی منفی وجود ندارد $\Longleftarrow ۴ - d > d_U$

بنابر این با احتمال ۹۸٪ خودهمبستگی وجود ندارد.

۷-نتیجه گیری

۷-۱- متوسط تعداد فرزند زنده بدنیا آمده هرزن ازدواج کرده ۱/۵ بدست آمده است

۷-۲- متوسط تعداد فرزند در حال حاضر زنده ،هرزن ازدواج کرده ۱/۲ بدست آمده است.

۷-۳- وضع سواد مادر با تعداد فرزندان زنده بدنیا آورده ها مرتبط است : به طوری که نتایج حاصـــــــــــل نشان می‌دهد متوسط تعداد فرزند زنان با سواد ۲ وزنان بی سواد ۶/۳ است . میزان همبستگی بین ایـــــن دو متغیر ۰/۴۸ است.

A ADJUSTED COEFFICIENT OF DETERMINATION

** AUTOCORRELETION

*** DURBIN-WATSON TEST

۷ـ۲ـ محل سکونت مادر با تعداد فرزندان زنده بدنیا آمده او مرتبط است . نتایج حاصل از بررسی نشان می‌دهد که متوسط تعداد فرزند زنان شهری ۲/۲ وزنان روستایی ۱/۶ است . میزان همبستگی بین این دوصفت ۰/۲۲ ـ است .

۷ـ۵ـ سن مادر هنگام ازدواج با تعداد فرزندان او رابطه دارد . نتایج حاصل نشان می‌دهد که متوسط تعداد فرزند مادرانی که درسن ۲۶ تا ۳۵ سالگی ازدواج نموده اند ۳/۶ است ، درصورتیکه برای مادرانی که درسن کمتراز ۱۶ سال ازدواج کرده اند این متوسط ۵/۸ است . میزان بستگی بین این دوصفت ۰/۱۷ ـ است .

۷ـ۶ـ عوامل وضع مهاجرت مادر ، وضع فعالیت مادر برتعداد فرزندان او موثر نیستند .

۸ـ فهرست منابع ومآخذ

۱ـ عجمی ـ دکتر اسماعیل

رابطه باروری وعوامل اجتماعی ـ اقتصادی ، مرکز جمعیت شناسی دانشگاه شیراز ، مطالعه درشهر ده فارس ، سال ۱۳۵۳

۲ـ مرکزآمارایران

نتایج بررسی باروری زنان ایران ، در رابطه با طرح W.F.S ، سال ۱۳۵۶

۳ـ سازمان برنامه وبودجه

نگرشی جمعیت شناختی برتشکل خانواده وباروری درایران ، معاونت اموراقتصادی ، دفتر جمعیت ونیروی انسانی ، سال ۱۳۷۰

۴ـ پیمان ـ حبیب الله ، هوشور ـ زردشت ، افتخار اردبیلی ـ حسن

بررسی باروری زنان روستایی دررابطه با موقعیت اجتماعی ـ اقتصادی خانوار شهرستان تنکابن ، دانشکده بهداشت دانشگاه تهران ، سال ۱۳۶۲

۵ـ سازمان ثبت احوال

نتایج نمونه گیری باروری دراستانهای کردستان ، ایلام ، همدان ، آذربایجان غربی ، سمنان ، سال ۶۹ـ۱۳۶۸

۶۶ـ امانی ـ مهدی

باروری درتهران ، دانشگاه تهران ، سال ۱۳۵۲

۶۷ـ جعفری صمیمی ـ احمد

مبانی اقتصاد سنجی ، سال ۱۳۶۲

۶۸ـ خواجه نوری ـ عباسقلی

آمار پیشرفته وبیومتری ، سال ۱۳۲۷

معرفی طرح آمارگیری از قیمت فروش محصولات و هزینهٔ خدمات کشاورزی

فرخ یونسی

مرکز آمار ایران

المقدمه

اطلاع از کم‌وکیف فعالیتهای مرتبط باکشاورزی به دلیل گستردگی آن و دامنه محصول این فعالیت در مناطق روستایی کشور، از اطلاعات پایه‌ای برای برنامه‌ریزی به شمار می‌رود. موضوع این مقاله معرفی طرح آمارگیری ازقیمت فروش محصولات وهزینه خدمات کشاورزی در مناطق روستائی(۱) است. تحولات، تغییرات، نحوه تعیین اقلام، نحوه تعیین واحد آماری مورد مراجعه، چگونگی جمع‌آوری اطلاعات وسایر ویژگیهای این طرح مورد بحث قرار می‌گیرد.

براساس نتایج حاصل از اجرای سرشماریهای جمعیتی درسالهای ۳۵، ۴۵و۵۵ جمعیت ساکن مناطق روستائی به ترتیب حدود ۶۹٪، ۶۲٪ و ۵۲٪ بوده است. علی رغم تحولات جمعیتی وکاهش نسبی جمعیت روستائی، هنوز حدود نیمی ازجمعیت کشور ساکن مناطق روستائی می‌باشند. در سال ۱۳۶۵ بنابه نتایج سرشماری عمومی نفوس‌ومسکن، ازجمعیت ۴۹ میلیون‌نفری ایران حدود ۲۲ میلیون نفر ساکن مناطق روستائی بوده‌اند. دراین سال ازجمعیت ۱۰ ساله وبالاتر، ۱۱ میلیون نفر شاغل بوده‌اند که ۲/۲ میلیون‌نفرآنان در فعالیتهای کشاورزی، شکار و جنگلداری مشغول بوده‌اند که ۸/۲میلیون نفرآن ساکن مناطق روستائی بوده‌اند؛ به عبارت دیگر روستائیان کشورحدود ۴۵ درصد جمعیت را تشکیل می‌داده‌اند. ۲۹ درصد جمعیت شاغل ۱۰ ساله وبالاتر در فعالیتهائی که با کشاورزی (بمفهوم مرتبط بوده اشتغال داشته که ۹۰ درصد آن ساکن مناطق روستائی بوده‌اند. ازطرف دیگر، به طور متوسط طی سالهای ۶۰-۱۳۶۴ حدود ۲۱ درصد تولید ناخالص داخلی ازفعالیتهای کشاورزی ناشی شده است. طی

(۱): در سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ آبادی (نقطه روستائی) بمجموعه‌ای باچند مکان واراضی به‌هم پیوسته (اعم ازکشاورزی وغیرکشاورزی) گفته شده است که خارج ازمحدوده شهرها واقع شده واراضی محدوده شهری یا عرفی مستقل بوده‌است.

همین سالها، به طورمتوسط هزینه‌های خوراکی ودخانی سالانه یک خانوار درکل کشور ۴۷ درصد کـــــــل هزینه‌های خانوار را تشکیل می‌داده‌است . باتوجه به اینکه منبع اولیه تامین اقلام خوراکی ودخانـــــــی، بخش کشاورزی بوده وفعالیتهای کشاورزی عمدتا" درمناطق روستائی انجام می‌شود، نقش و اهمیـــــت فعالیتهای کشاورزی دراقتصاد کشور ودرجامعه روستائی ایران مشخص می‌شود . حاصل فعالیتهـــــای کشاورزی ازیکسو ماده اولیه غذائی ساکنین کشوراعم ازروستائی وشهری راتشکیل داده وازسوی‌دیگرمنبع تامین مواداولیـــه منابع غذائی به طورمستقیم وبسیاری ازفعالیتهای دیگر به طورغیرمستقیم اســـــــت . بدینجهت اطلاع ازکم و کیف این فعالیت نیازاولیه برای رونق هرچه بیشترآنرا محسوس می‌نمایـــد . دراین میان نه تنها پروسه انجام این فعالیتها اهمیت خاصی دارد، بلکه تفصیل آن نیازمند آگاهـــی بیشترازاین ارتباط است . بیان عددی یابه‌عبارتی آماری آن رامرابرای برنامه ریز هموار می‌سازد .تصویر آماری این فعالیت: وضع ومیزان تولید، سطح زیرکشت ، تفصیل هزینه‌های مربوط، ارزش تولید را گویـــا می‌سازد . ارزش تولید عامل مهمی دراین میان است زیرا نه تنها متاثرازمیزان تولید می‌باشدبلکـــه لزوم آگاهی ازقیمت‌های آنرانیز مشخص می‌سازد . دو زمینه‌هائی نظیر میزان تولید،سطح زیرکشـــت ، چگونگی تولیدو ... مطالبیاتی صورت گرفته و طی‌آمارگیری‌های متعدد، اطلاعاتی در زمینه‌هـــــای فوق‌الذکر جمع‌آوری شده است ولی به زمینه قیمت وهزینه محصولات وخدمات کشاورزی توجه زیـــــــادی نشده است . هرمحصول کشاورزی هنگام گذرازمرحله تولید تامصرف نهائی به علت تاثیرعوامل متعدد ، قیمتهای متفاوتی پیدا می‌کند که اصطلاحا" قیمت سرخرمن، قیمت عمده‌فروشی، ... وقیمت‌خرده فروشی نامیده می‌شود، به منظور دستیابی به‌قیمتی که صرفا" ناشی ازفعالیتهای کشاورزی‌است بایـــــــد تاثیرعواملی نظیرهزینه‌های حمل‌ونقل ، انبارداری و ... حذف شود . درچند ساله اخیر ، طـــــــرح آمارگیری از قیمت محصولات وهزینه خدمات کشاورزی درمرکزآماراایران که سازمان مرکزی موظف‌ومسئول درامرجمع‌آوری اطلاعات آماری است ، مجرد توجه قرارگرفته است . صرفی این طرح ، موضوع مقاله حاضر، درچند قسمت صورت می‌گیرد :

۱- محدوده فعالیتهای کشاورزی

درتعاریف وطبقه‌بندیهای بین‌المللی برای‌فعالیتهای‌کشاورزی‌تعریفی‌مشخص‌وقاطع‌ارائه‌نشده‌است .

ومرز بین این فعالیت و صنعت، خدمات ونیز سایرفعالیتهای اقتصادی همواره بحث‌انگیز بوده‌است . زیرا کاشت، داشت وبرداشت مثلا " شلتوك " (برنج) فعالیت کشاورزی تلقی می‌شود ولی باجداشدن ازمزرعه و ورود به کارخانه برنج‌کوبی، فعالیت صنعتی به حساب می‌آید . این موضوع در صور مختلف ودرارتباط با خدمات کشاورزی، استعمال آب و... نیز صادق است . درتعیین مرز برای شناخت کشاورزی ومحدوده فعالیتهای آن، با طراحی، کاشت ، داشت وبرداشت قلمرو فعالیت کشاورزی تلقی شده ومبنای تهیه طرح مورد بحث بوده است .

۳- دامنه شمول

باتوجه‌به‌نظرگرفتن‌اینکه‌فعالیتهای کشاورزی عمدتا " درمناطق روستائی انجام می‌شود دامنه شمول این آمارگیری محدوده روستائی در نظر گرفته شده وازاخذ وجمع‌آوری اطلاعات مربوط در محدوده شهرها (مناطق شهری) خودداری شده است و اینموضوع درعنوان طرح " قیمت فروش محصولات وهزینه خدمات کشاورزی در مناطق روستائی " نیز مشخص شده است .

۴- هدفهای طرح

۱- محاسبه متوسط فصلی وسالانه قیمت فروش محصولات وهزینه خدمات کشاورزی در مناطق روستائی کشور به تفکیك استان .

۲- تهیه شاخص سالانه قیمت فروش محصولات وهزینه خدمات کشاورزی در مناطق روستائی کشور به تفکیك استان .

۵- نحوه تعیین اقلام مورد سوال

براساس بررسیهای انجام شده درمورد نتایج حاصل ازآمارگیری‌های کشاورزی که توسط مرکزآمارایران درسالهای مختلف انجام شده‌است، نظرات تعدادی ازسازمانهای مصرف کننده آمارهای مذکور ونیز تجارب کارشناسی، اقلام مورد نظر تعیین وباتوجه به ویژگیهای آنها مورد سوال قرارگرفته‌اند . با درنظرگرفتن تنوع محصولات وخدمات کشاورزی، تعداد اقلام مورد سوال دراین طرح از ۱۷ قلم در سال ۱۳۵۴ (فقط محصولات کشاورزی) تا ۷۶ قلم درسال ۱۳۷۱ (محصولات، خدمات و کارگرکشاورزی) متغیر بوده است .

و- تعاریف مورد استناد

قیمت فروش محصول (قیمت سرخرمن) دراین طرح عبارتست ازمتوسط قیمت فروش یک کیلوگرم هریک ازمحصولات کشاورزی در زمان معمول حداکثر برداشت آن که توسط کشاورزان در آبادی نمونه (واحد آماری) تولید شده باشد . باتوجه به تعریف فوق مشخص است که :

۱- محصول اولا " تولید شده (به ثمر رسیده) وثانیا " به فروش رسیده است .

۲- کشاورزان درآبادی نمونه مبادرت به تولید محصول کرده اندوخارج ازآبادی نمونه در شمول تعاریف این طرح قرار نمی گیرد .

۳- جمع آوری قیمتها مستلزم این فرض است که زمان معمول حداکثر برداشت وزمان فروش محصولات درآبادی نمونه برهم منطبق یا به هم نزدیکند .

برای مشخص ترکردن حیطه آمارگیری، نکات زیرموردتوجه قرارگرفته است :

- محصولات کشاورزی به صورتی که حاصل فعالیت کشاورزی است ، قیمت گیری شده اند مگر درموارد یکه فروش محصول عرفا " پس از تغییر شکل صورت گیرد .

- ازجمع آوری قیمت محصولات پیش فروش شده، نوبرانه ومحصولاتی که بعدازبرداشت مدتی انبارشده وسپس به فروش رسیده است خودداری شده زیرا پیش فروش محصول ، قیمت واقعی آنرا مشخص نمی کند انبارکردن محصول به منظور عرضه آن بعداززمان حداکثر برداشت وباخارج ازفصل محصول آن ، برای فروش به قیمت گرانتراست . نگهداری محصول برای عمل آوردن یاخشک شدن که ممکنست به طورمعمول درمورد برخی ازمحصولات کشاورزی انجام شود، انبارکردن به حساب نمی آید . (نگهداری یونجه، گردو ... مثالهائی ازاین مقوله اند) .

- درمواردی که محصول موردنظر برداشت شده ولی به فروش نرسیده باشد، کشاورزان مشغول برداشت محصول باشند، هنوز زمان برداشت محصول فرانرسیده باشد ویا کاشت وبرداشت محصول بعدازمراجعه به واحدآماری (آبادی نمونه) باشد، قیمت گیری به عمل نیامده است .

- قیمتهای موردنظر چه در خصوص محصولات وچه درمورد خدمات کشاورزی ، قیمتهای علنی و قطعی (نقدی یاجنسی) بوده است، بدینجهت مظنه سوال نشده است .

- اگرمحصول آبادی نمونه درشهروبیاآبادی دیگری به فروش رسیده باشد، هزینه‌های حمل‌ونقل ازقیمت فروش کسرشده است .

- چنانچه محصول کشاورزی به فروش نرسیده باشد (اعم از مصرف خانوار یا ...) خود مصرفی تلقی شده و فاقد قیمت فروش به حساب آمده است .

- قیمت پایاپای (مبادله محصول بامحصولی دیگر ویا محصول باخدمت) مورد نظر نبوده است . دریافت سهمی‌ازمحصول به عنوان تمام یاقسمتی ازدستمزد انجام خدمت،باضرایطی به ریال برآوردشده است .

- برای پاره‌ای از محصولات که دولت قیمت تضمینی خرید تعیین ونرموواردی جایزه‌ای هم درنظر گرفته بوده است، قیمت نقدی خرید توسط دولت ملاک نظر قرار گرفته است .

- از محصولات دارای تنوع تولید، قیمت فروش نوعی که درآبادی نمونه بیشترین تولید راداشته، سوال شده است .

- محصولاتی که درواحدآماری نمونه، برداشت آنها طی سال درچندنوبت انجام می‌شود، می‌توانند درهر نوبت دارای قیمت فروش باشند .

بایدتوجه داشت که :

- گروهی ازکشاورزان محصول خودرا پیش فروش کرده وگروهی آنرابس ازبرداشت " تاروشن شدن وضع و قیمت بازار " انبارمی‌کنند . بدینجهت دراین‌گونه موارد دستیابی به قیمت واقعی فروش محصولات کشاورزی درزمان برداشت (قیمت سرخرمن) میسر نیست . دربرخی‌ازآبادیهای نمونه نیز تولید محصول فقط برای مصرف پرده وبه‌فروش نمی‌رسدکه دراینصورت محصول مورد بحث فاقدقیمت‌فروش‌است .

- به دلیل شرایط متفاوت منطقه‌ای ، محصولات وخدمات کشاورزی ازتنوع خاصی برخوردار است که این تنوع نه تنها زمان ونحوه تولید محصولات وانجام خدمات را متفاوت می‌سازد، بلکه قیمتها ودرآمدهای مختلفی رانیز باعث‌می‌شود که درنتیجه ارائه متوسط قیمت یاهزینه راندرموواردی با اشکال مواجه می‌سازد .

ز- ویژگیهای اقلام پرسشنامه

- دراین طرح قیمت فروش محصولات برحسب کیلوگرم ریال ، هزینه خدمات ماشینی برحسب هکتار ریال، دستمزد کارگر کشاورزی برحسب ریال روز کار سوال شده است . برای محصولات کشاورزی، قیمت ومقدار فروش محصول باتوجه به زمان حداکثر برداشت ، برای فراورده‌های دامی، قیمت و مقدار آخرین فروش طی فصل، برای خدمات، هزینه بارآخر هرخدمت درهرفصل درغیراینصورت بارآخر فصل قبل و مساحت مورد عمل وبرای کارگر کشاورزی، دستمزد انجام کار ونقر روزکار طی فصل مورد مراجعه سوال شده است .

- دراین طرح باتوجه به تعاریف ذکرشده ، قیمت فروش شیر ، قبل ازآنکه در آن تغییری داده‌شود اخذ شده و دام زنده وسالم (که به منظور استفاده ازگوشت آن در آبادی نمونه تولیدوپرورار شده و برای کشتار به فروش رسیده) سوال شده است .

- محصولاتی نظیرنخود، لوبیاسفید، ... معمولا " پس ازخشک شدن ، دانه گیری و بوجاری به فروش می‌رسند و اینموضوع ویژگی این گونه محصولات به شمار می‌رود .

- محصولاتی نظیر پسته، بادام و گردو بدون پوست سبز روئی آنها (غلاف) قیمت گیری شده‌اند .

- پشم نشسته (گوسفند) که پس از چیدن ازدام شسته نشده باشد، قیمت گیری شده است .

باآنکه یونجه و شبدر تر حاصل اولیه فعالیت کشاورزی است، معهذا قیمت فروش این محصولات به صورت خشک آنها، مورد نظر این طرح بوده است .

- برای خدمات ماشینی کشاورزی، هزینه معمول یکبار شخم، دیسک ... در زمینهای آبی* دیم جداگانه سوال شده است واین هزینه‌ها برحسب آنکه سطح مورد عمل مبنای پرداخت باشد یاهزینه براساس زمان انجام کار پرداخت شده باشد، بسته به مورد، سؤال شده است ، که درهرصورت این هزینه شامل کرایه دستگاه و دستمزد راننده آن بوده است . پرداخت تمام یاقسمتی ازاین هزینه به صورت سهمی ازمحصول باملحوظ نظرگرفتن تعداد روزهای انجام خدمت، وزن محصول عاید شده وقیمت خرده‌فروشی آن در روز دریافت محصول درآبادی نمونه به صورت ریال برآورد شده است . چنانچه خوراک مصرفی کار

* زمینهای آبی به زمینهایی اطلاق می‌شودکه زراعت انجام‌شده درآنها آبیاری می‌شودوعرف محل آنرا به‌نام زراعت آبی می‌شناسندوبه‌همین‌ترتیب ، زمینهای دیم به زمینهائی اطلاق می‌شودکه زراعت انجام‌شده درآنها آبیاری نمی‌شودوعرف محل آنرا به نام زراعت دیم می‌شناسند .

نیز در حین انجام خدمت تأمین شده باشد ، برآورد ریالی خوراک به نرخ روز، محاسبه ومنظور شـــــده است ، اینموضوع درمورد کارگرکشاورزی نیز صادق بوده است .

ج-کاربرد نتایج این طرح

خلاصه‌ای ازکاربردهای اطلاعات به دست آمده از اجرای این طرح به قرار است :

- تهیه شاخص‌قیمت فروش (قیمت سرخرمن) محصولات وهزینه خدمات کشاورزی

برآورد ارزش تولید محصولات کشاورزی

- برآورد ارزش افزوده بخش کشاورزی

- برآورد متوسط درآمد خانوارهای کشاورز

- بررسی تفاوت بین قیمت فروش (سرخرمن) با قیمت عمده فروشی وقیمت خرده‌فروشی (قیمـــــت مصرف کننده) محصولات کشاورزی

- تهیه شاخص رابطه مبادله بین بخش کشاورزی وسایر بخشها

- بررسی اثرات تغییر قیمت محصولات وهزینه خدمات کشاورزی بردرآمد وهزینه خانوارها

- بررسی اثرات تغییر قیمت محصولات وهزینه خدمات کشاورزی برتولیدات سایر بخشها

- برآورد ارزش تسعیتی ازمحصولاتیکه به عنوان ماده اولیه (مجددا " درهمین بخش) به مصرف‌رسیده است .

ط - تسلسل اجرای آمارگیری

این طرح ازسال ۱۳۵۳ شروع وهفت ساله اجراشده است ، به استثناء موارد زیر :

۱- سال ۱۳۵۷ فقط شامل اطلاعات نیمه اول سال است و نیمه دوم به علت وقوع انقلاب اسلامـــی این طرح اجرانشده‌است .

۲- نیمه‌اول سال۱۳۶۸ به علت ادامه تحولات نیمه دوم ۵۷ ونیز نیمه دوم ۵۹ به علت شروع جنـگ تحمیلی این طرح اجرانشدواز‌اینرو جمع‌آوری اطلاعات موردنظر فقط به نیمه دوم سال ۸ونیمـــــه اول سال ۵۹ مربوط می‌شود .

۳ـ درسال ۱۳۶۰ به علت ادامه جنگ تحمیلی ونیزمشکلات اجرائی حادث ، این طرح اجرانشد .

۴ـ درسال ۱۳۷۰ مقدمات مرحله جدید اجرای طرح شروع شد . دراین سال با عدم اجرای طرح ، عملا " مقدمات تهیه طرح مناسب فراهم شد که دوسال ۷۱ قرار است آزمایش شود .

ی ـ نحوه اجرای آمارگیری

ـ این طرح ازاولین سال اجرا یعنی ۱۳۵۳ تاحال حاضر به جهت استفاده از موقعیتی که اجرای طرح قیمت خرده فروشی کالاها وخدمات درمناطق روستائی به دست می‌داده است ، همراه با طرح مذکور اجراشده است . به این معنا که مامورآمارگیر پس از طی مراحل آموزش ، درهمان آبادیهای مورد مراجعه برای اخذ قیمت های خرده فروشی ، اقدام به تکمیل پرسشنامه قیمت فروش محصولات وهزینه خدمات کشاورزی نموده است .

ـ آمارگیری ازطریق مصاحبه با مطلعین محلی (دومطلع برای هرسوال) وتکمیل پرسشنامه در آبادی صورت گرفته‌است .

ـ آمارگیری این طرح چهاربار درسال (هرفصل یکبار) انجام شده است .

ـ تکمیل پرسشنامه توسط معاونت آمارو اطلاعات سازمان برنامه‌وبودجه استانها بوده وفرمراکزاستانها به فرمهای استخراج دستی منتقل شده است . تعداد پرسنل حدود ۱۵۰ نفر آمارگیر ، ۲۲ کارشناس مسئول که آمارگیران ، مسئول بازبینی واستخراج زیر نظر کارشناس مسئول طرح ودرنهایت زیر نظر معاون آماروا طلاعات فعالیت داشته‌اند .

ـ فرمهای استخراج پس ازتکمیل وبازبینی به مرکزآماروایران ارسال شده است . پرسشنامه‌های طرح نیز پس ازآخرین فصل اجرا به مرکز ارسال شده است .

ـ بازبینی وکنترل (درصورت لزوم بامراجعه به پرسشنامه‌های وصولی ازاستانها) ونیز انجام محاسبات دستی برروی فرمهای استخراج وکاربرگهای واسط وسرانجام کنترل نهائی در مرکزآماروایران صورت گرفته‌است .

ـ تدوین وانتشار نتایج این طرح بین ۸ تا ۱۰ ماه بعدازوصول نتایج آخرین فصل اجرا بوده است .

ک- فرمولهای محاسبه متوسط ها

اطلاعات جمع‌آوری شده ازآبادیهای نمونه درهرفصل پس از بازبینی کارشناسان مربوط دراستانها و اطمینان از صحت آنها استخراج پس از محاسبات اولیه به مرکزارسال شده است . در مرکز نیز پس از بررسی مجدد ورفع اشکالات احتمالی، متوسط ها محاسبه شده‌اند . فرمولهای برآورد در برگه ضمیمه درج شده است . باتوجه به اینکه محاسبه شاخص قیمت مستلزم استاندارد کردن خصوصیات و مشخصات اقلام مورد قیمت گیری است، به دلیل پاره‌ای ازمشکلات اجرائی، اقلام استاندارد نشده و بدینجهت بند دوم اهداف طرح عملی نشده است .

ل- تغییرات اقلام مورد سوال

تغییرات وتحولات این طرح رامی‌توان طی ۲ دوره : ۱۳۵۲ لغایت ۱۳۵۷ ونیمه دوم ۱۳۵۸ لغایت نیمه اول ۱۳۵۹ ، ۱۳۶۱ لغایت ۱۳۶۶ ، ۱۳۶۷ لغایت ۱۳۶۹ و ۱۳۷۱ مورد بررسی قرارداد و ضمن آن سه جامعه آماری، واحدآماری وتعداد آبادیهای مورد مراجعه اشاره کرد .

ا- دوره ۱۳۵۲ لغایت نیمه اول سال ۱۳۵۶

ا-ا- دوره ۱۳۵۲ لغایت نیمه اول ۱۳۵۷

جامعه آماری : کلیه آبادیهای بیش از ۵۰ خانوار درمناطق روستائی کشور (نقاط دارای جمعیت کمتراز دهزار نفر براساس نتایج سرشماری ۱۳۴۵)

چارچوب آماری : فهرست اسامی ومشخصات آبادیهای دارای بیش از ۵۰ خانوار براساس نتایج سرشماری ۱۳۴۵

واحد آماری : آبادی (۱۰ آبادی درهراستان، جمعا " ۴۶۰ آبادی- درسال ۵۲ با احتساب تهران دراستان مرکزی، جمعا " ۲۳ استان وجود داشته است) واحد آماری مرحله اول ومطلع محلی واحدآماری مرحله دوم بوده است .

تعداد اقلام : طی این سالها تعداد اقلام حداقل ۱۷ وحداکثر ۲۲ بوده است .

۲-۱- دوره نیمه دوم ۱۳۵۸ لغایت نیمه اول ۱۳۵۹

جامعه آماری : همانند دوره ۱۳۵۲ لغایت نیمه اول ۱۳۵۷

چارچوب آماری : همانند دوره ۱۳۵۲ لغایت نیمه اول ۱۳۵۷

واحدآماری : باتقسیم مناطق روستائی کل کشور به هفت حوزه روستائی وگروه بندی آبادیها ازنظرتعدادخانوار در سه طبقه ،تعداد۷۵۶آبادی به عنوان واحدآماری مرحله اول ودر داخل هرآبادی ، ۲نفر مطلع محلی به عنوان واحدآماری مرحله دوم انتخاب شدند .

تعدادواقلام : دراین دوره از ۲۰ قلم محصول کشاورزی قیمت گیری شد . طی دوره ۱۳۵۲ لغایت نیمه اول سال ۱۳۵۹ ، هرسال نتایج این آمارگیری همراه بانتایج آمارگیری ازقیمت خرده فروشی کالاها درنشریه ای واحد منتشرشده بود .

۲- دوره ۱۳۶۱ لغایت ۱۳۶۶

جامعه آماری : کلیه آبادیهای مسکونی واقع درمناطق روستائی کشور

چارچوب آماری : فهرست آبادیهای دائما" مسکونی مناطق روستائی کشور که حداقل دارای یک بقالی ویافروشگاه تعاونی روستائی بوده اند.

واحدآماری : درهراستان باتوجه به تعدادشهرستانها ،تعدادآبادی تعیین شد (حداقل ۲۲آبادی) که هرآبادی واحدآماری مرحله اول راتشکیل می دادودر داخل هرآبادی ، دونفرمطلعین محلی واحدهای آماری مرحله دوم بودند . طی این دوره تعدادآبادیهای نمونه ۳۹۰ آبادی درکل کشوربوده است که براساس جمعیت،این آبادیها در ۲ گروه طبقه بندی شده بودند .

تعدادواقلام : طی این دوره ، باانجام بررسیهای مختلف واستفاده ازنظرات کارشناسی صاحبنظران ، تعدادواقلام مال به سال افزایش داشته به طوریکه از ۲۹ قلم مورد سئوال به ۷۲ قلم رسید . ازاین دوره ببعد مدارک اجرای این طرح به طورمستقل تدوین و منتشرشد ونتایج آمارگیری هرسال نیز تحت نشریه ای جداگانه انتشاریافت .

۲- دوره ۱۳۶۷ لغایت ۱۳۶۹

جامعه آماری : کلیه آبادیهائی که براساس نتایج سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ مسکونی دائمی بوده وبه‌عنوان مراکز خرید روستائیان (آبادی های مسکونی دائمی که دارای چند کارگاه خرده فروشی بوده به طوریکه مایحتاج عمومی اهالی این آبادیها وبآبادیهای مجاور توسط آنها تامین می‌شده است) شناسایی شدند .

چارچوب آماری : فهرست اسامی ومشخصات کلیه آبادیهائی که براساس نتایج سرشماری عمومی نفوس ومسکن سال ۱۳۶۵ مسکونی دائمی بوده و در عین حال به عنوان مراکز خرید روستائیان شناخته شده‌اند .

واحدآماری : هریك ازآبادیهای واقع در مناطق روستائی که درسرشماری سال ۱۳۶۵ مسکونی دائمی بوده وبه عنوان مرکزخرید روستائیان شناخته شده بود، واحد آماری مرحله اول ودرداخل هریك ازآنها دونفر مطلع محلی واحدهای آماری مرحله دوم راتشکیل می‌دادند . برای تعیین تعداد آبادی ، از هردهستان حداقل یك آبادی به شرط دارابودن خصوصیت مرکز خرید روستائی انتخاب شد وبدین ترتیب تعداد آبادیها برابر ۲۰۰۴ آبادی بوده‌است .

درطی‌این دوره نیز باانجام بررسیهای کارشناسی ونیازبخشهای تحقیقاتی واجرائی ، تعداد اقلام مورد سؤال در زمینه محصولات کشاورزی ، خدمات کشاورزی‌ونیروی انسانی شاغل تعدیل شد وبه ۶۰ قلم رسید . برای دستیابی به نتایج بهتروزمینه ساز محاسبه شاخص قیمت ، ضرائب مورد نیاز مانند مقدار فروش، سطح زیرکشت وتعداد نفرات نیروی انسانی وابسته‌بموردبرای سئوالات مربوط درپرسشنامه‌ها گنجانده شده که نتایج آمارگیری بااستفاده ازاین ضرائب موزون شد .

۳- دوره ۱۳۷۱ به بعد

جامعه آماری : درمرحله اول (سال ۱۳۷۱)همانند دوره ۱۳۶۷ لغایت ۱۳۶۹ و درمراحل بعد کلیه آبادیهای دارای فعالیت تولید وفروش محصولات کشاورزی بااستفاده از نتایج سرشماری کشاورزی سال ۱۳۶۷ :

چارچوب آماری : درمرحله اول همانند دوره ۱۳۶۷ لغایت ۱۳۶۹ و در مراحل بعد لیست وشخصات کلیه آبادیهای دارای فعالیت تولید وفروش محصولات کشاورزی که با استفاده از نتایج سرشماری کشاورزی سال ۱۳۶۷ به دست می‌آید .

واحد آماری : درهردوره واحد آماری مرحله اول آبادی ودرمرحله دوم مطلع محلی خواهد بـــــود . دردوره ۱۳۷۱ ، تعداد آبادیها با استفاده از نتایج این آمارگیری در سالهای ۱۳۶۷ و ۱۳۶۸ تعدیل شد . به‌گونه‌ای که آبادیهای فاقد اطلاع ویاکم‌اطلاع دراستانهای مختلـــف حذف شد و درنتیجه تعداد آبادیها به ۱۱۷۶ آبادی تقلیل یافت . بررسی نتایـــــــج سرشماری کشاورزی تعداد آبادیهای مناسب طرح را مشخص خواهد کرد .

برای دستیابی به نتایج واقعی ، پاره‌ای تغییرات اساسی لازم می‌نماید زیرا :

آبادیهایی باید نمونه این طرح باشند که عوامل لازم برای تولید محصول وبه تبع آن انجام خدمـات کشاورزی در آنها فراهم باشد . درصورتیکه آبادیهای مراکز خرید طرح قیمت خرده فروشی الزاما" واجد این صفات نیستند . بررسی نتایج حاصله ازاجرای آن درسالهای گذشته واستفاده از نتایج سرشماری کشـــاورزی لازمه انجام این تغییر اساسی است . ثانیا" ماهیت محصولات وخدمات کشاورزی چنان است که داشـــت ، برداشت وفروش محصول وبانجام خدمت درهرمنطقه جغرافیائی طی زمانی خاص صورت می‌گیرد که لزومـــا" چهاربار مراجعه درسال نخواهد بود .

تجدیدنظر در زمینه‌های مختلف اجرای این طرح دردومرحله به شرح زیر پیش بینی شده است .

مرحله اول بررسی جوانب مختلف طرح درسالهای ۶۹ـ۱۳۶۳ برای اجرای مستقل آن درسالهای آینده

- مطالعه وبررسی آبادیهای نمونه مراکز خرید روستائیان برای تعیین آبادیهای نمونه مناسب بـرای این طرح

- تهیه وتنظیم پرسشنامه مناسب به منظور تسهیل آمارگیری واستخراج ماشینی

- حذف فرم استخراج

- مطالعه وبررسی برای تعیین زمان مناسب مراجعه به آبادیهای نمونه

مرحله ۲ـ مطالعه وبررسی برای تعیین چارچوب مناسب

مطالعه وبررسی برای تعیین روشهای مناسب محاسبه شاخص

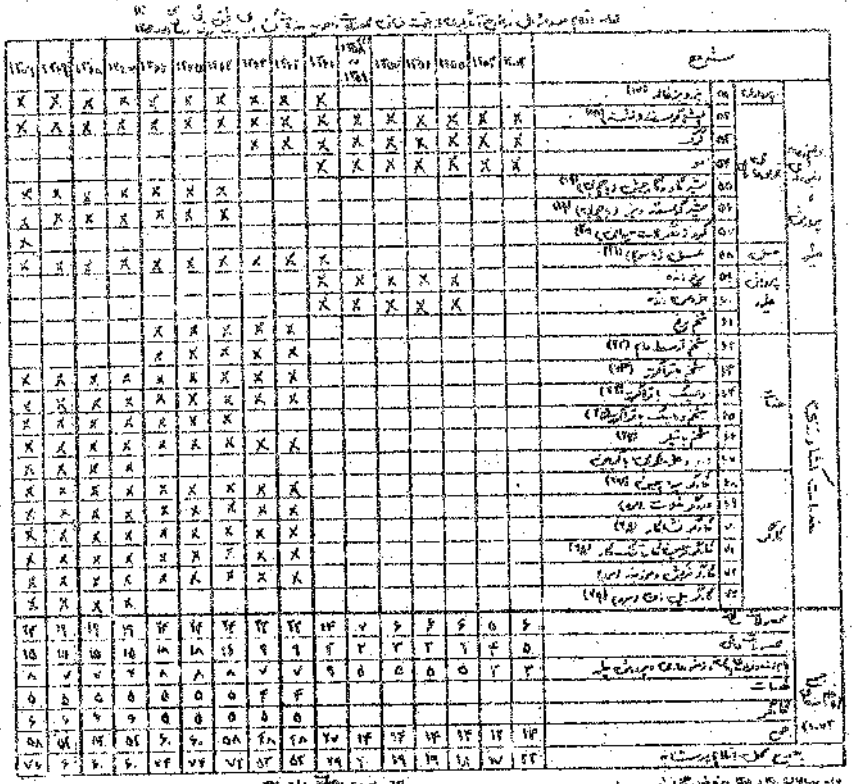
درمقطع کنونی ،مرحله ۱ انجام شده است وبااجرای این طرح درسه فصل باقیمانده سال ۱۳۲۱،مرحله ۲ جامه عمل خواهد پوشید ، دراین راستا به منظور حذف موارد غیرضروری آمارگیر پس از مصاحبه با چندین مطلع وبااستفاده از پاسخهای آنها فقط یک اطلاع را برای هرصفت درج می‌نماید ، همچنین ازآنجائیکه داشتن مشخصات محصول یاخدمت ازنیازهای اساسی واولیه برای محاسبه شاخص قیمت فروش محصولات وهزینه خدمات کشاورزی است ، برای هرمحصول بیش از یک سطر در پرسشنامه در نظر گرفته شده تاآمارگیر انواع یک محصول به فروش رفته یاخدمت انجام شده در سطح آبادی را باذکر مشخصات هرکدام جداگانه منظورنماید .

