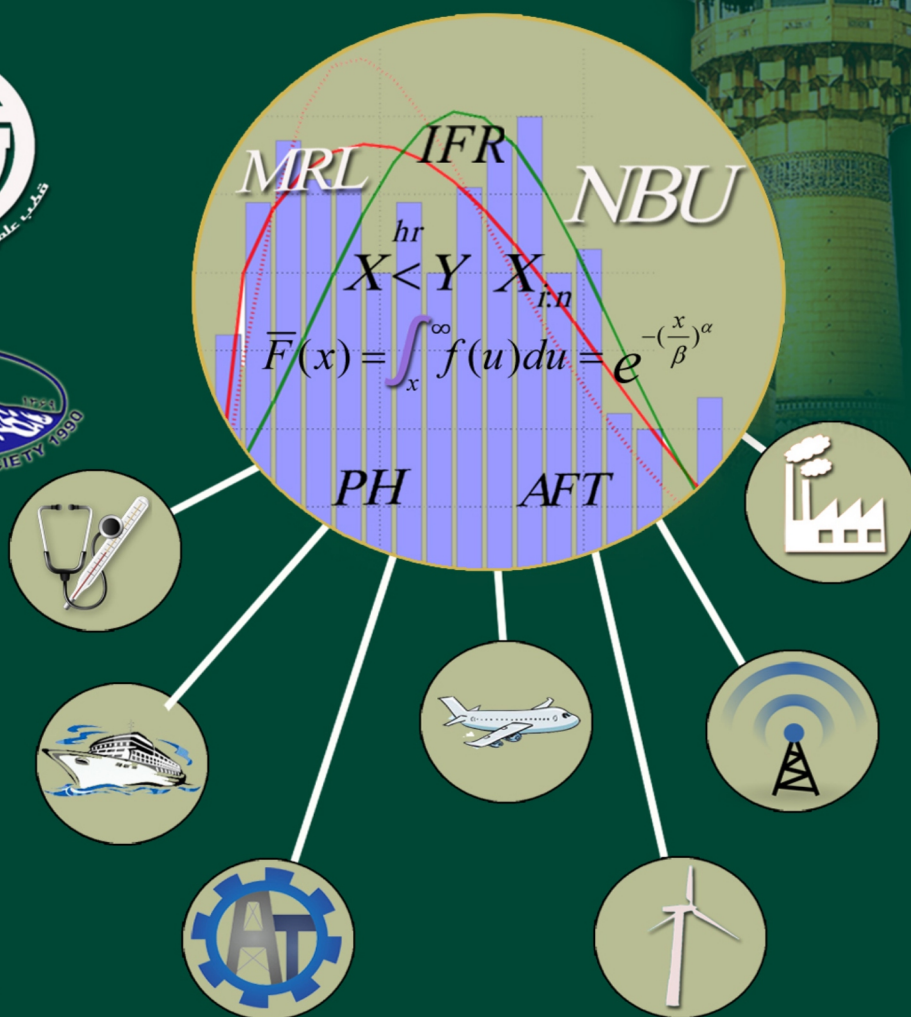


مجموعه مقالات سومین سمینار تخصصی

نظریه قابلیت اعتماد و کاربرد های آن

۲۶ و ۲۷ اردیبهشت ۱۳۹۶

گروه آمار دانشگاه فردوسی مشهد



باسمه تعالی



مجموعه مقالات فارسی

سومین سمینار تخصصی

نظریه قابلیت اعتماد و کاربردهای آن

گروه آمار، دانشگاه فردوسی مشهد

با همکاری

قطب علمی داده های ترتیبی و فضایی

۲۶ و ۲۷ اردیبهشت ماه ۱۳۹۶

این کتاب شامل مجموعه مقالات فارسی سومین سمینار تخصصی نظریه قابلیت اعتماد و کاربردهای آن است. بدیهی است که نویسندگان محترم مسئول صحت مطالب، اصالت و ویرایش علمی و ادبی مقالات خود بوده و کمیته علمی و اجرایی سمینار مسئولیتی در این موارد ندارد.

عنوان: مجموعه مقالات فارسی سومین سمینار تخصصی نظریه قابلیت اعتماد و کاربردهای آن

تدوین کننده: جعفر احمدی

صفحه آرا: فاطمه حوتی

طرح جلد: احمد دادرس جوان

ناشر: کمیته برگزاری سمینار

شمارگان: ۱۲۰

تاریخ انتشار: خرداد ۱۳۹۶

مقدمه

پیرو برگزاری سه دوره کارگاه آموزشی در دانشگاه های فردوسی مشهد (۱۳۷۸، ۱۳۹۱) و اصفهان (۱۳۹۰) و دو دوره سمینار تخصصی "نظریه قابلیت اعتماد و کاربردهای آن" در دانشگاه های اصفهان (۱۳۹۴) و تهران (۱۳۹۵)، خداوند منان را سپاسگزاریم که توفیق برگزاری سومین دوره این سمینار را نصیب گروه آمار دانشگاه فردوسی مشهد کرد. این سمینار با حمایت قطب علمی داده های ترتیبی و فضایی و همکاری انجمن آمار ایران با هدف زمینه سازی برای تبادل اطلاعات پژوهش های اعضای هیات علمی، دانشجویان تحصیلات تکمیلی و کارشناسان و کاربران در محورهای سمینار، روزهای سه شنبه و چهارشنبه ۲۶ و ۲۷ اردیبهشت ۱۳۹۶ در دانشکده علوم ریاضی دانشگاه فردوسی مشهد برگزار خواهد شد. امیدواریم که تلاش کمیته های مختلف برگزاری بتواند رضایت خاطر میهمانان و شرکت کنندگان گرامی را فراهم نماید. به منظور مشارکت بیشتر متخصصین این شاخه و معرفی پیشرفت های صورت گرفته در زمینه قابلیت اعتماد در داخل کشور به سایر پژوهشگران، تعدادی از آماردانان برجسته خارجی در زمینه قابلیت اعتماد به کمیته علمی دعوت شده اند.

پس از فراخوان سمینار، مقالات دریافت شده توسط اعضای کمیته علمی و کمیته داوران سمینار مورد ارزیابی و در نهایت ۵۰ مقاله به صورت ارائه شفاهی و ۲۳ مقاله به صورت ارائه پوستر مورد پذیرش قرار گرفت. در پایان لازم می دانم از کمیته های علمی، اجرایی و داوران سمینار نهایت تشکر و سپاسگزاری را داشته باشم. همچنین مراتب قدردانی خود را از حوزه معاونت پژوهشی دانشگاه فردوسی مشهد، قطب علمی داده های ترتیبی و فضایی، انجمن آمار ایران، پایگاه استنادی علوم جهان اسلام، مدیریت و کارکنان دانشکده علوم ریاضی و دانشجویان تحصیلات تکمیلی گروه آمار دانشگاه فردوسی مشهد اعلام می نمایم.

از خداوند منان آرزوی توفیق تمامی شرکت کنندگان محترم در این سمینار را دارم، امیدوارم اقامت دلپذیری در شهر مشهد مقدس داشته باشند.

جعفر احمدی (دبیر)

اردیبهشت ۱۳۹۶

محورهای سمینار

آزمون‌های طول عمر تسریع‌یافته	استنباط آماری داده‌های قابلیت اعتماد
قابلیت اعتماد سیستم‌های منسجم	الگوهای تعمیر و نگهداری سیستم‌ها
مدل‌های تنش-مقاومت	ترتیب‌های تصادفی در قابلیت اعتماد
قابلیت اعتماد شبکه‌ها	مفاهیم سالخوردگی
تحلیل بقاء	قابلیت اعتماد داده‌های فرسایشی
تحلیل ریسک در قابلیت اعتماد	وابستگی در مباحث طول عمر
بهینه سازی در قابلیت اعتماد	روش‌های بیزی در قابلیت اعتماد

اعضای کمیته علمی (به ترتیب حروف الفبا)

۱. دکتر جعفر احمدی، دانشگاه فردوسی مشهد
۲. دکتر ناصر رضا ارقامی، دانشگاه فردوسی مشهد
۳. دکتر مجید اسدی، دانشگاه اصفهان
۴. دکتر اکبر اصغرزاده، دانشگاه مازندران
۵. دکتر فیروزه حقیقی، دانشگاه تهران
۶. دکتر بهاءالدین خالدی، دانشگاه رازی
۷. دکتر احمد خدادادی، دانشگاه شهید بهشتی
۸. دکتر محمد خنجری صادق، دانشگاه بیرجند
۹. دکتر مهدی دوست پرست، دانشگاه فردوسی مشهد
۱۰. دکتر عبدالحمید رضایی رکن آبادی، دانشگاه فردوسی مشهد
۱۱. دکتر غلام رضا محتشمی برزادران، دانشگاه فردوسی مشهد
۱۲. دکتر علی همدانی، دانشگاه صنعتی اصفهان

اعضای کمیته علمی خارجی (به ترتیب حروف الفبا)

1. Prof. Balakrishnan, N., McMaster University (Canada)
2. Prof. Bayramoglu, I., Izmir University of Economics (Turkey)
3. Prof. Di Crescenzo. A., Università degli Studi di Salerno (Italy)
4. Prof. Eryilmaz, S., Atilim University (Turkey)
5. Prof. Navarro, J., Universidad de Murcia (Spain)
6. Prof. Singpurwalla, N. D., City University of Hong Kong (Hong Kong)

اعضای کمیته برگزار کننده (به ترتیب حروف الفبا)

۱. دکتر جعفر احمدی، دانشگاه فردوسی مشهد (دبیر علمی سمینار)

۲. دکتر مجید اسدی، دانشگاه اصفهان

۳. دکتر محمد امینی، دانشگاه فردوسی مشهد (مدیر گروه آمار)

۴. دکتر هادی جباری نوقابی، دانشگاه فردوسی مشهد (دبیر اجرایی سمینار)

۵. دکتر مهدی جباری نوقابی، دانشگاه فردوسی مشهد

۶. دکتر آرزو حبیبی راد، دانشگاه فردوسی مشهد

۷. دکتر مصطفی رزمخواه، دانشگاه فردوسی مشهد

۸. دکتر عبدالحمید رضایی رکن آبادی، دانشگاه فردوسی مشهد

۹. دکتر بهرام صادقیور گیلده، دانشگاه فردوسی مشهد

۱۰. دکتر مهدی عمادی، دانشگاه فردوسی مشهد

۱۱. دکتر معصومه فشندی، دانشگاه فردوسی مشهد

۱۲. دکتر غلام رضا محتشمی برزادران، دانشگاه فردوسی مشهد

کادر اجرایی سمینار (به ترتیب حروف الفبا)

رضا احمدئی (امور رایانه)، خیرا... اخلی (دانشجو)، مهری باقدان (کارشناس دفتر گروه)، زهره پاکدامن (دانشجو)، مطهره پارسا (دانشجو)، مهدی پورموسی (راننده)، علی پیوندی (کارشناس دفتر بین الملل دانشگاه)، حمیدرضا حافظی احمدی (مسئول امور عمومی)، فرزانه حکم آبادی (کارشناس سامانه دانشگاه)، فاطمه حوتی (دانشجو)، میلاد رسولی (دانشجو)، حسن رئیس المحدثین (حسابداری دانشکده)، محمدجواد سبک خیز (دانشجو)، فاطمه صفائی (دانشجو)، مهدیه عرفانیان (دانشجو)، محمدقنبری (خدمات)، فرانک گودرزی (دانشجو)، مرتضی محمدی (دانشجو)، حمیدرضا مومنی (امور عمومی)، انسیه نزاکی (دانشجو)، وحیده وحیدنیا (مسئول دفتر ریاست دانشکده)، علیرضا وطن دوست (مسئول امور رایانه)

فهرست مقالات فارسی

تخصیص اجزای مازاد در سیستم های سری موازی با ترکیب اجزای تعمیرناپذیر و تعمیرپذیر در هر زیر سیستم

اکبری، ز.، همدانی، ع.، ذوالفقاری، ح. ۱۰

تعیین سن شروع مصرف کراک با استفاده از آنالیز بقا و بررسی عوامل موثر بر مصرف آن با استفاده از مدل کاکس و رگرسیون لجستیک

برومند، ف.، باغستانی، ا. ۲۳

بررسی عوامل موثر بر بقای سرطان خون کودکان با استفاده از فرایندهای نیمه مارکوفی

بهرامپور، ع.، محمودی منش، م.، فرهمندی نیا، ز. ۲۸

سانسور پیش رونده نوع II توأم تطبیقی

بیات، م.، ترابی، ح. ۳۷

بررسی نرخ لگ-بخت و ارتباط آن با نرخ شکست و نرخ شکست معکوس

جانی، ن.، صالحی، ا.، خراشادیزاده، م. ۴۷

ارتباط بین بعضی از توزیع های طول عمر

جمشیدیان، ا. ر. ۵۴

مطالعه ای بر نگهداری و تعمیر پیشگیرانه سیستم های یک بار مصرف و برآورد پارامترها

چرکزی، ح.، احمدی، ج. ۶۰

مدل‌های شکنندگی و خواص توزیعی آنها

حوتی، ف.، احمدی، ج. ۶۸

بررسی عوامل موثر بر زمان شکست ایمپلنت با استفاده از دو مدل رگرسیون خطرات متناسب کاکس و مدل بقای چندسطحی با رویکرد بیزی