منابع : ۱ـ مدارک اجرای طرح آمارگیری ازقیمت فروش محصولات وهزینه خدمات کشاورزی درمناطق روستائی ، سالهای ۷۱ـ۱۳۵۳ ، مرکزآمارایران

۲ـ مدارک اجرای طرح آمارگیری ازقیمت خرده فروشی کالاهاوخدمات مورد مصرف خانوارهای روستائی ، سالهای ۷۱ـ۱۳۵۳ ، مرکزآمارایران

[illegible] (۱)

گروه	ردیف	شرح	۱۳۵۳	۱۳۵۴	۱۳۵۵	۱۳۵۶	۱۳۵۷	۱۳۵۸-۱۳۶۰ (۲)	۱۳۶۱	۱۳۶۲	۱۳۶۳	۱۳۶۴	۱۳۶۵	۱۳۶۶	۱۳۶۷	۱۳۶۸	۱۳۶۹	۱۳۷۰
محصولات سالانه — غلات	۱	گندم (۳)						X	X	X	X	X	X	X	X	X	X	X
	۲	جو	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	۳	شلتوک (۴)										X	X	X	X	X	X	X
	۴	ذرت (۵)							X	X	X	X	X	X				X
	۵	ارزن																X
	۶	سایر (۶)																X
حبوبات	۷	نخود							X	X	X	X	X	X	X	X	X	X
	۸	لوبیا سفید							X	X	X	X	X	X	X	X	X	X
	۹	لوبیا چیتی							X	X	X	X	X	X	X	X	X	X
	۱۰	لوبیا قرمز								X	X	X	X	X	X	X	X	X
	۱۱	عدس								X	X	X	X	X	X	X	X	X
دانه‌های روغنی	۱۲	آفتابگردان (روغنی) (۷)	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X
	۱۳	سویا [illegible] (۸)								X	X	X	X	X				X
	۱۴	گلرنگ								X	X	X	X	X	X	X	X	X
محصولات لیفی	۱۵	پنبه (وش) (۹)	X		X	X	X	X		X	X	X	X	X	X	X	X	X
محصولات [illegible]	۱۶	[illegible]	X	X	X	X	X	X										
محصولات جالیزی	۱۷	خربزه							X	X	X	X	X	X	X	X	X	X
	۱۸	هندوانه							X	X	X	X	X	X	X	X	X	X
	۱۹	طالبی								X	X	X	X	X				
	۲۰	خیار							X	X	X	X	X	X	X	X	X	X
	۲۱	[illegible]								X	X	X	X	X				
سبزیجات	۲۲	سیب زمینی (۱۰)							X	X	X	X	X	X	X	X	X	X
	۲۳	گوجه فرنگی							X	X	X	X	X	X	X	X	X	X
	۲۴	بادمجان								X	X	X	X	X	X	X	X	X
	۲۵	پیاز							X	X	X	X	X	X	X	X	X	X
	۲۶	سایر (۱۱)																X
نباتات علوفه‌ای	۲۷	[illegible] (۱۲)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	۲۸	[illegible] (۱۳)	X	X	X	X	X	X	X	X	X	X	X	X				
محصولات دائمی — میوه‌های هسته‌دار	۲۹	[illegible]													X	X	X	X
	۳۰	زردآلو										X	X	X	X	X	X	X
	۳۱	گیلاس											X	X	X	X	X	X
	۳۲	آلبالو											X	X				
	۳۳	آلو													X	X	X	X
میوه‌های دانه‌دار	۳۴	سیب							X	X	X	X	X	X	X	X	X	X
	۳۵	گلابی	X							X	X	X	X	X	X	X	X	X
	۳۶	به										X	X	X				
مرکبات	۳۷	پرتقال							X	X	X	X	X	X	X	X	X	X
	۳۸	نارنگی								X	X	X	X	X	X	X	X	X
	۳۹	لیموشیرین										X	X	X				
	۴۰	لیموترش								X	X	X	X	X	X	X	X	X
[illegible]	۴۱	انجیر										X	X	X				
	۴۲	انگور							X	X	X	X	X	X	X	X	X	
	۴۳	خرما	X	X						X	X	X	X	X	X	X	X	X
	۴۴	پسته [illegible] (۱۴)	X	X	X	X	X	X				X	X	X	X	X	X	X
	۴۵	بادام [illegible]	X	X	X	X	X			X	X	X	X	X	X	X	X	X
	۴۶	گردو [illegible]										X	X	X	X	X	X	X
	۴۷	[illegible]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	۴۸	فندق										X	X	X				
[illegible]	۴۹	[illegible] (۱۵)							X	X	X	X	X	X	X	X	X	X
	۵۰	[illegible] (۱۶)							X	X	X	X	X	X	X	X	X	X



شهرستان	۱۳۵۳ لغایت ۱۳۵۷	۱۳۵۸ پیش‌بینی لغایت ۱۳۵۹ نسبت به دوره ۱۳۵۳-۱۳۵۷	۱۳۶۱ لغایت ۱۳۶۲	۱۳۶۷ نسبت ۱۳۶۹	۱۳۷۱
کل کشور	۴۶۰	۷۵۶	۱۰۴۹	۲۰۰۴	۱۱۷۶
تهران	۲۰ *	حوزه مرکزی *	۲۸	۷۱	۴۴
مرکزی			۲۶	۵۲	۳۸
گیلان	۲۰	حوزه غرب	۶۰	۷۰	۳۶
مازندران	۲۰	حوزه غرب	۶۰	۱۷۰	۹۸
آذربایجان شرقی	۲۰	حوزه شمال غربی	۶۶	۱۲۷	۷۷
آذربایجان غربی	۲۰	حوزه شمال غربی	۵۲	۷۷	۴۵
کردستان	۲۰	حوزه غربی	۳۱	۷۰	۳۰
همدان	۲۰	حوزه جنوب غربی	۶۴	۱۲۱	۴۱
کرمانشاه	۲۰	حوزه جنوب شرقی	۸۴	۱۴۵	۱۱۴
لرستان	۲۰	حوزه جنوب شرقی	۴۸	۱۴۲	۶۲
خوزستان	۲۰	حوزه شمال شرقی	۹۶	۲۷۸	۱۳۵
اصفهان	۲۰	حوزه مرکزی	۶۶	۹۹	۶۴
سیستان و بلوچستان	۲۰	حوزه جنوب شرقی	۳۶	۷۸	۳۸
کرمان	۲۰	حوزه غرب	۲۵	۶۰	۲۷
یزد	۲۰	حوزه غرب	۲۸	۵۴	۴۲
چهارمحال و بختیاری	۲۰	حوزه جنوب غرب	۲۶	۲۴	۲۶
ایلام	۲۰	حوزه غربی	۲۷	۶۶	۲۹
کهگیلویه و بویراحمد	۲۰	حوزه غربی	۲۹	۵۱	۲۲
بوشهر	۲۰	حوزه جنوب غربی	۲۷	۲۷	۲۰
زنجان	۲۰	حوزه جنوب غرب	۲۶	۴۲	۲۸
سمنان	۲۰	حوزه شمال غربی	۲۹	۵۶	۴۹
یزد	۲۰	حوزه شمال شرقی	۲۸	۳۳	۲۴
هرمزگان	۲۰	حوزه جنوب شرقی	۲۹	۳۳	۲۸
		حوزه جنوب شرقی	۲۸	۳۷	۲۹

* ـ استان مرکزی شامل تهران بزرگ است .

متوسط قیمت محصول با خدمت درسطح آبادی

$$\bar{P}_{rsjh} = \frac{\sum_{i=1}^{l} P_{rsjhi}}{l}$$

P_{rsjhi} عبارت است از قیمت محصول باخدمت r ام اخذ شده از منبع i ام در آبادی h ام دراستان j ام درفصل s ام و l تعداد منابع (حداکثر ۲ منبع)

متوسط قیمت فصلی محصول با خدمت درسطح استان

$$\bar{P}_{rsj} = \frac{\sum_{h=1}^{m} \bar{P}_{rsjh} \bar{K}_{rsjh}}{\sum_{h=1}^{m} \bar{K}_{rsjh}}$$

$\bar{P}_{rsjh}$ عبارت است از متوسط قیمت محصول با خدمت r ام در آبادی h ام دراستان j ام درفصل s ام و m تعداد آبادیهای دارای اطلاع درهراستان است . در رابطه فوق $\bar{K}_{rsjh} = \frac{\sum_{i=1}^{l} K_{rsjhi}}{l}$ که K_{rsjhi} عبارت است از ضریب اهمیت اخذ شده (۱) برای محصول با خدمت r ام ازمنبع i ام درآبادی h ام دراستان j ام درفصل s ام و l تعداد منابع (حداکثر ۲ منبع)

متوسط قیمت فصلی محصول باخدمت درکل کشور

$$\bar{P}_{rs} = \frac{\sum_{j=1}^{t} \bar{P}_{rsj} K_{rsj}}{\sum_{j=1}^{t} K_{rsj}}$$

$\bar{P}_{rsj}$ عبارت است از متوسط قیمت محصول با خدمت r ام دراستان j ام درفصل s ام و K_{rsj} مجموع ضرایب اهمیت به دست آمده برای محصول باخدمت r ام دراستان j ام درفصل s ام . در رابطه فوق $K_{rsj} = \sum_{h=1}^{m} \bar{K}_{rsjh}$ که $\bar{K}_{rsjh}$ ضریب اهمیت برای محصول با خدمت r ام درآبادی h ام دراستان j ام درفصل s ام و m تعداد آبادیهای دارای اطلاع درهراستان و t تعداد استانهای دارای اطلاع (حداکثر ۲۲) است .

متوسط قیمت سال دراستان

$$\bar{P}_{rj} = \frac{\sum_{s=1}^{u} \bar{P}_{rjs} K_{rjs}}{\sum_{s=1}^{u} K_{rjs}}$$

$\bar{P}_{rs}$ عبارت است از متوسط قیمت محصول یا خدمت r ام در استان s ام در فصل d ام و n تعداد فصلهای دارای اطلاع در هر سال (حداکثر ۴) است .

متوسط قیمت سال در کل کشور

$$\bar{P}_r = \frac{\sum_{s=1}^{n} \bar{P}_{rs} K_{rs}}{\sum_{s=1}^{n} K_{rs}}$$

$\bar{P}_r$ عبارت است از متوسط قیمت محصول یا خدمت r ام در فصل d ام در کل کشور ، در رابطه فوق $\sum_{s=1}^{n} K_{rs} = K_r$ که K_{rs} عبارت است از ضریب اهمیت برای محصول یا خدمت r ام در استان s ام در فصل d ام

---

(۱) - ضریب اهمیت بدست آمده در هر آبادی عبارتست از وزن محصول فروخته شده در مورد محصولات کشاورزی ، مساحت زمینهای مورد عمل در مورد خدمات ماشینی و کل نفوس کارگران در مورد خدمات کارگری که توسط منابع مورد مراجعه در هر آبادی اعلام شده است.

چکیده پوسترها

سرشماریهای جمعیّتی کشور و اطّلاعات مسکن

نسرین افتخاری

مرکز آمار ایران

۵شنبه ۱۳:۳۰ بانل ۱

بررسی سرشماریهای نفوس و مسکن کشور از دیدگاه نحوهٔ برخورد با اصل قابلیت مقایسه در امر گردآوری و انتشار اطّلاعات مسکن، هدف این مقاله را تشکیل می‌دهد. برای رسیدن به این هدف موارد زیر دنبال شده است:

- بررسی توصیه‌های بین‌المللی (اهداف سرشماری نفوس و مسکن، جنبه‌های مشترک و موضوعات مورد نظر در این گونه سرشماریها)
- بررسی سرشماریهای نفوس و مسکن از نظر اطّلاعات مسکن. این بررسی نشان داده است که:
  - از نظر تعداد اقلام مسکن، سرشماریها رشد قابل ملاحظه‌ای داشته‌اند.
  - واحد آماری مورد نظر، یکسان نبوده است. به این ترتیب که جامعهٔ مورد نظر در ارتباط با خانوار و نیز در ارتباط با مسکن در هر سرشماری با سرشماری قبلی تفاوت دارد. در زمان انتشار اطّلاعات نیز قابلیّت مقایسه در اغلب موارد رعایت نشده است
- در پایان، مقاله به ارائهٔ پیشنهاداتی در زمینهٔ بهبود اطّلاعات مسکن و کاهش مشکلات موجود در امر مقایسهٔ این اطّلاعات در سرشماریها پرداخته است.

طرح شناسایی نیازها و امکانات آماری کشور

اکبر افشار

مرکز آمار ایران

۳شنبه ۱۳:۳۰ پانل ۱

کاربرد وسیع آمار در علوم مختلف و اهمیت ویژه آن در مراحل مختلف برنامه‌ریزی، مرکز آمار ایران را برآن داشت تا «طرح شناسایی نیازها و امکانات آماری کشور» را در برنامهٔ کار خود قرار دهد و به همین منظور اوّلین قدم را با اجرای «طرح شناسایی امکانات آماری کشور» در سال ۱۳۶۷ در سطح ۲۵۶۸ سازمان در کلّ کشور برداشت.

اهم اهداف طرح بشرح زیر است:

* شناسایی سازمانهای تهیه‌کنندهٔ آمار
* تعیین سازمانهای دارای واحد آمار
* شناسایی نحوهٔ مدیریّت و فعّالیّت اصلی سازمانهای تهیه‌کنندهٔ آمار
* تعیین تعداد و مشخّصات نیروی انسانی شاغل در تهیهٔ آمار
* تعیین سازمانهایی که برای تهیهٔ آمار از کامپیوتر استفاده می‌کنند و نوع کامپیوتر مورداستفاده
* تعیین تعداد و مشخّصات فعّالیّتهای آماری سازمانها

پرسشنامهٔ طرح به طریقی طرح‌ریزی شده است که نتایج حاصل از طبقه‌بندی اطّلاعات دریافتی، اهداف ذکرشده در فوق را در قالب جداول نهایی ارائه می‌دهد.

مقاله به منظور ارائهٔ نتایج حاصل از اجرای طرح فوق تهیه شده است.

آمارگیری از مرغداریهای صنعتی کشور

جلال افضل‌زاده نائینی

مرکز آمار ایران

۵شنبه ۱۴:۰۰ پانل ۲

در سی سال گذشته تعداد زیادی مؤسسهٔ مرغداری در سراسر کشور تأسیس و مشغول فعالیت شده‌اند، به‌طوری که امروزه این صنعت، نقش عمده‌ای در اقتصاد کشور ایفا می‌کند. بسیاری از صاحبنظران را عقیده بر این است که در حال حاضر از بعضی از رشته‌های این صنعت ظرفیت موجود برای نیاز کشور کافی و شاید افزون بر نیاز می‌باشد ولی به علت پاره‌ای کسبودها و ناهماهنگیها از سرمایه‌گذاری و ظرفیت موجود بهره‌برداری کامل به عمل نمی‌آید. در چنین شرایطی نیاز به برنامه‌ریزی همه‌جانبه جهت حلّ مسائل و مشکلات این فعّالیت محسوس است. مرکز آمار ایران برای پاسخگویی به درخواستهای مکرّر دستاندرکاران و برآوردن انتظارات و توقّعات دستگاههای برنامه‌ریزی و اجرایی، نسبت به تهیه و اجرای طرحهایی برای آمارگیری از مؤسّسات مرغداری اقدام کرده است. هدف کلّی اجرای این طرحها، جمع‌آوری و انتشار اطّلاعات آماری لازم برای شناخت جنبه‌های مختلف این فعّالیّت به منظور برنامه‌ریزی و انجام بررسیهای اقتصادی به خصوص محاسبهٔ ارزش افزوده بوده است. اهمّ اطّلاعاتی که از اجرای این طرحها بدست آمده است عبارتند از: مقدار و ارزش تولیدات، مقدار و ارزش مواد مورداستفاده، ظرفیتهای مرغداری، انواع هزینه و سرمایه‌گذاری و وضعیّت اشتغال.

توزیع شغلی و میزان تحصیلات زنان شاغل در سرشماری ۱۳۶۵

فرشته امیرفریار

مرکز آمار ایران

شنبه ۱۶:۳۰ بانل ۱

در این مقاله بررسی کوتاهی در مورد توزیع زنان شاغل در گروههای اصلی شغلی و میزان تحصیلات به عمل آمده است و نکات عمدهٔ آن به شرح زیر است:

- تمرکز عمدهٔ زنان شاغل در گروههای معلّمین و مدرّسین، کارمندان دفتری و اجرایی، خدمات بهداشت، بهره‌برداران کشاورزی و دامداران مالک، الیاف‌کاران و بافندگان است.
- بررسی سطح تحصیلات زنان شاغل در این گروهها نشان می‌دهد که در مناطق شهری و روستایی اکثریت زنان شاغل دارای دیپلم متوسط بوده و یا بی‌سواد می‌باشند.
- از کلّ زنان شاغل که در حدود ۹۸۷۱۰۳ نفر می‌باشند «۱۱۳۴۹۷» نفر یا ۱۱/۵ درصد دارای تحصیلات عالی می‌باشند.
- از کلّ شاغلین کشور تنها ۹/۸ درصد آنان را زنان تشکیل می‌دهند.

آمارگیری از فعّالیّت اقتصادی خانوار

بهرام امیراحمدی

مرکز آمار ایران

۳شنبه ۱۵:۰۰ پانل ۵

در کشورهای درحال‌توسعه، به خصوص کشورهایی که فاقد ساختار اقتصادی منسجم و پایدار هستند، در کنار فعّالیّتهای اقتصادی متشکّل، نوعی از فعّالیّت اقتصادی به نام «اقتصاد غیرمتشکّل» شکل گرفته است.

در بعضی از موارد و در پاره‌ای از کشورها، حدس زده می‌شود که سهم این گونه فعّالیّتهای غیرمتشکّل تا حدود ۳۰درصد از تولید ناخالص داخلی را تشکیل دهد. دسترسی به فهرست واحدهای انجام‌دهندهٔ این فعّالیّتها، همیشه یکی از معضلات سازمانهای آماری بوده است.

فعّالیّت اقتصاد غیرمتشکّل به دو صورت کلّی «فعّالیّت در مکان» و «فعالیت در خارج‌ازمکان» انجام می‌شود.

در این مقاله سعی شده است تا حدّ امکان و تسهیلات موجود سازمانهای آماری، به خصوص مرکز آمار ایران، روش قابل دسترسی برای جمع‌آوری اطّلاعات اقتصادی مربوط به فعّالیّت اقتصادی بخش غیرمتشکّل ارائه شود. در پایان، ضمن ارائهٔ سابقه جمع‌آوری آمارهای مربوط به فعّالیّت اقتصادی خانوار در سالهای مختلف، جمع‌بندی و پیشنهادهای لازم ارائه شده است.

برآورد افزودهٔ حمل‌ونقل جاده‌ای مسافر در اقتصاد ملی با استفاده از تکنیکهای آماری

محمدرضا انصاری

مرکز آمار ایران

۴شنبه ۱۲:۳۰ بانل ۵

در این مقاله نخست سعی شده است که زیرمجموعه‌های بخش حمل‌ونقل به تفکیک بیان و در حدّ معرّفی این مجموعه اقدام شود. سپس به اهداف توسعهٔ حمل‌ونقل پرداخته‌شده به‌طوری که در قسمت اوّل دورنمای توسعهٔ صنعت حمل‌ونقل ذکرشده است. بعداً وضعیّت کشور خودمان را تجسّم کرده و نظراتی ارائه شده است. در قسمت بعدی نقش افزوده‌های حمل‌ونقل و قیمت کالاها به همراه جدولی ارائه شده است. در قسمت آخر چگونگی برآورد افزودهٔ حمل‌ونقل جادّه‌ای مسافر با استفاده از طرحی که به همین منظور در مرکز آمار ایران تهیّه شده، مطالبی به‌طور خلاصه بیان شده است.

آمارگیری از منابع آب

فریده باقری

مرکز آمار ایران

۳شنبه ۱۴:۰۰ پانل ۲

تهیه و تدوین این بررسی براساس طرح آمارگیری ازکارگاههای تأمین، تصفیه و توزیع آب می‌باشد که برای اولین بار در مرکز آمار ایران در سال ۱۳۷۰ به اجرا در آمد. طبق نتایج سرشماری عمومی سال ۱۳۶۵ مجموع مکانهای دارای فعالیت تأمین، تصفیه و توزیع آب بالغ بر ۳۳۰۰۰ بوده است که به صورت نمونه‌گیری با ۱۷۵۱ مکان مورد آمارگیری قرارگرفت. هدف کلّی طرح جمع‌آوری و ارائهٔ اطّلاعات مورد نیاز جدول داده ـ ستانده‌ی اقتصاد کشور در ارتباط با فعالیت تأمین، تصفیه و توزیع آب جهت برنامه‌ریزی و اتّخاذ سیاستهای کوتاه‌مدّت، میان‌مدّت و بلندمدّت می‌باشد. اهداف تفصیلی طرح در سه بخش کلّی، تفصیلی، فعالیتهای فرعی شامل:

نحوهٔ مدیریّت، وضعیّت حقوقی، حجم استحصال آب، تعداد انتساب واگذارشده، مزد و حقوق پرداختی، ارزش سوخت مصرف‌شده، ارزش پرداختیها و دریافتیها، برآورد ارزش اموال سرمایه‌ای موجود در کارگاه، محاسبه و برآورد ارزش استهلاک و شناسایی فعالیّتهای فرعی بوده است.

مقایسهٔ تعداد آبادیهای خالی از سکنهٔ کشور در دو سرشماری و علّت افزایش آن

محمد تربتی صراف

مرکز آمار ایران

۵شنبه ۱۵:۰۰ پانل ۱

این گزارش اشاره‌ای مختصر و کوتاه به عناوین ذیل دارد:

- فراند سرشماریها و آمارگیریهای جمعیتی.
- موقعیّت جغرافیایی کشور نسبت به کرهٔ زمین، مساحت و غیره.
- افزایش شهرها و جمعیّت شهرهای ایران از سال ۱۳۵۵ لغایت ۱۳۶۵، علل افزایش شهرها و جمعیّت شهرنشینی.
- علل مهاجرت روستائیان به شهرها.
- تعاریف و مفاهیم مورد استفاده در این گزارش.
- تعداد کلّ آبادیهای کشور، تعداد آبادیهای دارای سکنه، تعداد آبادیهای خالی از سکنه در دو سرشماری ۱۳۵۵ و ۱۳۶۵ و مقایسهٔ درصدی آنها.
- مقایسهٔ هزینه و درآمد خانوارهای روستایی از سال ۱۳۵۵ لغایت ۱۳۶۵ و نسبت آن با هزینه و درآمد خانوارهای شهری.
- جمعیّت شاغل شهری و روستایی در سالهای ۱۳۵۵ و ۱۳۶۵ (جمعیّت شاغل ۱۰ ساله و بیشتر) در سه بخش کشاورزی، صنعت و خدمات و مقایسهٔ آنها.
- جمعیّت فعال روستایی و شهری در دو دهه سرشماری ۱۳۵۵ و ۱۳۶۵ و نسبت مقایسهٔ این جمعیّت.

ساختار سنّی جمعیّت در جامعهٔ عشایری ایران

فضل‌الله توکلی

مرکز آمار ایران

۴شنبه ۱۴:۰۰ بالل ۲

ترکیب و ساختار سنّی جمعیّت در جامعهٔ عشایری ایران تحت تأثیر عواملی متعدّد چون: درصد موالید و فوت زیاد (بخصوص مرگ‌ومیرهای نوزادان و کودکان)، پیری زودرس و عمر کوتاه ناشی از شرایط نامطلوب زندگی، همانند سایر جوامع توسعه‌نیافته بوده و دارای هرم سنّی با قاعده‌ای پهن و رأسی باریک است. در اثر زایمانهای متعدّد و پیری زودرس ناشی از شیوهٔ زندگی بیشتر جمعیّت در گروههای سنّی زیر ۲۵ سال قرار دارد. براساس نتایج سرشماری از عشایر کوچنده در سال ۱۳۶۶، حدود ۶۸درصد از کلّ جمعیّت را کمتر از ۲۵ساله‌ها تشکیل داده‌اند و فقط ۲/۸درصد از آنان در گروههای سنی بالاتر از ۶۵ سال قرار گرفته‌اند در صورتی که در همان سال در کشورهای توسعه‌یافته تعداد افراد زیر ۲۵ سال حدود ۴۰درصد جمعیّت را تشکیل می‌داده و تعداد سالخوردگان بالای ۶۵ سال بیش از ۱۰درصد از جمعیّت را به خود اختصاص داده‌اند. نسبت جنسی جمعیّت در جامعهٔ عشایری نیز یکی از موارد خاصّ و کم نظیر است. تقریباً در کلیّه کشورها و جوامع جمعیّت زنان سالخورده از مردان بیشتر است و گاهی تا ۳ برابر مردان نیز افزایش می‌یابد، در صورتی که در جامعهٔ عشایری چنین نیست و در گروههای سنی بالا و در میان سالخوردگان تعداد مردان از زنان خیلی بیشتر است. برای مثال نسبت جنسی در گروههای سنّی ۶۵ساله و بیشتر و ۸۵ساله و بیشتر، درکشور انگلیس به ترتیب ۶۵ و ۳۰ و در جامعهٔ عشایری به ترتیب ۱۵۰ و ۱۶۱ بوده است.

ارائهٔ تجربیّات مربوط به تهیّهٔ اطلس عشایری و تهیّهٔ پیش‌نویس اطلس سرشماری عمومی نفوس و مسکن و همکاری در تهیّهٔ اطلس ملّی

علی‌اصغر تینافر*، ماه‌منیر مجتهدی، مینا غفاری

علی‌محمد احمدپناهی

مرکز آمار ایران

۲شنبه ۱۳:۴۰ پانل ۵

از حدود سی سال پیش تاکنون، همواره تهیّه و تدوین اطلس، در زمینه‌های مختلف، مورد توجّه مسئولان بوده است ولی توفیق تهیّه و تدوین آن کمتر حاصل شده است تا جایی که در حال حاضر اطلسهایی که از نظر علمی قابل اعتنا باشند بسیار اندک است. از جمله علل اساسی این مسئله، عدم انتقال تجربیّات یک گروه به گروه دیگر است. در این مقاله تلخیصی از تجربیّات مربوط به تهیّه و تدوین اطلس عشایری و تهیّهٔ پیش‌نویس اطلس سرشماری عمومی نفوس و مسکن و همکاری در تهیّهٔ اطلس ملّی شرح داده شده است.

اشتغال زنان در ایران

مهرانگیز چنگیزی‌آشتیانی

مرکز آمار ایران

۴شنبه ۱۳:۳۰ پانل ۱

عامل نیروی انسانی در میان عوامل تولید، امروز جایگاه ویژه‌ای دارد و از مهمترین است. جای تردید وجود ندارد که کیفیت نیروی کار در فرآیند تولید از کمیّت آن مؤثرتر است. برای کشور درحال‌توسعه‌ای نظیر ایران مشارکت زنان در فرآیند توسعه، بخصوص در شرایط فعلی اقتصادی امری الزامی و ضروری است به منظور به‌کارگیری و استفاده مؤثر و مفید از نیروی عظیم زنان (که تقریباً نیمی از جمعیّت و نیروی کار کشور را تشکیل می‌دهند) تولید آمار و ارقام صحیح و قابل اعتماد به‌عنوان ابزاری برای تجزیه و تحلیل و کاربرد آن در راستای اهداف توسعهٔ اقتصادی و اجتماعی جامعه حائز کمال اهمیّت است.

در این مقاله آمارهای تولیدی حاصل از سرشماریهای عمومی نفوس و مسکن و آمارگیریهای میان‌دوره‌ای در زمینهٔ نیروی انسانی برای تحلیل و بررسی اوضاع اجتماعی اقتصادی زنان از دیدگاههای مختلف مورد استفاده قرار گرفته است. بررسی سابقهٔ آمار اشتغال زنان در ایران، توزیع سنّی، توزیع جغرافیایی، درجات تحصیلی، توزیع زنان شاغل در فعالیّتهای اقتصادی و وضع شغلی از اهمیّت خاصّی برخوردار است. علاوه بر آن در چارچوب زمانی قبل و بعد از انقلاب اسلامی به‌عنوان محوری برای انجام مقایسه‌ها در زمینهٔ تغییرات ساختار شغلی زنان قبل و بعد از انقلاب به‌کار گرفته شده است و همچنین مقایسهٔ پاره‌ای از داده‌ها با اطّلاعات سایر کشورها این مجموعه را پربارتر خواهد کرد.

شاخصهای آموزش عالی در ایران، ۱۳۶۹_۱۳۶۱

صغری حاج‌سیدجوادی

مرکز آمار ایران

در این بررسی سعی شده است تصویر روشنی از آموزش عالی در ایران در سالهای بعد از انقلاب فرهنگی، با محاسبهٔ شاخصهای مالی، ساختی، جمعیتی و بازدهی داده شود و دگرگونیها نیز تشریح شود. همچنین در پایان تحلیلی از ارتباط برخی شاخصهای آموزش عالی و رشد و توسعهٔ اقتصادی ارائه شود. اکثر شاخصهای مالی در سالهای بعد از خاتمهٔ جنگ در مقایسه با قبل از آن سیری افزاینده داشته‌اند. بحرانهای اقتصادی در زمان جنگ مهمترین عامل این تغییرات شناخته‌شده است. شاخصهای ساختی نشان می‌دهد تمرکز دانشجویان در دورهٔ لیسانس بوده است، گروههای تخصصی علوم انسانی و اجتماعی، پزشکی و مهندسی به ترتیب بیشترین تعداد دانشجو را زیر پوشش داشته‌اند و نسبت دانشجو به مدرّس در سالهای بعد از خاتمهٔ جنگ بیش از سالهای قبل از آن بوده است. شاخصهای جمعیتی افزایش بی‌وقفهٔ دانشجویان را نشان می‌دهد. در بین گروههای تخصصی، گروههای پزشکی و کشاورزی و در بین دوره‌های تحصیلی دوره‌های دکتری و فوق دیپلم بیشترین افزایش را داشته است. شاخصهای بازدهی نشانگر افزایش دانش‌آموختگان است. بیشترین افزایش به گروه فنّی و مهندسی تعلق داشته است. بازدهی سرمایه در این دوره در مقایسه با سالهای قبل به علت نرخ بالای تورم کاهش داشته است. افزایش درآمد سرانه در افزایش متغیرهای آموزشی و ساخت تخصصی مؤثر است.

بررسی روند مزد و حقوق و مقایسهٔ آن با بهره‌وری کار در کارگاههای بزرگ صنعتی در سالهای مختلف

وجیهه خلیلی

مرکز آمار ایران

مزد و حقوق به‌عنوان هزینهٔ عامل کار از شاخصهای مهم اقتصادی در برنامه‌ریزی، تصمیم‌گیری و ارزیابی اجرای برنامه‌های توسعهٔ اقتصادی و اجتماعی است. این شاخص یکی از ابزارهای مهم اندازه‌گیری رفاه شاغلان نیز محسوب شده و مورد استفاده قرار می‌گیرد.

در مقابل آنچه به‌عنوان مزد و حقوق پرداخت می‌شود، شاخصهای دیگری برای اندازه‌گیری کمیّت کار انجام‌شده وجود دارد که برای قضاوت و تصمیم‌گیری در اختیار کارفرمایان قرار دارد. یکی از عمده‌ترین این شاخصها، بهره‌وری کار کارکنان است. بهره‌وری کار و پرداختهای انجام‌شده به مزد و حقوق‌بگیران در یک فعالیت اقتصادی و بررسی روند تغییرات این دو شاخص، تصویر روشنی از عملکرد آن فعالیت را نمایان می‌سازد. بررسی روند این دو شاخص طیّ سالهای ۶۵ـ۱۳۶۱ نشان می‌دهد که:

- رشد شاخص مزد و حقوق به قیمت جاری طیّ دوران مورد بررسی از رشد بهره‌وری به قیمت جاری در همان دوره بالاتر بوده است.
- برعکس، رشد شاخص مزد و حقوق سرانهٔ واقعی از رشد شاخص بهره‌وری به قیمت ثابت در همان دوره کندتر بوده است. به عبارت دیگر علی‌رغم تلاش کارکنان در بالا بردن بهره‌وری کار در کارگاههای بزرگ صنعتی، میزان مزد و حقوق سرانه واقعی آنان کاهش داشته است یعنی این کارکنان به دلیل بالا بردن شاخص هزینهٔ زندگی به‌تدریج قدرت خرید خود را به میزان قابل توجّهی از دست داده‌اند.

روش جمع‌آوری آمار در فعّالیّت پرورش کرم ابریشم

غلامعلی خیری

مرکز آمار ایران

پرورش کرم ابریشم یکی از فعّالیتهای نسبتاً مهّم بخش کشاورزی است و در حال حاضر از جمله منابع درآمد و اشتغال در شمال کشور محسوب می‌شود و زمینه‌های توسعهٔ آن از هرجهت وجود دارد. برای تهیهٔ آمارهای مورد نیاز ابتدا لازم است بررسی کاملی برای شناخت جنبه‌ها و ابعاد گوناگون این فعّالیّت و وضع موجود آن در ایران انجام گیرد. این بررسیها نشان می‌دهد که از میان روشهای مختلف تهیهٔ آمار و اطّلاعات، اجرای طرح آمارگیری نمونه‌ای از طریق مراجعهٔ مستقیم به بهره‌برداریهای پرورش‌دهندهٔ کرم ابریشم مناسبترین روش است. نتایج حاصل از اجرای طرح سرشماری عمومی کشاورزی ۱۳۶۷ چارچوب مناسبی را برای تهیهٔ این طرح فراهم کرده است. این اطّلاعات چارچوبی دربرگیرنده فهرست آبادیها و نقاط شهری دارای فعّالیت نوغانداری، تعداد خانوارهای بهره‌بردار، تعداد جعبه، تخم نوغان مصرفی و مقدار تولید پیله است. با استفاده از این چارچوب و روشهای آماری، ابتدا تعدادی واحد نمونه (آبادی/موزه‌شهری) انتخاب می‌شود و کلیهٔ مکانها و خانوارهای آن مورد فهرست‌برداری قرار می‌گیرد. از بین خانوارهای فهرست‌شده تعدادی نمونه انتخاب و پرسشنامهٔ بهره‌برداری برای آنها تکمیل می‌شود. اطّلاعات این آمارگیری به روش ماشینی استخراج و در قالب نشریه آماری ارائه خواهدشد.

روند مصرف انرژی در کشور طیّ سالهای اخیر و بررسی مصرف انرژی در صنعت سیمان

محمود درودی

مرکز آمار ایران

روند مصرف انرژی نهایی درکشور طیّ دهه‌های اخیر سیری صعودی طیّ کرده و از ۵۲/۸ میلیون بشکه معادل نفت خام در سال ۱۳۴۶ به ۳۹۲/۸ میلیون بشکه معادل نفت خام در سال ۱۳۶۹ افزایش یافته است و فرآورده‌های نفتی عمده‌ترین حاملهای انرژی نهایی مصرفی بوده‌اند. مهمترین عوامل رشد سریع مصرف انرژی را بشرح زیر می‌توان طبقه‌بندی کرد:

۱ـ رشد جمعیت ۲ـ مهاجرت روستائیان به شهرها ۳ـ رشد کمّی اقتصاد کشور ۴ـ تحولات ساختاری در سیستمهای تولیدی کشور ۵ـ تغییر الگوی مصرف انرژی در بخش خانگی ۶ـ گسترش حمل‌ونقل و توسعهٔ شهرنشینی

بررسی همه جانبهٔ تحولات سیستم انرژی و مطالعهٔ امکانات و محدودیتهای بخش انرژی در چارچوب توسعهٔ اقتصادی و اجتماعی کشور مستلزم بکارگیری آمار و اطّلاعات جامع و منسجم می‌باشد. مقالهٔ تهیه‌شده، کاربرد آمار در سیستم انرژی کشور و مصرف انرژی در بخشهای مهّم اقتصادی را به‌طور اعمّ و بخش صنعت بخصوص صنعت انرژی‌بر سیمان را مورد توجّه قرار می‌دهد.

بررسی آماری صنعت چرم ایران

پرویز رحمانی

مرکز آمار ایران

اطلاعات مندرج در این بررسی شامل پنج فصل است که بشرح زیر بیان می‌شود. در فصلهای اول و دوم تاریخچهٔ استفاده از پوست و چرم در ایران و جهان و همچنین مراحل ساخت و پرداخت آن مورد بررسی قرار می‌گیرد. در فصل سوم مقدار تولید کارگاههای بزرگ چرمسازی کشور در سالهای ۱۳۶۳ـ۱۳۵۳ بررسی می‌شود و سپس به روند تغییرات مقدار تولید می‌پردازد و همچنین در این فصل آمار تعداد دام ذبح‌شده کشور به تفکیک نوع دام در سالهای فوق ارائه می‌شود که با استفاده از اطلاعات کمکی در آن می‌توان مقدار تولید چرم را در سالهای مختلف برآورد کرد. در این فصل نیز مسائل اقتصادی این صنعت در سال ۱۳۶۳ در زمینهٔ تعداد کارگاه، کارکنان، ارزش افزوده، بهره‌وری کار، ارزش افزودهٔ سرانه کارگاه، مزد و حقوق پرداختی و درصد وابستگی این صنعت به خارج بررسی می‌شود. فصل چهارم اختصاص به واردات و صادرات انواع پوست و چرم در سالهای ۱۳۶۳ـ۱۳۰۳ دارد و همچنین در این فصل نسبت صادرات و واردات پوست و چرم به صادرات و واردات کلّ کشور بجز مواد نفتی و میزان وابستگی آن به خارج و نوسانات آن طیّ سالهای ۱۳۰۳ به بعد مورد کاوش قرار می‌گیرد. درفصل پنجم مسائل و مشکلات مربوط به این صنعت بررسی و پس از نتیجه‌گیری برای حلّ مشکلات پیشنهاداتی نیز ارائه می‌شود.

بررسی آگاهی، بینش و عملکرد زنان همسردار ۴۹ـ۱۵ ساله روستایی ایران نسبت به برنامهٔ تنظیم خانواده (۱۳۶۹)

منصور رضایی

دانشگاه علوم پزشکی باختران

تعداد ۵۰۰۰ خانوار روستایی ایران به تناسب جمعیت در ۲۴ استان کشور با استفاده از روشهای نمونه‌گیری خوشه‌ای، سیستماتیک و احتمال متناسب با اندازه انتخاب و مورد مصاحبهٔ حضوری قرار گرفتند و چند نتیجهٔ مهم بشرح زیر بدست آمد.

- جمعیتی: میانگین سنّ جاری زنان ۲۰/۸ سال و مد آن ۲۵ سال و اختلاف سنّ زن از همسر به‌طور متوسط ۶/۷۷ سال است و ۷۸/۶٪ از زنان بی‌سواد می‌باشند.
- باروری: میانگین سنّ اولین ازدواج ۱۶/۷۶ و مد آن ۱۵ سال است. ۲۲/۸٪ از دختران کمتر از ۱۵ سالگی و ۱۱/۵٪ بالای ۲۰ سالگی ازدواج کرده‌اند. ۳۱٪ از زنان در ۱۶ سالگی و پایینتر حامله شده‌اند. میانگین زنده‌زایی، مرده‌زایی و فرزندان موجود بترتیب ۵/۲۲، ۰/۱۶ و ۴/۳۶ فرزند می‌باشد.
- آگاهی: ۸۶٪ حداقل یک روش پیشگیری را می‌شناسند که قرص با ۹۹/۲٪ مشهورترین و دورهٔ ایمنی با ۲/۵٪ ناشناخته‌ترین روش می‌باشد.
- عقیده: ۸۷٪ آگاهی دارندگان موافق اجرای برنامهٔ تنظیم خانواده و ۷/۷٪ مخالف آن هستند.
- استفاده: ۵۱/۶٪ تاکنون و ۳۱/۱٪ در حال حاضر از روشهای پیشگیری استفاده کرده و یا می‌کنند.

بررسی صنعت فرش ایران

سعید زارعدوست

مرکز آمار ایران

این بررسی شامل دوازده بخش است که بشرح زیر عنوان می‌شود:

- بخش یک، تاریخچهٔ صنعت فرش ایران و جهان مورد بررسی قرار می‌گیرد
- بخش دو، سازمانهای عمومی کشور که در تولید فرش دستباف دخالت مستقیم می‌کنند
- بخش سه، جایگاه فرش در طبقه‌بندیهای بین‌المللی
- بخش چهار، تولید شامل مراحل تولید و عوامل تولید
- بخش پنج، انواع بافت فرش دستباف در کشور
- بخش شش، اندازه‌های تولید
- بخش هفت، مقدار تولید فرش دستباف در کشور و سایر کشورهای تولیدکننده
- بخش هشت، هزینهٔ تولید و عوامل متشکلهٔ آن در صنعت فرش دستباف
- بخش نه، صادرات فرش دستباف کشور و سایر کشورهای تولیدکننده و مقایسهٔ آن
- بخش ده، نتیجه‌گیری شامل بررسی اهمّ مسائل و مشکلات موجود در صنعت فرش ایران و ارائهٔ راه‌حلها و پیشنهادها
- بخش یازده، جداول آماری مربوط به مقدار تولید فرش و صادرات آن و مقایسهٔ آن با کلّ صادرات غیرنفتی
- بخش دوازده، نمودار آماری درصد ارزش صادرات کالاهای غیرنفتی کشور به تفکیک تولید کالا از جمله قالی و قالیچه

طرح آمارگیری از محلّ خدمات بیمارستانی

ثریا زاهد

مرکز آمار ایران

طرح خدمات بهداشتی و درمانی شامل در زیر گروه:

- محل خدمات بیمارستانی
- محلّ خدمات غیربیمارستانی

و یکی از ۵۲ رشته فعالیت بخش خدمات است که در سال ۱۳۶۸ محلّ خدمات بیمارستانی به صورت سرشماری و محلّ خدمات غیربیمارستانی به صورت نمونه‌گیری آمارگیری شد. به دلیل آماده نبودن جداول محلّ خدمات غیربیمارستانی تنها طرح آمارگیری محلّ خدمات بیمارستانی ارائه خواهد شد.

مطالب مقاله شامل:

مقدمه، سابقه طرح، اهداف طرح، اقلام آماری، جامعهٔ آماری، واحد آماری، پوشش جغرافیایی طرح، دوره آماری، چارچوب آماری، روش آمارگیری، تعاریف و مفاهیم و خلاصه یافته‌ها می‌باشد.