رحیمی فروشانی، ع.، اکبری شارک، ن.، رازی فرد، م. ج. ۷۵

توزیع نسبت متغیرهای مستقل کاماراسوامی و کاربرد آن در مباحث قابلیت اعتماد

رسولی، ع.، ایمانی، م.، سعدی، س. ۸۲

حفظ برخی ترتیب‌های تصادفی وابسته در سیستم‌های منسجم

شریفی پیرکوهی، آ.، امینی، م.، رزمخواه، م. ۸۹

سیستم‌های مرکب دومؤلفه‌ای و سه‌مؤلفه‌ای با مؤلفه‌های وایبول

صابرزاده، ز.، رزمخواه، م. ۹۷

خانواده‌ای از توزیع‌های دو متغیره وایبول تعمیم یافته نمایی شده-سری توانی

عابدینی، آ.، روزگار، ر. ۱۰۶

برخی ویژگی‌های توزیع‌هایی با نرخ خطر خطی تعمیم یافته

غلامی گردونک، ز.، صالحی، ا. ۱۱۳

اندازه نادرستی باقیمانده آماره‌های ترتیبی

فهم شریفی، م.، عباس نژاد، م. ۱۲۱

توزیعی جدید برگرفته از خانواده وایبول

قاسمیان، ف.، هاشمی، ر. ۱۲۶

اندازه‌ی نادرستی باقیمانده تجمعی پویا

کشمیری، ز.، عباس نژاد، م. ۱۳۴

توابع محدب، مقعر و ارتباط آن با مفاهیم قابلیت اعتماد و مدل های اقتصاد

محمدی پور، م.، فشندی، م.، احمدی، ج. ۱۴۰

برآورد توزیع لگ- گاما تحت داده های سانسور شده فزاینده نوع II تطبیقی

مرادیان الوار، س.، شرفی، م. ۱۴۶

تعیین حجم نمونه بهینه با استفاده از حدود تحمل برای طول عمر سیستم های k از n

نقی زاده، م.، وحیدیان، م. ۱۵۴



تخصیص اجزای مازاد در سیستم های سری موازی با ترکیب اجزای تعمیرناپذیر و تعمیرپذیر در هر زیر سیستم

اکبری، ز^۱ همدانی، ع^۲ ذوالفقاری، ح^۳

دانشکده مهندسی صنایع، دانشگاه صنعتی اصفهان

چکیده

با افزایش پیچیدگی سیستم ها و سرعت گرفتن فرآیند تکنولوژی با شکست اجزا، سیستم دچار عیب فنی گردیده و خطرات جدی می تواند بروز پیدا کند. بنابراین افزایش قابلیت اعتماد سیستم ها به یکی از دغدغه های مهم طراحان تبدیل شده است. مساله تخصیص اجزای مازاد از متداول ترین روش ها برای بهینه سازی قابلیت اعتماد سیستم ها است. در بسیاری از تحقیقات گذشته سیستم ها صرفا دارای اجزای تعمیرناپذیر یا اجزای تعمیرپذیر بودند؛ نوآوری این پژوهش در نظر گرفتن سیستمی است که در آن اجزای تعمیرناپذیر زیرسیستم ها در کنار اجزای تعمیر پذیر قرار گرفته و لذا برای بهینه سازی چنین سیستمی از مفهوم قابلیت دسترسی استفاده می شود و هدف اصلی عبارت است از ارائه مدل و حل مدل ارائه شده جهت بهینه سازی قابلیت دسترسی سیستم در کنار کمینه کردن هزینه ها با توجه به محدودیت های وزن و حجم که در نتیجه تعداد و نوع اجزای هر زیر سیستم برای تعیین ساختار سیستمی با قابلیت دسترسی بالا تعیین میگردد.

کلمات کلیدی: قابلیت اعتماد، قابلیت دسترسی، تخصیص اجزای مازاد.

^۱ zahra.akbari1@in.iut.ac.ir

^۲ hamadani@cc.iut.ac.ir

^۳ h.zoulfaghari@in.iut.ac.ir

۱ دست‌آوردهای پژوهش

۱- مقدمه

متداول‌ترین تعریف قابلیت اعتماد (۱۹۹۱) توسط IEC^۱ به این صورت که قابلیت اعتماد^۲ به عنوان مشخصه‌ای از یک موجود^۳، عبارتست از احتمال این‌که موجود کار مورد نظر را تحت شرایط معینی در فاصله زمانی مشخص بدون خرابی انجام دهد؛ ارائه شده است. یک موجود می‌تواند تعمیرناپذیر یا تعمیرپذیر باشد. مفهوم قابلیت اعتماد برای موجودات تعمیرناپذیر بیان می‌شود و برای موجودات تعمیرپذیر از مفهوم دیگری به نام قابلیت دسترسی^۴ استفاده می‌شود. قابلیت دسترسی عبارتست از احتمال این‌که موجود در لحظه‌ی مشخص تحت شرایط معین برای انجام کار مشخص در دسترس باشد. با افزایش پیچیدگی سیستم‌ها و سرعت گرفتن فرآیند تکنولوژی شکست یکی از اجزا موجب عیب فنی و خطرات جدی می‌شود و این امر در صنایع حساسی مانند صنایع دفاعی و صنایع هوایی خطرات جبران‌ناپذیری به بار می‌آورد. همچنین طراحان با افزایش قابلیت اعتماد (دسترسی) سیستم رضایتمندی مشتری را که موجب حضور مستمر در صحنه رقابت و دستیابی به نام ماندگار می‌شود فراهم می‌کنند. بنابراین طراحی و تولید سیستم‌هایی با حداکثر قابلیت اعتماد (دسترسی) به یکی از مهمترین دغدغه‌های طراحان تبدیل شده است. یافتن بالاترین سطح قابلیت اعتماد (دسترسی) سیستم براساس محدودیت‌های موجود مانند (هزینه، وزن، حجم و...) بهینه‌سازی قابلیت اعتماد (دسترسی) نامیده می‌شود. طراحان برای افزایش قابلیت اعتماد (دسترسی) سیستم چهار گزینه زیر را مطرح می‌کنند [۱].

۱- افزایش قابلیت اعتماد (دسترسی) اجزا: در اینجا قابلیت اعتماد اجزا بعنوان یک متغیر تصمیم‌گیری در نظر گرفته می‌شود و با استفاده از تکنیک‌هایی نظیر استفاده از مواد اولیه بهتر، طراحی‌های مناسب‌تر، بازدید فنی و ... قابلیت اعتماد اجزای سیستم را افزایش می‌دهند.

۲- استفاده از اجزای مازاد به صورت موازی: این روش نسبت به افزایش قابلیت اعتماد (دسترسی) اجزا اقتصادی‌تر می‌باشد زیرا هزینه هر جزء ممکن است با افزایش قابلیت اعتماد (دسترسی) به صورت نمایی افزایش یابد [۲]. این مساله را مساله تخصیص اجزای مازاد^۵ می‌نامند (RAP).

^۱ International Electrotechnical Commission

^۲ Reliability

^۳ Entity

^۴ Availability

^۵ Redundancy Allocation Problem

۳- ترکیبی از افزایش قابلیت اعتماد (دسترسی) اجزا و استفاده از اجزای مازاد به صورت موازی: این مسائل به مسائل قابلیت اعتماد- تخصیص اجزای مازاد (RRAP) یا برای قابلیت دسترسی به مسائل قابلیت دسترسی- تخصیص اجزای مازاد (ARAP) مشهور هستند [۳].

۴- انتقال اجزای تعویض پذیر

به علت محدودیت های اقتصادی و تکنولوژیکی در حال حاضر از تخصیص اجزای مازاد بعنوان بهترین و کاربردی ترین روش برای افزایش قابلیت اعتماد سیستم ها یاد می شود [۴، ۵]. در مدلسازی مسائل تخصیص اجزای مازاد فرضیاتی در نظر گرفته می شود: فرضیات عمومی مساله RAP [۶] :

سیستم یا اجزای آن دو حالت دارند یا کاملاً سالم یا کاملاً خراب.

اجزا به تعداد نامحدود در دسترس هستند.

خصوصیات اجزا از قبیل قابلیت اعتماد و هزینه و وزن مشخص و قطعی هستند.

خرابی اجزا معیوب تاثیری در قسمت های دیگر ندارد و قطعات تعمیر ناپذیر هستند.

اجزای مازاد همواره در حال کار محسوب می شوند (یعنی نرخ خرابی همه اجزا یکسان در نظر گرفته می شود و به محض به کار افتادن هم ممکن است خراب شود).

در بسیاری از تحقیقات صورت گرفته محققان روی یک هدف مانند افزایش قابلیت اعتماد (دسترسی) سیستم یا مینیم کردن هزینه تمرکز کرده اند. اما در عمل یک تصمیم گیرنده یا یک طراح سیستم مجبور است که گاهی چند هدف متفاوت و در عین حال متضاد را در کنار هم در نظر بگیرد. بنابراین مسائل چندهدفه یک جنبه حیاتی از مسائل مربوط به بهینه سازی قابلیت اعتماد (دسترسی) سیستم ها تبدیل شدند. ثابت شده است که مسائل تخصیص اجزای مازاد (RAP) چه در حالت تک هدفه و چه در حالت دو هدفه و بیشتر از دسته مسائل NP-HARD هستند [۷]. در مسائل NP-HARD با افزایش اندازه مساله و محدودیت ها حجم محاسبات به صورت نمایی افزایش پیدا می کند و این امر محاسبات کامپیوتری را پیچیده می کند. از نمونه این مسائل می توان به مساله فروشنده دوره گرد و مساله یافتن بزرگترین خوشه نیز اشاره کرد. راه های مختلفی از جمله روش های ابتکاری^۶ و فراابتکاری^۷ برای حل این مسائل وجود دارند. عمدتاً در مسائل تخصیص اجزای مازاد (RAP) دو رویکرد در نظر گرفته می شود [۸]:

- مشخصات فیزیکی از جمله قابلیت اعتماد، وزن، حجم و... ثابت هستند. در این دسته از مسائل تعیین تعداد و

^۶Heuristic

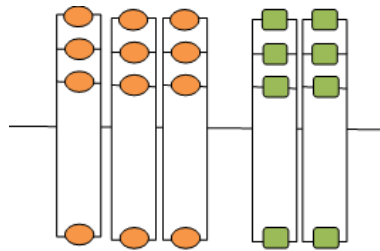
^۷Meta Heuristic

نوع اجزای مازاد به عنوان متغیر تصمیم در نظر گرفته می شود. در این حالت تابع هدف را به صورت ماکسیم سازی قابلیت اعتماد (دسترسی) و کمینه کردن هزینه مطرح می کنیم و محدودیت ها به صورت خطی است. - در دسته دوم از مسائل، قابلیت اعتماد اجزا مشخص نیست و به عنوان یک متغیر تصمیم در نظر گرفته می شود و سایر مشخصات فیزیکی همانند وزن، هزینه، حجم و غیره بعنوان یک تابع افزایشی از قابلیت اعتماد است. در این دسته تابع هدف برای مینیم کردن هزینه و همچنین محدودیت ها غیرخطی هستند. این دسته مسائل RRAP یا ARAP نامیده می شوند. در سال های اخیر تحقیقات گسترده ای روی مساله تخصیص اجزای مازاد صورت گرفته است. راوی و همکاران [۹] در سال ۲۰۰۰ مساله بهینه سازی قابلیت اعتماد سیستم های پیچیده با اهداف مختلف را مطرح نمودند. در سال ۲۰۰۲ هسیه [۱۰] روش برنامه ریزی خطی برای حل مساله تخصیص اجزای مازاد که به صورت مدل عدد صحیح غیرخطی طراحی شده بود پیشنهاد کردند. دو سال بعد کیت و همکاران [۱۱] مساله تخصیص اجزای مازاد دو هدفه را با اهداف بیشینه کردن قابلیت اعتماد و مینیم کردن واریانس قابلیت اعتماد کل سیستم در شرایط عدم قطعیت مطرح کردند و از روش جمع اوزان برای تخمین مجموعه بهینه پارتو استفاده کردند. رامیرز و همکاران [۱۲] مدلی برای سیستم های سری موازی با هدف بیشینه کردن حداقل قابلیت اعتماد زیر سیستم ها مطرح کردند. در سال ۲۰۰۵ ناهاس و نوافاس [۱۳] مدلی برای ماکزیم کردن قابلیت اعتماد یک سیستم سری بدون تخصیص اجزای مازاد با در نظر گرفتن محدودیت بودجه ارائه کردند. در سال ۲۰۰۶ چن [۱۴] یک مدل برای مساله تک هدفه بیشینه کردن قابلیت اعتماد با در نظر گرفتن محدودیت های وزن و حجم و هزینه ارائه کرد و برای حل آن از یک روش فراابتکاری استفاده کرد. ژائو و همکاران [۱۵] در سال ۲۰۰۷ مساله بهینه سازی تک هدفه قابلیت اعتماد را با در نظر گرفتن دو محدودیت وزن و حجم در نظر گرفتند و از الگوریتم کلونی مورچگان^۸ برای حل آن استفاده کردند و نتایج را با الگوریتم ژنتیک^۹ مقایسه کردند و به این نتیجه رسیدند که الگوریتم کلونی مورچگان جواب بهتری تولید کرده است. در همان سال تابودا و همکاران [۱۶] یک مساله بهینه سازی سه هدفه با اهداف بیشینه کردن قابلیت اعتماد و کمینه کردن هزینه و وزن با محدودیت تعداد اجزای مازاد انتخابی در نظر گرفتند و از روش NSGA-II برای تولید مجموعه بهینه پارتو^{۱۰} استفاده کردند و برای کاهش سائز این مجموعه از دو روش استفاده کردند. در همان سال کوناک و همکاران [۱۷] دو نوع مختلف از مسائل بهینه سازی سه هدفه را در نظر گرفتند و الگوریتم جست و جوی ممنوعه^{۱۱} برای حل آن ارائه کردند و سپس از روش شبیه

^۸Ant Colony Optimization^۹Genetic Algorithm^{۱۰}Pareto^{۱۱}Tuba Search

سازی مونتو کارلو^{۱۲} برای کاهش سائز مجموعه پارتو استفاده کردند. در سال ۲۰۰۷ لیانگ و همکاران [۱۸] یک مساله تک هدفه بیشینه کردن قابلیت اعتماد را با محدودیت های هزینه و وزن در نظر گرفتند و روش جست و جوی همسایگی متغیر^{۱۳} را برای حل آن پیشنهاد کردند و نشان دادند این روش نسبت به سایر الگوریتم های فراابتکاری مانند الگوریتم ژنتیک، الگوریتم کلونی مورچگان و الگوریتم جست و جوی ممنوعه نتایج بهتری را فراهم می کند. در سال ۲۰۰۹ لیانگ و همکاران [۱۹] سه نوع مساله متفاوت چند هدفه را در نظر گرفتند و از روش جست و جوی چندمتغیره چندهدفه برای حل آن استفاده کردند. برخلاف قابلیت اعتماد تحقیقات اندکی روی بهینه سازی قابلیت دسترسی صورت گرفته است.