ویژگی شاغلان بخش صنعت در کشور

علیرضا سحرخیز

مرکز آمار ایران

با رشد و گسترش رشته‌های مختلف صنایع در ایران در فاصلهٔ سالهای ۴۵ تا ۵۵ تعداد شاغلان بخش صنعت در کلیهٔ گروههای عمده، فعالیت این بخش بخصوص در گروههای عمدهٔ صنعت و ساختمان افزایش چشمگیری داشته است امّا پس از انقلاب اسلامی ایران و فرار سرمایه‌دارانی که در بخش صنعت سرمایه‌گذاری کرده بودند و محاصرهٔ اقتصادی کشور توسط کشورهای صنعتی و متعاقب آن آغاز جنگ تحمیلی عراق علیه ایران، فعالیتهای مربوط به بخش صنعت دچار رکود شد که این رکود باعث کاهش تعداد شاغلان در بخش صنعت شد بطوری که در سال ۱۳۶۵ با وجود افزایش تعداد ۲۲۰۲۱۳۱ نفر به تعداد شاغلان ۱۰ ساله و بیشتر کشور نسبت به سال ۱۳۵۵، از تعداد شاغلان بخش صنعت، ۲۳۱۲۸۸ نفر کاسته شد که این کاهش از نظر گروههای عمده، فعالیت مربوط به گروههای استخراج معادن و صنعت و از نظر وضع شغلی مربوط به کاهش تعداد شاغلان بخش خصوصی در فاصلهٔ زمانی مذکور بوده است.

بررسی آماری نیروی انسانی شاغل در بخش ساختمان کشور

اسکندر سروش

مرکز آمار ایران

مقالهٔ اینجانب شامل بخشهای زیر می‌باشد:

۱ـ مقدمه ۲ـ تعاریف و مفاهیم مورد استفاده در سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ ۳ـ شرح کدهای مورد استفاده در بخش ساختمان ۴ـ شرح روش ۵ـ جداول آماری ۶ـ شکلها ۷ـ نمودارها ۸ـ یافته‌های مهم ۹ـ فهرست منابع و مآخذ

در مقدمه به اهمیت و سهم نیروی انسانی در کلیهٔ فعالیتها و نقش مؤثر آن در توسعهٔ اقتصادی و اجتماعی کشور اشاره شده است. در ادامهٔ مقاله تعاریف و مفاهیم مورد استفاده در سرشماری سال ۱۳۶۵ و همچنین کدهای مورد استفاده در بخش ساختمان به صورت فعالیت و مشاغل شرح داده شده است. روش تحقیق نیز در قالب جامعهٔ آماری، صفات مورد بررسی، تحقیق برمبنای اطلاعات قبلی، جمع‌آوری داده‌ها، عملیات اجرایی تحقیق و نتیجهٔ تحقیق مورد بررسی قرار گرفته است. در این مقاله برای توضیح بیشتر از شکلها و همچنین نمودارهای آماری استفاده شده است. دو مقاله حاضر اطلاعات مربوط به شاغلان بخش ساختمان برحسب سکونت آنان در مناطق شهری و روستایی استانهای کشور و به تفکیک سن، جنس، سواد، وضع شغلی و گروههای اصلی، فعالیت طبقه‌بندی، محاسبه و مورد بررسی قرار گرفته است. سپس یافته‌های مهم آماری در ارتباط با شاغلان بخش ساختمان تشریح شده ضمناً جداول آماری مربوط در قالب ۱۶ جدول ارائه شده است. در پایان فهرست منابع و مآخذ مورد مراجعه نوشته شده است.

بررسی آماری مشخّصات بیماران مسموم بستری‌شده در بخش مسمومین

محمود رحیمی - بهرام سلیمانی*

دانشگاه علوم پزشکی اصفهان

در این تحقیق اطلاعات مربوط به ۹۷۴۲ بیمار مسموم که در فاصلهٔ سالهای ۱۳۶۴ تا ۱۳۶۶ به اورژانس مسمومین واقع در بیمارستان خورشید اصفهان مراجعه کرده‌اند از نظر سن، جنس، نوع مسمومیّت، مواد مسمومیّت‌زا و تلفات مورد بررسی آماری قرار گرفته و ضمن برآورد پارامترها و رسم نمودار، با استفاده از آزمونهای آماری از جمله F, X^2, Z, t و آزمون ناپارامتری کولمگروف-اسمیرنوف، تفاوت نسبت جنسی مسمومین در سالهای مختلف، نسبت مسمومین عمدی و اتفاقی، همگن‌بودن توزیع جنسی در سالهای مختلف و برحسب نوع مسمومیّت، میانگین و واریانس سنّ فوت‌شدگان زن و مرد مقایسه‌شده و در هر مورد نتیجه‌گیری بعمل آمده است.

شاخصهای کمّی مسکن

فاطمه سمیعی

مرکز آمار ایران

مسکن یکی از نیازهای اساسی انسان است. به جهت رسیدن به هدف نهایی تأمین مسکن مناسب برای آحاد جامعه، شناخت وضعیّت موجود و سمت‌گیری آن و تغییرات بوجودآمده در طول زمان، لازم است. اینکه تا چه حدّ به استانداردهای بین‌المللی نزدیک شده‌ایم و یا از آن فاصله گرفته‌ایم فقط از طریق بررسی داده‌های آماری میسّر است. در این زمینه، سرشماریها اطّلاعات ارزنده‌ای در اختیار قرار می‌دهند. این اطّلاعات به منظور قابل استفاده‌شدن باید به پارامترهای گویایی تبدیل شوند.

جهت بررسی و شناخت مسئله مسکن و شناسایی زمینه‌های کمبود آن یکی از روشهای مورد استفاده محاسبهٔ شاخصهای تراکم در واحدهای مسکونی است. این مقاله با استفاده از روش فوق‌الذّکر و با بهره‌گیری از نتایج سرشماریهای عمومی نفوس و مسکن سال ۱۳۵۵ و ۱۳۶۵ به محاسبهٔ شاخصهای تراکم در واحدهای مسکونی به تفکیک نقاط شهری و روستایی کشور پرداخته است. شاخصهای تراکم مورد استفاده در مقاله عبارتند از متوسّط نفر در اتاق، متوسّط نفر در واحد مسکونی، متوسّط خانوار در واحد مسکونی و متوسّط اتاق در واحد مسکونی. اطّلاعات بدست‌آمده نشانگر بهبود وضعیّت مسکن خانوارها در فاصلهٔ دو سرشماری بوده است.

تحلیل علّتهای مرگ‌ومیر در ایران براساس اطلاعات ثبتی سالهای ۶۵ تا ۶۹

حجت‌ا... سیف‌الهی تنکران

دانشگاه شهید بهشتی

برای تحلیل جدولهای مختلف با هدفهای زیر تجزیهٔ عاملی شده‌اند:

- تعیین تفاوت در دو گروه زن و مرد از لحاظ علل مرگ‌ومیر (ارتباط علل مرگ‌ومیر با جنس)
- تعیین تفاوت میان ساکنان شهر و روستا از لحاظ علل مرگ‌ومیر (ارتباط علل مرگ‌ومیر با محلّ سکونت)
- تعیین تفاوت میان مردها و زنان ساکن در شهرها و روستاها از حیث مرگ‌ومیر (اثر محیط زندگی و جنس در مرگ‌ومیر)
- تعیین ارتباط بیماریها با گروههای سنّی زن و مرد و تغییرات حاصل در سالهای یادشده
- تعیین ارتباط ماههای سال در مرگ‌ومیر مرد و زن
- تعیین ارتباط مرگ‌ومیر با گروههای سنّی و جنس در سالهای مختلف
- بررسی علّتهای مرگ‌ومیر در شهرهای مختلف در سالهای متفاوت

بررسی آماری حمل‌ونقل در ایران

علی‌محمد شاه‌محمدی

مرکز آمار ایران

در این مقاله راجع به مطالب زیر بحث شده است:

- نقش حمل‌ونقل در رشد و توسعهٔ اقتصادی
- سهم هزینهٔ حمل‌ونقل در سبد هزینهٔ زندگی خانوارهای شهری و روستایی
- نسبت سهم ارزش افزودهٔ حمل‌ونقل، سرمایه‌گذاری در حمل‌ونقل و تولید حمل‌ونقل در ایران
- زیرجامعه‌های حمل‌ونقل شامل حمل‌ونقل زمینی، هوایی، دریایی
- ارائهٔ تصویری آماری از حمل‌ونقل با راه‌آهن شامل، طول شبکهٔ راه‌آهن، بار و مسافر حمل‌شده، ظرفیتهای موجود و مقایسه با سایر کشورها
- بررسی آماری حمل‌ونقل جاده‌ای بار و مسافر (بین‌شهری و درون‌شهری)
- بررسی حمل‌ونقل دریایی شامل ناوگان حمل بار و مسافر دریایی
- تحلیلی از حمل‌ونقل هوایی شامل تعداد هواپیماهای مسافربری، بار و مسافر حمل‌شده توسط هواپیما، ظرفیتهای موجود

طرح آمارگیری از کارگاههای بازرگانی و تجزیه و تحلیل نتایج آن

نسرین شعبانی

مرکز آمار ایران

اقتصاد هر کشور از سه جزء اصلی صنعت، کشاورزی و خدمات تشکیل شده است که سهم هریک از این سه جزء در کشورهای مختلف در تولید ناخالص ملّی آن کشورها به تناسب درجهٔ توسعه‌یافتگی آنها متفاوت است. طبق شواهد و گزارشهای موجود بانک مرکزی، در کشور ما در سالهای قبل از انقلاب و بعد از آن بیشترین سهم تشکیل‌دهندهٔ تولید ناخالص ملّی مربوط به بخش خدمات بوده است و بخش بازرگانی به‌تنهٔ خود سهمی اساسی در تولید ناخالص بخش خدمات داشته است. به‌طوری که سهم بخش بازرگانی در تولید بخش خدمات برابر با ۲۹/۵ درصد بوده است.

بعد از پیروزی انقلاب، جمهوری اسلامی ایران تصمیم گرفت که دو بخش کشاورزی و صنعت را محور قرار دهد امّا به دلایل مختلف مشاهده می‌شود که گرایش به سوی سرمایه‌گذاری در بخش خدمات و علی‌الخصوص در بخش بازرگانی که بهره‌دهی آن تقریباً بلافاصله می‌باشد بوده است.

بررسی راههای روستایی کشور

منوچهر صالحی

مرکز آمار ایران

همانطوری که هر نقطه از بدن انسان به وسیلهٔ شبکه‌های ارتباطی مویرگها، رگها و شاهرگها به مرکز بدن یعنی قلب مربوطند و این ارتباط باعث ادامهٔ حیات شخص می‌شود یک کشور هم به‌عنوان یک ارگان زنده برای ادامهٔ حیات و توسعهٔ روابط اقتصادی، اجتماعی و فرهنگی خود باید دارای شبکهٔ گسترده‌ای از راههای فرعی، اصلی و بزرگراه باشد که بتواند تمام نقاط کشور را به هم متصل کند. لذا آبادیهای کشور براساس شرایط اقتصادی، اجتماعی و جغرافیایی خود از طریق راههای مختلف (زمینی، آبی و راه‌آهن) با یکدیگر و به شبکهٔ راههای اصلی و فرعی کشور مربوط شده‌اند. در این مقاله سعی شده است که توزیع درصد آبادیهای دارای راه برحسب استان و نوع راه، تعداد آبادیهای دارای راه برحسب نوع راه، توزیع درصد آبادیهای دارای راه برحسب تعداد خانوار، براساس نتایج سرشماری سال ۱۳۶۵ بررسی و با نتایج سرشماری سال ۱۳۵۵ مقایسه شده است. مقاله شامل قسمتهای زیر می‌باشد.

۱ـ مقدمه ۲ـ تعریف راههای مختلف و مشخصات فنّی آنها ۳ـ تاریخچهٔ راه در ایران ۴ـ خلاصهٔ یافته‌های آماری ۵ـ جداول ۶ـ فهرست منابع و مآخذ

کاربرد سنجش از دور در آمارگیری کشاورزی و منابع طبیعی

نصرت‌الله صالحی

مرکز آمار ایران

تنوع کشت و پراکندگی مزارع، فاصلهٔ زیاد روستاها و تعداد زیاد بهره‌برداران کشاورزی و وسعت بسیار جنگلها و مراتع کشور از خصوصیات کشاورزی و منابع طبیعی (جنگلها و مراتع) کشور است. هرگونه جمع‌آوری اطلاعات از پدیده‌های مختلف کشاورزی و منابع طبیعی (جنگلها و مراتع) مستلزم صرف وقت، هزینه و نیروی انسانی زیاد جهت تهیهٔ چارچوب، طرح فنی، راهنماها، پرسشنامه‌ها، جمع‌آوری داده‌ها، فرایری داده‌ها و انتشار نتایج است. بدیهی است که آمار بایستی دارای دو خصیصه دقت و به هنگام بودن باشد که تأکید بر روی هر یک از آنها باعث لطمه زدن به دیگری می‌شود.

علی‌رغم تمامی کوششهایی که تاکنون به عمل آمده است انتشار نتایج طرحهای آماری ماهیتاً در زمانی نسبتاً طولانی بعد از جمع‌آوری داده‌ها صورت می‌پذیرد. چیزی که هر روز آن از اعتبار اطلاعات می‌کاهد. با محدوده در آمارگیری از جنگلها و مراتع با واحد آماری مشخصی روبرو نیستیم بدین لحاظ استفاده از سایر ابزار ضروری به نظر می‌رسد.

فنّ سنجش از دور که در سالیان اخیر رشد و پیشرفت چشمگیری داشته است و در اغلب کشورهای جهان با موفقیت مورد استفاده قرار گرفته است و امکانات آن نیز در ایران فراهم است می‌تواند در جمع‌آوری داده‌های زمینی مورد استفاده قرار گیرد.

با استفاده از قابلیتهای سنجنده‌های متفاوت و سایر موضوعات مانند تقویم زراعی می‌توان به برآورد سطح زیر کاشت و پیش‌بینی محصول و طبقه‌بندی جنگلها و مراتع پرداخت.

جمعیّت و آثار اقتصادی

احمد صفایی‌نیک

مرکز آمار ایران

با توجّه به نتایج سرشماریهای چهار دههٔ اخیر و نیز داده‌های مقدّماتی طرح آمارگیری جاری جمعیّت سال ۱۳۷۰، ملاحظه می‌شود که جمعیّت ایران از سال ۱۳۳۵ تا ۱۳۷۰ یعنی طیّ سی‌وپنج سال بیش از سه برابر شده است. روند رشد جمعیّت در فاصلهٔ چهار سرشماری دچار افزایش و کاهشهایی شده است و بدنبال آن رشد متوسّط درآمد سرانه در کشور نیز دچار نوسان بوده است به‌طوری که در سال ۱۳۶۵ با کاهش شدیدی مواجه گشته است. میزان بالای رشد جمعیّت و سطح پایین تولید ناخالص ملّی امروزه به‌عنوان مشخّصه مهّم کشورهای توسعه‌نیافته به‌شمار می‌رود. بررسی در این زمینه و مقایسهٔ داده‌های قبلی با نتایج مقدّماتی طرح آمارگیری جاری جمعیّت ۱۳۷۰ بیانگر بهبودی محتصری در میزانها و شاخصهای اقتصادی-اجتماعی است که در نخستین برنامهٔ پنجساله توسعهٔ اقتصادی-اجتماعی و فرهنگی کشور پیش‌بینی شده است.

شناسایی وضعیّت فعّالیّت آماری سازمانهای تحت پوشش سازمان امور اداری و استخدامی کشور

علی اصغر صفایی نیک

مرکز آمار ایران

شناسایی وضعیّت فعّالیّت آماری سازمانهایی که تحت پوشش سازمان امور اداری و استخدامی کشور قرار دارند برای ۳۵۲ سازمان مشمول جامعه طرح تا آخر سال ۱۳۶۹ مورد بررسی و شناسایی قرار گرفت. طرح مزبور برای سازمانهایی تهیّه شده که دارای تشکیلات حقوقی و سازمانی می‌باشند (بجز بخش خصوصی و تعاونی) تا بتوان تصویری روشن از سازمانهای تهیّه‌کنندهٔ آمار در جهت سامان دادن به بخشی از تشکیلات نظام آماری کشور را ارائه داد. در جهت تحقق هدف مزبور و با توجّه به اینکه اکثریّت سازمانهای دولتی براساس یکی از سه حالت

- مشمول مقررات قانون استخدام کشوری
- مشمول مقررات استخدامی شرکتهای دولتی
- مشمول مقررات استخدامی خاصّ

با سازمان امور اداری و استخدامی کشور ارتباط دارند، اقدامات لازم برای اجرا و شناسایی سازمانهای مزبور به عمل آمد و در نتیجه سازمانهایی که دارای واحد آمار مصوّب در مرکز یا استانها و همچنین سازمانهایی که در تشکیلات سازمانی خود واحد آمار و یا پست سازمانی و یا وظایف آماری دارند و همچنین سازمانهایی که در تشکیلات خود واحد آمار، پست سازمانی آماری و وظایف آماری پیش‌بینی نکرده و یا تشکیلات آنها به تصویب سازمان امور اداری و استخدامی کشور نرسیده است مورد بررسی قرار گرفتند.

طرح آمارگیری از خدمات آموزشی

پروین طوفانیان

مرکز آمار ایران

در این طرح کارگاههای خدمات آموزشی کشور در هشت زیرگروه بشرح ذیل مورد مطالعه قرار گرفته است.

۱ـ کلاس آمادگی ۲ـ دبستان و کلاس نهضت سوادآموزی ۳ـ دبستان دانش‌آموزان استثنایی ۴ـ مدرسه راهنمایی و دبیرستان دانش‌آموزان استثنایی ۵ـ مدرسه راهنمایی، دبیرستان، مدرسه آموزش بزرگسالان و دانشسرای تربیت معلم روستایی ۶ـ هنرستان و آموزشگاه ۷ـ مرکز آموزش عالی ۸ـ مدرسه علوم دینی و حوزهٔ علمیه. در این مقاله به مراحل مختلف تهیهٔ طرح خدمات آموزشی شامل مقدمه، سابقهٔ طرح فهرست ۵۲ رشته فعالیتهای بخش خدمات، هدف کلی، اهداف تفصیلی، اقلام آماری، جداول نهایی، جامعهٔ آماری، واحد آماری، پوشش جغرافیایی طرح آزمایش طرح، دورهٔ آماری، پرسشنامهٔ آماری، چارچوب آماری، معرّفی راهنماها، روش آمارگیری، روش نمونه‌گیری، سازمان اجرایی طرح، نحوهٔ جمع‌آوری اطّلاعات، نیروی انسانی، آموزش، زمان اجرای طرح، استخراج، تدوین و انتشار و نظارت و پیگیری پرداخته شده است.

نتایج این طرح در دست تهیّه می‌باشد که در صورت آماده شدن نتایج بدست‌آمده به صورت یافته‌های مهم ارائه خواهد شد.

بررسی آماری مؤسّسات تحقیقاتی کشور

مولود کامکار

مرکز آمار ایران

این بررسی براساس اجرای طرح آمارگیری از خدمات پژوهشی و تحقیقی کشور می‌باشد که برای اولین بار در مرکز آمار ایران در سال ۱۳۷۰ به اجرا درآمده است. تعداد کارگاههای قابل تعمیم در این آمارگیری براساس سرشماری عمومی نفوس و مسکن سال ۱۳۶۵، ۶۲۲ کارگاه است. هدف کلّی این بررسی گردآوری و ارائهٔ اطّلاعات و آمار مورد نیاز جدول داده‌ستانده اقتصاد کشور جهت برنامه‌ریزی و اتّخاذ سیاستهای کلّی در مقالیتهای اقتصادی است. طبق تعریف: خدمات پژوهشی و تحقیقی شامل تحقیقات، بررسی، مطالعه و گردآوری اطّلاعات در زمینه‌های خاص علمی، فرهنگی و اجتماعی است که عمدتاً به تحقیقات اساسی و کلّی در زمینهٔ زیست‌شناسی، فیزیک و علوم اجتماعی اشتغال داشته‌اند. در این طرح کلیهٔ واحدهای پژوهشی و تحقیقی کشور در سه گروه مراکز مراشناسی، آزمایشگاهی و دیگر مؤسّسات تحقیقاتی مدّ نظر می‌باشد. اطّلاعات ارائه‌شده در ارتباط با اهداف تفصیلی طرح شامل شناخت نوع فعالیّت، نحوهٔ مدیریت، وضعیّت حقوقی، وضعیّت نوع بیمه، تسهیلات و امکانات موجود و همچنین ارزش پرداختیها و دریافتیها، ارزش افزودهٔ ناخالص، تشکیل سرمایه، سوخت مصرف‌شده، حجم اشتغال، حجم مزد و حقوق پرداختی بهره‌وری کار، ارزش استهلاک و ارزش اموال سرمایه‌ای کارگاه.

آمار در ایران

مجیدرضا کتابدار

دانشگاه تربیت مدرس و وزارت جهاد سازندگی

این مقاله شامل موارد زیر می‌شود:

(فصل ۱) ریشه لغوی و معانی:

- معنی لغوی آمار
- کلمه آمار در لغت نامه‌ها
- تعریف آمار
- تاریخچهٔ احتمال

(فصل ۲) آمار نظری در ایران:

- اوّلین مدرّس، کلاس و کتاب آماری ایران
- آمار در آموزش و پرورش
- آمار و آموزش عالی
- اوّلین مؤسسه آموزش عالی آمار در ایران
- وضعیّت فعلی آمار در آموزش عالی
- شناسایی دوره‌های کارشناسی و کاردانی آمار

(فصل ۳) آمار کاربردی در ایران

بخش ۱: سازمانهای آماری

* تعاریف و نیاز کشورها به سازمانهای آماری
* تاریخچهٔ سازمانهای آماری در ایران
* فعالیتهای آمار عمومی سابق
* وضع کارمندان هنگام تاسیس مرکز آمار ایران
* نمودار تشکیلاتی مرکز آمار ایران و توزیع کارکنان برحسب مدرک تحصیلی
* نمودار تشکیلات مرکز آمار ایران
* وظائف مرکز آمار ایران

بخش ۲: آمار اقتصادی

* اصطلاح آمارهای اقتصادی
* آمار و حسابداری
* رابطه علم آمار با رونق بازرگانی و صنعت
* اهمیت آمار در اقتصاد ممالک

بخش ۳: ادارهٔ آمار اقتصادی بانک مرکزی جمهوری اسلامی ایران

* تاریخ تاسیس
* وظایف
* موقعیت پرسنلی
* آموزش کادر
* نحوهٔ اجرا و وظایف فعلی
* فهرست رشته‌های خدماتی تحت بررسی ادارهٔ آمار اقتصادی
* اسامی شهرهای مورد بررسی ادارهٔ آمار اقتصادی

بررسی کوتاه در مورد ساختار جمعیّتی اقلیّتهای مذهبی ایران براساس سرشماری سالهای ۱۳۳۵، ۱۳۴۵، ۱۳۵۵ و ۱۳۶۵

حسن کیانی خوزستانی

مرکز آمار ایران

این مطالعه که براساس نتایج سرشماری ۱۳۶۵ و با مقایسهٔ نتایج سرشماری سالهای ۱۳۳۵، ۱۳۴۵ و ۱۳۵۵ صورت گرفته است، اقلیّتهای مذهبی مورد قبول در سرشماری یعنی کلیمی، زرتشتی و مسیحی را شامل می‌شود. براین‌اساس جمع کلّ اقلیّتهای مذهبی در سال ۱۳۶۵ بالغ بر ۲۱۴۸۰۲ نفر بوده که حدود ۰/۴۳ درصد کلّ جمعیّت کشور است. مقایسهٔ این جمعیّت بین سالهای ۱۳۵۵-۱۳۴۵ افزایشی برابر ۹/۷ درصد و همین جمعیّت در طیّ سالهای ۱۳۶۵-۱۳۵۵ کاهشی برابر با ۳۷۴۴۹ نفر یعنی در حدود ۱۴/۸ درصد را نشان می‌دهد که این کاهش، به احتمال زیاد مربوط به مهاجرت خارجی این جامعه در سالهای مذکور است. مهمترین عواملی که در این مطالعه در نظر گرفته شده نسبت جنسی، میزان پراکندگی، درصد شهرنشینی، توزیع سنّی و اهمیّت نسبی هرگروه از کلیمیان، زرتشتیان و مسیحیان نسبت به کلّ جامعهٔ اقلیّتهای ایران است. در نتیجه‌گیری کلّی مشخص می‌شود در سال ۱۳۶۵ حدود ۶۲ درصد جمعیّت غیرمسلمان کشور را گروه مسیحیان، کلیمیان و زرتشتیان تشکیل داده‌اند و تمایل به شهرنشینی و اشتغال و سرمایه‌گذاری در کارهای آزاد و فعّالیّتهای اقتصادی، سبب شده تا شهرهای بزرگ و مراکز صنعتی و اقتصادی مهم کشور درصد بیشتری از این جامعه را در خود جای دهد. دراین تحقیق نتایج بررسی در قالب سه جدول شامل جمعیّت اقلیّتهای مذهبی به تفکیک گروه مذهبی، شهری و روستایی، نسبت جنسی و اهمیّت نسبی ارائه شده است.

آمارگیری کشاورزی در ایران

سیداحمد محدث

مرکز آمار ایران

از آنجا که بخش کشاورزی رکن عمدهٔ توسعه در کشور است از آن بعنوان محور توسعه یاد می‌شود. با توجه به اینکه اطلاعات و آمارهای کشاورزی اولین ابزار برنامه‌ریزی و ارزیابی پیشرفت برنامه‌هاست، اطلاع برنامه‌ریزان از کم‌وکیف آمارهای موجود در این بخش می‌تواند مفید باشد.

این مقاله به دو بخش تقسیم می‌شود:

- بخش نخست، تاریخچهٔ آمارگیری کشاورزی در ایران: شاید اولین آمارگیری نمونه‌ای که در ایران انجام شده است مربوط به آمارگیری نمونه‌ای کشاورزی در منطقهٔ ورامین باشد که توسط آقای دکتر خواجه‌نوری در سال ۱۳۲۷ انجام شده است. امّا اولین آمارگیری سراسری کشاورزی در سال ۱۳۳۹ توسط ادارهٔ آمار عمومی انجام شده است. در سالهای بعد هم آمارگیریهای نمونه‌ای انجام شده است. اولین سرشماری کشاورزی در ایران در سال ۱۳۵۲ در مناطق روستایی کشور انجام شده است. لیکن سرشماری عمومی کشاورزی (شامل مناطق روستایی، شهری و غیرساکن) در سال ۱۳۶۷ توسط مرکز آمار ایران انجام شده است.
- بخش دوم، به معرّفی دو طرح جاری کشاورزی، زراعت و دام پرداخته است.

نقش و پایگاه اجتماعی-اقتصادی زن در جوامع پیشرفته و درحال‌توسعه با تکیه بر کار و تحصیل

مهناز محسن‌پور سی‌سی

مرکز آمار ایران

مطالعه و بررسی دربارهٔ وضع و موقعیت شغلی زنان نشان می‌دهد که در گذشته بیشترین کسانی که در علوم و صنایع پیشرفت کرده و سایر مقاماتی شده‌اند از مردان بوده‌اند و نام زنان مشهور در این زمینه کمتر به چشم می‌خورد.

- اشتغال: در مورد اشتغال زنان به‌طورکلی می‌توان گفت که امکان اشتغال بیشتر زنان در طی ۳۰ سال گذشته یعنی از ۱۳۳۵ تا ۱۳۶۵ در فعالیتهای مختلف چندان افزایش نیافته است. (یعنی افزایشی حدود ۴ درصد در طول ۲۰ سال اول و کاهشی در حدود ۲/۸ درصد در طول ۱۰ سال اخیر)
- تحصیلات: در مورد تحصیلات زنان به‌طورکلی می‌توان گفت از کل جمعیت باسواد ۶ ساله و بیشتر در کشور در سال ۱۳۵۵ معادل ۷/۳٪ باسواد و جمعیت ۶ ساله و بیشتر در سال ۱۳۳۵ در زنان معادل ۱۷/۹٪ و در سال ۱۳۵۵ به ۳۵/۵٪ و در سال ۱۳۶۵ به ۵۲/۱٪ رسیده که افزایش آن چشمگیر بوده است و لذا این مقایسه‌ها را در چند کشور انتخابی پیشرفته و درحال‌توسعه مقایسه‌هایی انجام شده که مورد توجه می‌باشد. بیکاری (که در کشورهای توسعه‌یافته و نیافته تحت عوامل متعددی متفاوت می‌باشد در ایران در طول سرشماریهای متعدد افزایش یافته خصوصاً این افزایش و اختلاف در بین زنان به علتی بیشتر بوده که به آن پرداخته شده است.

طرّاحی نظامهای آماری SIS به‌عنوان نمونه، نظام آماری شرکت سهامی شیلات ایران

علی‌اوسط مظفری

مرکز آمار ایران

در این مقاله در مورد عناوین زیر توضیح مختصری داده شده است:

- آشنایی با نظام آماری
- لزوم طرّاحی نظام آماری برای وزارتخانه‌ها، سازمانها، نهادها، ارگانها و فعالیتهای بزرگ که فاقد نظام آماری هستند.
- اثرات ناشی از طرّاحی یک نظام آماری مطلوب و بهره‌برداری مناسب از آن
- معرفی مرکز آمار ایران بعنوان طرّاح نظام آماری درکشور
- کلیات روش طرّاحی نظام آماری شرکت سهامی شیلات ایران
- نتایج بررسی وضع موجود شبکهٔ اطّلاعات و آمار شرکت سهامی شیلات ایران
- برخی از تعاریف و مفاهیم و معیارهای مورد استفاده در شیلات ایران
- روش شناخت نیازهای آماری شرکت سهامی شیلات ایران
- طرّاحی شبکهٔ مطلوب آمار و اطّلاعات شرکت سهامی شیلات ایران
- نتایج حاصل از اجرای طرح شناخت نیازهای آماری شرکت سهامی شیلات ایران
- چارچوب مورد استفاده در طرّاحی نظام آماری شرکت سهامی شیلات ایران

مقایسهٔ نسبت سواد جامعهٔ عشایری، روستایی و شهری کشور و بررسی علّت آن از دیدگاه آمار

میراحسان معاون اسلامی

مرکز آمار ایران

دراین بررسی نسبت درصد مردان مشغول به تحصیل در جامعهٔ شهری معادل ۸۰/۳۶ درصد، روستایی ۵۹/۹۵ درصد، و جامعهٔ عشایری ۳۸/۸ درصد است. در همین بررسی نسبت درصد زنان مشغول به تحصیل جامعهٔ شهری معادل ۶۵/۴۳ درصد، روستایی ۳۶/۳۳ درصد و جامعهٔ عشایری معادل ۱۴ درصد است. با عنایت به نسبتهای ارائه‌شده، به تفاوت فاحش موجود و شکاف ایجادشده در این سه جامعه مورد بررسی، می‌توان دلایل عدیده‌ای که باعث بوجود آمدن این تفاوت فاحش شده، از دیدگاه آماری مورد بررسی قرار داد. از جمله مواردی که می‌توان در این بررسی به صورت جامع با آن برخورد کرد، اشتغال زنان شاغل در جامعهٔ عشایری، شهری و روستایی است. برطبق نتایج سرشماری عمومی نفوس و مسکن سال ۱۳۶۵ و سرشماری عشایر کوچنده سال ۱۳۶۶، زنان شاغل در جامعهٔ عشایری ۳۶٪، در جامعهٔ روستایی معادل ۴۰/۵٪ و در جامعهٔ شهری ۱/۱۴٪، نشانگر فعالیت پایه‌ای زنان همراه مردان عشایری در آن جامعه است. عهده‌دار بودن کارهای سنگین خانوار، جهت کمک به گردش چرخ اقتصادی خانوار، مراقبت از کودکان، شیردوشی، قالی‌بافی، زیلو و گلیم بافی، تولید فرآورده‌های دامی، بافت چادرهای عشایری و غیره، فرصتی جهت تحصیل علی‌الخصوص به کودکان سنین ۶ساله و بالاتر را نمی‌دهد تا بتوانند فراغتی برای تحصیل داشته باشند این نسبت درجامعهٔ روستایی به مراتب کمتر است و به مقدار ۹ برابر کاهش نشان می‌دهد و در جامعهٔ شهری به میزان ۴ برابر این نسبت کمتر از جامعهٔ روستایی است.

بررسی وضعیّت مرگ‌ومیر در کشور برحسب جنس

مسعوده ملک‌پور

مرکز آمار ایران

براساس اطّلاعات موجود، زنان و مردان کشور در میزان مرگ‌ومیر، میزان تلفات و به تبع آن امید به زندگی تفاوت دارند. در این مقاله با استفاده از اطّلاعات سرشماری سال ۶۵ ، آمار تلفات ۲۴ شهر انتخابی وزارت بهداشت و درمان، و بررسیهای دموگرافیک دانشگاه شیراز این اختلاف در قالب شاخصهای فوق ارائه می‌شود.

میزان مرگ‌ومیر تصحیح نشده براساس اطّلاعات سرشماری ۶۵، برای مردان و زنان به ترتیب برابر ۳/۶ و ۲/۰ در هزار و میزان تصحیح‌شده برابر ۱۳/۴۶ و ۱۰/۶۸ در هزار بوده است.

میزان تلفات در ۲۴ شهر انتخابی کشور، در سال ۱۳۶۹ معادل ۳۰۵۴۶ زن در برابر ۴۵۸۳۳ مرد بوده است. این اختلاف در مورد هفده گروه بین‌المللی بیماری و همچنین در سال ۱۳۶۶ وجود داشته است.

امید به زندگی در مورد زنان و مردان نیز طبیعتاً متفاوت بوده است. به طوری که در جدول عمر منعکس است امید به زندگی در بدو تولّد برای زنان در کلّ کشور ۶۰/۹ سال و برای مردان ۵۵/۶ سال بوده است. در نقاط شهری به ترتیب ۶۴/۸۹ و ۵۹/۳ سال و در نقاط روستایی به ترتیب ۵۶/۶ و ۵۱/۹ سال بوده است.

این مقاله همچنین نظری گذرا بر اختلاف میزان مرگ‌ومیر زنان و مردان در سطح جهانی و برخی علل و عواقب آن داشته است.

نظری به سالخوردگی جمعیّت

احمدعلی منصوریان

مرکز آمار ایران

پیش از قرن بیستم امید زندگی در بدو تولّد در کشورهای جهان در سطح پائینی قرار داشت و تعداد افرادی که به سنین سالخوردگی می‌رسیدند اندک بود، این تعداد مختصر هم مسئله مهّمی برای اجتماع ایجاد نمی‌کرد. ولی در چند دههٔ اخیر تعداد سالخوردگان و سهم آنان در کلّ جمعیّت جهان به مقدار قابل ملاحظه‌ای افزایش یافته است. این افزایش نه تنها در کشورهای توسعه‌یافته حتّی در کشورهای در حال رشد هم مشاهده می‌شود. به‌طوری که در کشورهای توسعه‌یافته تعداد سالخوردگان با میزان رشد جمعیّت افزایش یافته است. آنچه مسلّم است این است که پدیدهٔ سالخوردگی در آینده‌ای نه چندان دور به‌عنوان یک مسئله در کشور ما هم بروز خواهد کرد. بنابراین لازم است که از هم‌اکنون در اندیشه حلّ این مسئله و ابعاد آن باشیم.

در این بررسی حجم سالخوردگان، تفاوتهای جنسی آنان، روشهای مطالعهٔ سالخوردگی جمعیّت و به‌طور کلّ مسائل جمعیّتی سالخوردگان از جمله مهاجرت، مرگ‌ومیر، وضع ازدواج، دین، وسعت خانوار در سالخوردگان و همچنین سالخوردگی جمعیّت در روستا مورد بحث و بررسی قرار گرفته است.

عشایر کوچنده در گذشته، حال و آینده

محمدصادق ناصح‌منش

مرکز آمار ایران

در این مقاله سعی شده است در مورد گذشته، حال و آیندهٔ عشایر کوچنده مطالبی به صورت خلاصه ارائه شود.

کوچ‌نشینی در ایران سابقهٔ کهن دارد. اغلب پادشاهان و سران بعد از اسلام از عشایر برده‌اند. انسان تا قبل از ۸۰۰۰ سال پیش که توانست منابع غذایی خود را در اختیار گیرد بدنبال شکار و جمع‌آوری دانه کوچ می‌کرد. براساس سرشماری سال ۱۳۶۶، ۹۶ ایل و ۵۴۷ طایفه با جمعیتی بحدود ۱۸۰ هزار خانوار و ۱۱۵۲ هزار نفر عشایر کوچنده داریم. نسبت جنسی جامعهٔ مذکور ۱۰۸ و ۵۱/۳ درصد در گروه سنی کمتر از ۱۵ ساله و باسوادان آنها ۲۷ درصد بوده است. نسبت باسوادی جامعهٔ شهری و روستایی کشور در سال ۱۳۶۵ حدود ۶۱/۸ درصد می‌باشد. خصوصیات عشایر کوچنده به‌ویژه شاخصهای جمعیتی آنها کاملاً متمایز از جامعهٔ شهری و روستایی کشور می‌باشد. جامعهٔ عشایری که ۴/۹ درصد جمعیت کشور هستند ۲۳/۲ درصد از کل گوسفند و بزه و ۳۰/۴ درصد از بز و بزغاله مملکت را پرورش می‌دهند. وضعیت عشایر کوچنده از نظر میزان سازگاری با شیوهٔ مرسوم زندگی خودش نه تنها از جایی به جای دیگر فرق می‌کند بلکه در یک ناحیه و حتی در یک طایفه از خانوار به خانوار دیگر نیز متفاوت می‌باشد. حضور روستاهای دامداری و مراکز بزرگ صنعتی در کنار آبادیهای کشاورزی برای اسکان عشایر پیشنهاد می‌شود.

مروری بر افزایش و رشد جمعیّت شهرهای کشور طیّ دورهٔ سی‌ساله (۱۳۶۵ـ۱۳۳۵)

محمداسماعیل ناصحی

مرکز آمار ایران

توازن جمعیّت در ایران ۲۶ سال جلوتر از جهان به سود جامعهٔ شهری بهم خورده است. این تغییر که رشد شتابان شهرنشینی را در کشورمان نشان می‌دهد، در همهٔ طبقات جمعیّتی یکسان نبوده است. برخی طبقات از نظر تعداد شهرها و برخی از نظر تعداد جمعیّت افزایش چشمگیر و در مقابل برخی کاهش داشته است. هدف این مقاله عمدتاً نشان دادن جابه‌جاییهای جمعیّت در طبقات عمدهٔ جمعیّت شهری و پیامدهای آن است. برای نیل به این هدف سرانجام جمعیّتی شهرها و رشد جمعیّت آنها در طبقات مختلف جمعیّتی طیّ سه دوره سرشماری در فاصلهٔ سالهای ۶۵ـ۱۳۳۵ مورد بررسی قرار گرفته است. این بررسی نشان می‌دهد که بخش عمدهٔ جمعیّت شهری (حدود ۷۰درصد) در معدودی (۸درصد) از شهرها ساکن بوده‌اند. در مقابل، ۱۲درصد جمعیّت شهری در ۷۰درصد شهرهای کشور سکونت داشته‌اند که بیانگر توزیع نامتعادل جمعیّت در مراکز جمعیّتی کشورمان است. مهاجرفرستی شهرهای کشور با تعداد جمعیّت آنها رابطه عکس دارد و جهت اصلی جابجاییهای جمعیّت در طبقات جمعیّت شهری، از شهرهای کوچکتر به سمت شهرهای بزرگتر است. این روند که منطقاً در پی توزیع نامتعادل امکانات آنها در گذشته به وجود آمده است، در آینده به قطبیتر شدن معدود مراکز پرجمعیت شهری خواهد انجامید که اگر در برنامه‌ریزیهای توسعه به آن توجّهی نشود نتیجه‌ای جز رویارو شدن با مشکلاتی از قبیل آنچه کلاً شهر تهران درگیر آن است نخواهد داشت.

بررسی وضع زناشویی با نگاهی به وضع سواد و فعّالیّت

مجید نوروزی

مرکز آمار ایران

زناشویی به لحاظ اثری که در زمینه‌های گوناگون زندگی اجتماعی و اقتصادی جامعه دارد از دیدگاههای مختلف اهمیّت یافته، مورد بررسی قرار می‌گیرد. این پدیده از طرفی باعث تغییر در متغیرهای جمعیّتی می‌شود و از طرف دیگر شرایط اجتماعی، اقتصادی جامعه اثر و نشان خود را در آن به خوبی نشان می‌دهد. در این مقاله سعی می‌شود با محاسبهٔ میزانها و شاخصهای گوناگون، قابلیت ازدواج و چگونگی وضع زناشویی جمعیّت در نقاط شهری و روستایی کشور نشان داده شود. به‌طور طبیعی بین میزان باروری و طول مدت زناشویی رابطه مستقیم و نزدیکی وجود دارد. با بالا رفتن سنّ ازدواج، باروری زن به سرعت رو به کاهش می‌گذارد و سنّ متوسّط ازدواج (به‌ویژه در زنان) در میزان باروری، نقش مهمّی را ایفا می‌کند. الگوی زناشویی در فاصلهٔ سالهای ۵۵ تا ۶۵ تفاوت چندانی نداشته و هنوز همان نسبت از زنان در معرض احتمال باروری قرار گرفته‌اند. در سال ۱۳۶۵، نسبت جمعیّت «هرگز ازدواج‌نکرده» در میان جمعیّت باسواد بیش از بی‌سوادان است. میانگین سنّ در اوّلین ازدواج برای مردان باسواد ۲۴/۲ و برای مردان بی‌سواد ۲۱/۴ سال و در جامعهٔ زنان به ترتیب ۲۰/۶ و ۱۸/۳ سال بوده است. در جمعیّت مردان شاغل ۱۰ ساله و بیشتر، عمومیّت ازدواج بیش از جمعیّت زنان شاغل بوده است. با بررسی و مقایسهٔ تعداد خانوار و مردان دارای همسر می‌توان اذعان کرد که بخشی از این وصلتها با اتکای اقتصادی به والدین صورت می‌پذیرد.

مصرف نهایی دولت و تشکیل سرمایهٔ ثابت ناخالص دولت در ساختمان و ماشین‌آلات و تجهیزات

هما نیکنام

مرکز آمار ایران

در این گزارش ابتدا مصرف نهایی دولت و سپس تشکیل سرمایهٔ دولت در ساختمان و ماشین‌آلات و تجهیزات عمده در سالهای ۱۳۴۱ الی ۶۹ ارائه شده است. شمول دولت که در این گزارش مورد بحث قرارگرفته است عبارت است از وزارتخانه‌ها و سازمانهای دولتی، سازمان تأمین اجتماعی و شهرداریها.

هزینهٔ مصرف نهایی خدمات دولتی برابر با ارزش کالاها و خدماتی است که آنها به منظور استفاده و مصرف خود در عملیات جاری تولید می‌کنند تولید ناخالص تولیدکنندگان خدمات دولتی یا مخارج تولید آنها یعنی مجموع مصارف واسطه، جبران خدمات کارکنان، مصرف سرمایهٔ ثابت و مالیاتهای غیرمستقیم برابر است.

تشکیل سرمایهٔ ثابت ناخالص دولت در ساختمان و ماشین‌آلات و تجهیزات عبارت است از افزایش دارایی دولت در مورد کالاهای ملموس قابل تولید مجدّد که عمر پیش‌بینی‌شده استفاده از آنها یک سال یا بیشتر باشد و شامل هزینه‌های سرمایه‌ای دولت در مورد ایجاد ساختمان، احیای زمین و ماشین‌آلات و تجهیزات عمده می‌باشد.

آگاهی، بینش و عملکرد زنان ۴۹ـ۱۵ ساله همسردار شهری ایران نسبت به برنامه تنظیم خانواده

قاسم یادگارفر*

دانشگاه علوم پزشکی اصفهان

حسین ملک افضلی

دانشگاه علوم پزشکی تهران

این بررسی در سال ۱۳۶۸ روی ۳۸۹۷ زن ۴۹ـ۱۵ ساله همسردار ساکن شهرهای کشور براساس نمونه‌گیری خوشه‌ای انجام گرفت. حجم نمونه در هر شهرستان متناسب با جمعیت آن شهرستان می‌باشد. هدف از انجام بررسی شناخت آگاهی، بینش و عملکرد زنان نسبت به برنامهٔ تنظیم خانواده است تا بر پایهٔ آن بتوان گامهای مؤثری در جهت توسعهٔ تنظیم خانواده و کنترل جمعیت برداشت. نتایج بررسی نشان می‌دهد که در ۲۵ سال گذشته همواره میانگین سنّ ازدواج رو به افزایش بوده بنحوی که میانگین ازدواج در زنانی که در ۵ سال اخیر ازدواج کرده‌اند به ۲۰ سال رسیده است. تعداد حاملگیها به شرط ثابت بودن سالهای ازدواج در باسوادان و شاغلین کمتر از بی‌سوادان و خانه‌دارها می‌باشد و در مجموع ۱۰ درصد از حاملگیها منجر به سقط و مرده‌زایی شده است. ۹۵ درصد خانمهائیکه بیش از ۵ سال پیش ازدواج کرده‌اند یکی از روشهای جلوگیری از حاملگی را می‌شناسند. این نسبت برای خانمهایی که در ۵ سال گذشته ازدواج کرده‌اند به ۸۳ درصد می‌رسد. خانمها ۹۷ درصد و شوهرانشان ۹۰ درصد با برنامهٔ تنظیم خانواده موافقند. در زمان بررسی ۶۴ درصد خانمها از یک روش جلوگیری از حاملگی استفاده می‌کرده‌اند و تنها نصف آنها یعنی حدود ۳۰ درصد از سه روش قرص، کاندوم، آی‌یودی استفاده می‌کرده‌اند. در بین استفاده‌کنندگان از این روشها درصد استفاده از قرص، کاندوم و آی‌یودی بترتیب ۶۰، ۲۵ و ۱۵ درصد است. دسترسی به قرص آسان، لیکن در تهیهٔ کاندوم و آی‌یودی مشکلاتی وجود دارد.

توجه بیشتر به سوادآموزی زنان، اشتغال خانمها، آموزش مردان در زمینهٔ تنظیم خانواده، گنجاندن مطالب مربوط به تنظیم خانواده درکتب درسی دبیرستانی و دانشگاهی، استفاده از آموزشیاران نهضت سوادآموزی در جهت رساندن پیام تنظیم خانواده و کنترل جمعیت به اقصی نقاط کشور، آموزش پرسنل پزشکی جهت معرفی روش مناسب به مراجعین، توجّه به رفاه سالمندان، رواج راههای آسان جلوگیری مانند لوله‌بستن در مردان و زنان و آزمایش روشهای جدید از توصیه‌های این بررسی است.

تجزیه و تحلیل داده‌های آماری نمرات دانش‌آموزان در درسهای علوم ریاضی

هدایت یاسائی میبدی

دانشگاه صنعتی شریف

در مسئله آموزش و پرورش به‌طور کلی یک مسئله مهم این است: تقسیم کلاس و درسهای تکراری چگونه برگزار شود؟ آیا نمرات دانش‌آموزان تحت تأثیر نوع تقسیم درسها قرار می‌گیرند؟ آیا درسهایی مانند ادبیات و دینی و علوم و ریاضیات هریک در نمرات دانش آموزان نفوذ دارند؟

برای پاسخ به این سؤالات در چند دبستان و دبیرستان شهرهایی مانند تهران در دو منطقه و کاشان در سه ناحیه طرحهای آزمایشی برگزار شد. آمار و نمرات امتحان دانش آموزان در دو نوبت امتحان معلّم و توسط خود ما تهیه شد. سپس با طرحهای آزمایش یک‌طرفه و دوطرفه و آزمون فرضها نتیجه گرفتیم. بعضی درسها هریک به تنهایی در نمرات دانش آموزان مؤثر هستند امّا در جمع این درسها در نمرات دانش‌آموزان اثر ندارند. بعضی درسها اثر منفی دارند برخی دیگر اثر مثبت دارند لکن در مجموع بی‌اثر هستند. آمارهای خیلی زیاد جمع‌آوری کرده‌ایم. قسمتی از اینها را ارزشیابی کرده‌ایم و بقیه در دست بررسی است که در این مورد ان شاءا... در آینده برایشان جواب کامل خواهیم داشت.

بررسی آمار انرژی (برق)

افسانه یزدچی

مرکز آمار ایران

صنعت برق به‌عنوان یکی از اساسیترین عوامل تولید مطرح بوده و نقشی که برق در فعالیتهای اقتصادی به‌طور مستقیم و یا غیرمستقیم ایفا می‌کند مشهود است، لذا در این مقاله سعی شده است که آمارهای مربوط به نیروگاههای تولید، ظرفیت، تعداد مشترکین، فروش به تفکیک نوع مصرف‌کننده، تعداد روستاهای برقدارشده و خطوط انتقال نیرو در سالهای اخیر مورد بررسی قرار گیرد.

Farsi Papers and Posters

On Robustness Property of MEP.1R Plans

H. Talebi
Department of Mathematics
University of Isfahan
Isfahan, Iran

Wed. 15:30 Hall 6

The search linear model is given by

$$E(y) = X_1\beta_1 + X_2\beta_2\ ,\ V(y) = \sigma^2 I.$$

Let β_1 be the vector of the general mean and main effects and β_2 be the vector of two and three factor interactions for 2^m factorial experiments. Main effect plans that allow search and estimation of one non zero element of β_2 are called MEP.1R plans.

Ghosh and Talebi (1990) developed some necessary and sufficient conditions for a plan to be an MEP.1R plan and obtained series of MEP.1R plans for $m \geq 4$, where m is the number of factors. As Ghosh (1979), Ghosh and Talebi (1990) defined. "An MEP.1R plan is said to be the robust against unavailability of any s (a positive integer) runs if the resulting plan after deletion of s runs remain an MEP.1R plan." In this work we show that MEP.1R plans are obtained by Ghosh and Talebi are all robust designs.

A Generalizable Formulation of Conditional Logit with Diagnostics

Ehsan S. Soofi
School of Business Administration
University of Wisconsin
Milwaukee, U. S. A.

Tue. 14:30 Hall 6

The conditional logit model is a multinomial logit model that permits the inclusion of choice specific attributes. This article shows that the conditional logit model will maximize entropy given a set of attribute value preserving constraints. A correspondence between the maximum entropy (ME) and maximum likelihood (ML) estimates for logit probabilities is established. Some easily computable and useful diagnostics for logit analysis are provided, and it is shown that an evaluation of the relative importance of attributes can be made using the ME formulation. The ME formulation is also generalized to accommodate initial choice probabilities into the logit model. An example is given.

References

[1] Brockwell, P.J. and Davis, R.A. (1987) *Time Series: Theory and Methods*. New York: Springer-Verlag.

[2] Cambanis, S. and Miller, G. (1981) Linear problems in pth order and stable processes. *SIAM J. Appl. Math.* 41, 43-69.

[3] Cambanis, S. and Soltani, A.R. (1982) Prediction of stable processes: Spectral and moving average representations. Tech. Rpt. 11, Center for Stochastic Processes, University of North Carolina, Chapel Hill.

[4] Cline, D.B.H. and Brockwell, P.J. (1986) Linear prediction of ARMA processes with infinite variance. *Stoch. Proc. & Appl.*, 19, 281-296.

[5] Kanter, M. (1975) The L^p norm of sums of translates of a function. *Trans. Amer. Soc.*, 207, 35-47.

[6] Kuelbs, J. (1973) A representation theorem for symmetric stable processes and stable measures on H. *Zeitschrift Wahrsch. verw. Geb.*,26, 259-271.

[7] Paulauskas, V.J. (1976) Some remarks on multivariate stable distributions. *J. Multivariate Anal.*, 6, 356-368.

[8] Samorodnitsky, S. and Taqqu, M.S. (1991) Conditional moments and linear regression for stable random variables. *Stoch. Proc. & Appl.*, 39, 183-199.

[9] Soltani, A.R. (1991) On spectral representation of multivariate stable processes. To appear in *Theory of probability and its Applications*.

[10] Soltani, A.R. (1992) More on conditional moments of stable random variables. Tech. Rpt., Department of Statistics, Shiraz University, Shiraz Iran.

The following theorem is a consequence of the Theorem 4.3, which provides a recursive formula for doing the prediction.

4.13 Theorem Let $X_t, t = 0, \pm 1, \cdots$ be a n-variate $S\alpha S$ ARMA(p,q) process given by (4.1), then based on the infinite past $X_m, X_{m-1}, \cdots$ the future value X_{m+k} has a unique minimum dispersion linear predictor $X^*_{m+k}, k \geq 1$, which is given by

$$X^*_{m+k} = \sum_{j=0}^{\infty} \Psi_{k+j} Z_{m-j}\ , \ k \geq 1, \tag{4.13}$$

and satisfies the recursive relation

$$X^*_{m+k} = \sum_{j=1}^{k-1} \Pi_j X^*_{m+k-j} + \sum_{j=k}^{\infty} \Pi_j X_{m+k-j}\ , \tag{4.14}$$

where Π_j are $n \times n$ matrices given by 4.6.

Proof: Use the linearity of the P_m.

where B is the backward shift operator. But note that

$$\eta(B)\Theta(B) = \sum_{j=0}^{\infty} \beta_j B^j$$

with

$$\beta_j = \begin{cases} \sum_{l=0}^{j} \Theta_l \eta_{j-l} & \text{if } j < q \\ \sum_{l=0}^{q} \Theta_l \eta_{j-l} & \text{if } j \geq q. \end{cases}$$

Thus

$$\begin{aligned} \sum_{j=1}^{\infty} \|\beta_j\|^a &\leq M \sum_{j=1}^{q} \sum_{l=0}^{j} j \|\eta_l\|^a + Mq \sum_{j=q}^{\infty} \sum_{l=j-q}^{j} \|\eta_l\|^a \\ &\leq Mq^2 (\sum_{l=0}^{q} \|\eta_l\|^a + \sum_{l=0}^{\infty} \|\eta_l\|^a) < \infty , \end{aligned}$$

where M is an upper bound for the absolute value of the entries of all matrices Θ_j, $j = 0, \cdots, q$. Now by applying Lemma 3.1 we obtain that X_t is a $S\alpha S$ vector, $t = 0, \pm 1, \cdots$. Clearly (1.2) is satisfied with X_t given by (4.9).

Uniquesness: As discussed by Brockwell and Davis(1985) page 85 all we need is to conclude from

$$\sum_{j=0}^{q} \Theta_j z_{t-j} = \sum_{j=0}^{\infty} \xi_j z_{t-j} \ , \text{ with } \ \sum \|\xi_j\|^a < \infty \ , \tag{4.10}$$

that

$$\Theta_j = \xi_j \ , \ j = 0, \cdots, q, \ \xi_j = 0, \ j > q \ .$$

It follows from(4.10) and Lemma 3.1 that

$$\sum_{j=q+1}^{\infty} \xi_j z_{t-j} \ , \ \sum_{j=0}^{q} (\Theta_j - \xi_j) z_{t-j}$$

are $S\alpha S$ random vectors which are independent. Thus

$$\begin{aligned} \sum_{j=0}^{q} (\Theta_j - \xi_j) Z_{t-j} &= 0 \\ \sum_{q+1}^{\infty} \xi_j Z_{t-j} &= 0 \ . \end{aligned} \tag{4.11}$$

Since $\{Z_t\}$ is an i.i.d. sequence, it follows from(4.11) and (1.2) that

$$\begin{aligned} \int_{S_{n-1}} |r.(\Theta_j - \xi_j)s|^a \, \Gamma(ds) &= 0 \ , \ \forall r \in R^n, j = 0, \cdots, q, \\ \int_{S_{n-1}} |r.\xi_j s|^a \, \Gamma(ds) &= 0 \ , \ \forall r \in R^n, j > q \ . \end{aligned} \tag{4.12}$$

Since Θ_j, ξ_j are full rank it follows very easily from (4.12) that

$$\Theta_j = \xi_j, \ j = 0, \cdots, q, \text{ and } \xi_j = 0, \ j > q \ .$$

The proof is complete.

The following theorem provides a unique solution for (4.1).

4.3 Theorem If $\det \Phi(z) \neq 0$ and $\det \Theta(z) \neq 0$ for all $z \in C$ such that $|z| \leq 1$, then (4.1) has exactly one solution,

$$X_t = \sum_{j=0}^{\infty} \Psi_j Z_{t-j} , \tag{4.4}$$

where the matrices Ψ_j are determined uniquely by

$$\sum_{j=0}^{\infty} \Psi_j z^j = \Phi^{-1}(z)\Theta(z) , \ |z| \leq 1 . \tag{4.5}$$

Furthermore the process Z_t will be recovered from the process X_t as

$$Z_t = \sum_{j=0}^{\infty} \Pi_j X_{t-j} , \tag{4.6}$$

where the matrices Π_j are determined uniquely by

$$\sum_{j=0}^{\infty} \Pi_j z^j = \Theta^{-1}(z)\Phi(z) , \ |z| \leq 1 , \tag{4.7}$$

and

$$\sum_{j=0}^{\infty} \|\Psi_j\|^a < \infty , \ \sum_{j=0}^{\infty} \|\Pi_j\|^a < \infty \qquad 0 < a \leq 2.$$

Proof: The existence part is very similar to the Gaussian case. Indeed by an argument as in Theorem 11.3.1 in Brockwell and Davis(1985) we obtain that under the assumption on $\det \Phi(z)$, the $\Phi^{-1}(z)$ has the power series expansion

$$\Phi^{-1}(z) = \sum_{j=0}^{\infty} \eta_j z^j = \eta(B)$$

where η_j are $n \times n$ matrices with

$$\sum_{j=0}^{\infty} \|\eta_j\|^a < \infty , \ 0 < a \leq 2 . \tag{4.8}$$

Now let

$$X_t = \eta(B)\Theta(B)Z_t , \ t = 0, \pm 1, \cdots \tag{4.9}$$

for each $Y \in \mathcal{L}_1 \oplus \mathcal{L}_2$.

Proof: Since (3.7) follows from Lemma (2.8), all we need is to show that P_∞ is linear. Let $Y^1 = Y_1^1 + Y_2^1$, $Y^2 = Y_1^2 + Y_2^2$ with $Y_1^1, Y_1^2 \in \mathcal{L}_1$, $Y_2^1, Y_2^2 \in \mathcal{L}_2$. Then

$$AY^1 + BY^2 = (AY_1^1 + BY_1^2) + (AY_2^1 + BY_2^2) ,$$

where A and B are $n \times n$ matrices. Clearly $AY_1^1 + BY_1^2 \in \mathcal{L}_1$ and $AY_2^1 + BY_2^2 \in \mathcal{L}_2$. Thus

$$\begin{aligned} P_\infty(AY^1 + BY^2) &= AY_2^1 + BY_2^2 \\ &= AP_\infty Y^1 + BP_\infty Y^2 , \end{aligned}$$

given the proof.

The Theorem 3.5 leads us to the following prediction procedure.

3.8 The Prediction Procedure. Let $\{Z_t\}_{t=-\infty}^{+\infty}$ be a sequence of independent $S\alpha S$ random vectors in R^n. Suppose the class A is as in (3.1) and $\mathcal{L}_1$, $\mathcal{L}_2$ are two independent subclasses of A . If

$$Y = Y_1 + Y_2 , \; Y_1 \in \mathcal{L}_1 , \; Y_2 \in \mathcal{L}_2 ,$$

then based on $\mathcal{L}_2$ the best linear predictor with minimum dispersion error of Y is Y_2 and the dispersion error is $disp(Y_1)$.

4 Applications

The prediction procedure given in section three have certain applications in predicting the future values of processes which are generated by independent sequences. Among these are ARMA models. A $S\alpha S$ process $X_t, t = 0, \pm 1, \cdots$ with values in R^n is an n-variate ARMA(p,q) if it satisfies the difference equation

$$X_t - \Phi_1 X_{t-1} - \cdots - \Phi_p X_{t-p} = Z_t + \Theta_1 Z_{t-1} + \cdots + \Theta_q Z_{t-q} , \qquad (4.1)$$

where $\Phi_1, \cdots, \Phi_p, \Theta_1, \cdots, \Theta_q$ are $n \times n$ matrices and $Z_t, t = 0, \pm 1, \cdots$ is an independent n-variate $S\alpha S$ process for which each Z_t , $t = 0, \pm 1, \cdots$ is a n-variate $S\alpha S$ random vector with the same spectral measure Γ. (4.1) can be written in the compact form

$$\Phi(B)X_t = \Theta(B)Z_t , \qquad (4.2)$$

with matrix-valued polynomials $\Phi(z) = I - \Phi_1 z - \cdots - \Phi_p z^p$ and $\Theta(z) = I + \Theta_1 z + \cdots + \Theta_q z^q$ on the complex plane; where I is the $n \times n$ identity matrix and B is the backward shift operator.

where the closure is w.r.t. convergence in probability. We have the following lemma.

3.2 Lemma . Let A_j be a sequence of $n \times n$ matrices and Γ_{t_j} be the spectral measure of $Z_{t_j}, j = 1, 2, \cdots$ for which

$$\sum_{j=1}^{\infty} \int_{S_{n-1}} \|A_j s\|^{\alpha} \Gamma(ds) < \infty \tag{3.3}$$

then $\sum_{j=1}^{\infty} A_j Z_{t_j}$ is a $S\alpha S$ random vector in A, and

$$disp(\sum_{j=1}^{\infty} A_j Z_{t_j}) = \sum_{j=1}^{\infty} \int_{S_{n-1}} \|A_j s\|^{\alpha} \Gamma_{t_j}(ds) . \tag{3.4}$$

Proof: Let $Y_{kl} = \sum_{j=k}^{l} A_j Z_{t_j}$, then since Z_{t_j} , $j = 1, 2, \cdots$ are independent we have

$$\begin{aligned} E \exp(ir.Y_{kl}) &= \exp(-\textstyle\sum_{j=k}^{l} \int_{S_{n-1}} |A_j' r.\nu|^{\alpha} \Gamma_{t_j}(d\nu)) \ , r \in R^n, \\ &= \exp(-\textstyle\sum_{j=k}^{l} \int_{S_{n-1}} |r.A_j \nu|^{\alpha} \Gamma_{t_j}(ds)), \end{aligned}$$

where A' is the transpose of A. But for each fixed r,

$$\sum_{j=k}^{l} \int_{S_{n-1}} |r.A\nu|^{\alpha} \Gamma_{t_j}(dr) \le \|r\|^{\alpha} \sum_{j=k}^{l} \int_{S_{n-1}} \|A_j \nu\|^{\alpha} \Gamma_{t_j}(d\nu) .$$

Thus under (3.3) $\sum_{j=1}^{l} A_j Z_{t_j}$ is fundamental in probability. Therefore it converges in probability and it is well known that the limiting value, $\sum_{j=1}^{\infty} A_j Z_{t_j}$, is a $S\alpha S$ random vector which clearly lies in A, see Kuelbs(1973). For (3.4) note that by part (b) of Lemma 2.6 and (2.4),

$$disp(\sum_{j=1}^{n} A_j Z_{t_j}) = \sum_{j=1}^{n} \int \|A_j s\|^{\alpha} \Gamma_{t_j}(ds) .$$

Now apply Lemma 2.6(c). The proof is complete.

The following theorem is for the existence of a linear predictor with minimum dispersion error.

3.5 Theorem. Suppose A is as (3.1) and $\mathcal{L}_1$ and $\mathcal{L}_2$ are two independent subspaces of A. Let $\mathcal{L}_1 \oplus \mathcal{L}_2 = \{Y_1 + Y_2 : Y_1 \in \mathcal{L}_1 , Y_2 \in \mathcal{L}_2\}$. Define $P_{\infty} : \mathcal{L}_1 \oplus \mathcal{L}_2 \to \mathcal{L}_2$ by

$$P_{\infty}(Y_1 + Y_2) = Y_2 , \tag{3.6}$$

Then P_{∞} is linear and

$$disp(Y - P_{\infty}Y) = \min\{disp(Y - X) : X \in \mathcal{L}_2\}, \tag{3.7}$$

(c) If Y_j converges to Y in probability, as $j \to \infty$, then

$$disp(Y_n) \to disp(Y)\ , as\ j \to \infty\ .$$

Proof

$$\begin{aligned} -\int_{S_{n-1}} \log \phi(t)\omega(dt) &= \int_{S_{n-1}} \int_{S_{n-1}} |\,t.s\,|^\alpha\, \Gamma(ds)\omega(dt) \\ &= \int_{S_{n-1}} \int_{S_{n-1}} |\,t.s\,|^\alpha\, \omega(dt)\Gamma(ds) \\ &= c\int_{S_{n-1}} \Gamma(ds) = c\Gamma(S_{n-1}) \end{aligned}$$

because the inner integral is independent of s. This gives (a). For (b) apply (2.7). Part (c) also follows from (2.7) and the fact that $\phi_n(t) \to \phi(t)$ uniformly on bounded sets.

The following lemma plays an essential role in our prediction procedure. Its idea is inherited from Lemma 2.1 given by Cline and Brockwell(1985).

2.8 Lemma. Let $Y = Z_1 + Z_2$ with Z_1 , Z_2 two independent $S\alpha S$ vectors in R^n. Let $P_\infty Y$ be a $S\alpha S$ vector which is independent of Z_1 and minimizes $disp(Y - \hat{Y})$ among all $S\alpha S$ random vectors $\hat{Y}$ which are independent of Z_1, then $P_\infty Y = Z_2$ and $disp(Y - P_\infty Y) = disp(Z_2)$.

Proof: Let ϕ_X denote the characteristic function of the $S\alpha S$ random vector X, then it follows from assumptions that

$$\phi_{Y-\hat{Y}}(t) = \phi_{Z_1}(t)\phi_{Z_2-\hat{Y}}(t)\ ,\ t \in R^n$$

thus it follows from (2.7) that

$$\begin{aligned} disp(Y - \hat{Y}) &= -c\int_{S_{n-1}} (\log \phi_{Z_1}(t) + \log \phi_{Z_2-\hat{Y}}(t))\omega(dt) \\ &= -c\int_{S_{n-1}} \log \phi_{Z_1}(t)\omega(dt) - c\int_{S_{n-1}} \log \phi_{Z_2-\hat{Y}}(t)\omega(dt) \\ &= disp(Z_1) + disp(Z_2 - \hat{Y})\ . \end{aligned}$$

Therefore $P_\infty Y = Z_2$ and $disp(Y - P_\infty Y) = disp(Z_2) \le disp(Y - \hat{Y})$ for other $\hat{Y}$. The proof is complete.

3 Prediction

In this section we are concerned with the existence of the best linear predictor, the one which minimizes the dispersion. To be more precise let $\{Z_t\}_{t=-\infty}^{\infty}$ be a sequence of independent $S\alpha S$ random vectors in R^n. Let A be the closed linear span of elements Z_t with $n \times n$ matrices as the scalar field, i.e.

$$A = sp\{AZ_t,\ A : n \times n\ \text{matrices}, t = 0, \pm 1, \pm 2, \cdots\} \tag{3.1}$$

(i) $\Gamma(S_{n-1})^{1/\alpha} \leq 2c(n/m)^{1/\alpha}\frac{d(Y)}{1-d(Y)}$

(ii) $d(Y) \leq 2n^2h^{-1}(M^{1/\alpha}\Gamma(S_{n-1})^{1/\alpha})$,

where $m = \min \Sigma_{i=1}^n \mid s_i \mid^\alpha$ on S_{n-1}, $M = \max \Sigma_{i=1}^n \mid s_i \mid^\alpha$ on S_{n-1}, $c = \int_{-\infty}^{\infty} e^{-|y|^\alpha} d(y)$, $h(x) = x|\log(1-x)|^{1/\alpha}$, $0 < x < 1$, and $d(Y) = \inf\{\epsilon > 0; P(\mid Y \mid > \epsilon) < \epsilon\}$ is the Lévy distance of Y to the origin.

It follows from Theorem 2.1 that $d(Y) \to 0$ if and only if $\Gamma(S_{n-1})$, the total variation of Γ, tends to zero. This leads us to the following definition.

2.2 Definition. Let Y be a $S\alpha S$ random vector with spectral measure Γ, $0 < \alpha < 2$, then the dispersion of Y (relative to the origin) is defined to be $\Gamma(S_{n-1})$:

$$disp(Y) = \Gamma(S_{n-1}) . \tag{2.3}$$

An application of the following theorem of Soltani(1992) gives the $disp(AY)$ for a $S\alpha S$ random vector Y and a $n \times n$ matrix A namely,

$$disp(AY) = \int_{S_{n-1}} \|As\|^\alpha \Gamma(ds) . \tag{2.4}$$

2.5 Theorem. Let $Y = (Y_1, \cdots, Y_n)'$ be a $S\alpha S$ random vector with spectral measure Γ, $0 < \alpha < 2$, and let A be a $n \times n$ matrix, then

$$\Gamma_{AY}(E) = \int_{T^{-1}(E)} \|As\|^\alpha \Gamma(ds) \quad , E \in B(S_{n-1}),$$

where Γ_{AY} is the spectral measure of AY and $T : S_{n-1} \to S_{n-1}$;

$$T(s) = \left(\frac{(As)_1}{\|As\|^\alpha}, \cdots, \frac{(As)_n}{\|As\|^\alpha}\right).$$

The following lemma gives an alternative expression for $disp(Y)$ which will be used in subsequent results.

2.6 Lemma. (a) Let Y be a $S\alpha S$ random vector with spectral measure Γ, $0 < \alpha < 2$, then

$$disp(Y) = -c \int_{S_{n-1}} \log \phi(t) \omega(dt) \tag{2.7}$$

where $c = \int_{S_{n-1}} \mid t.s \mid^\alpha \omega(dt)$, ω is the Lebesgue measure on S_{n-1}, and ϕ is the characteristic function of Y given by (1.2).

(b) If Y_1 and Y_2 are independent then

$$disp(Y_1 + Y_2) = disp(Y_1) + disp(Y_2).$$

method for multivariate stable processes. The first step, in the multivariate case, is to define a dispersion, as its definition in the univariate case is very restricted to that case. In section two we will define a dispersion for stable random vectors. In section three, by following the line of Cline and Brockwell(1986) we develop the ingredients of our prediction technique. In section four we apply our method to certain time series models. Before closing this section let us introduce some notions and notations.

A random variable X is a symmetric α stable if its characteristic function is of the form

$$\phi_X(t) = \exp(-c|t|^\alpha) \quad , t \in R, \tag{1.1}$$

where R is the set of real numbers. A random vector $X = (X_1, ..., X_n)$ is symmetric α stable ($S\alpha S$) if every linear combination of its components is a $S\alpha S$ random variable. The characteristic function of a $S\alpha S$ random vector is specified by a symmetric measure on S_{n-1}, the unit circle in the Euclidian space R^n, namely,

$$\phi_X(t) = \exp\{-\int_{S_{n-1}} |t.s|^\alpha \Gamma(ds)\} \quad , t \in R^n, \tag{1.2}$$

where $t.s = \sum_{i=1}^n t_i s_i$ Kuelbs(1973). The measure Γ in the case that $0 < \alpha < 2$ is unique Kanter(1975). Since the measure Γ uniquely specifies the characteristic function ϕ_X it plays an essential role in studying the statistical behaviour of the random vector X. Some of such considerations are included in the works of Paulauskas(1976), Cambanis and Miller(1981) and Samorodnitsky and Taqqu(1991).

2 A Dispersion for Stable Random Vectors

The criterion of minimizing error dispersion in predicting a random quantity Y rests on finding a predictor $\hat{Y}$ that among all possible predictors of Y minimizes the probabilities of large deviations $P(|Y - \hat{Y}| > k)$ for every k. In the case that the error $Y - \hat{Y}$ is a $S\alpha S$ random variable its characteristic function is given by (1.1) and is uniquely specified by the scale parameter c, and the minimization is equivalent to the minimization of the quantity c. This motivates to define the dispersion of Y (relative to the origin) to be its scale parameter c. Of course the way that dispersion is defined for stable random variables can not be generalized to stable random vectors, as the law of a $S\alpha S$ random vector is specified by its spectral measure. But note that the Lèvy distance of a $S\alpha S$ random variable Y, $d(Y)$, converges to zero as its scale parameter $c \to 0$. This and the following theorem of Soltani (1991) are our motivations for our definition of dispersion of a random vector,

2.1 Theorem Let $Y = (Y_1, ..., Y_n)$ be a $S\alpha S$ random vector, $0 < \alpha < 2$ with spectral measure Γ, then for $0 < \alpha < 2$,

On dispersion of stable random vectors and its application in the prediction of multivariate stable processes

A. Reza Soltani
Shiraz University, Iran
and
R. Mocanoddin
Shiraz University, Iran

Abstract

Our aim in this article is to derive a recipe formula for the best linear predictor when the underlying stochastic model is a multivariate symmetric α stable process. For doing this first we introduce a dispersion for stable random vectors, then adopt the technique of minimizing the dispersion error for deriving our recipe formula in various cases.

Keywords: *Best Linear Predictor; Multivariate Stable Processes; Spectral Measure; Symmetric α Stable ARMA Processes.*

1 Introduction

The well known method of minimizing the expected square error in the prediction theory of second order processes fails when one is dealing with the strictly stable processes, as the elements of a strictly stable process can be viewed as the elements of a certain L^p space, $p < 2$. Two methods have been examined in predicting the future values of a univariate stable sequence which are based on minimizing the error dispersion. One is the minimization of a certain L^p distance employed by Cambanis and Soltani (1982), the second is to minimize the error dispersion in time domain, introduced by Stuck (1978) and applied to ARMA(p,q) stable processes by Cline and Brockwell(1985). The work of Cline and Brockwell(1985) provides a time domain technique while the work of Cambanis and Soltani(1982) gives a spectral domain technique in the prediction theory of multivariate stable processes. In this article we are concerned with the time domain error dispersion

statistics based on these samples, respectively, then for a given reliability R_Y of the distribution of Y, the R-Reliable interval for Y is $(0, T_Y)$ and the R-Reliable interval for Z is $(0, T_Z)$ for a given reliability R_Z of the distribution of Z.

One-Sided R-Reliable Intervals and Their Associated Confidence on the Sum of Two Random Variables

Mohammad R. Kafai
Mathematics Department
San Francisco State University
San Francisco, U. S. A.
Hamid Shahnasser*
Electric Engineering Department
San Francisco State University
San Francisco, U. S. A.

Thur. 15:30 Hall 6

Let L_Y and U_Y be statistics based on a random sample of size m from the distribution of random variable Y with probability density function (pdf) $f(y)$. The interval (L_Y, U_Y) is called the tolerance interval and the Neyman-Pearson confidence is defined as

$$C_Y(R) = P_{L_Y,U_Y}\left[\int_{L_Y}^{U_Y} f(y)dy \geq R\right].$$

R is in the interval $[0,1]$ and may be interpreted as the reliability of the distribution of Y. For this reason the above tolerance interval will be referred to as the R-Reliable interval. In this study the more convenient notation

$$C_Y(R) = P_{L_Y,U_Y}[P(L_Y \leq Y \leq U_Y) \geq R]$$

will be used.

Suppose Y and Z are two random independent, continuous, random variables defined on $(0,\infty)$. If $Y_1, Y_2, \cdots, Y_m$ is a random sample from the distribution of Y and $Z_1, Z_2, \cdots, Z_n$ is random sample from the distribution of Z; and if T_Y and T_Z are

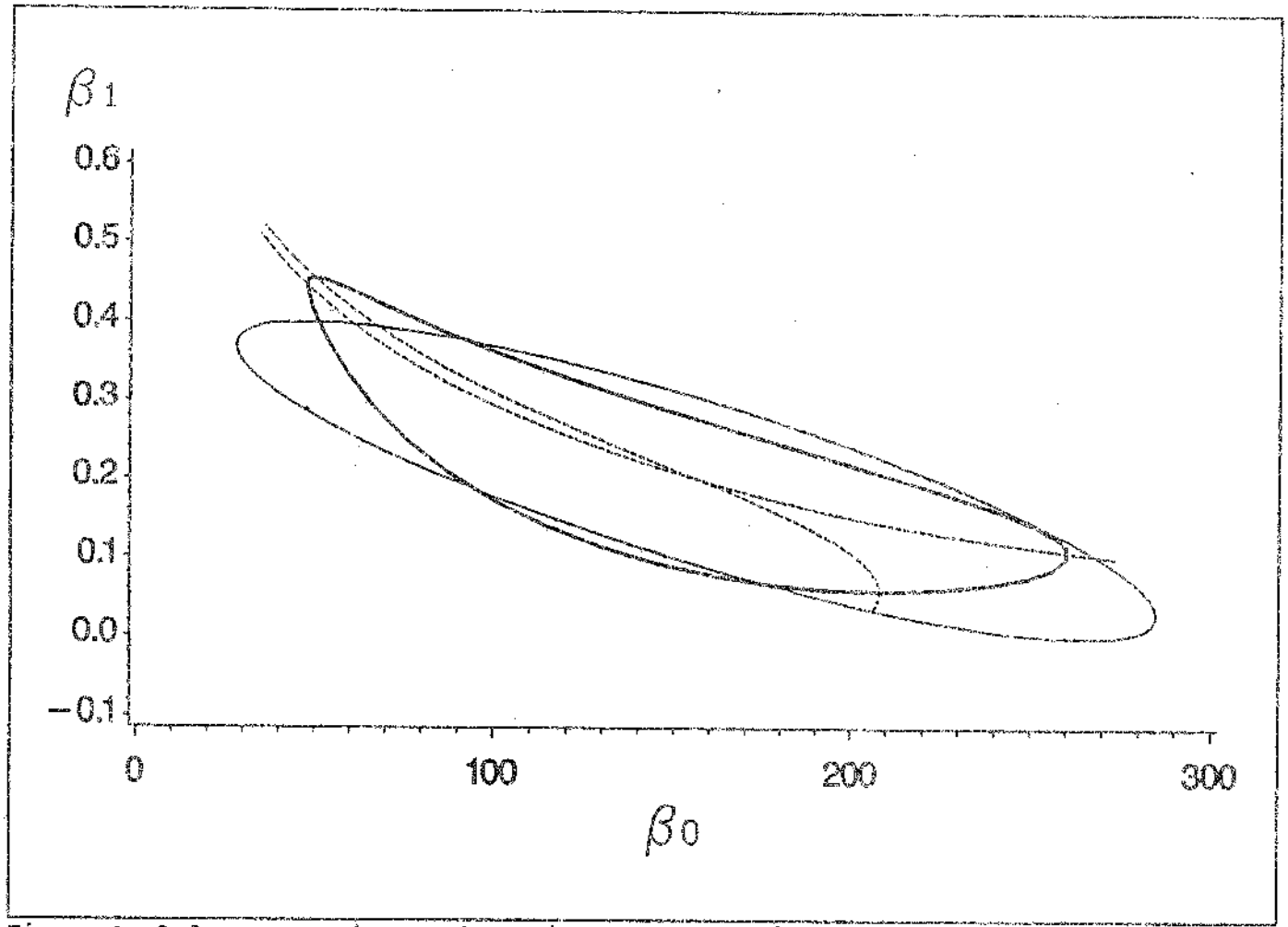


Figure 2 Inference regions and profile traces for β_0 and β_1.