۲- تعریف مساله: در اغلب تحقیقات صورت گرفته همواره یک سیستم تعمیرناپذیر یا تعمیرپذیر در نظر گرفته شد؛ اما در دنیای واقع کمتر سیستمی یافت می شود که صرفاً تعمیرناپذیر یا صرفاً تعمیرپذیر باشد. برای اولین بار ذوالفقاری و همکاران [۲۰] یک سیستم سری موازی را در نظر گرفتند (شکل ۱) که شامل s زیر سیستم است و در آن همانطور که ملاحظه می شود برخی زیر سیستم ها دارای اجزای تعمیرناپذیر و برخی دارای اجزای تعمیرپذیر هستند. اجزا نیز درون هر زیر سیستم متفاوت در نظر گرفته شد و فرض شد در هر زیر سیستم m_i جز وجود دارد. با



شکل ۱: سیستم سری موازی با زیرسیستم های تعمیرناپذیر و تعمیرپذیر

در نظر گرفتن این فرض مساله به دنیای واقع نزدیکتر شد اما سیستم هایی وجود دارند که در آن در یک زیر سیستم اجزای تعمیرپذیر و تعمیرناپذیر وجود دارند. بنابراین در دیدگاه سیستمی در یک زیر سیستم از یک دستگاه اجزای تعمیرپذیر و اجزای تعمیرناپذیر در کنار هم قرار دارند. به طور مثال خطوط لوله انتقال نفت از این نوع سیستمها می باشند. با توجه به شکل زیر، خطوط انتقال نفت همواره از قطعات تعمیرپذیر مانند شیرهای کنترلی و قطعات

^{۱۲}Monto-carlo Simulated

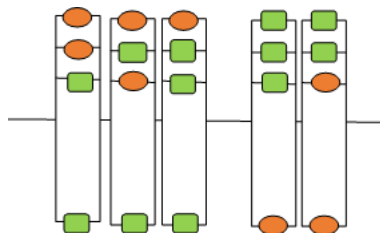
^{۱۳}Variable Neighborhood Search

تعمیرناپذیر مانند لوله‌های واسطه تشکیل می‌شوند که قسمتی از خط انتقال که یک زیرسیستم در نظر گرفته می‌شود، متشکل از تعدادی شیرکترلی و لوله واسطه می‌باشد.



شکل ۲: خطوط انتقال نفت

در شکل زیر شمایی چنین سیستمی را مشاهده می‌کنیم.



شکل ۳: سیستم سری موازی با اجزای تعمیرناپذیر و تعمیرپذیر در هر زیرسیستم

مدل ریاضی: برای مدلسازی مساله چون اجزای تعمیرپذیر و اجزای تعمیرناپذیر در کنار یک دیگر قرار دارند دیگر قابلیت اعتماد مطرح نیست و بهینه سازی قابلیت دسترسی مطرح می‌شود. از جمله محدودیت های مهم که در اغلب سیستم ها در نظر گرفته می‌شود می‌توان به هزینه کل طراحی سیستم، وزن و حجم کل سیستم اشاره کرد. گاه برخی از محدودیت ها به دلیل اهمیت به صورت یک هدف در می‌آیند و تصمیم گیرنده مجبور است علی‌رغم متضاد بودن هدف ها مسائل چندهدفه را برای یافتن حداقل یک راه حل برتر که می‌تواند در طراحی سیستم اجرا شود حل کنند.

$$\max f_1 = \max A_{sys} = \prod_{i=1}^s AV_i = \prod_{i=1}^s [1 - (\prod_{j=1}^{m_i} (1 - R_{ij}) \wedge (n_{ij})) * (\prod_{k=1}^{d_i} (1 - A_{ik}) \wedge (n'_{ik}))] \quad (1)$$

$$\min f_{\text{r}} = \sum_{i=1}^s \left[\sum_{j=1}^{m_i} c_{ij} * n_{ij} + \sum_{k=1}^{d_i} c'_{ik} * n'_{ik} \right] \quad (2)$$

به طوریکه:

$$\sum_{i=1}^s \left[\sum_{j=1}^{m_i} w_{ij} * n_{ij} + \sum_{k=1}^{d_i} w'_{ik} * n'_{ik} \right] \leq W \quad (3)$$

$$\sum_{i=1}^s \left[\sum_{j=1}^{m_i} v_{ij} * n_{ij} + \sum_{k=1}^{d_i} v'_{ik} * n'_{ik} \right] \leq V \quad (4)$$

$$P_i \leq \sum_{j=1}^{m_i} n_{ij} + \sum_{k=1}^{d_i} n'_{ik} \leq N_i \quad \forall i \in \{1, \dots, s\} \quad (5)$$

$$n_{ij}, n'_{ik} \in Z^+ \quad \forall i \in \{1, \dots, s\} \quad j \in \{1, \dots, m_i\} \quad k \in \{1, \dots, d_i\} \quad (6)$$

A_{sys} : قابلیت دسترسی کل سیستم

Av_i : قابلیت دسترسی زیر سیستم i ام R_{ij} : قابلیت اعتماد جز تعمیرناپذیر j در زیر سیستم i

A_{ik} : قابلیت دسترسی جز تعمیرپذیر k در زیر سیستم i

s : تعداد زیرسیستم ها

n_{ij} : تعداد اجزای تعمیرناپذیر نوع j تخصیص داده شده به زیر سیستم i

n'_{ik} : تعداد اجزای تعمیرپذیر نوع k تخصیص داده شده به زیر سیستم i

c_{ij} : هزینه اجزای تعمیرناپذیر نوع j ام در زیر سیستم i ام

c'_{ik} : هزینه اجزای تعمیر پذیر نوع k ام در زیر سیستم i ام

m_i : انواع اجزای تعمیرناپذیر موجود در زیر سیستم i ام

d_i : انواع اجزای تعمیرپذیر موجود در زیر سیستم i ام

w_{ij} : وزن جز تعمیرناپذیر نوع j ام در زیر سیستم i ام

w'_{ik} : وزن جز تعمیرپذیر نوع k ام در زیر سیستم i ام

v_{ij} : حجم جز تعمیرناپذیر نوع j ام در زیر سیستم i ام

v'_{ik} : حجم جز تعمیرپذیر نوع k ام در زیر سیستم i ام

W : حداکثر وزن مجاز سیستم

V : حداکثر حجم مجاز سیستم

P_i : حداقل تعداد قطعات در زیر سیستم i ام برای کارکرد سیستم

N_i : حداکثر تعداد قطعات در زیر سیستم i ام

شرح محدودیت ها

رابطه (۱) بیانگر ماکسیم سازی قابلیت دسترسی می باشد. در رابطه (۲) تابع هدف دوم که به صورت مینیم سازی هزینه است مشاهده می شود. رابطه های (۳) و (۵) بیانگر محدودیت های وزن و حجم سیستم هستند که از یک وزن و حجم از پیش تعیین شده W و V نمی تواند فراتر رود. و رابطه (۸) بیانگر محدودیت برای تعدادی اجزای انتخابی در هر زیر سیستم است که همواره این مقدار باید بین دو مقدار P_i و N_i برای هر زیر سیستم i باشد. در این مقاله بیشینه کردن قابلیت دسترسی و کمینه کردن هزینه به عنوان دو هدف در نظر گرفته می شود. همچنین قابلیت اعتماد و قابلیت دسترسی اجزا ثابت در نظر گرفته شده است و روابط بین محدودیت ها نیز خطی در نظر گرفته شده است. فرض می شود درون هر زیر سیستم m_i جز تعمیرناپذیر و d_i جز تعمیرپذیر وجود دارد. مساله تعیین تعداد اجزای مازاد از هر نوع درون هر زیر سیستم است.

۳- مثال عددی و حل مساله

برای شبیه سازی یک مثال عددی از اطلاعات موجود در ادبیات موضوع برای انواع مختلف جز ها با خصوصیات متفاوت (قابلیت اعتماد، هزینه، وزن و حجم) استفاده کردیم و یک سیستم با شش زیر سیستم را با اطلاعات موجود در جدول زیر شبیه سازی کردیم.

همانطور که در جدول مشخص شده است سیستمی را با شش زیر سیستم در نظر گرفته ایم؛ که درون هر زیر سیستم حداقل یک نوع جز تعمیرپذیر و تعمیرناپذیر و حداکثر سه نوع جز تعمیرناپذیر و چهار نوع جز تعمیرپذیر مختلف می تواند وجود داشته باشد. که از هر کدام به هر تعدادی که انتخاب شود فقط یکی از اجزا در حال کار است و سایر اجزا نیز در حالت فعال هستند تا به محض شکست جز اول جایگزین آن شوند. برای حل مثال عددی از نرم افزار گمز استفاده کردیم. نرم افزار گمز قادر به حل مسائل در حالت تک هدفه می باشد. به این دلیل از روش ϵ محدودیت برای حل مساله مورد نظرمان استفاده کردیم. بدین ترتیب که تابع هدف دوم یعنی تابع هزینه را به صورت یک محدودیت در نظر گرفتیم و چون به صورت مینیم سازی می باشد، کوچکتر مساوی یک ϵ دلخواه قرار می دهیم و هربار با تغییر این مقدار، مقادیر مجهول (که در اینجا تعداد اجزای انتخاب شده از هر نوع جز در هر زیر سیستم می باشد) را به گونه ای که مقدار تابع هدف اول در حالت ماکسیم قرار بگیرد، محاسبه می کنیم.

جدول ۱: مثال

۳				۲				۱				
V	W	C	R	V	W	C	R	V	W	C	R	
۸	۶	۱۱	۰/۹۱	۱۲	۹	۱۸	۰/۹۴	۴	۶	۱۸	۰/۹۷	نوع ۱
۱۰	۷	۱۵	۰/۹۵	۷	۵	۲۲	۰/۹۷	۶	۸	۲۰	۰/۹۸	نوع ۲
—	—	—	—	۸	۶	۲۱	۰/۹۶	۵	۷	۱۶	۰/۹۶	نوع ۳
۷	۵	۴	۰/۸۶	۱۱	۸	۷	۰/۸۸	۷	۵	۹	۰/۹	نوع ۱
۹	۵	۲	۰/۸۵	۵	۴	۶	۰/۸۶	۴	۹	۷	۰/۸۵	نوع ۲
—	—	—	—	۷	۴	۸	۰/۹	۸	۸	۸	۰/۸۷	نوع ۳
—	—	—	—	—	—	—	—	۵	۷	۱۰	۰/۹۱	نوع ۴

در جدول زیر نتایج حاصل از حل مساله شامل مقدار تابع هدف با ϵ های مختلف و همچنین مقادیر تعداد اجزای انتخابی از هر نوع جز در هر زیر سیستم آورده شده است. لازم به ذکر است که حداکثر وزن و حجم مجاز برای سیستم ۲۰۰ واحد در نظر گرفته شده است و در هر زیر سیستم حداقل ۱ جز و حداکثر ۵ جز می تواند انتخاب شود. در جداول زیر نتایج حاصل از حل مثال عددی با ϵ های مختلف و مقادیر تابع هدف و تعداد اجزای انتخابی آورده شده است:

جدول ۲: مثال

۶				۵				۴				
V	W	C	R	V	W	C	R	V	W	C	R	
۵۳	۴۲/۱	۷۵/۵	۰/۹۷۱	۳۲	۲۶/۵	۳۰/۵	۰/۹۹۵	۱۱	۸	۱۲	۰/۹	نوع ۱
۵۰	۴۰/۸	۹۱	۰/۹۹۷	۳۰	۲۵/۴	۲۸	۰/۹۷۷	—	—	—	—	نوع ۲
—	—	—	—	۲۸	۲۲/۴	۲۸/۹	۰/۹۹۶	—	—	—	—	نوع ۳
۵۰	۳۹/۶	۲۴/۲	۰/۷	۲۹	۲۴/۷	۱۸	۰/۹۷۱	۱۰	۵	۴	۰/۸۱	نوع ۱
۵۱	۴۱/۷	۵۰	۰/۸	۲۸	۲۰/۳	۱۴	۰/۹	—	—	—	—	نوع ۲
—	—	—	—	۲۳	۱۵	۱۶	۰/۹۳	—	—	—	—	نوع ۳
—	—	—	—	۲۶	۱۴/۹	۱۵	۰/۹۱	—	—	—	—	نوع ۴

جدول ۳: نتایج

	€	f_1
$n_{53} = 1, n_{61} = 1, m_{11} = 1, m_{22} = 3, m_{32} = 3, m_{41} = 4, m_{12} = 4$	۲۰۰	۰,۹۶۱۵۸۴۶۰
$n_{11} = 2, n_{31} = 5, n_{53} = 1, n_{62} = 1, m_{22} = 3, m_{41} = 4$	۲۵۰	۰,۹۶۹۹۷۲۵۴
$n_{11} = 2, n_{22} = 1, n_{31} = 1, n_{32} = 1, n_{41} = 1$	۳۰۰	۰,۹۷۵۵۲۶۷۶
$n_{53} = 1, n_{61} = 1, m_{22} = 1, m_{41} = 2, m_{61} = 1$	۳۰۰	۰,۹۷۵۵۲۶۷۶
$n_{11} = 2, n_{22} = 2, n_{31} = 3, n_{41} = 2, n_{53} = 1, n_{62} = 2$	۳۵۰	۰,۹۸۳۰۲۸۱۰
$n_{11} = 2, n_{22} = 2, n_{31} = 3, n_{41} = 2, n_{53} = 1, n_{62} = 2, m_{12} = 1$	۴۰۰	۰,۹۸۳۷۸۰۷۹

۴- نتیجه گیری

در دنیای واقعی کمتر سیستمی یافت می شود که صرفاً تعمیرپذیر یا تعمیرناپذیر باشد. در این تحقیق با در نظر گرفتن یک سیستم سری موازی که در هر زیر سیستم آن ترکیبی از اجرای تعمیرناپذیر و تعمیرپذیر وجود دارد، گام بلندی جهت نزدیک کردن سیستم های مهندسی به شرایط دنیای واقع برداشته می شود. در نتیجه طراحان با طراحی سیستم های انعطاف پذیرتر و با قابلیت دسترسی بیشتر از یک سو حضور مستمر خود را در صحنه رقابت تضمین می کنند و از سوی دیگر از خطرات جدی و جبرانناپذیری که برای سیستم ها به خصوص در صنایع حساسی مثل

صنایع نظامی و هوایی به وجود می آید جلوگیری می کنند و از هزینه های زیاد ناشی از خرابی سیستم ها جلوگیری می کنند. همواره افزودن جز به صورت موازی به سیستم باعث افزایش قابلیت اعتماد (دسترسی) سیستم می شود. یک راه مهم برای افزایش قابلیت اعتماد (دسترسی) سیستم ها افزودن اجزای مازاد به صورت موازی به سیستم است تا در صورت شکست یکی از اجزا جز مازاد جایگزین جز شکست خورده شود و سیستم کار خودش را به بهترین نحو انجام دهد. با اینکه در مساله تخصیص اجزای مازاد فرض بر این است که اجزا به تعداد نامحدود در دسترس هستند اما برای افزودن این اجزا با محدودیت هایی از قبیل هزینه و وزن و حجم روبرو هستیم. بنابراین در این تحقیق مدلسازی مساله را با در نظر گرفتن این محدودیت ها انجام می دهیم. برخی از محدودیت ها مانند هزینه به دلیل اهمیت به صورت یک هدف در می آیند و مساله دو هدفه را برای این دست مسائل در نظر می گیریم.

مراجع

- [1] Ardakan, M. A., Hamadani, A. Z., Alinaghian, M. (2015). Optimizing bi-objective redundancy allocation problem with a mixed redundancy strategy. ISA transactions, 55, 116-128.
- [2] Misra, K. B., Ljubojevic, M. D. (1973). Optimal reliability design of a system: a new look. IEEE Transactions on Reliability, 22(5), 255-258
- [3] Garg, H., Rani, M., Sharma, S. P., Vishwakarma, Y. (2014). Bi-objective optimization of the reliability-redundancy allocation problem for series-parallel system. Journal of Manufacturing Systems, 33(3), 335-347.
- [4] Yun, W. Y., Kim, J. W. (2004). Multi-level redundancy optimization in series systems. Computers Industrial Engineering, 46(2), 337-346.
- [5] Utkin, L. V. (1994). Redundancy optimization by fuzzy reliability and cost of system components. Microelectronics Reliability, 34(1), 53-59.
- [6] Wu, J. S., Chen, R. J. (1994). An algorithm for computing the reliability of weighted-k-out-of-n systems. IEEE Transactions on Reliability, 43(2), 327-328
- [7] Chern, M. S. (1992). On the computational complexity of reliability redundancy allocation in a series system. Operations research letters, 11(5), 309-315.

- [8] Coit, D. W. (2003). Maximization of system reliability with a choice of redundancy strategies. IIE transactions, 35(6), 535-543
- [9] Ravi, V., Reddy, P. J., Zimmermann, H. J. (2000). Fuzzy global optimization of complex system reliability. IEEE Transactions on Fuzzy systems, 8(3), 241-248.
- [10] Hsieh, Y. C. (2003). A linear approximation for redundant reliability problems with multiple component choices. Computers industrial engineering, 44(1), 91-103.
- [11] Coit, D. W., Jin, T., Wattanapongsakorn, N. (2004). System optimization with component reliability estimation uncertainty: a multi-criteria approach. IEEE transactions on reliability, 53(3), 369-380.
- [12] Ramirez-Marquez, J. E., Coit, D. W., Konak, A. (2004). Redundancy allocation for series-parallel systems using a max-min approach. Iie Transactions, 36(9), 891-898.
- [13] Nahas, N., Nourelfath, M. "Ant system for reliability optimization of a series system with multiple-choice and budget constraints", Reliability Engineering System Safety vol.87, pp.1-12, 2005.
- [14] Chen, T. C. (2006). IAs based approach for reliability redundancy allocation problems. Applied Mathematics and Computation, 182(2), 1556-1567.
- [15] Zhao, J. H., Liu, Z., Dao, M. T. (2007). Reliability optimization using multiobjective ant colony system approaches. Reliability Engineering System Safety, 92(1), 109-120.
- [16] Taboada, H. A., Baheranwala, F., Coit, D. W., Wattanapongsakorn, N. (2007). Practical solutions for multi-objective optimization: An application to system reliability design problems. Reliability Engineering System Safety, 92(3), 314-322. system. Journal of Manufacturing Systems, 33(3), 335-347.
- [17] Kulturel-Konak, S., Coit, D. W., Baheranwala, F. (2008). Pruned Pareto-optimal sets for the system redundancy allocation problem based on multiple prioritized objectives. Journal of Heuristics, 14(4), 335-357.
- [18] Liang, Y. C., Chen, Y. C. (2007). Redundancy allocation of series-parallel systems using a variable neighborhood search algorithm. Reliability Engineering System Safety, 92(3), 323-331.

- [19] Liang, Y. C., Lo, M. H. (2010). Multi-objective redundancy allocation optimization using a variable neighborhood search algorithm. *Journal of Heuristics*, 16(3), 511-535
- [20] Zoulfaghari, H., Hamadani, A. Z., Ardakan, M. A. (2014). Bi-objective redundancy allocation problem for a system with mixed repairable and non-repairable components. *ISA transactions*, 53(1), 17-24.



تعیین سن شروع مصرف کراک با استفاده از آنالیز بقا و بررسی عوامل موثر بر مصرف آن با استفاده از مدل کاکس و رگرسیون لجستیک

برومند، ف^۱ باغستانی، ا^۲

^۱ کمیته تحقیقات دانشجویی، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی مشهد

^۲ گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی

چکیده

در این مقاله تلاش می شود علاوه بر برآورد سن شروع مصرف کراک، به این سوال پاسخ داده شود که آیا جنس، سطح تحصیلات و وضعیت ازدواج در مصرف کراک موثر است یا خیر. در این مطالعه، ۱۰۰۰ نفر با تکمیل پرسشنامه مورد بررسی قرار گرفتند و تجزیه و تحلیل داده‌ها به وسیله آنالیز بقا، رگرسیون کاکس و رگرسیون لجستیک انجام پذیرفت. عوامل موثر بر مصرف کراک در مدل کاکس، وضعیت ازدواج و سطح تحصیلات به ترتیب مقدارهای $p=0.022$ و $p=0.035$ معنی دار شدند. همچنین در گام بعدی مقایسه مدل های کاکس و رگرسیون چند متغیره لجستیک در دستور کار قرار گرفت. متغیرهای وضعیت ازدواج (0.043) و سطح تحصیلات (0.011) معنی دار شدند. با توجه به مدل کاکس سن شروع مصرف، ۱۷ سال گزارش شد. بر اساس مدل های برازش شده، وضعیت ازدواج و سطح تحصیلات بر مصرف کراک تاثیرگذار شناخته شدند.

کلمات کلیدی: مدل کاکس، عوامل موثر بر مصرف کراک، رگرسیون لجستیک.

^۱ farzaneh.boroumand@hotmail.com

^۲ baghestani.ar@gmail.com

۱ مقدمه

در حال حاضر اعتیاد به مواد مخدر یکی از مسایل مهم و اصلی سلامت جسمی و بهداشت روانی در ایران و سراسر دنیاست. (Teimouri et al., ۲۰۱۱). در مقاله پیش رو که به طور خاص بر مصرف کراک تاکید دارد تلاش می شود با تشخیص درست سن اولیه مصرف، عوامل خطر ساز برای استفاده کراک را شناسایی کرد. در ایران هر روزه شاهد پایین آمدن سن مصرف کراک هستیم. پایین آمدن سن شروع مصرف کراک حتی از مصرف آن نیز خطرناک تر است، (بهزادفر et al., ۲۰۱۳) نوجوانان در هر جامعه ای نیروی بلقوه رشد و شکوفایی آن جامعه به ویژه در زمینه های علمی و ورزشی هستند. حال فرض کنیم درگیر شدن نوجوانان در این سن حساس به موادی چون کراک چه صدمات جبران ناپذیری را بر پیکره جامعه وارد خواهد کرد. پس برانیم تا با تعیین درست سن شروع استفاده از کراک تمهیداتی برای کاهش زمینه های خطر در این سن را فراهم کرد، همچنین با بررسی عوامل موثر بر مصرف کراک، خطر استفاده جوانان از کراک را کاهش داد.

مواد و روش ها: این پژوهش از نوع مشاهده ای است که در آن ۱۰۰۰ نفر از مردم شهر تهران در سال ۹۱ مورد بررسی قرار گرفتند. اطلاعات مربوط به این ۱۰۰۰ نفر از طریق تکمیل پرسشنامه در پارک های شهر تهران بدست آمده است.

داده های جمع آوری شده، کد گذاری و وارد کامپیوتر شد سپس با استفاده از نرم افزار SPSS ۲۰ مورد تجزیه و تحلیل قرار گرفت و سطح معنی ۰/۰۵ در نظر گرفته شد. برای تعیین سن شروع مصرف منحنی بقا مورد استفاده قرار گرفت و مدل حاصل با استفاده از رگرسیون کاکس برازش داده شد و در گام آخر مدل کاکس و لجستیک از نظر تعیین عوامل موثر بر مصرف مورد مقایسه قرار گرفتند.

یافته ها: در این مطالعه ۱۰۰۰ نفری ۴۷۹ نفر مرد و ۵۲۱ نفر زن حضور داشتند که در این جمع ۵۱۸ نفر مجرد و ۴۴۹ نفر متاهل و ۳۳ نفر طلاق گرفته بودند.

همچنین در بررسی سطح تحصیلات این افراد مشخص شد، ۱۳ نفر تحصیلات ابتدایی، ۵۰ نفر تحصیلات راهنمایی، ۱۳۴ نفر تحصیلات متوسطه، ۵۴۶ نفر دیپلم و فوق دیپلم و در نهایت ۲۵۷ نفر تحصیلات لیسانس و بالاتر داشته اند.

در گام اول برای مشخص کردن سن شروع مصرف منحنی بقا رسم شد. (نمودار شماره ۱) با توجه به نمودار سن شروع مصرف ۱۷ سالگی بدست آمد.

در گام دوم برای برازش مدل کاکس به داده‌ها برای تعیین عوامل موثر بر مصرف کراک، متغیرهای جنس وضعیت ازدواج و سطح تحصیلات وارد مدل شدند. (جدول شماره ۱) متغیرهای وضعیت ازدواج (مقدار احتمال برابر ۰/۰۲۲) و سطح تحصیلات (مقدار احتمال برابر ۰/۰۳۵) معنی دار شدند. بدین ترتیب متغیرهای وضعیت ازدواج و سطح تحصیلات در عدم گرایش به استفاده از ماده مخدر کراک تاثیر دارد. در نهایت با مقایسه مدل کاکس و لجستیک عوامل موثر را می یابیم. در جدول شماره ۲ متغیرهای معنی دار مدل رگرسیون لجستیک را می توان مشاهده کرد. متغیر وضعیت تاهل (مقدار احتمال برابر ۰/۰۴۳) و سطح تحصیلات (مقدار احتمال ۰/۰۱۱) معنی دار شدند.

۲ بحث

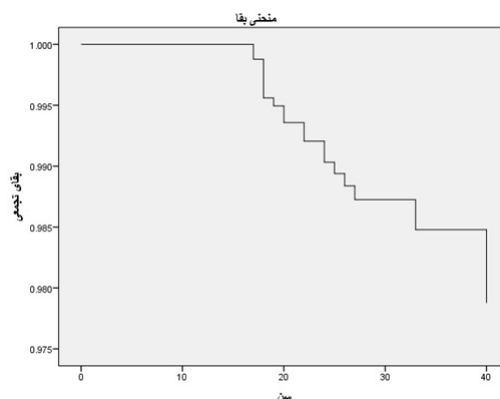
مقاله پیش رو به اولین مقاله درباره تعیین سن شروع مصرف کراک و یافتن عوامل موثر بر گرایش به مصرف با رویکرد انالیز بقا است. در گذشته افرادی چون سراج زاده و همکاران (۱۳۸۵) مقاله ای با عنوان عوامل اجتماعی موثر بر مصرف تریاک و مشروبات الکلی در بین دانشجویان (زاده و همکاران ۲۰۰۷) و نقی زاده و همکاران (۱۳۹۲) با مقاله ای با عنوان "بررسی الگوی مصرف مواد مخدر در زنان خودمعرف شهرستان فسا" کارهای مشابهی انجام دادند. (زاده و همکاران ۲۰۱۴)

پژوهشگرانی چون منصوریان و همکاران (۱۳۸۹) در مقاله ای تحت عنوان "بررسی اپیدمیولوژیکی علل گرایش به مصرف سیگار و دخانیات در جمعیت بالای ۱۵ سال شهرستان دامغان در سال ۱۳۸۸" سن شروع به مصرف سیگار با استفاده از میانگین حسابی برآورد کردند.