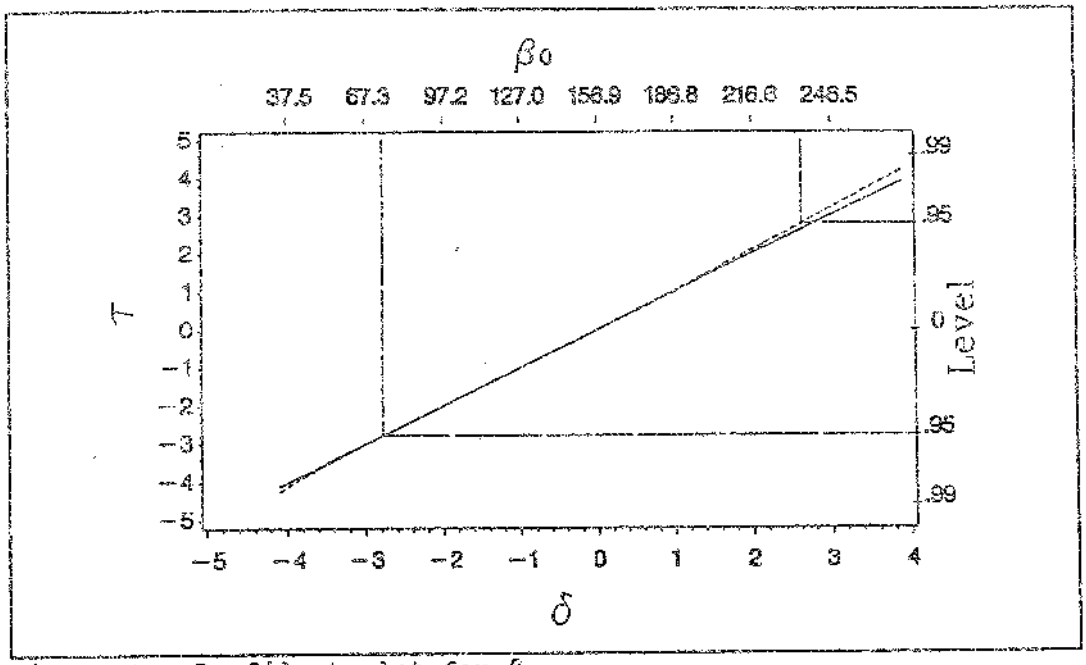


Figure 1a Profile t plot for β_0.

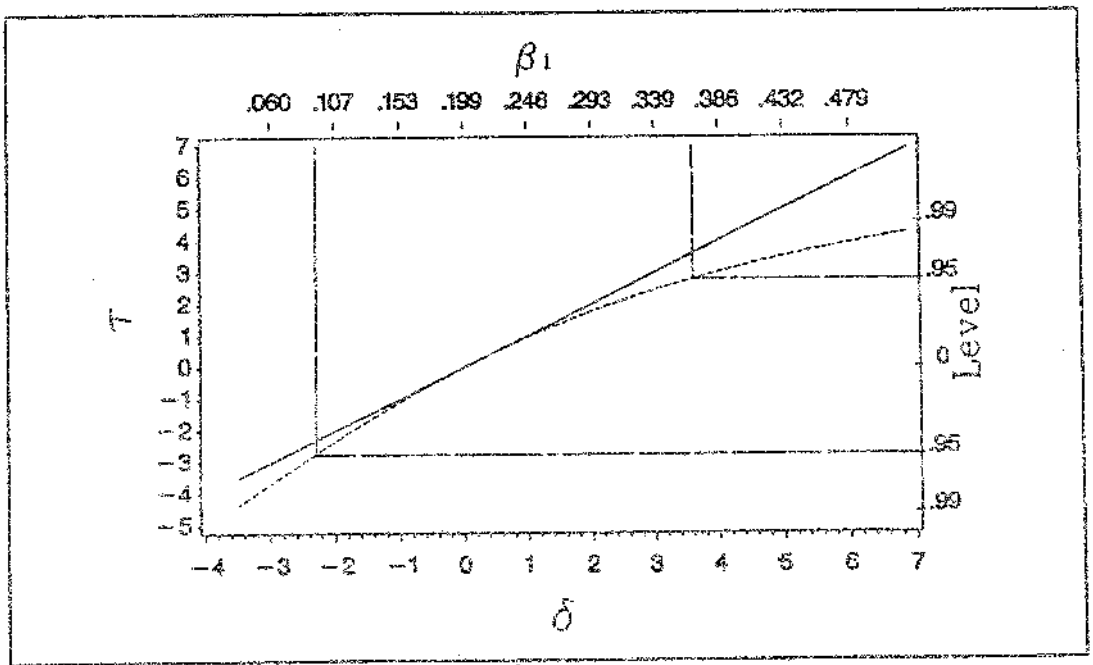


Figure 1b Profile t plot for β_1.

```
  END;
  FIT=FIT(|2:NROW(FIT),1|);
  FREE IND VAL YF;

FINISH;
```

REFERENCES

Bates, D.M., and D.G. Watts. 1988. Nonlinear Regression Analysis and its Applications. John Wiley and Sons. New York.

Bliss, C.I., and A.T. James. 1966. Fitting the rectangular hyperbola. Biometrics. 22:573-602.

Press, W.H., B.P. Flannery, S.A. Teukolsky, W.T. Vetterling. 1988. Numerical Recipes in C, The Art of Scientific Computing. Cambridge University Press. New York.

SAS Institute Inc. 1989. SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2, SAS Institute Inc., Cary, NC.

SAS Institute Inc. 1990. SAS/IML Software: Reference, Version 6, First Edition, SAS Institute Inc., Cary, NC.

SAS Institute Inc. 1990. SAS/GRAPH Software: Reference, Version 6, First Edition, Volume 2, SAS Institute Inc., Cary, NC.

```
      Y2(|I,1|)=Y2(|I,1|)*Y2(|I+1,1|)+W(|I,1|);
      Y2(|I,1|)=Y2(|I,1|)/2;
   END;

   FREE W SIG P I;
FINISH;


START SPLINE(X, Y, Y2, VAL, YF);
   IF VAL > MAX(X) THEN
          DO;
              YF=.;
              LINK DONE;
          END;

   IF VAL < MIN(X) THEN
          DO;
              YF=.;
              LINK DONE;
          END;

   DO LO=1 TO NROW(Y);
      IF VAL=X(|LO,1|) THEN
          DO;
              YF=Y(|LO,1|);
              LINK DONE;
          END;
      IF VAL>X(|LO,1|) THEN
          IF VAL<X(|LO+1,1|) THEN
              LINK CONT;
   END;

   CONT:
   HI=LO+1;

   H=X(|HI,1|)-X(|LO,1|);
   A=(X(|HI,1|)-VAL)/H;
   B=(VAL-X(|LO,1|))/H;
   YF=A*Y(|LO,1|) + B*Y(|HI,1|) +
      ((A*A*A - A)*Y2(|LO,1|) +
      (B*B*B - B)*Y2(|HI,1|))*(H*H)/6;
   DONE:
   FREE HI LO H A B;

FINISH;

START INTERP(X, Y, Y2, NEWIN, FIT);

   FIT=J(1,1,.);

   DO IND=1 TO NROW(NEWIN);
      VAL=NEWIN(|IND,1|);
      RUN SPLINE(X, Y, Y2, VAL, YF);
      FIT=FIT//YF;
```

```
   FREE A D ATOD;

   TAUi=COS(ANGLES(|,1|) +
         ANGLES(|,2|)/2)*&SQRTF;
   TAUj=COS(ANGLES(|,1|) -
         ANGLES(|,2|)/2)*&SQRTF;
   FREE ANGLES;


   USE T2 VAR{TAU_2, T1_2, T2_2};
   READ ALL VAR{TAU_2} INTO TAU2;
   READ ALL VAR{T2_2} INTO T22;

RUN DERIV(TAU2, T22, TAU2TOT2);
RUN INTERP(TAU2, T22, TAU2TOT2, TAUj, T1T2);
   FREE T22 TAU2 TAU2TOT2 TAUj;

   USE T1 VAR{TAU_1, T1_1, T2_1};
   READ ALL VAR{TAU_1} INTO TAU1;
   READ ALL VAR{T1_1} INTO T11;

RUN DERIV(TAU1, T11, TAU1TOT1);
RUN INTERP(TAU1, T11, TAU1TOT1, TAUi, T1STAR);
   FREE T11 TAU1 TAU1TOT1 TAUi;

T1T2=T1STAR||T1T2;

CREATE PERM.BETA FROM T1T2;
APPEND FROM T1T2;
```

PROGRAM 2b - Cubic Spline Routines

```
   START DERIV(X, Y, Y2);
      W=J(NROW(Y)-1,1,0);
      Y2=J(NROW(Y),1,0);

      DO I=2 TO (NROW(Y)-1);
         SIG=(X(|I,1|)-X(|I-1,1|))/
            (X(|I+1,1|)-X(|I-1,1|));
         P=SIG*Y2(|I-1,1|)+2;
         Y2(|I,1|)=(SIG-1)/P;
         W(|I,1|)=(Y(|I+1,1|)-Y(|I,1|))/
                  (X(|I+1,1|)-X(|I,1|)) -
                 (Y(|I,1|)-Y(|I-1,1|))/
                  (X(|I,1|)-X(|I-1,1|));
         W(|I,1|)=(6*W(|I,1|))/(X(|I+1,1|)-
                 X(|I-1,1|)) - SIG*W(|I-1,1|)/P;

      END;

      DO I=(NROW(Y)-1) TO 1 BY -1;
```

```
RUN SPLINE(TAU1, TAU2STAR, TAU1TAU2, -&SQRTF, FIT);
TEMP(|2,2|)=FIT;

RUN DERIV(TAU2, TAU1STAR, TAU2TAU1);
RUN SPLINE(TAU2, TAU1STAR, TAU2TAU1, &SQRTF, FIT);
TEMP(|3,1|)=FIT;
RUN SPLINE(TAU2, TAU1STAR, TAU2TAU1, -&SQRTF, FIT);
TEMP(|4,1|)=FIT;

TEMP=TEMP/&SQRTF;
FREE TAU1 TAU2STAR TAU2 TAU1STAR FIT;

   CREATE NEW FROM TEMP;
   APPEND FROM TEMP;

DATA NEW(KEEP=AVG DIFF);
   SET NEW;
   A=ARCOS(COL1);
   D=ARCOS(COL2);
   AVG=MEAN(A,D);
   DIFF=A-D;
   IF DIFF<0 THEN DO;
                AVG=-AVG;
                DIFF=-DIFF;
               END;

PROC SORT;
   BY AVG;

PROC IML;
   USE NEW VAR{AVG, DIFF};
   READ ALL VAR{AVG} INTO A;
   READ ALL VAR{DIFF} INTO D;

%INCLUDE 'D:SPLINE.MOD';

START PERIODIC(X, Y, DER, FINAL);

   FINAL=J(1,2,.);

   DO TRY=MIN(X) TO MAX(X) BY ((MAX(X)-MIN(X))/50);
      RUN SPLINE(X, Y, DER, TRY, FIT);
      FINAL=FINAL//(TRY||FIT);
   END;
   FINAL=FINAL(|2:NROW(FINAL),|);
   FREE TRY FIT;

FINISH;

   A=(MAX(A)-8*ATAN(1))//A//(MIN(A)+8*ATAN(1));
   D=D(|NROW(D),1|)//D//D(|1,1|);

RUN DERIV(A, D, ATOD);
RUN PERIODIC(A, D, ATOD, ANGLES);
```

```
DATA EXP;
   INFILE 'D:EXP.DAT';
   INPUT X Y;

%LOOP;
```

PROGRAM 2a - Profile Pair Sketches

```
OPTIONS NOSOURCE NONOTES;
LIBNAME PERM 'D:';

DATA T1;
   INFILE 'D:AB0PLOT.DAT' ;
   INPUT DEL_1 TAU_1 T1_1 T2_1;

PROC SORT;
   BY TAU_1;

DATA T2;
   INFILE 'D:AB1PLOT.DAT' ;
   INPUT DEL_2 TAU_2 T1_2 T2_2;

PROC SORT;
   BY TAU_2;

%LET SQRTF=3.72673;

PROC IML;
   USE T1 VAR{TAU_1, T1_1, T2_1};
   READ ALL VAR{TAU_1} INTO TAU1;
   READ ALL VAR{T1_1} INTO T11;
   READ ALL VAR{T2_1} INTO T21;

   USE T2 VAR{TAU_2, T1_2, T2_2};
   READ ALL VAR{TAU_2} INTO TAU2;
   READ ALL VAR{T1_2} INTO T12;
   READ ALL VAR{T2_2} INTO T22;

%INCLUDE 'D:SPLINE.MOD';

RUN DERIV(T11, TAU1, T1TOTAU1);
RUN INTERP(T11, TAU1, T1TOTAU1, T12, TAU1STAR);
FREE T11 T1TOTAU1 T12;

RUN DERIV(T22, TAU2, T2TOTAU2);
RUN INTERP(T22, TAU2, T2TOTAU2, T21, TAU2STAR);
FREE T22 T2TOTAU2 T21;

TEMP = {&SQRTF 0,-&SQRTF 0,0 &SQRTF, 0 -&SQRTF};

RUN DERIV(TAU1, TAU2STAR, TAU1TAU2);
RUN SPLINE(TAU1, TAU2STAR, TAU1TAU2, &SQRTF, FIT);
TEMP{1,2}=FIT;
```

PROGRAM 1 - t PLOTS

```
%LET B0HAT=-156.9399006;
%LET B1HAT=-0.1996716;
%LET SE=29.857774612;
%LET SS=13390.09313;
%LET DF=4;
%LET sqrtF=4.24;
%LET T=0;

%MACRO TAU(DELT);

     DATA _NULL_;
        VAL=(&DELT*&SE) + &B0HAT;
        CALL SYMPUT('B0TRY',VAL);

     PROC NLIN METHOD=GAUSS BEST=1 DATA=EXP OUTEST=EST;
        PARMS B1= &B1HAT B0= &B0TRY;
        BOUNDS &B0TRY <=B0<= &B0TRY;

        MODEL Y = -B0*EXP(-B1*X);
        DER.B0 = -EXP(B1*X);
        DER.B1 = B0*X*EXP(B1*X);

     DATA EST;
        SET EST;
        IF _TYPE_='FINAL';
        DEL = &DELT;

        TAU = SIGN(&B0TRY - &B0HAT) *
            SQRT((_SSE_ - &SS)/(&SS/&DF));
        TABS=ABS(TAU);
        FILE 'D:B0.DAT' MOD;
        PUT @1 DEL @10 TAU @20 B0 @30 B1;

        CALL SYMPUT('B1HAT',B1);
        CALL SYMPUT('T',TABS);

%MEND;

%MACRO LOOP;
   %DO HALF=1 %TO 2;
      %DO I=0 %TO 6;
         %DO J=0 %TO 9;
            %IF &HALF=1 %THEN %LET DELT=-&I..&J;
            %ELSE %LET DELT=&I..&J;
            %TAU(&DELT);
            %IF &T> &sqrtF %THEN %GOTO LABEL1;
         %END;
      %LABEL1: %END;
   %END;
%MEND;
```

interpolated will also be the number of points defining the likelihood region and thus, may be increased or decreased depending on the desired continuity of the final region. The last steps convert these interpolated (a, d) pairs to the τ_i and τ_j and finally θ_i and θ_j scales, using the same splining routine.

Figure 2 illustrates the continuation of the exponential example using Programs 2a and 2b. The profile traces appear as dashed lines which intersect each other at $\hat{\theta}$ (+) and the likelihood region (shaded area) at its vertical and horizontal tangents. One should note how the traces are related to the shape of the likelihood region. The straighter the traces are, the more elliptical the region will be. The angle of the traces to each other also reflects the correlation between the parameters, with the condition of no correlation resulting in right angles. The linear approximation to the likelihood region is an ellipse (solid line) which is centered at $\hat{\theta}$ (note: codes for this ellipse are not provided but can be obtained easily by applying equation (3) to the appropriate θ pairs). As was seen with the univariate intervals, the linear approximation is not perfect but is probably acceptable. In this case a conservative approach to joint inferences on β_0 and β_1 could be taken by increasing the level of α, which would reduce the size of the linear approximation ellipse, giving more overlapping of the likelihood region.

confidence interval on β_1 is .071 to .329 which is somewhat different than the nominal interval of .093 to .367. Although this is more of a difference than was demonstrated in β_0, inferences made on β_1 using the linear approximation will still be reasonable if the α level is set conservatively.

Profile pair sketches for β_0 and β_1 were obtained using the output files of the previous t plot program in Program 2a. This program relies heavily on cubic splines to translate from the θ scale to the τ scale (and visa versa) for which SAS/IML was appropriate. Although a spline function exists in IML, it was not utilized here because it did not compute splines in the manner required. The existing function is also memory intensive and is not currently implemented in all versions of SAS. The spline routines used are listed as IML modules in Program 2b. These spline routines were adapted from Press et al (1988), but any similar cubic spline routine would suffice.

The initial steps of Program 2a involve splining the profile trace values from the θ scale to the τ scale. When this is accomplished, the τ values are splined as functions of each other, i.e.: $\tau_i = f(\tau_j)$ and $\tau_j = f(\tau_i)$, and used to interpolate the four points corresponding to the vertical and horizontal tangents of the likelihood region. The values of the tangents to interpolate from are given in the macro variable SQRTF which has been previously set to $\sqrt{p * F(p, n-p; \alpha)}$, where $1-\alpha$ is the confidence level of the likelihood contour desired. After all four τ_i, τ_j pairs are determined, they are scaled by dividing by SQRTF and output to a data set for ease in sorting and further calculations. This step also involves changing the sign of some values which is done to account for the periodic nature of the cosine function. IML is reinvoked with sorted (a, d) pairs and a spline of period 2π is used to interpolate several new values of d (50 in this case) as a function of a. The number of values

The heart of the program lies in macro TAU(.) where incoming values of $\delta(\theta_p)$ are transformed to the θ_p scale and used to constrain θ_p with macro variable MTRY. $\tau(\theta_p)$ is calculated from equation (1) using output from PROC NLIN and updated along with $\delta(\theta_p)$ and $\hat{\theta}_p$ to an external file. Additionally, the absolute value of $\tau(\theta_p)$ is calculated to later test for the termination of the routine. The final step updates the parameter starting values excluding the parameter being worked on, in this case β_0. This is an important step because it speeds execution time and helps prevent abnormal results which could arise from an irregular likelihood surface.

The driver for TAU(.) is macro LOOP which increments values of $\delta(\theta_p)$ over a given range and passes them to TAU(.). The size of the range for δ may be modified to suit the problem at hand. In this case values of $\delta(\theta_p)$ are incremented in two halves, first from 0 to -6.9 and then 0 to 6.9, which facilitates testing the stopping condition. This test compares the latest value of T to the set value of sqrtF and if true, the macro either proceeds to the next half or terminates the program.

Figures 1a and 1b illustrate the profile t plots for β_0 and β_1, respectively. β_0 appears to have slight nonlinearity because its profile sum of squares function (short dashed line) deviates little from the linear estimation situation (solid line). β_1 on the other hand, exhibits more nonlinearity since its profile sum of squares function is quite distinct from the reference line. The difference in the nonlinearity of β_0 and β_1 becomes evident with a comparison of nominal (eq. 2) and linear approximate intervals for each parameter. The linear approximation for a 95% confidence interval on β_0 is 74.0 to 239.8 which agrees well with the nominal interval of 73.9 to 235.2 (long dashed lines). Thus, inferences on β_0 using this linear approximation will be very reliable. In contrast, the linear approximation for a 95%

This provides a good approximation when parameters are close to linear and although increased nonlinearity may adversely affect its usefulness, it will provide better inferential results than the linear approximation in most practical situations.

APPLICATION

To illustrate the techniques described above, a two- parameter exponential function of the form:

$$Y = -\beta_0 \exp(-\beta_1 X)$$

was fitted to the data listed in Table 1. All programs were run under MSDOS version 4 and written using SAS/STAT,IML version 6.03 (1989, 1990). Plots were produced with SAS/GRAPH (1990).The Gauss-Newton method of PROC NLIN resulted in convergence on the parameter estimates: $\hat{\beta}_0 = 156.9$, $\hat{\beta}_1 = .199$.

Table 1 - Exponential Data

X	Y
-5	-396.2983
-3	-372.2983
-1	-144.2983
1	-102.2983
3	-63.2983
5	-97.2983

Profile t plots were produced for each parameter using Program 1 as follows. Starting values for the parameters are defined in macro variables along with the error sum of squares, error degrees of freedom, and the standard error of the specified parameter of interest, all of which are from the previous unconstrained estimation. Variables sqrtF and T are used subsequently to test for the termination of the calculations.

approximations.

As with likelihood intervals, a linear approximation is available for pairwise likelihood inference regions and is defined as:

$$(\theta - \hat{\theta})'\hat{V}'\hat{V}(\theta - \hat{\theta}) \leq Ps^2F(p, n\text{-}p;\alpha), \qquad (3)$$

where θ and $\hat{\theta}$ are the parameter and estimate vectors respectively and $\hat{V}$ is the derivative matrix evaluated at $\hat{\theta}$. A two dimensional pairwise inference region is expressed geometrically from equation (3) as an ellipsoid in the parameter space centered at $\hat{\theta}$.

Nonlinearity in the parameter estimates can limit the value of the approximate likelihood regions, by bending and distorting the ellipsoid. Graphing the actual nominal likelihood contours for parameter pairs is computationally too intense for most practical purposes. Bates and Watts (1988) however, have suggested a method for approximating nominal likelihood contours using information acquired from the calculation of the profile t function. When the ith parameter is constrained to θ_i, least squares estimates for the remaining parameters are found, $\hat{\theta}_i(j)$, for j=1 to P; $j \neq i$. A plot of $\hat{\theta}_i(j)$ versus θ_i is called the profile trace, and it reflects the path followed by the estimate of θ_j when θ_i is systematically constrained. Overlaid plots of profile traces for two parameters, θ_i vs $\hat{\theta}_i(j)$ and θ_j vs $\hat{\theta}_j(i)$; $i \neq j$, will give two crossing lines which have the special property of passing through the vertical and horizontal tangents of the likelihood contours. Using these tangent points for a given 100(1-α)% contour, four θ_i, θ_j pairs are obtained. These are transformed to the τ scale and the relationships $\tau(\theta_i) = \cos(a + d/2)$ and $\tau(\theta_j) = \cos(a - d/2)$ are used to solve for the four (a, d) pairs. Interpolating on d = f(a) provides several intermediate (a, d) pairs which are then transformed back to τ and finally θ coordinates for contour plotting.

remaining parameters. This sum of squares is used in the profile sum of squares function (Bliss & James, 1966; Bates & Watts, 1985), $\tau(\theta_p)$, which is defined as:

$$\tau(\theta_p) = \text{sign}(\theta_p - \hat{\theta}_p)\sqrt{\frac{SS_2 - SS_1}{MS_1}}, \qquad (1)$$

where $\hat{\theta}_p$ is the unconstrained least squares estimate of the pth parameter, SS_1 and MS_1 are, respectively, the error sum squares and mean square from the unconstrained least squares fit, and SS_2 is the profile sum of squares at θ_p. In the linear estimation situation this function is equivalent to the studentized difference,

$$\delta(\theta_p) = \frac{(\theta_p - \hat{\theta}_p)}{se(\hat{\theta})}.$$

A plot of $\tau(\theta_p)$ vs $\delta(\theta_p)$ is called a profile t plot for parameter p and usually includes a reference line at 45° to both axes which represents the linear case where $\delta(\theta_p) = \tau(\theta_p)$. The more the profile sum of squares function deviates from the referenoe line for a particular parameter, the more nonlinearity is associated with that parameter. A suggested range for calculating τ and δ is those values of θ_p such that $|\tau(\theta_p)| \leq \sqrt{F(p, n-p; \alpha)}$ where α is set to a small value like .01.

Profile t plots can also provide nominal 100(1-α)% likelihood intervals by applying the relationship,

$$-t(n\text{-}p,\alpha/2) \leq \tau(\theta_p) \leq t(n\text{-}p,\alpha/2), \qquad (2)$$

to the $\tau(\theta_p)$ axis and extrapolating back through the profile sum of squares curve to the $\delta(\theta_p)$, or equivalently θ_p, axis. Comparison of this interval with the linear approximation interval,

$$\hat{\theta} \pm t(n\text{-}p,\alpha/2)se(\hat{\theta}),$$

will also help determine the extent of nonlinearity and the reliability of such linear

INTRODUCTION

Nonlinear regression can be a useful tool in modeling. Ordinary nonlinear least-squares techniques are iterative, involving updating of parameter estimates until a convergence criteria is reached. This process may result in nonlinearity in the estimation situation with respect to one or more parameters. Nonlinearity of this type can affect the validity of statistical inferences made on the model and its components. Such inferences are based on linear approximations which assume a planar expectation surface near the solution while the actual surface in nonlinear regression is curved. The degree of curvature is influenced by the specified expectation function, its parameterization, and the data to be fitted. Depending on the severity of the curvature, this may result in inference intervals and regions that are asymmetric and distorted when compared to the corresponding linear estimation situation. It is of interest therefore, to determine the degree of this distortion as part of the model assessment and validation process. Bates and Watts (1985) have suggested some graphical approaches to address this problem. These procedures lead to exact likelihood intervals and regions which can be compared to their linear counterparts to determine departure from linearity and the reliability of inferences made based on the linear approximation. SAS® codes to produce the required graphics are also provided and demonstrated using a two-parameter exponential function.

METHODS

The graphical techniques to be described here are based on a statistic known as the profile sum of squares, which is defined as the error sum of squares reached by constraining one parameter to a constant value, θ_p, and achieving least squares estimates, θ_{p_*} for the

Assessing Parameters Estimates in Nonlinear Regression Using Profile Plots

Bahman Shafii
Statistical Programs, College of Agriculture
University of Idaho
Moscow, U.S.A.

ABSTRACT

Nonlinear regression analysis requires iterative estimation procedures to achieve least squares solutions. Depending on the form and parameterization of the specified expectation function and the data being considered, these techniques may result in nonlinearity of the estimation situation with respect to one or more parameters. The degree of nonlinearity for each estimated parameter can be assessed by examining profile t plots, whereas the pairwise behavior of parameter estimates is examined using profile pair sketches. The extent of nonlinearity is determined by comparing these plots to those of the corresponding linear estimation situation. These plots may also be used to develop exact likelihood intervals and inference regions for the parameters. Computational procedures to produce the required graphics are reviewed and the application is demonstrated using a two-parameter exponential function.

Key Words: Nonlinear Estimation, Profile Sum of Squares, Pairwise Inference Regions.

REFERENCES

[1] Johnston, J (1984). Econometric Methods. McGaw-Hill Book Co.

[2] Draper, N R and Smith, H (1981). Applied Regression Analysis. 2nd edition. Wiley & Sons.

[3] Atkinson, A C (1985). Plots, Transformations and Regression. Oxford University Press.

[4] McPherson, G (1990). Statistics in Scientific Investigation. Springer-Verlag.

[5] Microfit data 1989 obtained from the Central Statistical Office, London.

[6] Pesaran, M H and Pesaran, B (1989). Microfit Econometric Computer Package. Version 2. Oxford University Press.

where the values in parentheses are the t-ratios. Using these ratios, we discard the non-significant variables, and regress DLC_t this time on the significant variables DLY_t, DLP_t and DLP_{t-1}. As a result we have:

$$\text{fitted } DLC_t = \underset{(2.6133)}{0.0137} + \underset{(8.4609)}{0.6372}DLY_t - \underset{(-2.9233)}{0.1755}DLP_t + \underset{(1.5591)}{0.09133}DLP_{t-1} \quad (t=1965-1984)$$

with the adjusted coefficient of determination $\bar{R}^2 = 0.8461$, indicating a strong association between the rates of growth of consumption, disposable income and inflation for current and previous years.

Now we attempt to increase our sample of observations and consider the data from 1950 to 1984. Then, using the same multiple regression equation, we derive the following fitted model:

$$\text{fitted } DLC_t = \underset{(2.7396)}{0.01064} + \underset{(9.1150)}{0.6337}DLY_t - \underset{(-2.8090)}{0.1546}DLP_t + \underset{(1.6997)}{0.0886}DLP_{t-1} \quad (t=1950-1984)$$

Where $\bar{R}^2 = 0.7976$ and Durbin - Watson statistic = 2.1665 implying the suitability of the fitted model to the data. Having derived an appropriate model, we are now in a position to predict the future rates of consumption on the basis of information available at the time. To demonstrate, and for comaparative purposes, prediction for the rate of growth of consumption for 1985, using the last fitted model, is obtained as fitted $DLC_{1985} = .0242$. This prediction, when compared to the actual value 0.0345 DLC_{1985}, shows a very low standard deviation of error of .0091.

CONCLUSION

From statistical methodology and inferential points of view, computers can be used for detailed examination of the data both before and after analysis. Multiple regrassin analysis is a powerful statistical technique which is widely applied to various scientific and industrial problems, and has particular importance in economic forecasting field. The use of computer software catering for this technique is illustrated to be immensly beneficial both in model building and forecasting areas in econometrics.

ACKNOWLEDGEMENT: The author would like to express their gratitude to Prof. M H Pesaran and Dr. B Pesaran for the use of Microfit econometric package.

to economists. To illustrate, consider the data in the following table on the real consumer's expenditures (C), real disposable income (Y) and consumer's price index (P, relative to 1975), and their corresponding rate of growth, DLC, DLY and DLP in the United Kingdom from 1948 to 1985 (in millions of Pounds sterling). (Data from CSO [5])

Year	C	Y	P	Year	C	Y	P
1948	34279.0	34307.0	.24948	1967	54383.0	59340.0	.46964
1949	34925.0	35254.0	.25503	1968	56026.0	60459.0	.49124
1950	35914.0	36407.0	.26174	1969	56312.0	61079.0	.51912
1951	35421.0	35951.0	.28655	1970	57938.6	64043.9	.55167
1952	35238.0	36477.0	.30339	1971	59740.2	64862.7	.59914
1953	36831.0	38300.0	.30958	1972	63400.5	70360.2	.63826
1954	38347.0	39559.0	.31530	1973	66671.7	75179.8	.69074
1955	39904.0	41372.0	.32691	1974	65700.9	74588.1	.80790
1956	40289.0	42480.0	.34163	1975	69336.9	74839.8	1.0000
1957	41118.0	42163.0	.35311	1976	69541.1	74424.1	1.1568
1958	42095.0	43866.0	.36161	1977	69209.0	73599.9	1.3292
1959	42911.0	46134.0	.36706	1978	68798.3	78659.3	1.4517
1960	45623.0	49143.0	.37128	1979	71864.2	82899.3	1.6479
1961	46680.0	52131.0	.38220	1980	71666.8	84376.6	1.9149
1962	47653.0	51499.0	.39725	1981	71511.2	82540.0	2.1331
1963	49725.0	53760.0	.40497	1982	72171.8	82810.5	2.3189
1964	51274.0	55767.0	.41934	1983	75091.0	84824.7	2.4354
1965	52131.0	57182.0	.43951	1984	76796.3	86671.3	2.5514
1966	53184.0	58467.0	.45747	1985	79587.8	89293.4	2.6853

In general, the rate of growth of consumption (without multiplying by 100) is defined as $DLC = \log_e C - \log_e [C(-1)] = \log_e \frac{C}{C(-1)}$, where C(-1) is the consumption lag derived from the previous year. Similarly, for DLY and DLP. In particular DLP, the rate of growth of consumer's price index, is generally known as 'the rate of inflation'.

The multiple regression model we assume here is:

$$DLC_t = \beta_0 + \beta_1 DLC_{t-1} + \beta_2 DLY_t + \beta_3 DLY_{t-1} + \beta_4 DLP_t + \beta_5 DLP_{t-1} + \varepsilon_t \quad , \quad (t=1965 - 1984)$$

Applying the multiple regression computer routine from Microfit econometric software [5] to the above model, we arrive at the following results:

$$\underset{t}{\text{fitted DLC}} = \underset{(2.3534)}{0.0157} - \underset{(-0.9970)}{0.2270}DLC_{t-1} + \underset{(7.7753)}{0.6534}DLY_t + \underset{(0.9660)}{0.1717}DLY_{t-1}$$

$$- \underset{(-2.5343)}{.1860}DLP_t + \underset{(1.0948)}{0.0837}DLP_{t-1} \quad ,$$

Where x_{ij} is the setting of the jth independent variable ($j = 1,2,..,k$) for the ith observation. The value of the coefficient β_j determines the contribution of the independent variable x_j to the model. These coefficients are usually unknown as they represent population parameters, and are estimated using the observed values y_i's and x_{ij}'s.

We can write the set of n simultaneous equations above in matrix notation as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

or,

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad ,$$

Where $\underset{\sim}{X}$ is a known 'design matrix', and $\underset{\sim}{\varepsilon}$ is a vector of errors assumed to be normally distributed with zero mean and variance $\sigma^2 I$, with I an n x n identity matrix, and the variance of error terms σ^2 constant and generally unknown - homoscedasticity (Johnston [1]).

Using the principle of least squares we arrive at the following estimates of β's, namely, $\hat{\beta}$'s:

$\underset{\sim}{\hat{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}$, with variances of $\underset{\sim}{\hat{\beta}}$ as $\text{var}(\underset{\sim}{\hat{\beta}}) = \sigma^2(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$. It is important to note that $\hat{\beta}$ s above are in fact proved to be the "best estimators" of β's - Gauss - Markov Theorem (Draper [2]).

Now, the $(k + 1)$ simultaneous equations that must be solved to determine the estimators $\hat{\beta}_0$, $\hat{\beta}_1$, ..., $\hat{\beta}_k$ (or, equivalently, the evaluation of the inverse of the square matrix $\underset{\sim}{X}'\underset{\sim}{X}$) are extremely difficult and at times impossible to solve and subsequently the use of statistical software can be infinitely beneficial here. Various computer packages have been developed to cater for this purpose. These provide the user with interactive facilities to produce a model which 'best' fits the data, and identifies and eliminates the effects of non-significant variables in the model (Atkinson [3]). Amongst other important inferences, we then would be able to assess the strength of relationship between the variables involved (Mc.Pherson [4]), and forecast future values of Y from the fitted model.

APPLICATIONS AND RESULTS

Applications of regression analysis in economic forecasting are of valuable importance

Computer Applications in Statistical Methodology and Inference and Econometrics

Nasrollah Saebi
Kingston Polytechnic
School of Mathematics
London, U.K.

ABSTRACT

The increasing availability of computers and development of statistical packages offers a great deal of advantage in inferential analysis of experimental and observational data. The capability of computers allied to the improving graphical facilities allows the statistical analyses and the assumptions made about the data to be examined in much more detail. The fundamental advantage which the computer offers is the sheer speed of its calculations. The other principal superiority is the use of its memory for storing data for future references.

In this paper an important area of statistical analysis where computers are found to be particularly beneficial is studied and discussed.

INTRODUCTION

During the past two decades, many computer programs and packages for conducting the calculation of statistical analyses have been developed. These range from single programs to evaluate statistical measures to more complex and extensive packages able to handle data from designed experiments with various blocking structures, and advanced forecasting methods using multiple regressional techniques, with widespread applications in science and industry.

THEORETICAL ANALYSIS

Regressional analysis is a technique for constructing and developing models of relationshipsbetween several variables in the presence of a random experimental error term ε. In a regression model, one of the variables which is generally referred to as the dependent or response variable Y, is expressed as a linear function of the remaining k (positive integer) variables. $x_1, x_2, \ldots, x_k$ known as the independent or explanatory variables. The general form of these probabilistic models is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \quad ;$$

Where we make n independent observations $y_1, y_2, \ldots, y_n$ on Y, and on x's. We can then write the observation y_i as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \quad , (i = 1,2,\ldots,n)$$

References

(1) Bickel,P.J.and Doksam,K.A.(1977)."Mathematical Statistics",Holden_Day,Inc. San Francisco
(2) Gibbons,Jean D.(1974)."Estimation of the unknown limit of a uniform distriliution. Sankhya,B.36(1),29_40
(3) Gibbons,J.D.and Litwin,S.(1974):Simultaneous Estimation of the unkown upper and lower limits in a TWO_Parameter uniform Distriliution,Sankhya,B.36(1), 39_54
(4) Hogg,R.V.and Craig,A.T.Introduction to "Mathematical statistics",Macmillan publishing,FOS,New York.
(5) Lindgren,B.W."statistical Thory",Macmillan Publishing co.New York.
(6) Mood,M.A.and Graybill,F.A.and Bose,D.C."Introduction to the Theory of stristics",Mcgraw_Hill
(7) Rohatgi,V.K."An Introduction to Probability Theory and Mathematical Statistics.J.W.New York

	(1)	(2)	(3)
Point estimate	T_0	B	T
Interval estimate	$T_0 \mp b\epsilon$	$B \mp b\epsilon$	$T \mp b\epsilon$
Confidence coefficient	$Eq(2.7)$	$Eq(2.8)$	$Eq(2.9)$
or ϵ given by			
Approximation given by	$Eq(3.1)$	$Eq(.2)$	$Eq(3.3)$

It can be easily showed that (3) results in a better estimate theoretically in the sense of precision for a fixed confidence coefficient.

In practical applications of confidence interval estimation, the usual procedure is to state the probability $(1-\alpha)$ and find the corresponding value of ϵ.

For testing $H_0 : \theta = \theta_0$ with a two sided alternative $H_1 : \theta \neq \theta_0$ if we suppose $\hat{\theta}$ be any statistic ($\hat{\theta}$ can be T_0,B or T), then the size of the test is

$$P(|\frac{\hat{\theta} - \theta_0}{b}| > \epsilon) = 1 - P(|\frac{\hat{\theta} - \theta_0}{b}| \leq \epsilon)$$

In this case for $\hat{\theta} = T$, we also have the most powerful test.

We conclude this discussion to notice that, on the basis of mean squared error criterion, for unbiased estimators, we have .

$$Var(S_{w(X)}) = [2(n-1)\lambda^2 - 2(n-1)\lambda + n]\frac{b^2}{(n+1)^2(n+1)} \geq 0$$

This Variance is minimum for $\lambda = 1/2$, which is corresponding to the estimator $T = \frac{Y_1+Y_n}{2}$.

$$z_1 = n\left(\frac{T_0-\theta}{b}\right) \rightarrow \lim_{n\to\infty} f_{Z_1}(z) = e^{z}, \qquad z < 0$$
$$z_2 = n\left(\frac{B-\theta}{b}\right) \rightarrow \lim_{n\to\infty} f_{Z_2}(z) = e^{-z}, \quad \frac{-n}{n+1} < z < \frac{n}{n+1}$$
$$z_3 = n\left(\frac{T-\theta}{b}\right) \rightarrow \lim_{n\to\infty} f_{Z_3}(z) = e^{-2|z|}, \quad z \in R$$

The large sample approximations to (2.7-2.9), the relative accuracies of these estimators are then given by

$$P\left(\left|\frac{T_0-\theta}{b}\right| < \epsilon\right) \approx 1 - e^{-n\epsilon} \quad , \epsilon > 0 \tag{3.1}$$

$$P\left(\left|\frac{B-\theta}{b}\right| < \epsilon\right) \approx \begin{cases} e^{-1}\left(e^{n\epsilon} - e^{-n\epsilon}\right) & , 0 < \epsilon < \frac{1}{n+1} \\ 1 - e^{-(n\epsilon+1)} & , \epsilon \geq \frac{1}{n+1} \end{cases} \tag{3.2}$$

$$P\left(\left|\frac{T-\theta}{b}\right| < \epsilon\right) \approx 1 - e^{-n\epsilon} \quad , \epsilon > 0 \tag{3.3}$$

These simple and convenient results provide excellent approximations even for small samples.

4.Inference applications of estimators

Any of these three estimator may be employed for inference, for either hypothesis testing or confidence interval estimation for parameter θ.

Properties of the confidence interval estimates of θ, based on the point estimate T_0, B and T are summaried as follows

an other objection to use them as estimators for θ. Opposed to these, the distribution of T is symmetric about θ; overestimates and underestimates of the same absolute magnitude,have the same probability of occurance. Figure 2.2 shows how the areas under the respective density functions comprise the error probability in (2.7) and (2.9).

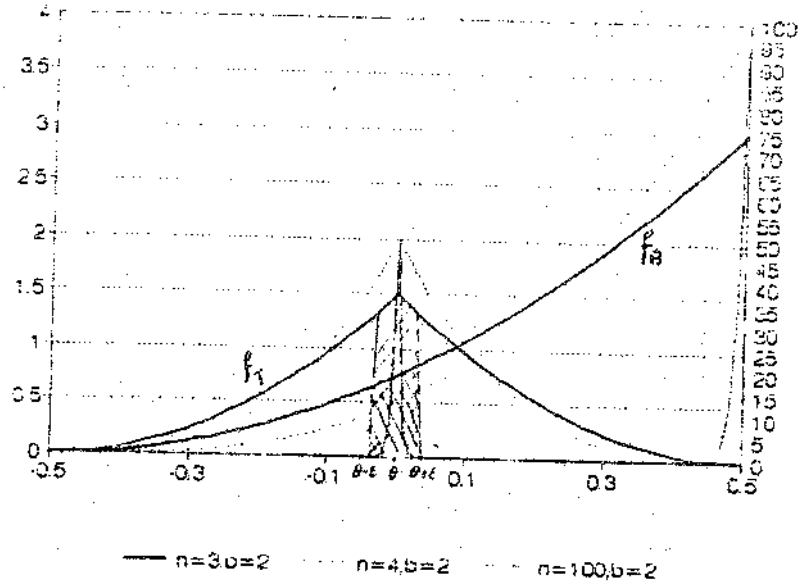


Fig 2.2.probability density function of the estimators B and T, with shaded areas representing the respective probabilities that the estimators is within ϵ of θ.

For all ϵ, T is clearly the best estimator of the three using this criterion.

3.Asymptotic Aprroximations to relative accuracy

Since the approximations of probabilties from (2.7-2.9), for all n, and specially for n large, is not easy, the limiting distributions of the standardized variables are easily found and are given as follows.

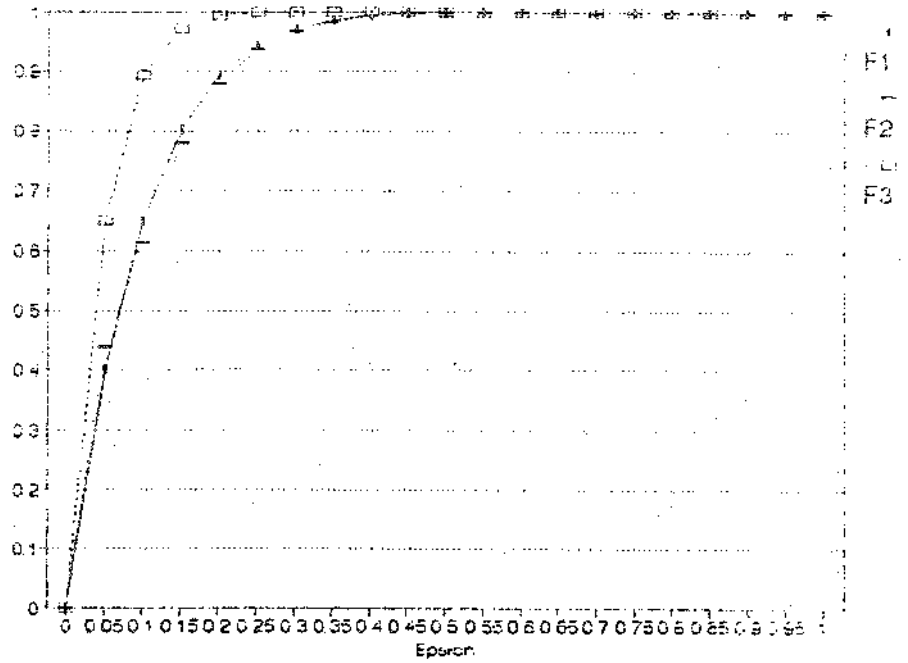


Fig. 2.1. Comparison of probabilities that relative absolute error of the estimate does not exceed ϵ,for estimators $B(F_3)$, $T_0(F_2)$ and T, all based on samples of size $n = 10$.

The fact that the estimators T_0 and T have exactly equal probabilities of relative error not exceding ϵ does not necessarily imply that their "goodness" as estimator is equivalent. In the case of $T_0 = Y_n - b/2$,since $P(T_0 > \theta) = 0$, these estimators approache to θ from left. on the other hand, since the distribution of T_0 and also B are highly skewed about θ, is

Both of these probability functions are continous everywhere.They increase with ϵ for n fixed and increase with n for ϵ fixed.It is clear that (2-8) is larger than (2-7) for $\frac{1}{n+1} < \epsilon < \frac{n}{n+1}$ and all n.At the end point $\epsilon = \frac{1}{n+1}$, the quantities and their limits as $n \to \infty$ are

$$P(|\frac{T_0 - \theta}{b}| < \frac{1}{n+1}) = 1 - (1 - \frac{1}{n+1})^n \to 1 - e^{-1}$$

$$P(|\frac{B - \theta}{b}| < \frac{1}{n+1}) = 1 - (1 - \frac{2}{n+1})^n \to 1 - e^{-2}$$

The density function of T is

$$f_T(t;\theta) = \frac{n}{b}(1 - 2|\frac{t-\theta}{b}|)^{n-1} \quad \text{if} \quad |\frac{t-\theta}{b}| \leq 1/2$$
$$= 0 \quad \text{otherwise}$$

Thus we have

$$P(|\frac{T - \theta}{b} < \epsilon) = \begin{cases} 0 & \text{if} \quad \epsilon \leq 0 \\ 1 - (1 - 2\epsilon)^n & \text{if} \quad 0 < \epsilon < 1/2 \\ 1 & \text{if} \quad \epsilon \geq 1/2 \end{cases} \qquad (2.9)$$

The value of ϵ for which (2.7),(2.8) and (2.9) are equivalent is not easily expressed in closed form. These probabilities were evaluated for $n = 10$ and all $0 < \epsilon \leq 1$ in increments of 0.05.Figure 2.1 show the graphical comparison between them.

It is well known that the statistic (Y_1, Y_n) is a minimal sufficient for θ, but not Compelet, thus, we do not expect to have a uniformly minimum variance unbiased(UMVU) estimator for θ.

Now we consider the estimators obtained for $\lambda = 0$ (both biased and unbiased) and $\lambda = 1/2$, that are :

$$T_0 = Y_n - b/2, B = Y_n - \frac{b(n-1)}{2(n+1)}, T = \frac{Y_1 + Y_n}{2}$$

and then try to form the probability statements concerning the relative accuracy of these estimators. On one hand,the relative accuracy of any estimator $\hat{\theta}$ is measured by the magnitude $|\frac{\hat{\theta}-\theta}{b}|$, on the othere hand,the density of Y_n is

$$g_n(y;\theta) = \frac{n}{b^n}(y - \theta + b/2)^{n-1} I_{\theta-b/2,\theta+b/2}(y) \quad (2.6)$$

If we put $U = \frac{Y-\theta}{b}$, then we have $h(u) = n(u + 1/2)^{n-1} I_{-1/2,1/2}(u)$.

It follows from (2.6) that the accuracies of T_0 and B are given respectively by the following probabilities.

$$P(|\frac{T_0 - \theta}{b}| < \epsilon) = \begin{cases} 0 & \text{if} \quad \epsilon \leq 0 \\ 1 - (1-\epsilon)^n & \text{if} \quad 0 < \epsilon < 1 \\ 1 & \text{if} \quad \epsilon > 1 \end{cases} \quad (2.7)$$

$$P(|\frac{B - \theta}{b}| < \epsilon) = \begin{cases} 0 & \text{if} \quad \epsilon \leq 0 \\ (\frac{n}{n+1} + \epsilon)^n - (\frac{n}{n+1} - \epsilon)^n & \text{if} \quad 0 < \epsilon < \frac{1}{n+1} \\ 1 - (\frac{n}{n+1} - \epsilon)^n & \text{if} \quad \frac{1}{n+1} \leq \epsilon < \frac{n}{n+1} \\ 1 & \text{if} \quad \epsilon \geq \frac{n}{n+1} \end{cases} \quad (2.8)$$

Each estimator can also be used for hypothesis testing and interval estimation.

The measure of relative accuracy used here is the absolute value of a linear function of the "normalized error".

2. The estimators,their properties and relative accuracy

Let $X = (X_1 \dots X_n)$ be a random sample of size n from a one-parameter uniform distribution.

$$f_X(x,\theta) = 1/b I_{[\theta - b/2, \theta + b/2]}(x) \qquad (2.1)$$

If we suppose $Y_1 = \min_{1 \le i \le n}(X_i), Y_n = \mathrm{Max}_{1 \le i \le n}(X_i)$ it is well known that every statistics $S(X)$ such that

$$Y_n - b/2 \le S(X) \le Y_1 + b/2 \qquad (2.2)$$

is a maximum likelihood (ML) estimator of θ. By considering the lenght of the random interval $[Y_n - b/2, Y_1 + b/2]$,and parameter $\lambda, 0 \le \lambda \le 1$, the set of ML estimator can be written as follows:

$$S(X) = \lambda Y_1 + (1 - \lambda) Y_n + b/2(2\lambda - 1) \qquad (2.3)$$

with expectation

$$E(S(X)) = \theta + (2\lambda - 1)\frac{b}{n+1} \qquad (2.4)$$

Therefore, the subset of unbiased ML for θ is

$$S_u(X) = \lambda Y_1 + (1 - \lambda) Y_n + (2\lambda - 1)\frac{b(n-1)}{2(n+1)} \qquad (2.5)$$

ESTIMATION OF THE UNKNOWN PARAMETER OF A UNIFORM DISTRIBUTION

by

ALI MESHKANI

Faculty of Sicences,Department of Statistics

Ferdowsi University of Mashhad,IRAN

Summary

The goal of this paper is a study of three Maximum likelihood estimators for the parameter θ in a population which is uniformly distributed over $(\theta - b/2, \theta + b/2)$ where b is known and positive.For a sample of n observations,the maximum minus $b/2$,the maximum as corrected for bias,and the midrange,are compared as regard unbiasedness variance,symmetry,and several possible evaluations of relative error. procedures for hypothesis testing and interval estimation $of\ \theta$ are described and compared.

key words:estimator,maximum,midrange,relative accuracy,relative error, unbiasedness

1. Introduction

In this paper we have concerned with the sample estimation of the unknown parameter θ in a population which is uniformaly distributed over the interval $(\theta - b/2, \theta + b/2)$.The three point estimators considered are the maximum minus $b/2$,the minimum as corrected for bias,and the midrange. They are compared using such criteria as mean,Varince,relative accuracy,average error,symmetry of error,and asymptotic approximation to relative accuracy.

Spiegelhalter, D.J. and Smith, A.F.M.(1982) Bayes factors for linear and log-Linear models with vague prior information. J.Roy.Statist.Soc., B, 44, 377-387.

West, M(1986) Bayesian model monitoring. J.Roy. Statist. Soc.,B,48,70-78.

* REFERENCES *

Akaike, H.(1973) Information theory and an extension of the maximum likelihood principle. 2nd International symposium on Information Theory, 267-281 . Akademaa Kiado,Budapest.

Akaike, H.(1978) A Bayesian analysis of the minimum AIC procedure. Ann. Inst. Statist. Math., 30, A, 9-14.

Akaike, H.(1979) A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. Biometrika, 66, 237-242.

Atkinson, A.C.(1978) Posterior probabilities for choosing a regression model. Biometrika, 65, 39-48.

Bhansali, R.J. and Downham, D.Y.(1977) Some properties of the order of an autoregressive model selected by a generalization of Akaik's EPF criterion, Biometrika, 64. 547-551.

Box,G.E.P.and Tiao, G.C.(1973) Bayesian Inference in statistical ana - lysis. Addison wesley, Readang, M.A.

Kheradmandnia, M.(1991) Aspects of Bayesian Threshold Autoregressive Modelling, Ph.D. Thesis, Institute of Mathematics and statistics, University of kent, canterbury, England.

Lempers, F.B.(1971) Posterior probabilities of Alternative Linear Models, Rotteradam university press.

Schwarz,G.(1978) Estimating the dimension of a model, Ann. Statist.,6 , 461-464.

Smith, A.F.M.(1973) A general Bayesian linear model. J.Roy.Stutist.Soc., B, 35, 67-75.

Smith, A.F.M. and spiegelhalter, D.J.(1980) Bayes factors and choice criteria for linear models. J.Roy. Statist. Soc., B, 42, 213 - 220.

this figure, corresponds to the simulated AR(1) series of length N=100, the second point corresponds to the simulated AR(2) series of length N=200, the third point corresponds to the simulated AR(5) series of length N=100 and the fourth point corresponds to the simulated AR(10) series of length N=100. From this figure we see that the percentage corresponding to the simulated AR(1) series of length N=100 (i.e. the first point) is very high (i.e. 96%) while the percentage corresponding to the simulated AR(10) series of length N=100 (i.e. the fourth point) is very low (i.e.31). This high degree of variability do not appear in other three criteria.

* (6) Finally we examined the very large sample size of N=1000. The results are given in Table 2 (see page 16). From this table we see that with such a large sample size the overall percentages corresponding to *OPA*, *NOPA*, *AIC* and *BIC* are 95.75, 95.75, 73.25 and 98.50 respectively. Note that the 2400 simulated series mentioned at the beginning of this section do not include the 400 simulated series of lengths 1000 related to Table 2 . So that the total number of simulated series analysed in this section is 2800.

From Table 1(6) the overall percentages corresponding to the sample size of N=600 are 94.00, 94.75, 73.50 and 96.75 respectively. We see that by increasing the sample size from N=600 to N=1000 the overall percentage corresponding to the four criteria respectively increased by 1.75, 1.00, −0.25 and 1.75 percent.

8.Acknowledgements

I wish to thank Dr I.T jolliffe for his careful reading of the earlier drafts of this work and for his valuable comments.

Figure (1) The 24 percentages of correct automatic decisions.
Data : Table 1

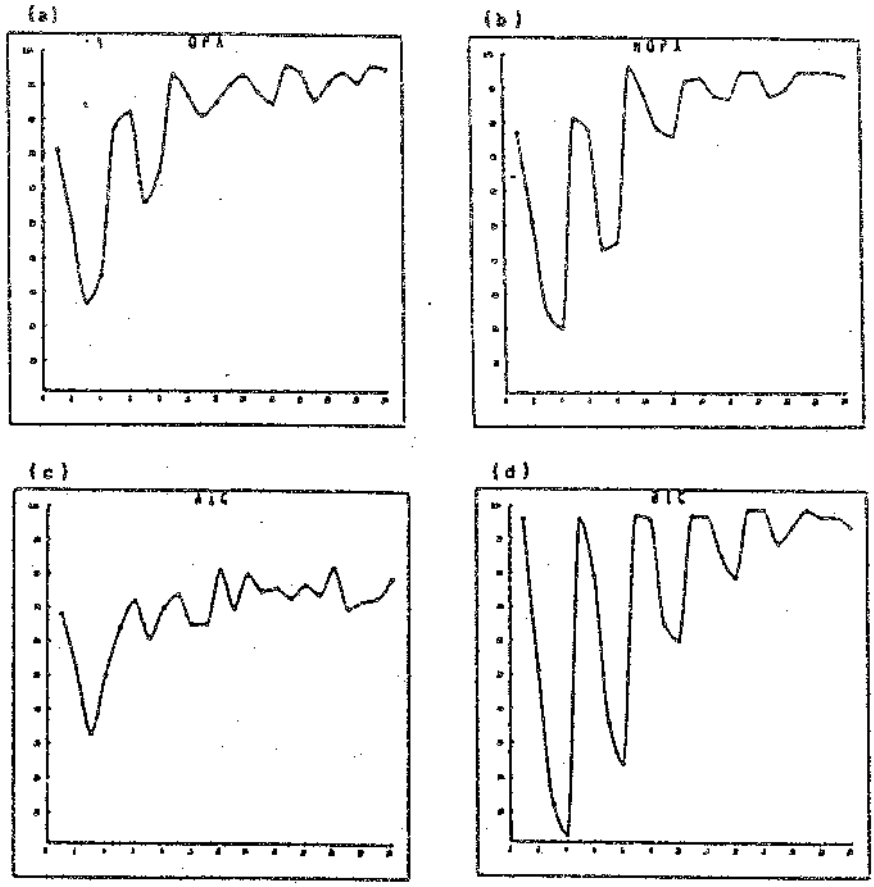


Table (2) Percentages of correct automatic decisions,
for all cases N = 12 , n0 = 24

| N = 1000 |

	Criteria			
series	OPA	NOPA	AIC	BIC
AR(1)	94	94	69	98
AR(2)	96	96	72	98
AR(3)	96	97	70	99
AR(10)	97	96	82	99
overall	95.75	95.75	73.25	98.50

the facility of incorporating informative priors, which may exist.The priors we have adopted have been, what may be termed as "*automatic data priors* ".

* [4] If the sample size is large the three criteria *GPA*, *NOPA* and *BIC* do not seem to be significantly different. Disregarding the percentages corresponding to the small training sample sizes 100 and 200, the overall percentages of correct automatic decisions for the four criteria are 85.63, 83.25,72.13 and 82.26 respectively. Disregarding the sample sizes 100, 200 and 300, the overall percentages for the four criteria are 91.58, 92.25, 75.08 and 93.83 respectively.

* [5] In Talbe 1, to any of the four criteria, there corresponds 4 different orders of AR models i.e. 1,2, 5 and 10, and 6 different sample size 100, 200, ... and 600. So that, all together, to any of the four criteria, there corresponds 24 percentages of correct automatic decision. From Table 1 the 24 percentages corresponding to each of the four criteria may be tabulated in the following form:

	1 , 2 , 3 , 4 , 5 , ... , 23 , 24
GPA :	71 , 50, 27, 33, 78, ... , 96 , 95
NOPA :	77 , 51, 24, 20, 81, ... , 95 , 94
AIC :	68 , 52, 33, 50, 64, ... , 75 , 79
BIC :	96 , 51, 12, 3 ,96, ... , 97 , 94

Figure 1 gives the plot of the above 24 percentages for each of the four criteria. The different degree of *consistency* of the approaches and different degree of *variability* amongst the observed percentages are apparent in this Figure. Also from this figure it is clear that, the *BIC* has the highest degree of variability amongst the four.In particular consider the first four points in Figure 1(d) ; the first point , in

the first row of Table 1[1] which are 71,77,68 and 96 respectively . Since the number of simulated series was 100 these percentages show the actual number of correct automatic decisions. We repeated the above experiment, but this time for the case of the above AR(2) model . The percentages of correct automatic decisions were 50, 51, 52 and 51 respectively, which are given in the second row of Table 1[1]. Similarly the third row of Table 1[1] gives the percentages based on 100 simulated series from the above AR(5) model , each series of length N=100 . Table 1[2] gives the percentages when the length of each of the 100 simulated series is N=200. For example the 4^{th} row of Table 1[2]gives the percentages of correct automatic decisions for 100 simulated AR(10) series, each series of length N=200 which are 65, 45, 70 and 24 respectively. Similarly Tables 1[3], 1[4], 1[5]and 1[6] give the corresponding results for the cases where the length of each series are N=300 , N=400, N=500 and N=600 respectively.

In the light of this simulation study we may conclude that as far as AR automatic modelling is concerned;

* [1] The importance of the Akaike Information Criteria(AIC) relates to its original stimulation which lead to a great deal of research and hence to some more efficient criteria such as BIC and OPA.

* [2] The overall percentages of correct automatic decisions out of the above 2400 simulated series for the cases of *OPA*, *HOPA*, *AIC* and *BIC* are 79.54, 77.67, 69.00 and 78.71 respectively. The *OPA* has the highest overall percentage of correct automatic decision. The differences between the overall percentages of *OPA*, *HOPA* and *BIC* do not seem to be highly significant. The difference between the *AIC* and each of the other three criteria seems to be highly significant.

* [3] The advantage of *BIC* is its simple formulation which makes it easy to use while the advantage of *OPA* is that, the modeller by using *OPA* has

accuracy). There may be a number of slightly incorrectly identified orders which are equal to one of the two closest permissible orders to the true order. For example, with 100 simulated AR(5) series, each of length N=500 and with an upper limit of U=12 to the order, the percentages of correct automatic decisions, based on *GPA*, *HQPA*, *AIC* and *BIC* were 86, 88, 74 and 69 respectively (see the 3rd row of Table 1 [5]), whereas the percentages of the cases in which the identified order was in the interval [4,6] were 95, 94, 83, and 93 respectively. So that if the definition of correct decision is extended to cover the cases in which the identified order is equal to one of the two closest permissible orders the percentages of correct decisions will be increased! In other words, the probability of a correct decision, to some extent, is a matter of definition. In this section by correct decision we mean "*pinpoint accuracy*" in detection of the true model.

Now here are the details concerning the analysis of 2400 series from the following AR models :

$x_t = 0.9x_{t-1} + e_t$

$x_t = 0.5x_{t-1} + 0.25x_{t-2} + e_t$

$x_t = 1.2x_{t-1} - 0.5x_{t-2} + 0.3x_{t-3} - 0.4x_{t-4} + 0.2x_{t-5} + e_t$

$x_t = x_{t-1} - 0.5x_{t-2} + 0.3x_{t-3} + 0.4x_{t-4} + 0.1x_{t-5} - 0.2x_{t-6} + 0.1x_{t-7} - 0.1x_{t-8} + 0.2x_{t-9} - 0.3x_{t-10} + e_t$ with $e_t \sim N(0,1)$ all cases.

Throughout this section we assume an upper limit of 12 to the order, and for calculation of *GPA(k)* and *HQPA(k)* we adopt a training sample size of $n_0 = 24$. We first simulated 100 series from the above AR(1) model each series of length N=100. We then used each of these 100 series to identify the order of the AR model using the four crateria *GPA*, *HQPA*, *AIC* and *BIC*. The percentages of correct automatic decisions are given in

Table (1) Percentages of correct automatic decisions.

for all cases U = 12 , n0 = 24

(1) N = 100

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	71.00	77.00	68.00	96.00
AR(2)	50.00	51.00	32.00	91.00
AR(5)	27.00	24.00	33.00	12.00
AR(10)	35.00	20.00	50.00	3.00
overall	45.75	43.00	54.75	40.50

(4) N = 400

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	90.00	92.00	78.00	99.00
AR(2)	93.00	93.00	80.00	97.00
AR(5)	88.00	86.00	75.00	89.00
AR(10)	85.00	83.00	78.00	99.00
overall	89.00	90.00	79.25	89.50

(2) N = 200

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	78.00	81.00	64.00	96.00
AR(2)	82.00	78.00	72.00	79.00
AR(5)	56.00	43.00	61.00	36.00
AR(10)	65.00	45.00	70.00	24.00
overall	70.25	61.75	66.75	58.75

(5) N = 500

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	96.00	95.00	73.00	25.00
AR(2)	94.00	95.00	77.00	99.00
AR(5)	86.00	88.00	74.00	89.00
AR(10)	91.00	90.00	82.00	94.00
overall	91.75	92.00	76.50	85.25

(3) N = 300

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	93.00	96.00	74.00	97.00
AR(2)	87.00	88.00	65.00	96.00
AR(5)	81.00	78.00	65.00	55.00
AR(10)	85.00	76.00	81.00	60.00
overall	86.50	84.50	71.75	79.50

(6) N = 600

series	Criteria			
	OPA	NOPA	AIC	BIC
AR(1)	94.00	95.00	70.00	99.00
AR(2)	91.00	95.00	72.00	97.00
AR(5)	96.00	95.00	73.00	97.00
AR(10)	95.00	94.00	79.00	94.00
overall	94.00	94.75	73.50	96.75

7. A comparative study

In this section we summarize the results of a large simulation study which involves 2400 simulated AR series of various orders 1,2,5 and 10 and of various lengths 100,200,300,400,500 and 600 and with various criteria *OPA*, *NOPA*, *AIC* and *BIC*. The summary results are given in Table 1 . This section is concerned with explanations related to this table of summaries and further analysis and comparisons of the percentages given in this table. With obvious notation these four criteria are given by;

$$AIC(k) = n\ Log_e(MSE(k)) + 2k$$

$$BIC(k) = -n\ Log_e(MSE(k)) - (Log_e\ n)k$$

$$NOPA(k) = Log_{10} P(X \mid k);\ \text{with Non-Overlap data prior}$$

$$OPA(k) = Log_{10} P(x \mid k);\ \text{with Overlap data prior}$$

where $P(X \mid k)$ is the predictive density of X which is given by Equation 3.5 . We shall use the terminology correct automatic decision. By automatic decision we mean a decision that is actually made by a certain criterion rather than by the statistician from the context.

We do not, in general, recommend any purely automatic decision in real world applications. It is often the case that external information, expert opinion and prior information are available, which may be combined with *automatic information* to improve the identified model. In many cases, from the context, the time series may simply be decomposed into some components such as level, trend, seasonality and error components. Note that by *automatic information* we mean the order and level of goodness of fit that a certain criterion (rather than the context and the common sense) attaches to the competing models.

In the context of AR modelling, by *correct decision* we mean a case in which the identified order is equal to the true order (i.e. pin- point

posterior probability) of the AR(k) model, $k \in PD(k)=\{1,2,\ldots,U\}$, we set aside the first $n_0 (>U+2)$ observations denoting by $Y=(x_1,\ldots,x_{n_0})'$, where U is a preassigned upper limit to the order. The choice of the training sample size is fully discussed by Kheradmandnia (1991), where he recommend the interval $U+2 \leq n_0 \leq 30$. Let B_k be the $(n_0 \times k)$ corresponding AR(k) design matrix associated with Y; then in order to construct data base priors, in 3.3 we set $V_0=(B'_k B_k)^{-1}$, $m_0=V_0 B'_k Y$, $\sigma_0^2 = R_0/n_0$ where $R_0 = Y'Y - \hat{Y}'\hat{Y}$ and $Y=B_k m_0$. We will refer to the above approach as the overlap training sample approach. Once the predactive abilities of the AR(1), AR(2),... and AR(U) models have been calculated from 4.1 ,the posterior probabilities corresponding to the each of the U competing models can be calculated using Equations 4.5 or 4.4, depending whether or not P(k) is uniform. In the simulation studies and the numerical examples in this paper a uniform prior for the order of the AR model has been adopted. Note that in the (non-overlap) *training sample approach* n,X and A_k in Section 3 are replaced by n_1, X_1 and A_{1k} respectively, where $X_1=(x_{n_0+1}, \ldots, x_n)'$ is the remaining (non-training) sample and $A_{1k}=(F'_{n_0+1},\ldots,F'_n)'$ is the $(n_1 \times k)$ design matrix corresponding to X_1 where $n_1=n-n_0$.

6. Posterior distribution of θ and σ^2, given $\hat{k}$

Once the mode of posterior distribution of k has been drived, $\hat{k}$ say, the AR model may be written as

$$X = A_{\hat{k}}\theta + \varepsilon \; ; \; \varepsilon \sim N(0, \sigma^2 I_n) \qquad (6.1)$$

On assuming the improper priors $(\sigma^2 | \hat{k}) = \sigma^{-2}$ and $P(\theta | \sigma^2, \hat{k}) = \sigma^{-\hat{k}}$ the marginal posterior distributions of $(\sigma^2 | \hat{k}, X)$ and $(\theta | \hat{k}, X)$ are given by

$$(\sigma^2 | \hat{k}, X) \sim IG\left(\frac{n}{2}, \frac{nMSE}{2}\right) \text{ and } (\theta | \hat{k}, X) \sim T_n(m, V\,MSE) \qquad (6.2)$$

respectively where $m=VA'_{\hat{k}}X$, $V=(A'_{\hat{k}}A_{\hat{k}})^{-1}$, $MSE=R/n$, $R=X'X-\hat{X}'\hat{X}$ and $\hat{X}=A_{\hat{k}}m$.

PA(k) is a suitable measure of goodness of fit of the AR(k) model. We denote the Bayes Criterion for an AR(k) model by BC(k) and define it to be the base 10 logarithm of $P(X \mid k)P(k)$ so that

$$BC(k) = \text{Log}_{10} P(X \mid k) + \text{Log}_{10} P(k) = PA(k) + \text{Log}_{10} P(k) \qquad (4.2)$$

If some information was available a priori about k and just a point estimate of k was of interest then, BC(k) is a suitable measure of goodness of fit of an AR(k) model. The third and most important measure of goodness of M is the posterior probability of that model, denoted by $P(M \mid X)$, and in particular for an AR(k) model denoted by $P(k \mid X)$ which is given by :

$$P(k \mid X) = C\, P(X \mid k)\, P(k) \qquad (4.3)$$

where C is a normalizing constant. For practical computational reasons it is easier (if not necessary) to first calculate the predictive ability or the Bayes criterion. It is therefore useful to write $P(k \mid X)$ in terms of BC(k) or PA(k) as follows

$$P(k \mid X) = C\, 10^{BC(k)} = \left[\sum_{j=1}^{U} 10^{BC(j)}\right]^{-1} 10^{BC(k)} \qquad (4.4)$$

If a uniform prior for k be adopted then we have

$$P(k \mid X) = C\, 10^{PA(k)} = \left[\sum_{j=1}^{U} 10^{PA(j)}\right]^{-1} 10^{PA(k)} \qquad (4.5)$$

5. Overlap training sample device

In this section we are concerned with the construction of data priors i.e. assigning values to n_0, m_0, V_0, σ_0^2 and hence $R_0 = \sigma_0^2/n_0$ of Equations 3.3 and 3.8. Let $X=(x_1,\ldots,x_n)'$ be the effective observation vector and $A_k=(F'_1,\ldots F'_n)'$ be the $(n \times k)$ corresponding AR(k) design matrax. In order to calculate the predictive abilaty (and hence the

$$P(X|\ k) = C_k\ \Gamma\left(\frac{n+n_0+k}{2}\right) \int \left(\frac{Q_k}{2}\right)^{-\frac{n+n_0+k}{2}} d\theta$$

$$= C_k \Gamma\left(\frac{n+n_0+k}{2}\right) 2^{\frac{n+n_0+k}{2}} \int Q_k^{-\frac{n+n_0+k}{2}} d\theta \qquad (3.9)$$

and then integrated out θ from 3.9 to get 3.8.

If a vague prior for k is adopted i.e. $P(k) = \frac{1}{U}$, $k=1,2,\ldots,U$, then the posterior distribution of k is simply given by;

$$P(K \mid X) = \frac{P(X \mid k)}{\sum_k P(X \mid k)} \quad ; \quad k \in PD(k) \qquad (3.10)$$

4. Predictive Ability, Bayes Criteria and Posterior Probability

Here we introduce three related measures of goodness of fit of a model. The first is the predictive ability, measured by the observed value of the predictive density. The terminology *"predictive ability"* was first introduced by West (1986). It is more practical to define the predictive ability of a model as the observed value of the (base 10) logarithm of the predictive density, as we do in this paper . We take logs to produce more managable numbers. We denote the predictive ability of a model M by PA(M) and in particular the predictive ability of an AR(k) model by PA(k), so that:

$$PA(k) = \text{Log}_{10}\ P(X \mid k) = \text{Log}_{10} \iint P(X|\theta, \sigma^2, k) P(\theta, \sigma^2 |\ k)\, d\theta\, d\sigma^2 \qquad (4.1)$$

where $P(X \mid k)$ is given by equation 3.8. If just a point estimate of k was of interest and no information was available a priori about k, then

this sum can be written as a quadratic form in θ plus another term not involving θ. Once written in this form it is easy to integrate out θ to get.

$$P(X|\ k) = \frac{(R_0/2)^{\frac{n_0}{2}}}{\Gamma(n_0/2)} (2\pi)^{-\frac{n}{2}} \{|V_0\ ||\ V_0^{-1} + \quad V^{-1}|\}^{-\frac{1}{2}} \times$$

$$\int (\sigma^2)^{-\frac{n_0+n}{2}-1} \exp\{-\frac{R_0 + R + (m-m_0)'(V+V_0)^{-1}(m-m_0)}{2\sigma}\}\ d\sigma^2 \tag{3.7}$$

where $R_0 = n_0\ \sigma_0^2$, and $R = n$ MSE is the residual sum of squares from a least squards fit, $MSE = (X'X - \hat{X}'\hat{X})/n$, $\hat{X} = A_k\ m$, $m = V\ A'_k X$ and $V = (A'_k A_k)^{-1}$. The above integrand has the form of an inverse Gamma distribution so that this integral is the inverse of the missing normalizing constant of that distribution. Thus:

$$P(X\ |\ k) = \frac{R_0^{n_0/2}}{\pi^{n/2}} \frac{\Gamma(\frac{n_0+n}{2})}{\Gamma(\frac{n_0}{2})} \{|V_0||\ V_0^{-1} + V^{-1}|\}^{-\frac{1}{2}} \times \tag{3.8}$$

$$\{R_0 + R + (m-m_0)'(V_0+V)^{-1}\ (m-m_0)\}^{-(n_0+n)/2}$$

For relevant discussion see Lempers (1971), Atkinson (1978) and Smith and Spiegelhalter (1980).

Note that, since 3.4 has the form of an Inverse Gamma distribution;

$$IG\ (\frac{n+n_0+k}{2},\ \frac{Q_k}{2})$$

we could have, first, integrated out σ^2 from 3.4 to get

$$P(X|\ k) = \iint P(X \mid k,\ \theta,\ \sigma^2)\ P(\ \theta,\ \sigma^2 \mid k)\, d\theta\, d\sigma^2 \qquad (3.2)$$

The first term in the intergrand is the likelihood defined by 2.3. The second term is the joint prior of θ and σ^2 conditional on k;

$$P(\theta\ ,\sigma^2 \mid k) = P(\ \theta|\sigma^2,\ k)P\ (\sigma^2 \mid k)$$

On assuming a priori

$$(\sigma^2 \mid k) \sim IG\left(\frac{n_0}{2},\ \frac{n_0\ \sigma_0^2}{2}\right) \text{and } (\theta|\sigma^2,k) \sim N(m_0, \sigma^2 V_0) \qquad (3.3)$$

and combining these priors with the likelihood the integrand is obtained as :

$$C_k(\ \sigma^2)^{-\frac{n+n_0+k}{2}-1}\ \exp\left(-\frac{Q_k}{2\ \sigma^2}\right) \qquad (3.4)$$

where :

$$Q_k = n_0\ \sigma_0^2 + (\theta - m_0)'\ V_0^{-1}(\ \theta - m_0) + (X - A_k\ \theta)'(X - A_k\ \theta) \qquad (3.5)$$

and

$$C_k = (2\ \pi)^{-\frac{k}{2}} \left(\frac{n_0\ \sigma_0^2}{2}\right)^{\frac{n_0}{2}} / \ \Gamma\left(\frac{n_0}{2}\right) \qquad (3.6)$$

Note that by IG we mean Inverse Gamma distribution. The random variable x is said to have an inverse Gamma distribution with parameters α and β, written $x \sim IG(\alpha, \beta)$, if its reciprocal has a Gamma distribution i.e. if $\frac{1}{x} \sim \Gamma(\alpha, \beta)$. The P.d.f. of x is given by :

$$P(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\ x^{-\alpha-1}\ e^{-\beta/x}\ ;\ \ x > 0\ ;\ \alpha \geq 1;\ \beta > 0$$

In (3.5) Q_k involves the sum of two quadratic forms in θ plus another term not involving θ. On using the matrix identities of Smith (1973),

to represent an AR(k) model, where $F_t=(x_{t-1},x_{t-2},\ldots,x_{t-k})$, $(t=1,2,\ldots,n)$, and $\theta=(\phi_1,\phi_2,\ldots,\phi_k)'$. In matrix form, an AR(k) model may be defined by

$$X= A_k\theta + \varepsilon \; ; \quad \varepsilon \sim N(0, \sigma^2 I_n) \qquad (2.3)$$

where $A_k=(F_1',F_2',\ldots,F_n')'$ and k is the unknown order of AR model. Suppose that there are U competing AR models AR(1), AR(2),...and AR(U).

We then use the first U observations as initial values, so that we represent a time series of size N by :

$$\{x_{-U+1}, \ldots, x_{-1}, x_0, x_1, x_2, \ldots, x_n\}$$

where $N=n+U$ and U is a preassigned upper limit to the order. We will refer to $X=(x_1, \ldots, x_n)'$ as the (effective)observation vector.

3. Posterior distribution of the order of an AR model

Given the stationary AR(k) model defined by (2.3),

Let p (k) be the prior distribution of the order of the AR model with permissible domain $PD(k)= \{1,2,\ldots,U\}$,where U is a preassigned upper limit to the order. Specifically we have ;$0< Pr(k=j)< 1$ and $\sum_{j=1}^{U} Pr(k=j)=1$, $(j=1,2,\ldots,U)$. P(j) is the prior protability that X is generated by an AR(j)model.By Bayes theorem the posterior distribution of k is given by

$$P(k \mid X) \propto P(X \mid k)\, P(k) \qquad (3.1)$$

where $P(X \mid k)$ is the predictive (marginal) density of X (given the AR(k) model), which is a likelihood function for k given by :

sample. Furthermore, in the context of AR modelling, we know of no method that can satisfactorily pick the true order of an AR model out of a simulated AR series if the sample size is rather small and no prior information is available about the distribution of the order.

In order to illustrate the practical utility of the overlap approach, in this paper we apply the overlap training sample device to obtain the posterior distributions of the order of AR models using some simulated AR(1), AR(2), AR(5) and AR(10) series.

In this paper we propose a new model selection criterion that we shall refer to as overlap predictive ability (OPA). Through 2800 simulated AR series of various order and various length we compare the utility of the new criterion with the wellknown exisiting criteria AIC and BIC. Further details can be found in author's ph.D thesis (kheradmandnia (1991)).

2. Notations and definitions

An autoregressive model of order k, AR(k), is defined by :

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_k x_{t-k} + e_t ; \quad e_t \sim N(0,\sigma^2) \qquad (2.1)$$

where x_t is the observation at time t,$(t=1,2,\ldots,n)$, ϕ_j,$(i=1,2,\ldots,k)$, are unknown autoregressive parameters and σ^2 is the unknown noise variance.

We sometimes use the alternative notation :

$$x_t = F_t \theta + e_t \; ; \; e_t \sim N(0, \sigma^2) \qquad (2.2)$$

we adopt improper priors for model parameters,the posterior distribution of the competing models involves an arbitrary, unspecified constant (. see e.g. Spiegelhalter and Smith 1982) so that the use of improper priors is not justified. If however the competing models have the same dimension, in certain cases, the use of improper priors leads to satisfactory results. A satisfactory solution to the problem of assignment of the arbitrary constant resulting from the use of an improper prior,for linear & loglinear models, is the concept of an Imaginary training sample device introduced by Spiegelhalter and Smith (1982). Lempers (1971) and Atkinson (1978) suggested the idea of training sample device. They proposed that, part of the data (the training sample) be set aside and be used to provide proper data priors and that these priors be combined with the remaining (non-training sample) to provide posteriors. Spiegelhalter and Smith (1982) argue that the training sample approach is impractical if the complete data set is itself rather small or if the data are of such a nature that a training sample would be somewhat " untypical " . This criticism may be overcome by setting aside the first n_0 observations and using them to provide proper data priors, and then combining these priors with all the n observations (rather than just with the remaining $n_1=n-n_0$ observations) to provide posteriors. We will refer to this approach as the overlap training sample device in the sense that the actual sample overlaps the training sample, unlike the original training sample device where the actual sample is separate from the training

1. Introduction

In choosing one model amongst several models with different dimensions, such as AR models, the maximum likelihood principle leads to choosing the model with highest dimension. Akaike(1973) extended the maximum likelihood method by introducing his well-known criterion ,the Akaike Information Criterion (AIC) which is defined as follow.

AIC=(-2) Log(*maximum likelihood*)+2(*number of Parameters*)

He suggested choosing that model for which AIC is minimised. The AIC, for the case of AR models, takes (approximately) the form AIC(k) = $n \text{ Log}_e (\text{MSE } (k)) + 2k$, where n is the number of observations and MSE (k) is the mean square error corresponding to the k^{th} order AR model. Schwarz (1978), with some Bayesian arguments, proposed an alternative criterion (BIC). He suggested to choose the model for which log of maximum likelihood minus ($\frac{1}{2} k_j \text{ Log } n$) is the largest, where k_j is the dimension of the j^{th} model and n is the number of observations. The BIC, for the case of AR models, takes (approximately) the form $\text{BIC}(k) = -n \text{ Log}_e(\text{MSE}(k)) - (\text{Log}_e n) k$. In AIC (k) criterion if we replace the multiplier 2 by $\text{Log}_e n$ the resulting criterion will te identical to that of BIC(k). Various other similar criteria have appeared in the literature,see e.g. Bhansali and Downham (1977), Akaike (1978). Akaike (1979) gave a Bayesian extension of his minimum AIC procedure. Smith and Spiegelhalter (1980) introduced a unified view of a number of choice criterion for linear models in terms of Bayes factors.