در حالی که در این مقاله برای تعیین سن شروع به مصرف از منحنی بقا استفاده شد و عدد بدست آمده ۱۷ سال بود با این وجود اگر از روش کلاسیک، میانگین حسابی سن افراد، برای محاسبه استفاده کنیم سن شروع مصرف ۲۲ سال و ۵ ماه بدست خواهد آمد. بنابراین استفاده از میانگین حسابی برای بدست آوردن سن شروع مصرف گمراه کننده خواهد بود و برآوردی با اختلاف ۵ سال نسبت به سن واقعی شروع مصرف به دست می دهد. همچنین بهزادفر و همکاران (۱۳۹۲) در مقاله ای با عنوان "فراوانی اختلال بیش فعالی و عوامل خطر ساز در نوجوانان دختر و پسر مبتلا به سوء مصرف مواد" برای سن شروع به مصرف مواد مخدر بازه ۲۰ تا ۴۰ سال را گزارش کردند. (بهزادفر و همکاران ۲۰۱۳)

همانطور که در نتایج این پژوهش آمده وضعیت تاهل بر مصرف کراک معنی دار شده است این در حالی است که اکبری زردخانه و همکاران (۱۳۸۹) در مقاله ای تحت عنوان ”ویژگی‌های جمعیت شناختی و مصرف مواد در دانشجویان دانشگاه‌های دولتی“ این موضوع اشاره دارند که در گروه دانشجویان مجرد و جدا شده نسبت مصرف کراک بالاتر از گروه دانشجویان متأهل است. (اکبری زردخانه سعید، ۱۳۸۹) همین طور امانی و همکاران (۱۳۹۱) معنی داری وضعیت تاهل با مصرف مواد مخدر در مقاله ای تحت عنوان ”بررسی روند تغییرات الگوی سوء مصرف مواد مخدر در معتادان مراجعه کننده به مراکز ترک اعتیاد طی سال‌های ۸۷-۸۲“ نشان دادند.

همچنین در این مطالعه نشان دادیم جنس و مصرف کراک رابطه معنی دار ندارند این در حالیست که حاجیان و همکاران (۱۳۹۲) در مقاله ای با عنوان ”بررسی همه‌گیری شناختی اعتیاد در معتادان خودمعرف مراجعه کننده به مراکز اعتیاد“ نشان دادند که مصرف کراک در مردان بیش تر از زنان است. به نظر می‌رسد این تفاوت در نتایج به دلیل پایین بودن حجم نمونه مطالعه حاضر باشد زیرا حجم نمونه بر معنی داری عوامل بسیار موثر است.



۳ نتیجه‌گیری

در مطالعه حاضر براساس تحلیل‌های انجام شده سن شروع مصرف کراک در سال ۱۳۹۱ در تهران، ۱۷ سال گزارش شد. همچنین مدل‌های کاکس و لجستیک نشان می‌دهند وضعیت ازدواج و سطح تحصیلات در مصرف کراک موثرند.

مراجع

- [۱] ق. بابایی تهرانی و ا. ع. جمشیدیان، *مدل‌های آماری*، انتشارات پارسه، ۱۳۹۱.
- [۲] ع. س. علوی، ر. قنبریا و ع. قنبریا، *استفاده از نظریه صف در مخابرات*، اندیشه آماری ۵ (۱۳۹۰)، شماره ۱، ۱-۱۸.



بررسی عوامل موثر بر بقای سرطان خون کودکان با استفاده از فرایندهای نیمه مارکوفی

بهرامپور، ع^۱ محمودی منش، م^۲ فرهمندی نیا، ز^۳

^۱ مرکز تحقیقات مدل سازی در سلامت دانشگاه علوم پزشکی کرمان، گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه

علوم پزشکی کرمان

^۲ گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی کرمان

^۳ گروه آموزشی بیماری های کودکان، مرکز آموزشی درمانی افضل پور، دانشگاه علوم پزشکی کرمان

چکیده

در اکثر سرطان ها، پیشامدهایی همچون عود و متاستاز در طول مدت درمان برای بیمار رخ می دهد که می تواند نتیجه ی نهایی را تحت تاثیر قرار دهد. در نتیجه استفاده از مدل هایی که توزیع مدت زمان انتظار تا وقوع مرگ و زمان انتظار تا چنین پیشامدهایی را در نظر بگیرد می تواند ارزیابی جامع تر و بهتری از تابع بقا ارائه دهد. در این مطالعه از فرایند نیمه مارکوفی جهت تجزیه و تحلیل و تعیین اثر متغیرهای مختلف بر انتقال ها و زمان توقف در هر مرحله از زمان استفاده شده است.

کلمات کلیدی: سرطان خون، بقاء، فرایندهای نیمه مارکوفی.

^۱ abahrampour@yahoo.com

^۲ m_mahmudi69@yahoo.com

^۳ farahmandinia@yahoo.com

۱ مقدمه

یکی از کاربردهای نظریه احتمال، مربوط به داده هایی است که به طور دنباله ای در طول زمان جمع آوری می شوند. چنین مجموعه ای از مشاهدات را می توان به فرایندهای تصادفی منتسب کرد.

از جمله فرایندهای تصادفی، فرایندهای مارکوفی می باشد که اگر زمان را در این فرایند به سه دوره گذشته، حال و آینده تقسیم کنیم؛ آینده این فرایند بستگی به مسیری که در گذشته طی کرده است، ندارد و تنها به موقعیت آن در زمان حال وابسته است [۱، ۲].

بسیاری بیماری ها به گونه ای است که در طول مدت درمان، بیمار ممکن است چندین بار عود داشته باشد؛ بنابراین توانایی ارزیابی دوره های مختلف درمانی و مدت زمان رسیدن به هر حالت و احتمال و شانس بهبودی یا عود به عنوان تابعی از زمان می تواند ارزشمند باشد [۳].

محاسبه ی احتمالات بقا در بسیاری از مطالعات پزشکی حائز اهمیت است؛ در این مطالعات معمولاً فوت بیماران بر اثر عامل معین مورد توجه است، در صورتی که همانطور که گفته شد در اکثر بیماری ها؛ پیشامدهای مختلفی در طول مدت درمان اتفاق می افتد که می تواند بر نتیجه ی نهایی تاثیرگذار باشد و از جمله روش هایی هم که می تواند چنین پیشامدهایی را مورد بررسی قرار دهد، فرایندهای تصادفی و به خصوص فرایندهای مارکوفی است [۴].

به عنوان مثال در سرطان ها می توان حالات مختلفی مانند عود، متاستاز، مرگ پس از متاستاز، مرگ بدون متاستاز، بهبودی و ... و انتقالات بین این ها را در نظر گرفت و از مدل های مارکوفی جهت تحلیل بهره گرفت [۵].

در این گونه مطالعات، اطلاعات در دسترس، شامل دنباله ای از این حالت ها و مدت زمان ماندن در هریک از حالات می باشد و همچنین اطلاعاتی درباره ی متغیرهای مستقل و متغیرهای مداخله گر که همواره در دسترس است. انتظار می رود با داشتن این اطلاعات، ارزیابی جامع تر و بهتری از تابع بقا به دست آید [۵].

در فرایندهای نیمه مارکوفی، توزیع زمان ماندن در هر حالت توزیعی روی اعداد طبیعی ($\mathbb{N}$) در نظر گرفته می شود. زمانی که مسیر و حرکت فرایند وابسته به زمان شروع مطالعه باشد، یکی از مدل های مورد استفاده مدل نیمه مارکوفی است. در مدل های نیمه مارکوفی، توزیع زمان های ماندن در حالات مختلف فرایند، توزیع وایبل، نمایی یا وایبل نمایی در نظر گرفته می شود. این مدل های پارامتری اجازه ی ورود متغیرهای مداخله گر را در توزیع زمان ماندن با استفاده از مدل های رگرسیونی خطرات متناسب می دهد.

$(J_n, T_n)_{n \in \mathbb{N}}$ یک فرایند مارکوفی است به طوری که $0 < T_0 < T_1 < \dots < T_n < \infty$ زمان های ورود به حالات

$J_n \neq J_{n+1} ; ; n \in \mathbb{N}$ می باشد و به ازای هر $J_1, J_2, \dots, J_n$

و فضای حالت نیز متناهی و گسسته با احتمالات انتقال $P_{hj} = p(J_{n+1} = j | J_n = h)$ است.

اگر $S_n = T_n - T_{n-1}$ باشد و $N(t) = \sup\{n \in \mathbb{N} : T_n \leq t, t \in \mathbb{R}_+\}$ نیز یک فرایند شمارشی باشد که تعداد کل انتقالات مشاهده شده را در طول فاصله زمانی $[0, t]$ می شمارد؛ فرایند $J_{N(t)}$ را فرایند نیمه مارکوفی می نامند که نشان دهنده ی حالت فرایند در زمان t است [۶].

یک فرایند نیمه مارکوفی به صورت زیر است:

$$F_{hj}(d) = P(S_{n+1} \leq d | J_n = h, J_{n+1} = j) = \frac{Q_{hj}(d)}{P_{hj}}$$

اگر $G_{hj}(\cdot)$ تابع بقا و $f_{hj}(\cdot)$ تابع چگالی و $\alpha_{hj}(\cdot)$ نیز نرخ خطر باشد؛ تابع بقای زمان ماندن در حالت h به صورت زیر تعریف می شود:

$$G_h(d) = 1 - p(S_{n+1} \leq d | J_n = h) = \sum_{j \in E} P_{hj}(1 - F_{hj}(d))$$

نرخ خطر فرایند نیمه مارکوفی نیز به صورت زیر تعریف شده است:

$$\lambda_{hj}(d) = \lim_{\Delta d} \frac{P(J_{n+1} = j, d < S_{n+1} \leq d + \Delta d | J_n = h, S_{n+1} > d)}{\Delta d} = \frac{P_{hj} f_{hj}(d)}{G_h(d)} \quad h \neq j$$

$$\lambda_{hj}(d) = - \sum_{j \neq h} \lambda_{hj}(d)$$

بسیاری از مطالعات پزشکی به محاسبه ی احتمال بقای بیماران اختصاص یافته اند. در چنین مطالعاتی نتیجه ی نهایی که مربوط به فوت بیماران بر اثر عامل معین است، در مرکز مورد توجه قرار می گیرد؛ اما در موارد زیادی پیشامدهایی در طول مدت بررسی برای بیماران رخ می دهند که ممکن است نتایج نهایی را نیز تحت تاثیر قرار دهند [۴].

در مطالعه ای از کوروش سایه میری و همکاران، از مدل های چندحالتی نیمه مارکوفی در بررسی عوامل موثر بر بقا بیماران لوسمی بعد از پیوند مغزاستخوان استفاده شده است که چهار حالت: پیوند مغزاستخوان، بروز عوارض پیوند، بهبودی نسبی و مرگ به عنوان فضای حالت فرایند و توزیع زمان انتقال بین حالات نیز توزیع وایبل در نظر گرفته شده است و استفاده از این مدل ها هنگامی که چند پیامد وجود دارد و متغیرهای مستقل وابسته به زمان هستند، توصیه شده است [۷].

در مطالعه ی دیگری از رشیدی و همکاران، که از مدل های غیریکنواخت چندحالتی جهت تحلیل بقای بیماران مبتلا به سرطان پستان استفاده شده است و چهار حالت دریافت درمان اولیه، متاستاز، بهبودی و مرگ در نظر گرفته

شده است و در نهایت نشان داده شده که استفاده از مدل چندحالتی با متغیر مستقل برازش بهتری نسبت به سایر مدل‌ها دارد [۱].

همچنین در مطالعه‌ی زراعتی و همکاران نیز بر روی بیماران مبتلا به سرطان معده، از فرایندهای نیمه مارکوفی جهت بقای پس از جراحی بیماران استفاده شده که نتیجه شده که استفاده از چنین مدل‌هایی به جای مدل مخاطرات متناسب با در نظر گرفتن متغیرهای وابسته به زمان مناسب تر است زیرا می‌توان بدین وسیله شیوع وقوع حالات مختلف را محاسبه و اثر هر حالت پیشینی را بر احتمالات بعدی بررسی کرد [۴].

تعیین احتمال بقا برای انواع سرطان‌ها از اهمیت ویژه‌ای برخوردار است زیرا بیماری سرطان در ایران بعد از بیماری‌های قلبی و عروقی و سوانح و حوادث به عنوان سومین عامل مرگ و میر محسوب می‌شود [۱۰].

سرطان در بین کودکان نیز دیده می‌شود؛ هرچند که در این گروه سنی (کمتر از ۱۴ سال) رخداد آن بسیار نادر می‌باشد و کمتر از یک درصد از کل سرطان‌ها را شامل می‌شود [۱]. اما سرطان یکی از علل اصلی مرگ و میر کودکان در کشورهای توسعه یافته و در حال توسعه است. سرطان‌های پیش از ۱۵ سالگی، سرطان کودکان نام گرفته و شایع‌ترین سرطان در اطفال، سرطان خون می‌باشد. به طوری که حدود ۳۳ درصد از سرطان‌های کودکان مربوط به سرطان لوسمی می‌باشد و دورده‌ی لوسمی حاد لنفوبلاستیک (ALL) و لوسمی میلوئیدی (AML) شایع‌ترین شکل سرطان خون در کودکان می‌باشد [۲].

کودکانی که دچار عود بعد از این سرطان می‌شوند کسر عمده‌ای از کودکان مبتلا به سرطان را به خود اختصاص می‌دهند. درمان بیماران عودی، در مقایسه با بیماران تازه تشخیص داده شده بسیار مشکل بوده و میزان بقای این گروه از بیماران بسیار کمتر است [۵].

۲ مواد و روش‌ها

اطلاعات مورد نیاز جهت انجام این مطالعه مربوط به کودکان مبتلا به سرطان خون ALL و AML بوده است که گروه سنی زیر یک سال تا ۱۵ سال را شامل می‌شود و تمامی اطلاعات با بررسی کامل و ثبت دقیق موارد مورد نیاز پرونده‌های بیماران که از بخش مدارک پزشکی و بایگانی بیمارستان مرجع افضلی پور کرمان استخراج شد، به دست آمد.

جمعیت استفاده شده در این مطالعه شامل ۱۹۰ بیمار، ۴۵ نفر فوت کرده بودند و مابقی افراد نیز درمان آنها پایان

و بهبود یافته اند.

متغیرهای مورد استفاده در این مطالعه شامل: جنسیت، سن، اسپلنومگالی (دارد-ندارد)، هیپاتومگالی (دارد-ندارد) و تشخیص (ALL-AML) می باشد.

در این مطالعه از فرایندهای تصادفی چندحالتی جهت تجزیه و تحلیل استفاده شده و احتمالات انتقال افراد بین این حالات و نرخ های انتقال مربوطه براساس مدل نیمه مارکوفی محاسبه شد و سپس اثر متغیرهای مختلف بر این انتقال ها و زمان های توقف در هر حالت به دست آمد.

برای انجام کلیه ی مراحل آنالیز، از نرم افزار R استفاده شده است.

۳ یافته ها

در این مطالعه، میانگین سنی بیماران ۶ سال بود و ۶/۶۱ درصد آنها را پسران و ۴/۳۸ درصد را نیز دختران تشکیل می دهند.

۸۳/۲ درصد از بیماران مبتلا به ALL و ۱۶/۷ درصد نیز AML بودند.

همچنین ۴/۳۸ درصد از بیماران دارای گروه خونی O، ۵/۹ درصد AB، ۷/۲۳ درصد B و ۴/۲۸ درصد گروه خونی A داشتند.

۷/۴۴ درصد از بیماران دچار اسپلنومگالی و ۷/۵۴ درصد نیز دچار هیپاتومگالی شده بودند. نوع درمان به کاررفته برای بیماران متفاوت بود اما اکثر افراد (۴/۸۸ درصد) شیمی درمانی دریافت کرده بودند و ۱/۲ درصد تنها رادیوتراپی شده بودند و برای ۵/۹ درصد بیماران نیز هردو درمان شیمی درمانی و رادیوتراپی انجام شده بود.

برای بررسی اثر متغیرهای مختلف بر طول عمر بیماران در یک مدل مبتنی بر فرآیندهای نیمه مارکوفی ۶ حالت مختلف برای هر بیمار در نظر گرفته شد: تشخیص اولیه، متاستاز مغز، متاستاز بیضه در پسران و متاستاز تخمدان در دختران، عود، فوت و بهبودی.

بر اساس مدل مارکوفی، ابتدا احتمالات انتقال افراد بین این حالات و نرخ های انتقال مربوطه محاسبه و سپس اثر متغیرهای مختلف بر این انتقال ها مورد بررسی قرار گرفت.

طبق نتایج به دست آمده، متغیرهای هیپاتومگالی، اسپلنومگالی، جنسیت و نوع تشخیص بر زمان تاخیر ابتلا به متاستاز مغز پس از تشخیص اولیه تاثیرگذار بوده و خطر ابتلا به این متاستاز پس از تشخیص سرطان در آنهایی که

هپاتومگالی داشتند ۳/۴۵ برابر افرادی است که هپاتومگالی نداشتند و در بیمارانی که اسپلنومگالی داشتند ۳ درصد بیشتر از بیمارانی است که دچار اسپلنومگالی نشدند و همچنین در بیماران با تشخیص اولیه‌ی ALL تقریباً ۳۳ درصد بیشتر از بیماران AML است؛ همچنین طبق نتایج مشخص شد خطر ابتلا به متاستاز مغز پس از تشخیص سرطان، در پسران ۲/۸۲ برابر دختران است.

متغیرهای جنسیت و تشخیص، می‌تواند بر زمان تاخیر متاستاز بیضه (تخمدان) پس از تشخیص اولیه‌ی سرطان اثرگذار باشد و طبق نتایج، خطر ابتلا در پسران ۵/۳۸ برابر دختران است و در بیماران با تشخیص ALL نیز ۶۷ درصد بیشتر از بیماران AML است.

همچنین متغیرهای جنسیت و نوع تشخیص، بر زمان تاخیر عود پس از تشخیص اولیه سرطان تاثیرگذار بوده است و خطر عود سرطان پس از تشخیص اولیه، در پسران ۵۳ درصد بیشتر از دختران می‌باشد و در بیماران با تشخیص ALL نیز ۴/۸ برابر AML است.

متغیرهای جنسیت و تشخیص می‌تواند بر بهبودی یا فوت پس از تشخیص اولیه‌ی سرطان اثرگذار باشد (انتقال از حالت تشخیص اولیه‌ی سرطان به حالت بهبودی یا فوت؛ به گونه‌ای که هیچ یک از حالت‌های دیگر (متاستاز مغز، متاستاز بیضه یا تخمدان و عود) رخ نداده باشد).

خطر فوت پس از تشخیص اولیه، در پسران ۱/۹۵ برابر دختران است و در بیماران با تشخیص ALL نیز تقریباً ۱۵ برابر بیماران AML است.

متغیر اسپلنومگالی بر زمان تاخیر عود پس از متاستاز مغز تاثیرگذار بوده و خطر عود پس از این متاستاز در بیمارانی که اسپلنومگالی داشتند ۲/۰۹ برابر افرادی است که دچار اسپلنومگالی نشده‌اند.

چهار متغیر هپاتومگالی، اسپلنومگالی، تشخیص و جنسیت بر زمان تاخیر فوت بیمار پس از رخداد متاستاز مغز تاثیرگذار می‌باشد و خطر فوت پس از این متاستاز در بیمارانی که دچار هپاتومگالی شدند ۰/۵ درصد بیشتر از بیمارانی است هپاتومگالی نداشتند و در بیماران با اسپلنومگالی ۸/۴۹ برابر بیمارانی است که اسپلنومگالی نداشتند و در پسران نیز ۸/۴ برابر دختران و در بیماران با تشخیص ALL نیز ۴۰ برابر بیماران AML می‌باشد.

متغیر اسپلنومگالی بر زمان تاخیر متاستاز مغز پس از عود سرطان اثرگذار است و خطر متاستاز مغز پس از عود، در بیمارانی که اسپلنومگالی داشتند تقریباً ۱۲ برابر بیمارانی است که اسپلنومگالی نداشتند.

متغیرهای هپاتومگالی، اسپلنومگالی و جنسیت نیز بر زمان تاخیر متاستاز بیضه (تخمدان) پس از عود تاثیرگذار می‌باشند به گونه‌ای که خطر ابتلا به این متاستاز پس از عود سرطان، در بیمارانی که هپاتومگالی داشتند ۴ درصد

بیشتر از بیمارانی است که هیپاتومگالی نداشتند، همچنین در بیمارانی که دچار اسپلنومگالی بودند، ۵ درصد بیشتر از آنهایی است که اسپلنومگالی نداشتند و در پسران ۴/۲۲ برابر دختران است.

۴ نتیجه‌گیری و بحث

شایع‌ترین شکل لوسمی در بین کودکان، لوسمی حاد لنفوبلاستیک (ALL) می‌باشد که ۸۰ درصد لوسمی را شامل می‌شود.

در مطالعه‌ی الماسی حشیانی و همکاران، نشان داده شد که متغیر جنسیت از عوامل تاثیرگذار در عود بیماری می‌باشند [۵]. به علاوه بر اساس مطالعه‌ی حجازی و همکاران، میزان بروز سرطان خون در جنس مذکر بیشتر از مونث است [۲]. در این مطالعه نشان داده شد که متغیرهای جنسیت و تشخیص بر زمان تاخیر عود پس از تشخیص سرطان موثر است و خطر عود در پسران بیشتر از دختران و در بیماران با تشخیص ALL نیز بیشتر از بیماران AML است؛ همچنین همانطور که گفته شد این دو متغیر بر زمان تاخیر فوت پس از تشخیص اولیه‌ی بیماری نیز موثر است و در پسران و بیماران ALL خطر فوت بالاتر است که می‌تواند به علت شیوع بیشتر ALL و همچنین میزان بروز بیشتر سرطان در جنس مذکر باشد.

علاوه بر متغیرهای جنسیت و تشخیص، عواملی مانند داشتن هیپاتومگالی و اسپلنومگالی نیز می‌تواند بر زمان تاخیر متاستاز بیضه یا تخمدان و متاستاز مغز و همچنین فوت بیمار در هر مرحله از بیماری، موثر باشد و توجه به این موارد و کنترل آنها می‌تواند به عدم رخداد متاستازها و عود سرطان کمک کند و احتمال بقای بیمار افزایش یابد. علاوه بر متغیرهای ذکر شده ممکن است متغیرهای بیشتری مانند سابقه‌ی بیماری‌های قبلی، وزن، سن، گروه خونی، تغذیه، نوع درمان و بسیاری پارامترهای دیگر بر زمان تاخیر عود یا متاستاز و فوت بیمار تاثیرگذار باشد و بهتر است در مطالعات آتی اثر این عوامل نیز مورد بررسی قرار گیرد تا بتوان تصمیم‌گیری بهتر و دقیق‌تری در هر مرحله از درمان داشت.

اما آنچه که اهمیت دارد این است که استفاده از فرایندهای تصادفی چند حالتی می‌تواند با بررسی و تحلیل دقیق تمام مراحل و پیشامدهایی که در طول مدت درمان برای بیماران رخ می‌دهد، ارزیابی دقیق‌تری از اثر هریک از متغیرهای مستقل و مداخله‌گر بر تابع بقا و همچنین زمان تاخیر رخداد هریک از مراحل پیشرفته تر بیماری، داشته باشد و در نتیجه در چنین مواردی استفاده از این مدل‌ها در کنار مدل مخاطرات متناسب کاکس در تحلیل بقا توصیه می‌شود.

مراجع

- [۱] زراعتی دح، محمودی دم، کاظم‌نژاد دا، محمد دک، حداد دپ. بقای پس از جراحی بیماران مبتلا به ادنوکارسینوما و متاستاز غدد لنفاوی: یک روش مبتنی بر فرآیندهای تصادفی. فصلنامه پژوهش حکیم. ۲۰۰۶؛ ۸(۴): ۱۵-۲۰.
- [۲] کوروش سم، اسحاق ا، دیانا س، اله سمر، کامران عم، یاسمن ا، al. et بررسی عوامل موثر بر زمان بقاء بیماران لوسمی حاد بعد از پیوند مغز استخوان با استفاده از مدل‌های چندحالتی نیمه مارکوفی در بیمارستان شریعتی تهران.
- [۳] بیگلریان، حاجی‌زاده، کاظم‌نژاد، انوشیروان، زالی. تحلیل بقای بیماران مبتلا به سرطان معده با مدل کاکس: یک مطالعه پنج ساله. مجله دانشکده پزشکی دانشگاه علوم پزشکی تهران. ۲۰۰۹؛ ۶۷(۵): ۳۱۷-۲۵.
- [۴] حجازی، ساسان، غلامی، لک س، شاکر، خلخالی، al. et بررسی میزان بروز سرطان خون حاد در کودکان کمتر از ۱۵ سال استان آذربایجان غربی طی سال‌های ۱۳۸۲-۱۳۸۷. مجله پزشکی ارومیه. ۲۰۱۰؛ ۲۱(۲): ۲۴۳-۸.
- [۵] حشیانی ا، فرز، حسینی، دهقان، عزیزاله. تعیین عوامل موثر در عود سرطان خون در کودکان، استان فارس (۱۳۸۳-۸۸). مجله دانشگاه علوم پزشکی اراک. ۲۰۱۲؛ ۱۵(۲): ۱-۷.
- [6] Stroock DW. An introduction to Markov processes: Springer Science & Business Media; 2013.
- [7] Bass RF. Stochastic processes: Cambridge University Press; 2011.
- [8] Weiss GH, Zelen M. A semi-Markov model for clinical trials. Journal of Applied Probability. 1965;2(2):269-85.
- [9] Pérez-Oón R, Ruiz-Castro JE. A Multiple-Absorbent Markov Process in Survival Studies: Application to Breast Cancer. Biometrical journal. 2003;45(7):783-97.
- [10] Wu S-C. A semi-Markov model for survival data with covariates. Mathematical Biosciences. 1982;60(2):197-206.

- [11] Barbu V, Boussemart M, Limnios N. Discrete-time semi-Markov model for reliability and survival analysis. *Communications in Statistics-Theory and Methods*. 2004;33(11):2833-68.
- [12] Alireza S, Mehdi N, Ali M, Alireza M, Reza M, Parkin D. Cancer occurrence in Iran in 2002, an international perspective. *Asian Pacific journal of cancer prevention*. 2005;6(3):359.
- [13] Buka I, Koranteng S, Vargas ARO. Trends in childhood cancer incidence: review of environmental linkages. *Pediatric Clinics of North America*. 2007;54(1):177-203.
- [14] Mehrvar A, Rahiminejad MS, Hedayati Asl AA, Tashvighi M, Faranoush M, Alebouyeh M, et al. Features of Childhood Acute Myeloid Leukemia in Iran: a Report from Double Center Study. *Acta Medica Iranica*. 2015;53(12).