As far as the Bayesian approaches are concerned,it is now well known that,if

A Bayesian Approach to Autoregressive Modeling

M. Kheradmandnia
School of Mathematics
Isfahan University of Technology
Isfahan, Iran

Abstract

In choosing one model amongst several models with different dimensions, The maxamum likelihood principle invariably leads to choosing the model with highest dimension. In order to overcome this problem Akaike (1973) introduced an information criterion (AIC). Later Schwarz(1978) introduced an alternative criterion (BIC). Various other similar criteria have appeared in the literature. Smith and spiegelhalter (1980) introduced a unified view of a number of choice criterion.

In this paper we introduce a new criterion (OPA) which is based on calculation of posterior probabilities of competing models using what we term as * *overlap data based priors* * for model parameters. Through a large number of simulated AR series we compare the extent to which different criteria can identify the true underlying generating mechanism of a series.

Key words : *Model selection , predictive ability , posterior mode , AIC , BIC , OPA .*

Confidence Bounds on Ratios of Variances in Unbalanced Two-Way Crossed Models

Kazem Kazempour*
Department of Statistics
University of Central Florida
Florida, U. S. A.
Franklin A. Graybill
Department of Statistics
Colorado State University
Ft. Collins, U.S.A.

Wed. 15:00 Hall 5

Approximate Confidence bounds on ratios of variances for unbalanced two-way crossed models with interactions and with missing observations are derived. A simulation study was conducted to evaluate the effectiveness of the bounds and to compare the proposed intervals to the cell mean procedure [1]. Balanced incomplete block (BIB) designs are used as a special case of design with missing cell, and confidence bounds on the variances are constructed for various design configurations.

Trade-Off Theory and Its Application

A. S. Hedayat

department of Mathematics, Statistics, and
Computer Science
University of Illinois
Chicago, U. S. A.

Wed. 14:30 Hall 5

Two collections $T1$ and $T2$ of treatment combinations forms a trade based on a given model if $T1$ and $T2$ share no treatment combinations and $T1$ and $T2$ have identical Fisher information matrices. Thus if a design contains $T1$ we can replace it with $T2$ without changing any statistical information in the design. The theory of trade-off deals with the characterization and construction of trades in many settings. Existing results are reviewed. Some new results and several open problems are discussed.

On Estimating the Variance of the Horvitz-Thompson Estimator

Muhammad Hanif*
Institute of Statistics
University of the Punjab
Lahore, Pakistan
P. Mukhopadhyay & S. Bhattacharyya
Indian Statistical Institute
Calcutta, India

Wed. 14:30 Hall 6

In this paper a new model-based variance estimator is proposed. It is along the lines of those suggested by Hanif and Brewer (1980) and Kumar-Gupta-Agarwal (1985). Empirical investigations show that the amount of bias is fairly small for all values of g and that it is more stable than all other estimators considered.

Table (2):comparision maximum likelihood and bayesian estimates

of parameters of largest value distribution

		δ=1.978			U=3.950		
METHOD	SAMPLE SIZE	δ-μ[δ]	δ [δ]		U-μ [U]	δ[U]	
M.L	10	-.02	.017		-.02	.158	
BAR		-.02	.015	1.15	-.02	.163	.95
M.L	30	-.02	.012		.00	.094	
BAY		-.02	.010	1.18	.01	.099	.89
M.L	50	-.02	.009		.01	.067	
BAY		-.02	.007	1.16	.02	.072	.84

(PARENT DISTRIBUTION WEIBULL SHAPE=2 SCALE=2)
M.L MAXIMUM LIKELIHOOD METHOD
BAY BAYESIAN METHOD
NO OF TRIALS 10,000

TABEL(1):ESTIMATION THE PARAMETERS OF THE LARGEST VALUE DISTRIBUTION USING

M.L.E & BAYESIAN METHOD (WEIBUI) AS PARENT, SCALE=2 , SHAPE=2

LARGEST DIS PARS NO LAR OBS:N=50	SCALE 1.978		LOCATION 3.956	
SAMPLE SIZE	M.L.E	BAY EST	M.L.E	BAY EST
10	1.992	2.269	4.057	4.000
15	2.000	2.276	3.958	3.914
20	1.996	2.177	3.897	3.860
25	2.004	2.256	3.931	3.897
30	1.996	2.149	4.050	4.025
35	2.001	2.183	3.978	3.953
40	1.996	2.114	3.935	3.914
45	2.001	2.160	3.985	3.963
50	2.000	2.143	4.020	4.000

9. Mann,N.R. (1968). Point and interval estimation procedures for the two parameters of the Weibull and extreme value distribution.Tech.vol. 10:231-256

References

1. Fisher, R.A., and L.H.C. Tippett, (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample, Proc. Cambridge phil. soc. 244:180.

2. Galambos, J. (1978), The Asymptotic Theory of Extreme Order Statistics, Wiley-Interscience, New York.

3. Greenwood, J.A., Landwehr, J.M., Matals, N.C., and Wallis, J.R. (1979). Probability weighted moments:Definition and relation to parameter of several distributions expressible in inverse form.Water Resources res, 15,1049-1059.

4. Gumbel, E.J. (1949). The statistical Forecast of Floods. Ohio Water Resources Board, Bull.15.

5. Hakimi, S.J. (1988) Application of Extreme Value Theory, Ph.D. thesis, Universiy of Bradford.

6. Jeffreys. H. (1961). Theory of Probability. Oxford at the Claredon Press.

7. Kimball, B.F.(1946). Sufficient statistical estimation functions for the parameters of the distribution of maximum values, Ann. Math.Stats., 17:299.

8. Landwehr, J.M., Matalas, N.C., and Wallis,J.R.(1979) Probability Weighted Moments compared with some Traditional Techniques in Estimating Gumbel parameters and Quantiles, Water Resources Res, 15, 1055-1064.

The Bayesian method has advantage that like the (PWM) method it always yield feasible parameter values and had an advantage over the (MLE) that either iteration was not required or the number of iteration was small.

Sample sizes were taken as n=10 to 50 (steps of 5). The results are given in Table 1. It is clear that Bayesian estimators are close to MLE estimators.

The method explained in section 3 can now be used to compare two methods. The results of this comparison are given in Table 2.

From the analysing of these results one can deduce the following:

1-The estimation of the location by MLE is better than that of Bayes.

2-The estimation of the scale by Bayes is better than that by MLE.

5: CONCLUSIONS

In terms of the overall performance of the moment(MoM) probability weighted moment(MWM), maximum likelihood estimation (MLE) and Bayesian (BAY) estimation, the method of (MLE) was generally the most efficient method of the four; except for the case of small samples the method (PWM) and (BAY) are more efficient.

Using (PWM), the parameters and quantiles of (EV) can be estimated.The method have several advantages in comparison with (MoM) and (MLE). The estimators are fast and straightforward to compute and always yield feasible values for the estimated parameters.

Assume a Weibull distribution with shape β and scale α as the parent for the largest distribution. [Hakimi,1988] has shown that the distribution of the largest value is given by :

$$G(x)=\exp(-\exp(-h(x_o)(x-x_o)) \qquad (11)$$

where x_o and $h(x_o)$ are given by

$$R(x_o)=\exp\left[-\left(\frac{x_o}{\alpha}\right)^{\beta}\right]=\frac{1}{N}$$

i.e

$$x_o=\alpha\,(\log N)^{1/\beta} \qquad (12)$$

and

$$h(x_o)=\frac{\beta}{\alpha}\left(\frac{x_o}{\alpha}\right)^{\beta-1}=\frac{\beta}{\alpha}[-\log N]^{(\beta-1)/\beta} \qquad (13)$$

where N is the number of largest value. Accordingly, from (11),(12), and (13) the appropriate inverse probability function for the largest value is

$$x=x_o-\log[-\log(R)]/h(x_o)$$

$$=\alpha\,(\log N)^{1/\beta}-\log[-\log(R)]/\frac{\beta}{\alpha}(\log N)^{(\beta-1)/\beta} \qquad (14)$$

where R is uniformly distributed over[0,1]. Then for any value of α, β and N, the values of x_o and $h(x_o)$ can be determined. The largest value can be now generated using the Monte Carlo method.

One can now estimate the location and scale parameters of the largest value distribution both by using maximum likelihood and by Bayesian method.

The above comparison was carried out for $\alpha=2$ and $\beta=2$.

where M is a large number. In the present experiment M is taken equal to 1000.However M has to be chosen taking computer time and capacity limitation into account.

For sufficiently large M, $\mu[\hat{\zeta}]$ and $\sigma[\hat{\zeta}]$ should be very good approximations to the analytical values.

For each of the two methods and different sample size n the bias of the estimate of ζ can be computed as

$$\text{Bias} = \zeta - \mu[\hat{\zeta}]$$

and the mean square error is given by

$$\text{MSE}(\hat{\zeta}|\, n, \text{method}) = (\text{Bias})^2 + \sigma^2[\hat{\zeta}]$$

The relative efficiency of the two method in estimating ζ given a sample size of n is defined as

$$\phi(\hat{\zeta}, \text{method 1, method 2}\,|n) = \frac{\text{MSE}(\hat{\zeta}|\, n, \text{method 1})}{\text{MSE}(\hat{\zeta}|\, n, \text{method 2})}$$

If $\phi < 1$, then method 2 is less efficient than method 1 in estimating ζ.If $\phi > 1$, then the opposite relationship holds and $\phi=1$ means that the two method are compatible (Landwehr et al, 1979).

4 : COMPARISON MLE AND BAYESIAN METHODS OF ESTIMATION OF THE LARGEST VALUE DISTRIBUTION PARAMETERS(WEIBULL AS A PARENT)

$$P(\alpha, x_0, \underline{x}) \propto \alpha^{n-1} e^{-\alpha\Sigma(x_i - x_0)} e^{-\Sigma e^{-\alpha(x_i - x_0)}}$$

the normalising constant can be obtained in the form

$$C_n = \int_{-\infty}^{+\infty}\int_0^{\infty} \alpha^{n-1} e^{-\alpha\Sigma(x_i - x_0)} e^{-\Sigma e^{-\alpha(x_i - x_0)}} d\alpha dx_0$$

then by definition, the bayesian estimate of α and x_0 are given respectively by:

$$\hat{\alpha} = C_n^{-1} \int_{-\infty}^{+\infty}\int_0^{+\infty} \alpha\, P(\alpha, x_0 | \underline{x})\, d\alpha dx_0 \qquad (7)$$

$$\hat{x}_0 = C_n^{-1} \int_0^{+\infty}\int_{-\infty}^{+\infty} x_0\, P(\alpha, x_0 | \underline{x})\, dx_0\, d\alpha \qquad (8)$$

The latter integration cannot analytically be evaluated and numerical methods are required. The method used to deal with this limitation is that of the trapezoidal rule

3 : METHOD OF COMPARISON

If $\hat{\varsigma}$ denotes an estimate of ς, then on the basis of M samples of size n, the mean and standard deviation of $\hat{\varsigma}$ respectively is given by

$$\mu[\hat{\varsigma}] = \sum_{i=1}^{M} \hat{\varsigma}_i / M \qquad (9)$$

$$\sigma[\hat{\varsigma}] = \left\{ \sum_{i=1}^{M} \hat{\varsigma}_i^2 / M - \left(\sum_{i=1}^{M} \hat{\varsigma}_i / M \right)^2 \right\}^{1/2} \qquad (10)$$

$$\alpha = [M_{(o)} - 2M_{(1)}]/\ln(2).$$

It is now required to $\hat{M}_{(k)}$. An unbiased estimate of $M_{(k)}$ is given by (see appendix of Landwehr et al).

$$\hat{M}_{(k)} = \frac{1}{n}\sum_{(i)} x_{(i)} \frac{C_k^{a-i}}{C_k^{a+1}}$$

$x_{(i)}$'s are the ith order statistics. Cosequently, the estimation of the Gumbel parameters are given by:

$$\begin{cases} \hat{\alpha} = [\hat{M}_{(o)} - 2\hat{M}_{(1)}]/\ln(2) & (5) \\ \hat{x}_{\circ} = \hat{M}_{(o)} - \alpha\gamma & (6) \end{cases}$$

2.4 Bayesian Estimation Method

According to Jeffreiy's rule [Jeffreiy, 1961] a prior distribution can be assumed as:

$$\begin{cases} \pi(\alpha, x_{\circ}) \propto \frac{1}{\alpha} \\ \alpha \in]0, +\infty[\\ x_{\circ} \in]-\infty, +\infty[\end{cases}$$

Then the posterior distribution is given by :

(PWM). The PWM is defined as:

$$M_{l,j,k} = E[\, x^{l} F^{j} (1-F)^{k} \,]$$

l,j,k are real numbers see (Greenwood etal, 1979) for more detail.

The analytical form of PWM for Gumbel distribution is given by:

$$\begin{cases} F(x) = \exp\{-\exp[-(\dfrac{x-x_\circ}{\alpha})]\} \\ x \in\,]-\infty\ ,+\infty[\end{cases}$$

Thus for the Gumbel distribution the inverse is:

$$x(F) = x_\circ - \alpha \ln(-\ln F)$$

If $M_{(o)} = M_{1,o,o}$ then by definition,

$$M_{(o)} = \int_0^\infty (x_\circ - \alpha \ln(-\ln(F)))dF = x_\circ + \alpha\gamma$$

$$x_\circ = M_{(o)} - \alpha\gamma$$

$$M_{(1)} = \int_0^\infty x_\circ - \alpha \ln(1-\ln F)(1-F)dF$$

$$= M_{(o)} + \alpha \int_0^\infty -F\ln(-\ln F)dF$$

From the transformation $y=-\ln(-\ln F)$: it yields to

$$L(\underline{x}, x_o, \alpha) = \alpha^n e^{-\alpha(\bar{x} - x_o)} e^{-\alpha\Sigma(x_i - x_o)} \qquad (2)$$

Suppose that $\bar{x}$ and s are mean and standard deviation respectively of the sample. As the likelihood function cannot be expressed in terms of $\bar{x}$, s^2, α and x_o, then by definition the estimators are not sufficient. Kimball (1946) introduced a set of new sufficient statistics. The likelihood function can be given as:

$$f = \ln L(\underline{x}, x_o, \alpha)$$

$$= n \ln(\alpha) - \Sigma y_i - \Sigma e^{-y_i}$$

where $y_i = (x_i - x_o)/\alpha$ also called the reduced variate.

Maximum likelihood estimates can then be computed using the following equations:

$$\frac{\delta f}{\delta x_o} = \frac{n}{\alpha} - \frac{1}{\alpha} \Sigma e^{-y_i} = 0 \qquad (3)$$

$$\frac{\delta f}{\delta \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha} \Sigma y_i - \frac{1}{\alpha} \Sigma y_i e^{-y_i} = 0 \qquad (4)$$

The solution to these two equations has been discussed extensively in (Mann, N.R, 1968).

2.3 Probability Weighted Moments

If the inverse form of distribution function F=F(x) can be written in the form x=x(F), then the distribution can be characterised by probability weighted moments

For the extreme value distribution, there are two parameters to be specified x_o and α. These are related to the population moments μ'_1 and μ'_2 by the following equations:

$$\begin{cases} \mu'_1 = x_o + \Psi\alpha \\ \mu'_2 = \dfrac{\pi^2\alpha^2}{6} \end{cases}$$

$\Psi = 0.5772$ Ψ is Euler's constant

This leads to the following formulae for x_o and α

$$\begin{cases} \hat{\alpha} = \sqrt{\dfrac{6}{\pi^2}}\sqrt{\hat{\mu}'_2 - \hat{\mu}'_1} \\ \hat{x}_o = \hat{\mu}'_1 - \Psi\sqrt{\dfrac{6}{\pi^2}}\sqrt{\hat{\mu}'_2} \end{cases}$$

Where

$$\begin{cases} \hat{\mu}'_1 = \bar{x} = \sum_i x_i / n \\ \hat{\mu}'_2 = (\sum_i x_i^2 - n\bar{x}^2)/(n-1) \end{cases}$$

2.2 Maximum likelihood method

The maximum likelihood methodology is probably the most used method of estimation. For a complete sample, the likelihood function is expressed as :

The methodology presented deals with the parameters estimation techniques. A review of moments and probability weighted moments procedures is first given. In addition, an extensive study of MLE and Bayes is explored. Finally a comparison between MLE and Bayes is given.

The objective of the statistical theory of extrem value (E-V) is to analyse observed extreme values and ultimately forcast future extremes. Extreme value distribution have used in a wide range of application such as hydrology [Gumble,1949], reliability [Galambos, 1978] and quality control.

2;ON THE ESTIMATION OF THE GUMBEL DISTRIBUTION PARAMETERS

2.1 Method of Moments

The method of moments makes use of the fact that if all moments of a distribution are known then every thing about the distribution is also known.If there are k parameters to be estimated, the method consists of expressing the first k pupulation moments in terms of these k parameters.Then one can equate the parameters to the corresponding sample moments and taking the solutions of the resulting equations obtain estimates of the parameters.

Bayesian Method Compared with Some Traditional Techniques in Estimating Gumbel Distribution Parameters

J. Hakimi Sibooni
Department of Statistics
Shahid Chamran University
Ahwaz , Iran

ABSTRACT

Maximum Likelihood(MLE),Moments(MOM) and Probability Weighted Moment (PWM) techniques were developed for the estimation of the parameters of Gumbel distribution A Bayesian method for estimating the largest value distribution parameters was also developed. A comparison of the accuracy using Monte Carlo Simulation and advantages of these estimation parameters are given.

1. INTRODUCTION

Let F(x) be any distribution function. If F(x) belongs to the exponential family, and that a limiting distribution for the largest value exists, then the limiting of this distribution is given by

$$G(x)=\exp[-\exp(-\frac{x-x}{a}^{*})] \qquad (1)$$

This is called largest distribution Type I [Fisher & Tippett,1928].

References

CLATWORTHY, W.H. (1973). *Tables of Two Associate Class Partially Balanced Designs.* U.S. Department of Commerce. National Bureau of Standards.

HEDAYAT, A.S., JACROUX, M. & MAJUMDAR, D. (1988). Optimal designs for comparing test treatments with controls. *Statist. Sci.* 3, 462-91.

FREEMAN, G.H. (1976). A cyclic method of constructing regular group divisible incomplete block designs. *Biometrika* 63, 555-8.

GRAYBILL, F.A. (1983). *Matrices with Applications in Statistics.* California: Wadsworth.

JOHN, J.A. & TURNER, G. (1977). Some new group divisible designs. *J. Statist. Plann. Inf.* 1, 103-7.

JOHN, P.W.M. (1980). *Incomplete Block Designs.* New York: Marcel Dekker.

JONES, B. & ECCLESTON, J.A. (1980). Exchange and interchange procedures to search for optimal designs. *J. R. Statist. Soc. B* 42, 238-43.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. *In A Survey of Statistical Design and Linear Models,* Ed. J.N. Srivastava, pp. 333-53. Amsterdam: North-Holland.

MAJUMDAR, D. (1986). Optimal designs for comparisons between two sets of treatments. *J. Statist. Plann. Inf.* 14, 359-72.

MARSHALL, A.W. & OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* New York: Academic Press.

SEARLE, S.R. (1971). *Linear Models.* New York: Wiley & Sons.

Proof of Theorem 3.Necessity. If $d \in D^*(n,2,b,k)$ has part-balance then $H\Omega_d H'$ is given by (5) and its inverse can be formulated as (8). Setting the product of the matrices equal to I_{2p} and equating submatrices, it can be shown that $A_i = a_u I_p + a_v J_p$, where u=2i, v=u+1 for i=2,3,4 and a_i (i=1,..., 9) is a function of α_j, β_j, δ and ϕ (i=1,2). Since $A_d 1_t = 0$, it follows that $a_1' = (a_2 \;\; a_3) \otimes 1_p'$ where a_j (j=2,3) is a function of the a_i's $(4 \le i \le 9)$.

Sufficiency. If A_d is given by (6) then, in the lemma,

$$U = \begin{pmatrix} a_4 I_p + a_5 J_p & a_6 I_p + a_7 J_p \\ a_6 I_p + a_7 J_p & a_8 I_p + a_9 J_p \end{pmatrix}.$$

It can be shown that U^{-1} has the same form as U with b_i replacing a_i for i=1,..., 9, where each b_i is a function of the a_i's, see Graybill(1983, p195). From (9), $H\Omega_d H' = H_1 U^{-1} H_1'$ which can be expressed, after some algebra, in the same form as U with $a_4 = b_8$, $a_5 = b_9$, $a_6 = b_8 - b_6$, $a_7 = b_9 - b_7$, $a_8 = b_8 + b_4 - 2b_6$ and $a_9 = b_9 + b_5 - 2b_7$.

3. *Proof of Theorem 5*

For fixed values of n, b, k and r let

$Q(\lambda_2) = (v_A + v_B)/(k\sigma^2) = 2/s_1 + [\{r(k-1) + \lambda_2\}/s_1 + \lambda_2/r]/(ks_2)$,

where v_A and v_B are given in Corollary 4. The theorem is proved by establishing that $\partial Q/\partial \lambda_2 \ge 0$ for $k \ge 4$, with strict inequality when $k > 4$.

and θ_i (i=1, ..., t-1), where we choose u so that $\theta_0 \geq \theta_1$. The eigenvalues of Ω_d, in decreasing order, are λ_{t-1}^{-1} (i=1, ..., t-1) and $\alpha_0 = 0$. Again applying Marshall and Olkin (1979, Theorem A.4),

$$tr(C^{*}\Omega_d C^{*\prime}) \geq \theta_0\alpha_0 + \sum_{i=2}^{t} \theta_{i-1}\lambda_{t-1}^{-1}.$$

2. *Proof of Theorem 3*

Lemma . *Let $d \in D^{*}(n,2,b,k)$ and A_d be expressed as*

$$A_d = \begin{pmatrix} a_0 & a_1' \\ a_1 & U \end{pmatrix},$$

where a_0 is a scalar, a_1 is a $2p\times 1$ vector, and U is a nonsingular matrix formulated as

$$U = \begin{pmatrix} A_2 & A_3 \\ A_3' & A_4 \end{pmatrix},$$

for A_i (i=2,3,4) $p\times p$ matrices. Then, for H in (2) with m=2, and any generalized inverse Ω_d of A_d,

$$(H\Omega_d H')^{-1} = \begin{pmatrix} A_2 + A_4 + A_3 + A_3' & -A_2 - A_3' \\ -A_2 - A_3 & A_2 \end{pmatrix} \tag{8}$$

Proof. From Searle(1971,p2), a generalized inverse, Ω_d^0 of A_d is formed by adjoining a first row and first column of zeros to U^{-1}. Now

$$H\Omega_d H' = H\Omega_d^0 H' = H_1 U^{-1} H_1', \tag{9}$$

where $H_1 = \begin{pmatrix} 0_{p,p} & I_p \\ -I_p & I_p \end{pmatrix}$.

Hence $(H\Omega_d H')^{-1} = (H_1')^{-1} U H_1^{-1}$ and (8) follows.

less than 13.3%. Further group divisible designs can be used in the same way, for example the 19 regular designs in Freeman(1976) and the 14 regular or semi-regular designs in John and Turner(1977).

ACKNOWLEDGEMENTS

A. Gerami was supported by a University of Southampton Research Studentship. We are grateful to Dr A. Grieve of ICI Pharmaceuticals and Dr J.N.S. Matthews of the University of Newcastle-Upon-Tyne for helpful early discussions.

APPENDIX

1. *Proof of Theorem 1*

(i) $L \geq t$: Let ξ_i be a $t \times 1$ normalized eigenvector of A_d with eigenvalue λ_i (i=1, ..., t-1). Using the Moore-Penrose generalized inverse, $tr(C\Omega_d C') = \sum \lambda_i^{-1} \xi_i' C' C \xi_i = tr(\Gamma' C' C \Gamma)$, where $\Gamma = (\gamma_1, \ldots, \gamma_{t-1}, 1_t/(at))$, for $a > 0$ such that $a^2 t \leq \lambda_{t-1}$ and $\gamma_i = \lambda_i^{-1/2} \xi_i$, $(i = 1, \ldots, t-1)$.

From Marshall and Olkin (1979, Theorem A.4),

$$tr(\Gamma' C' C \Gamma) \geq \sum_{i=1}^{t-1} \theta_{t-i+1} \eta_i,$$

where $\eta_1 \geq \eta_2 \geq \ldots \geq \eta_t$ are the eigenvalues of $\Gamma'\Gamma = diag(\lambda_1^{-1}, \ldots, \lambda_{t-1}^{-1}, a^{-2}t^{-1})$. Hence $\eta_1 = a^{-2}t^{-1}$ and $\eta_i = \lambda_{t-i+1}^{-1}$ (i=2, ..., t). It follows that

$$tr(C\Omega_d C') \geq \sum_{i=1}^{t} \frac{\theta_{t-i+1}}{\lambda_{t-i+1}},$$

where the first term in the sum is zero .

(ii) $L=t-1$: If $C^{*'} = (u1_t, C')$, where u is an arbitrary constant, then $tr(C\Omega_d C) = tr(C^* \Omega_d C^{*'})$. It is easily shown that the eigenvalues of $C^* C^{*'}$ are $\theta_0 = u^2 t$

Group divisible designs are of three types: singular ($\lambda_1 = r$), semi-regular ($\lambda_1 < r$ and $t_1\lambda_2 = k_1 r$) or regular ($\lambda_1 < r$ and $t_1\lambda_2 < k_1 r$). The following theorem enables a choice to be made between the reinforced designs arising from the three types. Its proof is outlined in the Appendix.

THEOREM 5. *The average variance of the dual versus single contrast estimators for design d in Theorem 4 is a function of λ_2 which is strictly increasing when $k > 4$, and nondecreasing when $k=4$.*

Theorem 5 shows that the average variance is a non-increasing function of λ_1 since, on applying a well known property of group divisible designs,

$$r(k-2) = \lambda_1 + 2(n-2)\lambda_2. \tag{7}$$

Hence, for $k \geq 4$, if a singular group divisible design exists for the parameter values, it yields a reinforced group divisible design with smaller average variance than the other two types. In addition, for experiments for which both reinforced regular and semi-regular designs are available, the reinforced regular designs have smaller average variance.

An advantage of the designs is that they are readily available from published group divisible designs. To assess their performance a study was made of all the group divisible designs in Clatworthy(1973). Where there was a choice of design for reinforcing, Theorem 5 was used to choose the best design, giving a total of 160 designs. For k=3 it was found that one or more copies of the disconnected set of blocks {(i0 i1); i=1,...,n-1} was preferred to the Clatworthy design, and hence these designs were used for k=3 experiments in the study.

The 57 reinforced singular designs were found to have maximum discrepancy of 12.3%, with 72% of the designs having discrepancy of 7% or less. The 75 reinforced regular designs have maximum discrepancy 27.2%, and 49% of these have discrepancy less than 13.4%. For the 28 reinforced semi-regular designs, the maximum discrepancy is 27.2%, with 64% of the designs having a discrepancy

THEOREM 4. *Any design d, constructed by the above method, has*

(*i*) *part-balance for the dual versus single treatment comparisons, and*

(*ii*) $$H\Omega_d H' = \begin{pmatrix} x_1 I_p + y J_p & (x_1 - x_2) J_p \\ (x_1 - x_2) J_p & 2(x_1 - x_2) J_p \end{pmatrix},$$

where $x_1 = \{r(k-1) + \lambda_2\}/(s_1 s_2)$, $x_2 = (\lambda_1 - \lambda_2)/(s_1 s_2)$, $s_1 = r(k-1) + \lambda_1$, $s_2 = \{r + 2(n-1)\lambda_2\}/k$ *and* $y = \lambda_2/(r s_2)$.

Proof. The first part follows from Theorem 3 since

$$A_d = \frac{1}{k} \begin{pmatrix} b(k-1) & -r 1'_{2p} \\ -r 1_{2p} & F \end{pmatrix},$$

where p=n-1 and $F = \{\{r(k-1) + \lambda_1\} I_2 + (\lambda_2 - \lambda_1) J_2\} \otimes I_p - \lambda_2 J_2 \otimes J_p$.

To prove (ii) take Ω_d to be F^{-1} augmented by a first row and column with every entry zero, see Searle(1971, p2). F^{-1} is found by applying $(X \otimes I_p + Y \otimes J_p)^{-1} = X^{-1} \otimes I_p - \{(X + pY)^{-1} Y X^{-1}\} \otimes J_p$, which holds for X and Y any square matrices for which all the inverses exist.

A useful feature of the designs is that the variances of the contrast estimators can be easily found from the parameters of d_1.

COROLLARY 4. *For d in Theorem 4, $v_A = var(\hat{\tau}_{i1} - \hat{\tau}_{i0}) = 2k\sigma^2/s_1$ and $v_B = var(\hat{\tau}_{i1} - \hat{\tau}_{01}) = \{r^2(k-1) + \lambda_2(rk + \lambda_1)\}\sigma^2/(r s_1 s_2)$, for i=1,..., n-1.*

The design of Example 3 in §2 was constructed from S1 in Clatworthy(1973) which has parameters r=2, $\lambda_1 = 2$ and $\lambda_2 = 1$. Hence $v_A = \sigma^2$ and $v_B = 0.875\sigma^2$. In general, the designs estimate the dual versus B comparisons more precisely than the dual versus A comparisons.

COROLLARY 5. *For d in Theorem 4, $v_B \leq v_A$, where $v_A = var(\hat{\tau}_{i1} - \hat{\tau}_{i0})$ and $v_B = var(\hat{\tau}_{i1} - \hat{\tau}_{01})$, $1 \leq i \leq n-1$.*

The proof, which is not given, involves showing that $v_A - v_B \geq 0$ by using relationships amongst the parameters of d_1.

EXAMPLE 3. n=4, b=3, k=5. The design with blocks (01 10 11 20 21), (01 10 11 30 31) and (01 20 21 30 31) is in D_H and has only 2% discrepancy with B.

EXAMPLE 4. n=3, b=2, k=5. Two copies of (01 10 20 11 21) form a design in D_H which achieves B and hence is A-optimal.

Devising construction methods for designs in D_H is simplified by the following theorem.

THEOREM 3. *A necessary and sufficient condition for* $d \in D^*(n,2,b,k)$ *to have part-balance for estimating the dual versus single treatment comparisons is that*

$$A_d = \begin{pmatrix} a_1 & a_2 1_p' & a_3 1_p' \\ a_2 1_p & a_4 I_p + a_5 J_p & a_6 I_p + a_7 J_p \\ a_3 1_p & a_6 I_p + a_7 J_p & a_8 I_p + a_9 J_p \end{pmatrix}, \tag{6}$$

where a_i *(i=1,...,9) are constants.*

The proof is outlined in the Appendix.

4. DESIGNS FROM REINFORCED GROUP DIVISIBLE DESIGNS

We give a simple method of constructing part-balanced designs for $2 < k < t$ in three stages:

1. Take a group divisible design d_1 with $m_1 m_2$ treatments, where $m_1 = 2$ and $m_2 = n-1$, replication r, b blocks, block size $k_1 = k-1$ and jth associates occurring together in λ_j blocks (j=1,2), see for example John(1980,chapter 4).

2. Map the treatment labels in d_1 to i0, i1 (i=1,..., n-1) so that the pairs of first associates are i0 and i1 (i=1,..., n-1).

3. Reinforce each block by the inclusion of 01.

The resulting design has replications r, b and r for each of the A-alone, B-alone and dual treatment combinations respectively.

2.3. *Assessment of bounds*

A numerical comparison of $B(H)$ and B_F for the parameter values $2 \leq n, m \leq 10$, $b \leq 30$ and $2 \leq k \leq 15$ shows that $B_F > B(H)$ when $k \geq t$, as in Example 2. When $k < t$, one bound is not uniformly tighter and hence we use $B = \sigma^2 \max(B(H), B_F)/\{(n-1)(m-1)\}$ as a lower bound on the average variance.

DEFINITION. The *discrepancy* in a design d compared with a bound B for the estimation of $H\tau$ is $\{tr(H\Omega_d H') - B\}/B$.

For $k > 3$ there are many designs which have zero, or very small discrepancy compared with B, as in Examples 3 and 4 below. When k is very small relative to t, the discrepancy in the best design $d \in D^*$ found from the computer algorithm of Jones and Eccleston (1980) can be as large as 20%; this should be borne in mind when comparing designs for k=2 or 3 with the bound.

3. PART-BALANCE IN $n \times 2$ DESIGNS

We define a property which is possessed by a large number of efficient designs.

DEFINITION. A design $d \in D^*(n, 2, b, k)$ has *part-balance for estimating the dual versus single treatment comparisons* if

$$H\Omega_d H' = \begin{pmatrix} (\alpha_1 - \beta_1)I_p + \beta_1 J_p & (\delta - \phi)I_p + \phi J_p \\ (\delta - \phi)I_p + \phi J_p & (\alpha_2 - \beta_2)I_p + \beta_2 J_p \end{pmatrix}, \tag{5}$$

where $\alpha_i \neq \beta_i$ (i=1,2), δ and ϕ are constants, H is (2) with m=2 and J_p is a $p \times p$ matrix with every element unity.

This structure ensures equal precision and correlation for the estimators of the dual versus A-alone contrasts and for the estimators of the dual versus B-alone contrasts. It permits two values for the correlations amongst the contrast estimators, according to whether the contrasts are orthogonal or not.

The class of part-balance designs $D_{II} \subset D^*(n, 2, b, k)$ contains efficient and, for some experiments, A-optimal designs.

Majumdar's result further establishes that the condition in Lemma 1 (ii) is sufficient for a design to be universally optimal in the sense of Kiefer(1975). However, the condition is an unfruitful means of obtaining designs for our problem, because it is satisfied mainly for sparse and large values of k. The cases of practical interest are n=m=3 and k=8k_0 (k_0=1,2,...) for which Lemma 1(ii) shows that complete block designs are universally optimal. Nevertheless Lemma 1 is useful for establishing a lower bound in conjunction with the following lemma.

LEMMA 2. *Let $r_i (i=1,2,...,L)$ be positive integer values and $\sum r_i = r$ be regarded as fixed. Also let $\bar{r} = [r/L]$, where $[.]$ means " integer part of ", then*

$$\sum_{i=1}^{L} \frac{1}{r_i} \geq \frac{2L\bar{r} + L - r}{\bar{r}(\bar{r}+1)}.$$

We partition the set of treatment combinations into the subsets of *A-alone, B-alone* and *dual treatment combinations* defined by

$\{i0; i = 1,...,n-1\}$, $\{0j; j = 1,...,m-1\}$ and $\{ij; i = 1,...,n-1; j = 1,...,m-1\}$ respectively. The total numbers of units in d allocated to each subset are denoted by T_A, T_B and T_D respectively.

THEOREM 2. *For $d \in D^*$, let T_A and T_B be fixed and such that $T_A \geq p$, $T_B \geq q$ and $T_A + T_B \leq bk - pq$. Also let $\bar{r}_A = [T_A/p]$, $\bar{r}_B = [T_B/q]$ and $\bar{r}_D = [T_D/pq]$. Then $tr(HR^{-1}H') \geq F(T_A, T_B)$, where*

$$F(T_A,T_B) = q\{\frac{2p\bar{r}_A + p - T_A}{\bar{r}_A(\bar{r}_A+1)}\} + p\{\frac{2q\bar{r}_B + q - T_B}{\bar{r}_B(\bar{r}_B+1)}\} + 2\{\frac{2pq\bar{r}_D + pq - T_D}{\bar{r}_D(\bar{r}_D+1)}\}.$$

Proof. Since T_A and T_B are fixed, $T_D = bk - T_A - T_B$ is also fixed and each summation in (4) can be separately minimized using Lemma 2.

COROLLARY 3. *Let $B_F = \min\{F(t_A,t_B); (t_A,t_B) \in T\}$, where $T = \{(t_A,t_B); t_A \geq p, t_B \geq q; t_A + t_B \leq bk - pq, t_A, t_B \in N^+\}$, for N^+ the set of positive integers. Then B_F is a lower bound on $tr(H\Omega_d H')$ for all $d \in D^*$.*

EXAMPLE 1. $B_F = 5.333 < B(H)$.

EXAMPLE 2. $B_F = 4.833 > B(H)$.

When D is the class of designs having b blocks of size k, $tr(A_d) = bk - \sum\sum n_{ij}^2/k$, where n_{ij} is the (i,j)th element of the incidence matrix, N_d and is maximized when $\sum\sum n_{ij}^2$ is minimized, that is when the n_{ij}'s are as equal as possible, subject to the constraint that $\sum n_{ij} = k$ for j=1,...,b. It is well-known that for $t \geq k$, $c_{max} = b(k-1)$ and occurs when the design is binary.

In order to apply Corollary 2 to find the lower bound B(H) on $tr(H\Omega_d H')$ for $d \in D^*$, we need the eigenvalues of

$$H'H = \begin{pmatrix} pI_q & O_{q,p} & -1_p' \otimes I_q \\ O_{p,q} & qI_p & -I_p \otimes 1_q' \\ -1_p \otimes I_q & -I_p \otimes 1_q & 2I_t \end{pmatrix}. \qquad (3)$$

These are: 2 with multiplicity (p-1)(q-1), $[p+2 \pm \{(p+2)^2 - 4p\}^{1/2}]/2$ each with multiplicity (q-1), $[q+2 \pm \{(q+2)^2 - 4q\}^{1/2}]/2$ each with multiplicity (p-1), $[p+q+2 \pm \{(p+q+2)^2 - 4(pq+p+q)\}^{1/2}]/2$ and 0.

EXAMPLES. m=n=3, t=8

1. b=8, k=3. $B(H)=(\sum\theta_i^{1/2})^2/\{b(k-1)\}=6.319$,

2. b=3, k=9. Take seven of the n_{ij}'s equal to 1 and one equal to 2 in each block to obtain $c_{max}=23.3$ and B(H)=4.333.

2.2. *Bound B_F*

A second bound, B_F, is obtained by applying Theorem 3.1 of Majumdar(1986).

LEMMA 1. *For any $d \in D^*$ and H in (2)*

$$(i) \quad tr(H\Omega_d H') \geq tr(HR^{-1}H') = q\sum_{i=1}^{p}\frac{1}{r_{Ai}} + p\sum_{j=1}^{q}\frac{1}{r_{Bj}} + 2\sum_{i=1}^{p}\sum_{j=1}^{q}\frac{1}{r_{Dij}}, \qquad (4)$$

where r_{Ai}, r_{Bj} and r_{Dij} are the treatment replications of i0, 0j and ij and R is the diagonal matrix of treatment replications.

(ii) Equality occurs in (4) if $N'R^{-1}H' = 0$.

for the treatment combinations and their effects will be 01, ..., 0q, 10, ..., p0, 11, ..., 1q, ..., p1, ..., pq, where p=n-1 and q=m-1 so that

$$H = \begin{pmatrix} -I_q \otimes 1_p & O_{l,p} & E \\ O_{l,q} & -I_p \otimes 1_q & I_l \end{pmatrix}, \tag{2}$$

where I_u is an identity matrix of order u, $O_{u,v}$ is a $u \times v$ matrix with every entry zero, 1_u is a $u \times 1$ unit vector, $\otimes$ denotes Kronecker product, $l = pq$ and $E = (E_{ij})$ is an $l \times l$ matrix in which submatrix E_{ij} has size $p \times q$, entry 1 in the (j,i)th position and zeros elsewhere (i=1, ..., q ; j=1, ..., p).

The bound is found by taking the maximum of two lower bounds for $tr(H\Omega_d H')$, where Ω_d is a generalized inverse of the intra-block information matrix A_d for $d \in D^*$,and where $A_d = R - NN'/k$ for R the diagonal matrix of treatment replications and N the incidence matrix of the design.

2.1. *Bound B(H)*

The first bound is derived from the following theorem which has application in a variety of design problems and is proved in the Appendix.

THEOREM 1. *Let D be a class of designs having one or more blocking factors and t treatments. For any d∈D let C_T contain $L \geq t-1$ contrasts of interest, where rank(C)=t-1. If $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_{t-1} > \theta_t = 0$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{t-1} > \lambda_t = 0$ are the eigenvalues of $C'C$ and A_d respectively, then*

$$tr(C\Omega_d C') \geq \sum_{i=1}^{t-1} \theta_i/\lambda_i.$$

COROLLARY 1. *If $D_c = \{d \in D \ ; \ tr(A_d)=c\}$ where c is a constant then, for any $d \in D_c$, $tr(C\Omega_d C') \geq (\sum \theta_i^{1/2})^2/c$.*

The corollary is proved by minimizing the bound in Theorem 1 subject to $\sum \lambda_i = c$ using a Lagrange multiplier.

COROLLARY 2. *Let $B(C)=(\sum \theta_i^{1/2})^2/c_{max}$, where $c_{max} = \max\{tr(A_d); d \in D\}$, then B(C) is a lower bound on $tr(C\Omega_d C')$, for all $d \in D$.*

Previous trials have established that the drugs are effective individually; a trial is now needed to find out if dual administration of the drugs has greater efficacy. We consider the case of a single blocking factor, for example groups of patients of the same age-range and sex.

Let the two treatment factors be A and B, having respective levels 0,1,...., n-1, and 0,1,....,m-1. Each treatment combination is denoted by ij, where i and j are the levels of A and B respectively. In this paper we discuss the case of a single excluded treatment combination labelled 00. We seek to estimate the following treatment comparisons which will be called *the dual versus single contrasts:*

$$\tau_{ij} - \tau_{i0} \quad , \quad \tau_{ij} - \tau_{0j} \quad (i = 1, 2, ..., n-1;\ j = 1, 2, ..., m-1), \qquad (1)$$

where τ_{ij} is the effect of ij. We assume a linear model for the observations in which there is no block-treatment interaction and the errors are uncorrelated with constant variance, σ^2.

The paper has two main aims: to find a lower bound on the average variance of the dual versus single contrast estimators for an $n \times m$ experiment and, secondly, to identify a balance property which is possessed by many efficient designs. The property leads to identifying a simple design construction method for $n \times 2$ experiments with block size,k,less than the number of treatment combinations,t, based on reinforcing group divisible designs.It is shown that using singular group divisible designs, when available, leads to efficient designs for the dual versus single treatment problem. Efficient designs for $k \geq t$ will be derived in a future paper.

2. Lower Bounds

Let $D^*(n,m,b,k)$ be the class of all connected designs for an $n \times m$ experiment with 00 excluded, arranged in b blocks each of size k. For $d \in D^*$, we derive a lower bound for the trace of the variance-covariance matrix of the least squares estimator of $H\tau$, where τ is the $(nm-1) \times 1$ vector of treatment effects and H is the contrast matrix for (1). Throughout this paper the standard order adopted

Comparing dual with single treatments in block designs

By A. GERAMI* AND S.M. LEWIS
Department of Mathematics, The University, Southampton, SO9 5NH, U.K.

SUMMARY

Experiments having two treatment factors and one blocking factor are considered in which a particular combination of treatment factor levels must be excluded. The motivation for the work is a medical trial of two drugs in which a double placebo cannot be administered on ethical grounds. The contrasts of interest compare the effects of having both treatment factors at non-zero labelled levels with the effects of having only one treatment factor at a non-zero labelled level. A lower bound is derived on the average variance of the contrast estimators in $n \times m$ experiments. For $n \times 2$ experiments a balance property is defined which is possessed by many efficient designs. Designs having the property are characterized through the structure of the intra-block information matrix. Efficient reinforced group divisible designs are constructed and assessed against the lower bound.

Some key words: A-criterion; Dual Versus Single Treatment Contrasts; Factorial Experiment; Group Divisible Designs; Part-balance.

1. INTRODUCTION

In factorial experimentation it is sometimes necessary to exclude certain treatment combinations from a study. An example is a medical trial to investigate the joint effect of two drugs , each of which is either absent or given at a number of predefined dose levels, in which it is unethical to administer a double placebo.

* Current add. Inst. of Seeds Improv., Mard - Abad Rd, Karaj - IRAN

REFERENCES

[1] Buhrman, J. M. (1973), On order statistics when the sample size has a Binomial distribution. Statistica Neerlandica 27, 125-26

[2] David, H. A. (1981) Order Statistics. New York: John Wiley and sons.

[3] Gupta, D. and Gupta, R. C. (1984). On the distribution for order statistics for a random sample size. Statistica Neerlandica 38, 13-19.

[4] Hoel, P. G. and Port, S. C. and Stone, C. J. (1971) Introduction to probability Theory, Boston: Houghton Mifflin.

[5] Raghunandanan, K. and Patil, S. A. (1972). On order statistics for random sample size.Statistica Neerlandica 26, 121-126

[6] Rohatgi, V. K. (1997). Distribution of order statistics with random sample size. Commun. Statist. Theory Meth., 16(12) 3739-3743

$$E(X_{i:N}X_{j:N}) = \frac{1}{n_j}\sum_{k=j}^{\infty}\sum_{s=k-i+1}^{k}\sum_{t=k-j+1}^{k}\frac{t+js}{jts^2}\,n(k) \qquad (4.2)$$

and finally by direct use of (2.3) we get

$$\delta_g(r) = \frac{e^{-r}}{n_2}\left(\frac{d}{dt}\,\frac{n(t)}{t}\,\Big|_{t=1} - e^{-r}\right) \qquad r\geq 0 \qquad (4.3)$$

5. ACKNOWLEDGEMENT

The author wishes to thank the research consul of Shiraz University for financial support.

and also it is a recurrence relation. Also their result for product moments of i^{th} and j^{th} order statistics is in error.

EXAMPLE 3.3:(N IS POISSON)

Let N be a Poisson random variable with parameter λ, and $\pi_i = P(N \geq i) = P_i(\lambda)$, then

$$E(X^m_{i:N}) = \frac{(m+i-1)!\lambda^{-m}}{(i-1)!} \times \frac{P_{m+i}(\lambda)}{P_i(\lambda)}$$

$$E(X_{i:N}X_{j:N}) = \frac{i(j+1)}{\lambda^2} \times \frac{P_{j+2}(\lambda)}{P_j(\lambda)}$$

and

$$g_R(r) = \frac{(1-r)\lambda^2}{P_2(\lambda)} e^{-\lambda(1-r)} \qquad 0 \leq r \leq 1$$

REMARK 3.3: In this case the m^{th} raw moment of the i^{th} order statistic is strictly determined, while the result obtained by Raghunandanam and Patil (1972) is a recurrence relation. The other results are agree with those obtained by them.

4. EXPONENTIAL DISTRIBUTION.

If $X_1, X_2, \ldots, X_N$ is a random sample from an exponential distribution with parameter $\lambda=1$, then according to the relation (page 48, David (1981))

$$E(X^m_{i:k}) = \sum_{j=k-i+1}^{\infty} \binom{j-1}{k-i}\binom{k}{j}(-1)^{j-k+i-1}E(X^m_{1:j})$$

and by direct computation, we have

$$E(X^m_{1:j}) = \frac{m!}{j^m}$$

so

$$E(X^m_{i:N}) = \frac{m!}{\pi_i}\sum_{k=1}^{\infty}\sum_{j=k-i+1}^{k}\binom{j-1}{k-i}\binom{k}{j}(-1)^{j-k+i-1}\frac{\pi(k)}{j^m} \qquad (4.1)$$

and it is easy to show that:

$$E(X_{i:k}X_{j:k}) = \sum_{s=k-i+1}^{k}\sum_{t=k-j+1}^{k}\left(\frac{t+js}{jt}\right)^2$$

Hence by direct use of (2.2) we get

in (0,1) then relations (2.1), (2.2) and (2.3) become

$$E(X_{i:N}^{m}) = \binom{m+i-1}{i-1} \frac{1}{\pi_i} \sum_{k=i}^{\infty} \frac{m!k!}{(m+k)!} \pi(k) \tag{3.1}$$

$$E(X_{i:N}X_{j:N}) = \frac{i(j+1)}{\pi_j} \sum_{k=j}^{\infty} \frac{\pi(k)}{(k+1)(k+2)} \tag{3.2}$$

$$g_R(r) = \frac{1}{\pi_2}(1-r)p''(r), \quad r>0 \tag{3.3}$$

EXAMPLE 3.1:(N IS BINOMIAL)

Let N be a Binomial random variable with parameters n and p. Then by direct use of (3.1)-(3.3) we have

$$E(X_{i:N}^{m}) = \frac{n!p^{-m}}{B_i(n,p)} \sum_{i-1}^{n-1} \binom{}{}\Gamma(n+1,n-i+1)\Big]_p(m+i,n-i+1)$$

where $I_p(m+i,n-i+1)$ is the incomplete beta function,

$$E(X_{i:N}X_{j:N}) = \frac{i(j+1)}{(n+1)(n+2)p^2} \times \frac{B_{j+2}(n+2,p)}{B_j(n,p)}$$

and

$$g_R(r) = \frac{n(n-1)p^2(1-p)(pr+q)^{n-2}}{B_2(n,p)}$$

in which $B_i(n,p) = \sum_{k=i}^{n} \binom{n}{k} p^k q^{n-k}$. These results are precisely the experession of Raghunandanan and Patil (1972).

EXAMPLE 3.2:(N IS NEGATIVE BINOMIAL)

Let N be a Negative Binomial random variable with parameters α and p. Then by direct use of (3.1)-(3.3) and doing some calculation, we get

$$E(X_{i:N}^{m}) = \frac{(m+i-1)!}{(i-1)!} \times \frac{(\alpha-m-1)!}{(\alpha-1)![1-B_\alpha(\alpha+i-1,p)]} \times \frac{1}{p^m q^m} [1-B_{\alpha-m}(\alpha+i-1,p)]$$

$$E(X_{i:N}X_{j:N}) = \frac{i(j+1)}{(\alpha-1)(\alpha-2)} (p/q)^2 \frac{1}{[1-B_\alpha(\alpha+j-1,p)]} [1-I_p(j+2,\alpha-2)]$$

and

$$g_R(r) = \frac{1-r}{\pi_2} \times \frac{\alpha(\alpha+1)p^\alpha q^2}{(1-rq)^{\alpha+2}}$$

REMARK 3.2: The result of Raghunandanan and Patil (1972) for m^{th} raw moment is in error because for k=m their result is undefined, and so the result will not work for first moment in case that N is Geometric

in case the parent distribution is exponential.

2. MOMENTS, PRODUCT MOMENTS AND DISTRIBUTION OF THE RANGE

With notation of section 1, we have

$$E(X_{i:N}^m) = \frac{1}{\pi_i} \sum_{k=i}^{\infty} \pi(k) E(X_{i:k}^m) \tag{2.1}$$

$$E(X_{i:N}X_{j:N}) = \frac{1}{\pi_j} \sum_{k=j}^{\infty} \pi(k) E(X_{i:k}X_{j:k}) \tag{2.2}$$

and

$$g_R(r) = \frac{1}{\pi_2} \int_{-\infty}^{\infty} f(x)f(r+x)\psi''(F(r+x)-F(x))dx \tag{2.3}$$

where $g_R(r)$ is the probability density function of the range R.

Relations (2.1) and (2.2) follow immediately by conditioning on N. For (2.3) note that the probability density function of $X_{N-i+1:N}$ is

$$g_{N-i+1}(x) = \frac{1}{\pi_i (i-1)!} (1-F(x))^{i-1} f(x) \psi^{(i)}(F(x)).$$

Now, let us define $R_{i:N} = X_{N-i+1:N} - X_{i:N}$. Then the (conditional) distribution function of $R_{i:N}$ is

$$G_{R_{i:N}}(r) = \frac{1}{\pi_{i+1}} \sum_{k=i+1}^{\infty} \pi(k) G_{R_{i:k}}(r)$$

so

$$g_{R_{i:N}}(r) = \frac{1}{\pi_{i+1}} \sum_{k=i+1}^{\infty} \pi(k) g_{R_{i:k}}(r) \tag{2.4}$$

For i=1, we get $R_{1:N} = X_{N:N} - X_{1:N}$ which is range R, and hence

$$g_R(r) = \frac{1}{\pi_2} \sum_{k=2}^{\infty} \pi(k) g_{R_{1:k}}(r) \tag{2.5}$$

Since

$$g_{R_{1:k}}(r) = k(k-1) \int_{-\infty}^{\infty} f(x)f(r+x)[F(r+x)-F(x)]^{k-2}dx \tag{2.6}$$

so, by substituting (2.6) in (2.5), we get (2.3).

3. UNIFORM DISTRIBUTION.

If $X_1, X_2, \ldots, X_N$ is a random sample from a Uniform distribution

considered the case that N is Binomial, Negative Binomial or Poisson and population is Uniform, and derived the density function of i^{th} order statistic, m^{th} raw moments of i^{th} order statistic and the density function of the range. Buhrman (1973) simplified the form of the density function of the i^{th} order statistics for the case N is Binomial. Gupta and Gupta (1983) derived the distribution of i^{th} order statistic and the range, assuming the sample size to be generalized Negative Binomial, generalized Poisson or generalized Logarithmic series. Rohatgi (1987) derived the distribution function of the i^{th} order statistic in random sampling from a population, when the sample size is random.

Following Rohatgi (1987), the probability density function of i^{th} order statistic is given by

$$g_i(x) = \frac{1}{(i-1)!\pi_i} F^{i-1}(x)f(x)\psi^{(i)}(1-F(x)),\ x\in R \qquad (1.1)$$

in which, $\pi_i = P(N\geq i) = \sum_{k=i}^{\infty} \pi(k)$, and for $|s|<1$, $\psi^{(i)}(s)$ is i^{th} derivative of $\psi(s)$, the probability generating function of N. The joint probability density function of the i^{th} and j^{th} order statistics $(1\leq i<j)$ is

$$g_{ij}(x,y)=\begin{cases}\dfrac{F^{i-1}(x)[F(y)-F(x)]^{j-i-1}f(x)f(y)\psi^{(j)}(1-F(y))}{(i-1)!(j-i-1)!\pi_j}, & x<y \\ 0 & x\geq y\end{cases} \qquad (1.2)$$

In this paper we consider m^{th} raw moments of i^{th} order statistic, product moments of i^{th} and j^{th} order statistics, and probability density function of the range for arbitrary N and arbitrary parent distribution. This is done in section 2. In section 3 we employ the results of section 2, to derive the results of Raghunandanan and Patil (1972) in simpler derivation . Section 4 is devoted to the m^{th} raw moments, product moments and probability density function of the range,

On Order Statistics When the Sample Size is Random

A. Bazargan-Lari
Department of Statistics
Shiraz University
Shiraz, Iran

Keywords and phrases: order statistics; probability generating function; m^{th} raw moments; random sample size

ABSTRACT

The m^{th} raw moments of i^{th} order statistic and product moment of i^{th} and j^{th} order statistics, and the density function of the range R, in random sampling from a uniform distribution in (0,1) and exponential distribution is obtained, when the sample size is random. The results of Raghunandanan and Patil (1972) follow in simpler derivation.

1. INTRODUCTION

Suppose $X_1, X_2, \ldots, X_N$ is a random sample of size N from a population with probability density function f(x), x>0. We assume that N is a nonnegative integer valued random variable, with $\pi(k)=P(N=k)$, $k=0, 1, \ldots,$ Let $X_{1:N} \leq X_{2:N} \leq \ldots \leq X_{N:N}$ be the corresponding order statistics. Recently several authors have discussed the statistical behaviour of the order statistics. Raghunandanan and Patil (1972),

Estimation of Probabilities Associated with Multi-Type Branching Processes

A. Badamchizadeh
Department of Mathematics
University of Tabriz
Tabriz, Iran

Wed. 15:30 Hall 7

A discrete k-type Branching Process is assumed. Based on complete history of the process, maximum likelihood estimates of p_{jq}, for $j, q = 1, 2, 3, \cdots, k$ the probability that individual of type p born to individual of type j and the mean matrix of the process $M = (m_{ij})$, $i, j = 1, 2, 3, \cdots, k$ is given. The results are applied to A 3-type Branching process.

Some Generalizations of Weak Convergence Results on Multiple Channel Queues in Heavy Traffic

H. A. Azarnoosh
Mashhad University
Mashhad, Iran

Wed. 14:30 Hall 7

This paper extends certain results of Iglehart and Whitt on multiple channel queues to the case where the inter-arrival times and service times are not necessarily identically distributed. It is shown that the weak convergence results in this case are exactly the same as those obtained by Iglehart and Whitt.

English Papers

each other, to communicate, and to express an unbiased opinion about this gathering.

I hope that you all will come again for the second conference.

The Speech of the Chairman of the Scientific Committee

Javad Behboodian

Ladies and Gentlemen:

It is my great pleasure, on behalf of the scientific committee, to welcome all of you, and in particular our international guests, to the cultural city of Isfahan for the First Iranian Statistics Conference.

Statistics today is an important scientific discipline with applications to most fields of studies. Numerical literacy and the ability to follow and understand statistical arguments is important for every educated person. Statistics is essential for every organization to deal with uncertainty, to evaluate risk, and to make predictions.

The purpose of this conference is to bring us together to present and discuss our statistical findings and also to promote the use of statistics in research, education, health, and developement in Iran.

I would like to congradulate and express my sincere gratitude to the authorities of Isfahan University of Technology, specially the chancellor and the dean of school of mathematics, for organizing the first conference in statistics in our country. We are also very much indebt to the organizing committee for their effort during the past twelve months.

We also thank the Iran Statistics Center for their generous support, the Iranian Statistical Society for their kind cooperation, the Ministry of Culture and Higher Education, for being so much helpful.

The generous contribution by the officials of Isfahan, in particular, the Governorship, the Municipality, the Budget and Plan Organization, and the Steel Mill twards this conference is highly appreciated.

It is my hope that this first conference will be a great success. I wish you all a pleasent stay in this beautiful city. This is the best opportunity for our guests from abroad and their Iranian colleagues in this country to meet

than its geographical coverage, and 24% has had no applications other than its defined one. This is tantamount to saying that the approved statistical information had had problems regarding geographical and time period considerations. According to this survey, the information categorized as low quality has mainly come from industry, exactly the part with the highest rate of statistical activities.

On the whole, one can say that the available information and statistics do not meet the disired quantitive and qualitative standards, and therefore, await much reconsideration and attention to overcome the present problems. This necessitates a close collaboration between the major information - producing units, such as C. S. I., on the one hand, and the academic community at universities, on the other.

Fortunately, during the past few years, my colleagues at C. S. I. have made great efforts to make the statistics both user-oriented and cost effective. Though these attempts require more facilities and equipment, there is hope for future development, considering the fact that the authorities have come to a deep understanding of and have also felt the great need correct, consistent, and up-to-date information.

To end, I wish success for all the colleagues at the present conference and hope that the conference will become a source of future developments in the field.

Junior degrees, 54.3% highschool diplomas, and 15.6% held lower degrees. Only 56 people were Ph. D holders. Therefore, over 70% of those involved had received a very low education and, as a consequence, the statistics produced by them was of a very low quality. This is also suprising because if statistics production is to be a professional job, how come it is carried out by such people and why with so high a cost and such a low quality?

A further point to be mentioned is that the highest amount of such activities has been centered in industry; population and man-power statistical reviews standing second to industry. To provide you with figures, 23.5% of the activities have been carried out in industry, 16.8% in population and man-power, 9.7% in education, 8.4% in health and medical care, 7.1% in agriculture, and 5.9% in commerce.

On the whole, it is seen that the distribution of the statistical activities do not enjoy a suitable and well-modulated distribution among the different socio-economic sectors. This fact is also confirmed by another survey.

In a survey carried out in the two years of 1990 and 1991, the statues-quo of statistics-production by different executive and planning departments in the country was studied.

Based on this study, indices were developed to show the type and amount of the applications made of the statistics produced. Using the indices, we can determine how far the statistics production matches the present and future needs of users.

The survey was carried out under the title "The recognition of the Information Needed for the First Economic, Social and Cultural Development Plan of Islamic Republic of Iran." From the 1333 information types needed by the users, 1106 types (83%) are already produced. Yet, 7% of the statistical information produced have been rejected by the users due to such problems as geographical coverage, framework problems, ambiguites in definitions and concepts, lack of standard classifications, employing incorrect methods of sampling, and finally because of problems in execution and information proccessing.

Therefore, 1030 types of the information developed (77%) have been to some extent confirmed and approved of by users and the remaining have been problematic or totally useless.

What is more, in the 1106 types of information produced, the statistical period has been in accordance with the desired time period whereas the remaining 35% have not. 76% of the information has found applications other

workers in the case of a major statistical activity:

15,242,640,000	*personel costs*
11,976,360,000	*overhead*
27,219,000,000	*Total*

In Japan, where the population is about 125 million and is also rather highly-developed regarding the statistics-production, has a personnel of 12700 people for the production of current statistics consuming an annual badget of 46 billion Japaneses Yens.

The results from the survey of the 1630 statistical organizations reveals that 546 of them have been active in their profession through the 1983-84 period, completing 1856 statistical projects. From the 546, 301 organizations (55%) are reported to have finished their projects, 150(27.5%) have had unfinished ones, and 95(17.4%) have had both completed and underway projects.

From the 1856 projects cited, 72% are reported to have been finished during the period of 1983-84, the results of which have become available to users, and the remaining 28% have continued throughout the year 1987.

Of the completed statistical projects, 60% were of a long-term nature and the other 40% have been case studies and short-termed. We can say that the results from 1339 such projects in the period cited above have been published and available to users. Now we are in a position to calculate the costs for the statistical 5-years period in question. The total expenditure of the period amounts to 136, 100, 000, 000 Rls. Therefore, 100, 000, 000 Rls. has been spent for each statistical project carried out, which is too high a price to pay because the cost of a sensus for a population of 50 million is estimated to be around 7, 500, 000, 000 Rls. inclusive of the temporary employment of 30, 000 people, and using cars with drivers. This cost is only enough for 75 statistical projects. As a result, it is clearly evident that although the statistics produced across the nation has not had a wide and versatile application, it has imposed high costs and we are bound to declare it as a luxury.

Another point to be considered is the educational qualifications of those involved in the enterprise.

In an attempt to survey the degrees held by the people on statistics-producting duties, it was shown that 22.1% held a university degree, 7.6%

isolation or in combination with other information external tothe system, to secure the continuing of that particular activity in its proper route, we still face governmental or nongovernmental organizations in which no statistics whatever are produced. The number of such organization is not small; of every four, only three organizations of all the statistics-producing ones (51%) include a statistics unit in their charts and 797 (49%) are collecting information without having such a unit.

To put it in a different way, one halfof the organizations under survey are collecting data and figures ina an organized manner and the rest carry out such activities without any organized facilities, depending solely on manpower for their work. The lack of an organized framework for the task of data collection and statistical analysis creates numerous problems in determining the standards, definitions, concepts, and in the process of data collection through time, and finally, raises difficulties in inter-organizational comparison of the results derived. This will, in turn, render the compiling and production of statistics in the form of statistical annals an impossibility in organizations lacking a statistical unit every activity is stopped or faces a great difficulty once the person in charge is removed from his job.

In the 1630 organizations in which the statistics producing activity is carried out, 9073 people in all are involved in producing statistics. Thus, in each of them, 5.6 persons are working in the field. Further more, 43.2% of them are only working in the Province of Tehran.

It will not be out of place here if we do some calculations now. Let us assume that for every individual working in the field, the organization pays a monthly average of 140,000 Rls. which covers net salary, the perks, overtime, and bonuses like research pays, end-year reward and others.

$$\underset{salary}{90,000} + \underset{overtime}{30,000} + \underset{others}{20,000} = 140,000$$

If we further assume that our organization has to pay an amount of approximately 110, 000 Rls. per person for other things like assignments, transportation, office stationary, electricity, water, building maintenance, and finally, the costs of information processing and publication of the results from each statistical activity, the total amount spent throughout the country annually amounts to some 272, 219, 000, 000 Rls., neglecting the need for temporary

trust towards development; the wheels of technology is turning with such a bewildering speed, thanks to the mental and intellectual power behind it and information technology, that if we do not keep with its pace today, we are sure to be left behind. The discussion to follow is about the statistics which stands at the far end of the information chain.
Let us imagine this chain as:

$$Intelligence \rightarrow Wisdom \rightarrow Knowledge$$

$$Information \rightarrow Data \rightarrow Statistics$$

Let us choose statistics from among them stripped off from its details and let us further leave the theoretical issues to the academicians and only cocentrate on the production of statistics.

Among the issues of importance in statistics production, which is under speculation in the Islamic Iran and deserves discussion here, one is its method of production, the other its applications. In other words, there exist two fundamental points regarding the production of statistics: the efficacy of the statistics produced from an organizational point of view, and their applicability by users; we have to find out whether the statistics produced have been user-oriented or not?

As regards the first of the points raised; i.e., their efficacy, I would like to point to the study carried out in 1988 by S. C. I., under the title "Statistical Analysis of the Statistical Facilities in Iran" and later I will draw some conclusions from this study.

A survey of the results from the above study reveals that from the total 2140 governmental and non-governmental organizations under study, 1630 (76%) have been active in producing statistics, and in 510 agencies (24%) no such activity has been reported. It is worth noticing that of all the agencies surveyed, 92.4% were under government control and 7.6% were run by the private sector.

In the above study, any activity leading to the production of statistics, either by a statistical agency or by individual workers working continuously or discontinuously, is defined as statistical activity.

It is suprising, however, that, ith this broad definition, we still find that for 24% of the agencies surveyed no statistics production is reported. Where any socio-economic activity leads, directly or indirectly, to the production of statistics, and the statistics produced must, of necessity, be applied, in

The Efficacy and Applicability of the Statistics Produced

M. Jamshidi
The Head of the Statistical Center of Iran

I find it necessery, before proceeding with my speech, to express my deepest contentment and obligation to the present, esteemed audience, his Exellency the Minister of Culture and Higher Education, professors, teachers, researchers, and the foriegn guests to the Islamic Republic of Iran, and particularly to the highly distinguished and respected scientistics from Isfahan University of Technology on the occasion of two important events; i.e., the First Statistics Conference and the meeting of the First Iranian Public Council of the Iranian Society of Statistics, which, as one of those in the service of the Iranian Statistical Community, I have always been longing for I further, hope that the way opened today never be closed and that with the help of the Almightly and the contributors of all the workers in the field our aims which are not but scientific activities, both applied and theoretical, will soon come true to prepare the grounds for us to enter a new age of prosperity and construction.

As the Head of S. C. I., which is the greatest source of statistics production in Iran, I should acknowledge all my sincere colleagues at the Central Bank of the Islamic Republic of Iran, the Ministry of Employment, the Ministry of Agriculture, the Ministry of Heavy Industries, the Ministry of Education, the Ministry of Culture and Higher Education, and the Ministry of Health and Medical Training for their technical and professional support. Let me also hope that each of them has had their own share in what has been going on here at I. U. T.

It is common knowledge among us that the 20th centry has suitably come to be called the Age of Information Explosion every ten years now all the information of the human community doubles, and this signifies a spiralling

Isfahan University of Technology, as one of the scientific centers of Iran, has always attempted to provide for the advancement of all branches of science and technology, and in particular, the pure sciences. At I. U. T., in addition to all engineering branches and agricultural sciences, we have three faculties of Physic, Chemistry, and Mathematics, the last of which enjoys a special status among us. The research activities of this faculty, especially in the popularization of mathematical sciences among the people, have been considerable of such activities I can name the holding of mathematics exams named the Mathematics Olympiad, across Iran, the publication of a mathematical journal, Peyk-e-Riadhi, and collaboration in a more advanced journal called, Mathematical thought and Culture. The educational activities by the faculty are also of high quality and great importance.

Let me inform you of the statistics recently published by the Ministry of Culture and Higher Education, in which I. U. T. ranks the first, second, and third among all Iranian Universities in the entrance exams to graduate schools in almost every branch for which a degree is offered at I. U. T. This is where the facilities in the universities in Tehran are much more than those in other cities. Taking this into account, and also the mean of the G. P. A.'s scored by the applicants, we can better appreciate the significance of these ranks.

We have prepared everything for our Faculty of Mathematics to start a Ph. D. program next year.

I have no doubts that this conference has also had advantages for our university. One, I should mention, is the honor you have given us on this occasion to gain your acquaintance, especially the expatriate Iranians whom I hereby invite to join us at I. U. T.

I should also acknowledge the great support we won from many sources for holding this conference. Due to their abundance I cannot name all, but let me especially thank:

The Statistical Center of Iran (C. S. I.), which gave us a friendly hand. There are also many individuals who deserve mine and everybody's thanks, all the faculty at the Department of Mathematics, the Head of the Department Dr. Haghani, and his colleagues. Dr. Rejali deserves my most sincere thanks who had a hard-time organizing the conference.

I hope that all their sweating and hard work lead to more friendships and the flourishing of our nation.

we are not short of the human factors in science and technology at all, and that if we are behind it is only because we are kept behind. But now that the Islamic Iran has been known as the pioneer and guardian of the high Islamic and human values, she is rightly expected to fill the gaps in science and technology created by others as fast as possible with due commitment will power and faith in God. It's exactly here that we have to make a new policy for ourselves:

1. Our scientists and researchers, regarding their great responsibility, are supposed to find each other and communicate their ideas, away from any kind of selfishness or prejudice, in a concerted attempt to solve the problems.

2. The high status of the scientist must be recognized by the people and authorities alike and attempts should be made to support them fully.

3. At least, their minimum requirements and facilities must be provided for the scientists to allow them the chance to take their definitive roles.

4. We must cut down on our material expenditure and increase our efforts to keep up with the science in the world today. In a word, and frankly, the authorities must value and support science and technology to the extent that in a matter of a few decades the gap is filled. For a few decades, people should want less.

Keeping these views in mind, we can find the statue of such conferences and feel the responsibility we have now. Our country is going through a fundamental reconstruction and is in need of a multi-faceted planning, the corner stone of which is statistics and information.

In the modern world, statistics, like a language and as a foundation, is a necessity in all sciences, particularly the advent of computers has added to its importance. Fortunately, there is a workshop on computer applications in statistics along with the conference to familiarize the statistical community in Iran with the statistical packages used in the field.

An exhibition of computer softwares and books from domestic and foriegn publishers also accompanies the conference. I hope they can all help towards a better understanding in the field.

the faculty of Mathematics at the Isfahan University of Technology and the newly-established Iranian Statistical Society .

To me, the main motivation behind these meetings is not but preparing a framework for the advacement of sciences and hence, technology. I have only named the meetings at I. U. T., but there are far more across the country. Holding such meetings heralds a promising future. It seems as if a scientific revival is taking place in our society to bring in light and suppress darkness and ignorance.

It might seem optimistic and a bit sentimental to speak as I am doing now, but if one glances at the history of this land on the one hand, and on the other, tries to be realistic and see both our honerable, scientific past and our present lagging behind in science and technology, then he can grasp the significance of such meetings; as if flames of new hopes start glittering in him. To elaborate on my purpose, let me provide you with a brief history of our pride in the past:

Our country has been imbued with efforts and wisdom thruoghout its history. The word Algorithm meaning the calculation of figures and numbers, for instance; is it not derived from the name of Alkharazmi the greatest Iranian genius in mathematics, astronomy and geogrophy?

Did our knowledge and technology together with the fame of Avi Cena not pass the gates of the Eastern Rome and conquer the global markets as far as China? Now there were more like Sheikh-Bahaee, Abou Reihan, Khaje-Nassir, Khayam, $\cdots$ only to mention a few Iranian scientists and not those from the whole of the Islamic world. It will suffice to visit the great libraries and museums across the world to discover the treasures of Islamic and Iranian thought everywhere. Yet, I do not believe in admiring the past at the expense of the present. I am, quite on the contrary, of the belief that we have to be realistic and face the realities as they stare us in the face.

Recently a book was published by the American ex-president, Richard Nixon, where reviewing our honorable past, and achievments after the Islamic Revolution in 1978, he formulates a set of instructions for the present U. S. administration in facing Iran. I am not referring to these points to appease the pains we are suffering from our present predicaments.

On the contrary, as I said earlier, we must be realistic and admit the fact that we are far behind in the domains of science and technology. We may even have to move at a speed far greater than that of light to keep up with the pace in the world today. What I am trying to clarify is that

The Speech of the Chancellor of Isfahan University of Technology

S. M. Saghaian Nejad

"We take Refuge in God from the Satan"

In the Name of God, the Benificent , the Merciful

On behalf of Isfahan University of Technology, its administrative body, and myself, I would like to extend my most cordial welcome to our dear guests, respected professors, scientists, and researchers from Iran and abroad, the forerunners of the statistical community of Iran, authorities and experts at the statistical centers in Iran, teachres across the country, expatriate Iranians who have travelled home for the present conference, the Islamic Society of Statistics in Pakistan, students, and the local authorities and officials from the Province of Isfahan. I would, further, like to appologize in advance for taking the time of the conference as an excuse to express my thanks and deliver a few words of welcome.

Dearest audience; no more than two months have elapsed in the new Iranian year and we have witnessed three scientific events at the Isfahan University of Technology during such a short time. The first was the International Workshop on Geometry and Algebra from 5-25 May, which was held here in collaboration with I. C. T. P., Trieste, and only a week ago finished successfully.

The workshop was of great impact on the mathematical sciences in Iran. During the workshop, a one-day large seminar was held in this same place under the title. "The First Conference on Catalyst and their Application", a joint seminar by our faculty of Chemistry and the Oil Company Research Center. That one also was a great importance for our oil industries and the workers in the field of petroleum Engineering and today we are witnesses to a third event: The First Iranian Statistics Conference held jointly by

5. More attempts on the part of the faculty members in all universities towards the development of graduate studies.

6. Revision of the curricula with respect to the modern scientific findings and the needs of the society.

7. More efforts directed towards the activities by specialized societies in the field of Statistics and Informatics.

8. The expansion of relationships amoung the research and scientific centers across the country involved in statistics in order to achieve as much of the developments as possible.

9. The necessity for a greater employment of the new statistical methods and techniques in the planning of the development and construction of the State; Iran.

technological, and cultural understanding with the help of the information and equipment and waves, has faced the Third World with more difficult decision-making tasks. What is more than anything else, a serious need in all developing countries is paying heed to the plundering and dominating nature of the ruling system over the world against which these nations must, as a first step in development, try to discover their status and regain their identity. It is a fact that the foundamental cause of underdevelopment in the Third World is more cultural than technical. Belief in the power of the intimate nations (in the Third World) and the existence of a precious background and the great sources of science and civilization amoung the muslim nations, creates the healthy appetite, for the intake and benefiting from the modern science and technology. At this historical stage of their development, the muslim communities need more than any time in the past to concentrate seriously and practically on issues of planning, technical and scientific education, fortification of their cultural and scientific development foundations, and expansion of their research and educational relations in an attempt to fertilze their intellectual creativity. Statistics, data collection, data keeping, and data analysis along with the desired application of information occupies a particular place in this regard.

Due to this legal, cultural, and revolutionary responsibility for the development and support of research and education, educating the expert and efficient man power, and also for the promotion of the level of scientic understanding, the Ministry of Culture and Higher Education emphasis on the following points:

1. Planning for the training of more man-power in the field of statistics at all university levels.

2. Attention to applied statistics as used by other university studies and also that to be used in analysing the soical problems for the excutive branches and departments.

3. Special attention to statistical labs and providing for the required equipment and facilities.

4. The great responsibility of university professors of statistics and informatics for the purpose of expansion in these studie, regarding the increasing importance of informatics in the present age.

The Message of the Minister of Culture and Higher Education

M. Moeen

In the Name of God, the Beneficient, the Merciful.

"... Every single thing is before his sight, in (due) proportion." "He knew the unseen and that which is the Great , the Most High." The Holy Quran, Ra'ad (Thunder) 8, 9.

Peace be upon the Great Founder of the Islamic Republic of Iran, the Late Imam Khomeini; Paying my respect to our Great Leader, I take the holding of The First Iranian Statistics Conference as a highly-valued and necessery action in revealing the importance of statistics, as well as in evaluating the most recent global developments and achievements in the field of statistics and its decisive role in the planning of national development. I would like to extend my thanks to all those involved in organizing this important event.

Our Islamic country, with an old and valuable status in the eastern culture and civilization, must, no doubt, be the cradle for gatherings of this sort in the process of regaining its historical identity and grandeur. It should provide, in this way, for the committed scientists and researchers, the grounds to study the needs and priorities of the social, cultural, and economic development plan of Iran from every possible aspect.

Fast and varied technical and scientific changes and developments as well as the need for gaining access to the new educational and research findings, and a propriate and timely employment of the modern science and technology are of the serious phenomena which open up a new and demanding view to our academic community.

The commencement of the era of technology and the vastness and high speed of the informatics and communications brings about new needs and demands. The modern intricate technology which attempts to declare 'Knowledge' as a new source of power and also constantly changes our scientific,

Openning Ceremony Lectures

Foreword

The first Iranian Statistics Conference was held at Isfahan University of Technology with over 500 participants, May 26-28, 1992. This Proceedings Volume follows the Invited Papers Volume, published earlier. It includes the contributed papers and the abstracts of the posters presented in the conference, in two sections of Farsi and English. The contents have been organized by the last name of the author presenting the paper or the poster.

None of the papers has been submitted to a refereeing process and the authors are responsible for the contents. This Volume is not copyrighted, hence permission for reproduction must be obtained from the author(s). Authors in these proceedings are encouraged to submit their papers to any journal of their choice.

The organizing and scientific committees like to thank all organizations and people who helped us make this conference possible. They also like to thank Isfahan University of Technology Press for printing and production of this Volume.

Scientific Committee:

J. Behboodian	Professor, Shiraz University
A. Z. Hamedani	Assistant Professor, I. U. T.
M. Jamshidian	Assistant Professor, I. U. T.
A. Parsian	Associate Professor, Shiraz University
A. Rejali	Associate Professor, I. U. T.
S. M. Sadooghi Alvandi	Associate Professor, Shiraz University
M. Zokaee	Assistant Professor, I. U. T.

Elected by the Iranian Statistical Society.

Organizing Committee:

A. Haghani	A. Z. Hamedani	M. Jamshidian
F. Kherad Pazhuh	M. Marzban	S. Pooladsaz
G. A. Raissi Ardali	A. Rejali	M. Zokaee

Faculty Members of School of Mathematics, I. U. T.

One-Sided R-Reliable Intervals and Their Associated Confidence on the Sum of Two Random Variables	Kafai and ···	105
On Dispersion of Stable Random Vectors and its Application in the Prediction of Multivariate Stable Processes	Soltani and ···	107
A Generalizable Formulation of Conditional Logit with Diagnostics	Soofi	117
On Robustness Property of MEP 1R Plans	Talebi	118

Estimation of Probabilities Associated with Multi-Type Branching Processes	Badamchizadeh	17
On Order Statistics When the Sample Size is Random	Bazargan-Lari	18
Comparing Dual with Single Treatments in Block Designs	Gerami and ···	25
Bayesian Method Compared with Some Traditional Techniques in Estimating Gumbel Distribution Parameters	Hakimi Sibooni	38
On Estimating the Variance of the Horvitz-Thompson Estimator	Hanif and ···	53
Trade-Off Theory and Its Application	Hedayat	54
Confidence Bounds on Rations of Variances in Unbalanced Two-Way Crossed Models	Kazempour	55
A Bayesian Approach to Autoregressive Modeling	Kheradmandnia	56
Estimation of the Unknown Parameter of a Uniform Distribution	Meshkani	75
Computer Applications in Statistical Methodology and Inference and Econometrics	Sachi	84
Assessing Parameters Estimates in Nonlinear Regression Using Profile Plots	Shafii	89

Contents

Scientific and Organizing Committee		ii
Foreword		iii

Openning ceremony lectures

The Message of the Minister of culture and Higher education	Moein	1
The speech of the chancellor of I.U.T.	Saghaian Nejad	4
The efficacy and Applicability of the Statistics Produced	Jamshidi	8
The speech of the chairman of the scientific committee	Behboodian	14

Contributed Papers

Some Generalizations of Weak Convergence Results on Multiple Channel Queues in Heavy Traffic	Azarnoosh	16