سانسور پیش‌رونده نوع II توأم تطبیقی

بیات، م^۱ ترابی، ح^۲

^{۱،۲} گروه آمار، دانشکده علوم ریاضی، دانشگاه یزد

چکیده

یکی از روشهایی که در آزمایشات قابلیت اطمینان بسیار مورد استفاده قرار می‌گیرد سانسور پیش‌رونده نوع II است. این سانسور توانسته است نظر آزمایشگران زیادی را به خود معطوف نماید به طوریکه تعمیم‌های زیادی برای آن ارائه شده است. یکی جدیدترین آنها سانسور پیش‌رونده نوع II توأم می‌باشد. سانسور پیش‌رونده نوع II توأم دارای محدودیت و ایراداتی می‌باشد که در این مقاله سعی شده است تا با ترکیب یکی از تعمیم‌های سانسور پیش‌رونده نوع II با سانسور پیش‌رونده نوع II توأم یکی از محدودیت‌های سانسور پیش‌رونده نوع II توأم که ثابت و از پیش مشخص بودن تعداد کل برداشت در هر مرحله‌ی سانسور می‌باشد برطرف گردد.

کلمات کلیدی: سانسور پیش‌رونده، سانسور پیش‌رونده نوع II توأم تطبیقی، سانسور پیش‌رونده نوع II توأم، آزمایشات طول عمر، شبیه‌سازی.

معرفی نمادها

F = تابع توزیع احتمال جامعه

f = تابع چگالی احتمال جامعه

^۱ bayat.stat@chmail.ir

^۲ htorabi@yazd.ac.ir

$$(\theta \in \Theta \subseteq \mathbb{R}^k) = \theta$$

$$(a_1, \dots, a_i) =_i$$

$$A_i = \text{متغیر تصادفی } i \text{ ام}$$

$$a_i = \text{مقدار مشاهده شدمی متغیر تصادفی } i \text{ ام}$$

$$w_r = \text{زمان } r \text{ امین شکست مشاهده شده}$$

$$s_{h,r} = \text{تعداد آیتم های سانسور شده از نمونه ی } h \text{ ام در } r \text{ امین مرحله ی سانسور}$$

$$\sum_{i=1}^h a_{i,r} = a_{(h),r}$$

$$\sum_{j=1}^r a_{h,j} = a_{h,(r)}$$

$$\sum_{j=1}^r \sum_{i=1}^h a_{i,j} = a_{(h),r}$$

$$n_h - s_{h,(r-1)} - \delta_{h,(r-1)} = \gamma_{h,r}$$

$$I_b(a) = \begin{cases} 1 & \text{if } a \in b \\ 0 & \text{O.W.} \end{cases}$$

$$\begin{bmatrix} a_{1,i} & a_{1,i+1} & \dots & a_{1,r} \\ a_{2,i} & a_{2,i+1} & \dots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,i} & a_{k,i+1} & \dots & a_{k,r} \end{bmatrix} =_{[i,r]} \delta_{h,i} = \begin{cases} 1 & \text{if } w_i \in \{x_1^h, \dots, x_{n_h}^h\} \\ 0 & \text{O.W.} \end{cases}$$

۱ مقدمه

سانسورها در آزمایشات قابلیت اطمینان دارای کاربرد فراوانی می باشند. سانسورها را می توان به دو دسته ی کلی قابل کنترل و غیر قابل کنترل تقسیم نمود. به طور معمول آزمایشگران جهت کاهش هزینه و زمان از سانسورهای قابل کنترل و جهت استفاده ی بهینه از اطلاعات در دسترس از سانسورهای غیر قابل کنترل استفاده می نمایند. البته باید توجه نمود که نمی توان یک سانسور را فقط عضو یکی از این دسته ها دانست بلکه با توجه به آزمایش انجام شده می توان دسته بندی سانسور را تعیین نمود. در ادامه تعدادی از روش های سانسور بازگو می شوند.

سانسور پیشرونده نوع II

فرض کنید آزمایشگری قصد انجام آزمون قابلیت اطمینان بر روی n آیتم را داشته باشد. پیش از شروع آزمایش می بایست بردار $(r_1, \dots, r_m) =$ که بردار سانسور نامیده می شود تعیین گردد. اگر آزمایش بدین صورت انجام شود

که بلافاصله پس از اولین شکست r_1 واحد از آیت‌های تحت بررسی به طور تصادفی از آزمایش خارج شوند و پس از دومین شکست r_2 واحد از اقلام مورد آزمایش سانسور گردند و آزمایش به همین ترتیب ادامه یابد تا اینکه در زمان مشاهده‌ی شکست m ام همه $r_m = n - m - \sum_{i=1}^{m-1} r_i$ واحد تحت بررسی از آزمایش حذف شوند آنگاه این سانسور را سانسور پیشرونده نوع II نامند. این سانسور سال ۱۹۶۳ توسط هرد معرفی گردید. برای دانستن جزئیات بیشتر می‌توان به بالا کریشنان [۱]، بالا کریشنان و آگاروالا [۲]، بالا کریشنان و کرامر [۴]، کرامر و ایلویپلوس [۸]، هرد [۹] و بالا کریشنان و همکاران [۳] مراجعه نمود.

سانسور پیشرونده‌ی نوع II توأم

فرض کنید یک محصول توسط k خط تولید متفاوت تولید شده باشد و آزمایشگر قصد انجام آزمون قابلیت اطمینان را برای این محصولات داشته باشد. همچنین فرض کنید که از خط تولید i ام $i=1, \dots, k$ تعداد n_i آیت‌م را همزمان تحت آزمایش قرار دهیم. سپس در زمان مشاهده‌ی اولین شکست (که ممکن است از آیت‌های هرکدام از خطوط تولید باشد) تعداد r_1 آیت‌م از تمامی آیت‌های تحت آزمایش به تصادف انتخاب و حذف گردند. سپس در زمان مشاهده‌ی دومین شکست (که ممکن است از آیت‌های هرکدام از خطوط تولید باشد) تعداد r_2 آیت‌م از تمامی آیت‌های باقیمانده در آزمایش به تصادف انتخاب و کنار گذاشته می‌شوند. آزمایش با همین روند ادامه پیدا می‌کند تا زمان مشاهده‌ی m امین شکست که در این زمان تمامی آیت‌های موجود در آزمایش حذف گردیده و آزمایش پایان می‌یابد. این روش انجام آزمایش را سانسور پیشرونده‌ی نوع II توأم می‌نامند و اولین بار برای $k=2$ توسط رسولی و بالا کریشنان [۱۰] و برای $k \geq 2$ توسط بالا کریشنان و همکاران [۷] معرفی گردید. برای جزئیات بیشتر می‌توان به بالا کریشنان و رسولی [۶] و بالا کریشنان و فنگ [۵] مراجعه نمود. از ایراداتی که می‌توان به سانسور پیشرونده‌ی نوع II توأم گرفت این است که تعداد کل سانسورها در هر مرحله پیش از شروع آزمایش تعیین شده است که ممکن موجب بروز سه اشکال ذیل گردد:

- ۱- ممکن است بازه‌ی زمانی مشاهده‌ی شکستها بسیار طولانی باشد. برداشتن تعدادی از قبل تعیین شده در هر مرحله می‌تواند احتمال مشاهده‌ی شکست را کاهش داده باعث طولانی‌تر شدن دوره‌ی انجام آزمایش گردد.
- ۲- ممکن است در زمان رخ دادن اولین شکست‌های آزمایش، آیت‌های مربوط به نمونه‌ای که هیچ شکستی از آن مشاهده نشده باشد به طور کامل حذف گردند و این بر روی برآورد پارامتر طول عمر آن نمونه اثر نامطلوبی خواهد

گذاشت.

۳- ممکن است در زمان مشاهده شدن شکست شرایطی وجود داشته باشد که سانسور همزمان تعداد از پیش تعیین شده امکان پذیر نباشد.

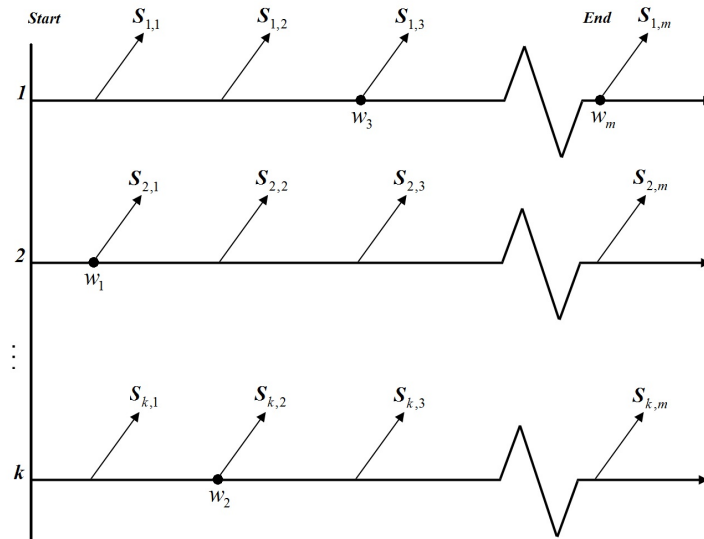
در بخش بعد سانسور جدیدی معرفی خواهد شد که دارای ایرادات بیان شده نباشد.

۲ معرفی طرح سانسور جدید و تابع درستنمایی آن

فرض کنید قصد انجام آزمایش قابلیت اطمینان برای k نمونه‌ی متفاوت با حجم‌های n_h ، $h=1, \dots, k$ با استفاده از سانسور پیش‌رونده‌ی نوع II توأم برای k نمونه‌ی مستقل را داریم. همچنین فرض کنید تعداد آیت‌هایی که از نمونه‌ی h ام $h=1, \dots, k$ در زمان مشاهده‌ی شکست i ام $i=1, \dots, m-1$ حذف می‌شوند از توزیعی گسسته پیروی نماید که وابسته به بردار سانسورهای قبلی از نمونه‌ی h ام و بردار زمان‌های مشاهدات آزمایش باشد و تابع جرم احتمال آن $g_{h,i}(s_{h,i} | i, [1, i-1], [1, i-1])$ باشد. فرض کنید همانند سانسور پیش‌رونده‌ی نوع II توأم تمامی آیت‌ها به طور همزمان در آزمایش قرار داده شود و در زمان مشاهده‌ی اولین شکست تعداد $s_{h,1}$ که یک متغیر تصادفی با تابع جرم احتمال $g_{h,1}(s_{h,1} | w_1, [1, 1])$ است از نمونه‌ی h ام حذف گردد. آزمایش با تمامی آیت‌های باقی مانده در آزمایش ادامه پیدا می‌کند و در زمان مشاهده‌ی شکست دوم تعداد $s_{h,2}$ که یک متغیر تصادفی با تابع جرم احتمال $g_{h,2}(s_{h,2} | 2, [1, 1], [1, 2])$ است از نمونه‌ی h ام $h=1, \dots, k$ سانسور می‌گردد. آزمایش به همین منوال ادامه پیدا می‌کند یعنی در زمان مشاهده‌ی شکست i ام $i=3, \dots, m-1$ تعداد $s_{h,i}$ که یک متغیر تصادفی با تابع جرم احتمال $g_{h,i}(s_{h,i} | i, [1, i-1], [1, i-1])$ است از نمونه‌ی h ام کنار گذاشته می‌شود و در پایان در زمان مشاهده‌ی شکست m ام تمامی آیت‌های باقی مانده از تمامی نمونه‌ها کنار گذاشته می‌شود. این روش سانسور "سانسور پیش‌رونده‌ی نوع II توأم تطبیقی" نام گذاری می‌شود. شکل ۱ روش انجام آزمایش با استفاده از این سانسور را نمایش می‌دهد.

داده‌های این آزمایش را می‌توان به صورت ذیل نمایش داد:

$$m = (w_1, w_2, \dots, w_m)_{[1, m] \times [1, m]}$$



شکل ۱: سانسور پیشرونده‌ی نوع II توأم تطبیقی

تابع درستنمایی سانسور پیشرونده‌ی نوع II توأم تطبیقی به صورت ذیل می‌باشد.

$$f(m_{[\cdot, m]} | [\cdot, m]) = C^m \prod_{i=1}^m \prod_{h=1}^k f_h(w_i)^{\delta_{h,i}} \{1 - F_h(w_i)\}^{s_{h,i}} \prod_{i=1}^m \prod_{h=1}^k g_{h,i}(s_{h,i} | i, [\cdot, i-1], [\cdot, m]) \quad (1)$$

که در آن

$$g_{h,1}(s_{h,1} | [\cdot, \cdot], [\cdot, m]) = g_{h,1}(s_{h,1} | [\cdot, m]), \quad h=1, \dots, k$$

همچنین

$$\sum_{C_m \cap C_m} \int f(m_{[\cdot, m-1]} | [\cdot, m]) = 1, \quad \sum_{C_{m,n_1, \dots, n_h}^s} g(\cdot) = 1$$

که در آن برای $h=1, \dots, k$ و $j=1, \dots, m$

$$C_{m,n_1, \dots, n_h}^s = \{s_{h,j} : \sum_{j=1}^m s_{h,j} \leq n_h, \sum_{h=1}^k \sum_{j=1}^m s_{h,j} = \sum_{h=1}^k n_h - m\}$$

و برای $r=1, \dots, m-1$

$$C_{r,n_1, \dots, n_h}^s = \{s_{h,j} : \sum_{j=1}^r s_{h,j} \leq n_h, \sum_{h=1}^k \sum_{j=1}^r s_{h,j} < \sum_{h=1}^k n_h - m\}$$

$$C_r = \{(w_1, \dots, w_r); w_i < w_{i+1}, i = 1, \dots, r-1\}$$

و

$$C_r = \{\delta_{h,j}; \sum_{j=1}^r \delta_{h,j} \leq \min\{r, n_h\}, \sum_{h=1}^k \sum_{j=1}^r \delta_{h,j} = r, h = 1, \dots, k, j = 1, \dots, r\}$$

و در آن $C^r = D_1 D_r$ یک عدد ثابت نرمال کننده می باشد. بالا کریشان و همکاران [۷] را ببینید.

$$D_1 = \prod_{i=1}^r [\sum_{h=1}^k \gamma_{h,i} \delta_{h,i}], \quad D_r = \prod_{i=1}^{r-1} \frac{\prod_{h=1}^k (\gamma_{h,i} - \delta_{h,i})}{(\gamma_{(k),i} - 1)} \frac{(\gamma_{(k),i} - 1)}{s_{(k),i}}$$

در بخش بعدی تعدادی قضیه در مورد سانسور پیشروندهی نوع II توأم تطبیقی ارائه می گردد. برای رعایت خلاصه نویسی از این به بعد بجای سانسور پیشروندهی نوع II توأم تطبیقی از AJT2PC و بجای سانسور پیشروندهی نوع II توأم از JT2PC استفاده خواهد شد.

۳ استنباط

نکته ۱.۳. تابع چگالی احتمال $w_1, \dots, w_r$ برای $r=1, \dots, m-1$ در AJT2PC برابر است با

$$f(r | [\cdot, r], [\cdot, r]) = C^r \prod_{i=1}^{r-1} \prod_{h=1}^k f_h(w_i)^{\delta_{h,i}} \{1 - F_h(w_i)\}^{s_{h,i}} \prod_{h=1}^k f_h(w_r)^{\delta_{h,r}} \{1 - F_h(w_r)\}^{\gamma_{h,r} - \delta_{h,r}} \quad (2)$$

همچنین تابع چگالی احتمال شرطی w_{r+1} به شرط $w_1, \dots, w_r$ برای $r=1, \dots, m-1$ در JT2PC برابر است با

$$f(r+1 | r, [\cdot, r], [\cdot, r]) = [\sum_{h=1}^k \gamma_{h,r+1} \delta_{h,r+1}] \frac{\prod_{h=1}^k (\gamma_{h,r} - \delta_{h,r})}{(\gamma_{(k),r} - 1)} \times \prod_{h=1}^k [\{\frac{f_h(w_{r+1})}{1 - F_h(w_r)}\}^{\delta_{h,r+1}}] \times \{\frac{1 - F_h(w_{r+1})}{1 - F_h(w_r)}\}^{\gamma_{h,r+1} - \delta_{h,r+1}}$$

همان طور که مشاهده می شود تابع چگالی فوق با تابع چگالی اولین مشاهده از JT2PC برای k نمونهی مستقل با حجم های $\gamma_{1,r+1}, \dots, \gamma_{k,r+1}$ که متغیرهای طول عمر آن ها دارای توزیعی با تابع چگالی f_h ، $h=1, \dots, k$ بریده شده از چپ در نقطهی w_r برابر است.

قضیه ۱.۳. تابع چگالی احتمال شرطی $w_{r+1, [r+1, r+1]}$ به شرط $r, [1, r]$ برای $r=1, \dots, m-1$ در JT2PC برابر است با تابع چگالی AJT2PC برای k نمونه‌ی مستقل با حجم‌های $\gamma_{1, r+1}, \dots, \gamma_{k, r+1}$ که متغیرهای طول عمر آن‌ها دارای توزیعی با تابع چگالی f_h ، بریده شده از چپ در نقطه‌ی w_k با $g_{h,i}$ هایی $i=r+1, \dots, m$ که به $r, [1, r]$ وابسته است.

با توجه به تعریف $\gamma_{h,i}$ ، $h=1, \dots, k$ و $i=1, \dots, m$ به راحتی می‌توان نشان داد که $\gamma_{h,i}$ و $s_{h,i}$ دارای رابطه‌ی یک به یک می‌باشد لذا تابع جرم احتمال متغیر تصادفی $s_{h,i}$ یعنی $g_{h,i}$ را می‌توان بر اساس $\gamma_{h,i}$ نوشت. فرض کنید که تابع جرم احتمال متغیر تصادفی $\gamma_{h,i}$ با $\tilde{g}_{h,i}$ نمایش داده شود.

قضیه ۲.۳. برای $h=1, \dots, k$ و $r=1, \dots, m-1$ اگر $a \geq 0$ وجود داشته باشد به طوریکه تابع توزیع شرطی $\Gamma_{h, r+1}$ و به شرط $r = r$ و $[1, r] = [1, r]$ تنها به $w_{r-a}, \dots, w_r$ و $[r-a, r]$ وابسته باشد آنگاه دنباله‌ی متغیرهای تصادفی $Z_j = (W_{j-a}, \dots, W_{j, [j-a, j]})$ یک زنجیره مارکوف تشکیل می‌دهند.

نکته ۲.۳. اگر در AJT2PC

$$\prod_{h=1}^k g_{s_{h,r}}(s_{h,r} | r-1, [1, r-1]) = \prod_{h=1}^k I_{h,r}(h, r) I_{sr}((k), r)$$

آنگاه معادله‌ی (۱) برابر با تابع درستنمایی JT2PC با طرح سانسور $(s_1, \dots, s_m)$ خواهد بود.

۴ شبیه سازی

این شبیه سازی با استفاده از نرم افزار R و روش مونت کارلو انجام پذیرفته است. فرض شده است که داده‌های هر سه نمونه از توزیع نمایی با پارامترهای به ترتیب ۲، ۴ و ۷ پیروی نمایند. همچنین فرض شده است تعداد سانسورها در سانسور پیش‌رونده نوع II توأم تطبیقی همانند سانسور پیش‌رونده نوع II توأم باشد با این تفاوت که تا زمانی که از نمونه‌ای شکستی مشاهده نشده باشد هیچ آیتمی از آن نمونه سانسور نمی‌شود. این داده‌ها برای n_h ها، m ها و طرح‌های سانسور کردن‌های متفاوت ۱۰۰۰ بار تولید شده است و هر بار با استفاده از به ترتیب سانسور پیش‌رونده نوع II توأم و سانسور پیش‌رونده نوع II توأم تطبیقی برآورد ماکسیمم درستنمایی پارامتر بدست آورده شده است و در پایان میانگین مربعات خطای تجربی برآوردها، اربیی تجربی و تعداد نمونه‌هایی که بدون حتی یک شکست بوده‌اند برای هر کدام از روش‌های سانسور در جداول ۱ و ۲ آورده شده است. که در آن:

$$MSE_1 = \text{میانگین مربعات خطای تجربی برآوردهای } \theta_1, \theta_2, \theta_3 \text{ با استفاده از T2PC,}$$

$$MSE_2 = \text{میانگین مربعات خطای تجربی برآوردهای } \theta_1, \theta_2, \theta_3 \text{ با استفاده از JT2PC,}$$

$$b_1 = \text{اربیی تجربی برآوردهای } \theta_1, \theta_2, \theta_3 \text{ با استفاده از T2PC,}$$

$b_2 =$ اریبی تجربی برآوردگرهای $\theta_1, \theta_2, \theta_3$ با استفاده از JT2PC،

$l_1 =$ تعداد نمونه‌هایی که هیچ شکستی از آن مشاهده نشده است با استفاده از T2PC، (از ۳۰۰۰۰ نمونه)،

$l_2 =$ تعداد نمونه‌هایی که هیچ شکستی از آن مشاهده نشده است با استفاده از JT2PC، (از ۳۰۰۰۰ نمونه).

در جدول ۱ برای رعایت خلاصه نویسی (a, a, a, b, b, b, b) به صورت (a_3, b_4) نوشته می‌شود.

جدول ۱: طرح‌های سانسور و حجم‌های متفاوت نمونه‌ها برای شبیه‌سازی سانسور پیش‌رونده نوع II توأم و سانسور پیش‌رونده نوع II توأم تطبیقی.

$no.$	r	m	(n_1, n_2, n_3)
۱	(۱۷, ۰)	۸	(۵, ۵, ۵)
۲	(۱۷, ۳)	۸	(۶, ۶, ۶)
۳	(۱۷, ۶)	۸	(۷, ۷, ۷)
۴	(۱۷, ۹)	۸	(۸, ۸, ۸)
۵	(۱۵, ۰۵)	۱۰	(۵, ۵, ۵)
۶	(۱۸, ۰۲)	۱۰	(۶, ۶, ۶)
۷	(۱۹, ۲)	۱۰	(۷, ۷, ۷)
۸	(۱۹, ۵)	۱۰	(۸, ۸, ۸)
۹	(۱۶, ۰۶)	۱۲	(۶, ۶, ۶)
۱۰	(۱۹, ۰۳)	۱۲	(۷, ۷, ۷)
۱۱	(۱۱۲)	۱۲	(۸, ۸, ۸)
۱۲	(۱۱۱, ۴)	۱۲	(۹, ۹, ۹)
۱۳	(۱۳, ۰۱۵)	۱۸	(۷, ۷, ۷)
۱۴	(۱۶, ۰۱۲)	۱۸	(۸, ۸, ۸)
۱۵	(۱۹, ۰۹)	۱۸	(۹, ۹, ۹)

جدول ۲: نتایج شبیه‌سازی سانسور پیش‌رونده نوع II توأم و سانسور پیش‌رونده نوع II توأم تطبیقی با ۱۰۰۰۰ مرتبه تکرار.

l_2	l_1	b_2	b_1	MSE_2	MSE_1	no.
۳۸۵	۳۹۰	(۰/۸۲، ۱/۴۱، ۲/۲۱)	(۰/۸۳، ۱/۴۷، ۲/۲۴)	(۷/۶۲، ۲۰/۹۷، ۴۴/۵)	(۸/۵۶، ۲۱/۲۱، ۴۶/۹۶)	۱
۱۱۰۴	۱۰۸۸	(۰/۵۵، ۱/۱۱، ۱/۷۱)	(۰/۵۷، ۱/۱۱، ۱/۷۴)	(۵/۸۱، ۱۴/۴۰، ۳۰/۵)	(۵/۸۷، ۱۴/۷۳، ۳۳/۳۸)	۲
۱۵۵۲	۱۵۸۶	(۰/۴۷، ۰/۸۹، ۱/۴۸)	(۰/۴۸، ۰/۹۳، ۱/۴۶)	(۵/۲۵، ۱۲/۰۵، ۲۷/۲۵)	(۵/۱۴، ۱۱/۶۷، ۲۷/۱۱)	۳
۱۶۵۹	۱۷۰۸	(۰/۴۶، ۰/۸۴، ۱/۱۹)	(۰/۴۸، ۰/۸۵، ۱/۲۲)	(۴/۹۵، ۱۱/۳۹، ۲۱/۶)	(۵/۰۸، ۱۱/۹۵، ۲۲/۷)	۴۷
۵	۱۱	(۰/۸۳، ۱/۵، ۲/۴۱)	(۰/۸۸، ۱/۵۷، ۲/۴۴)	(۶/۹۳، ۲۰/۰۷، ۴۷/۵۴)	(۷/۵۶، ۲۲/۸۳، ۵۰/۱۱)	۵
۸۲	۹۷	(۰/۶۹، ۱/۱۹، ۱/۷۴)	(۰/۶۹، ۱/۱۹، ۱/۸۱)	(۴/۱۷، ۱۲/۷۲، ۳۰/۰۱)	(۴/۲۲، ۱۳/۴۴، ۳۱/۱۳)	۶
۴۰۱	۴۱۲	(۰/۴۷، ۰/۷۹، ۱/۴۲)	(۰/۴۹، ۰/۸۷، ۱/۳۶)	(۳/۶۸، ۹/۰۴، ۲۲/۹۹)	(۳/۹۱، ۹/۶۵، ۲۳/۰۲)	۷
۶۲۳	۷۰۱	(۰/۴۴، ۰/۷۹، ۱/۲)	(۰/۴۵، ۰/۸۲، ۱/۲۴)	(۳/۷۸، ۸/۶۱، ۱۹/۲۶)	(۳/۹۵، ۱۱/۷، ۲۰/۴۲)	۸
۱	۳	(۰/۶۸، ۱/۱۴، ۱/۹۷)	(۰/۷۱، ۱/۱۵، ۲)	(۳/۰۸، ۱۰/۴۵، ۳۵/۴۱)	(۳/۷۵، ۱۱/۸۶، ۳۵/۹۶)	۹
۱۵	۱۶	(۰/۶۰، ۰/۹۸، ۱/۵۲)	(۰/۶۱، ۱/۰۱، ۱/۵۹)	(۲/۳۱، ۹/۰۸، ۲۴/۰۵)	(۲/۹۷، ۹/۷۶، ۲۶/۴۸)	۱۰
۱۱۷	۱۳۰	(۰/۴۰، ۰/۷۹، ۱/۲۴)	(۰/۴۲، ۰/۷۸، ۱/۲۶)	(۲/۶، ۷/۵۲، ۱۸/۴۱)	(۲/۶۸، ۷/۹، ۱۹/۸۹)	۱۱
۲۹۰	۳۱۸	(۰/۴۰، ۰/۶۶، ۰/۹۵)	(۰/۴۱، ۰/۷۱، ۱)	(۲/۱، ۶/۶۷، ۱۴/۶)	(۲/۸، ۶/۹۱، ۱۵/۰۲)	۱۲
۰	۰	(۰/۴۰، ۰/۷۷، ۱/۴)	(۰/۴۱، ۰/۷۷، ۱/۴)	(۱/۴۶، ۵/۸، ۱۷/۰۷)	(۱/۶۲، ۵/۹۹، ۱۸/۰۹)	۱۳
۰	۰	(۰/۳۱، ۰/۷۵، ۱/۲۶)	(۰/۳۸، ۰/۷۶، ۱/۳)	(۱/۵۰، ۵/۴۳، ۱۶/۸۴)	(۱/۵۲، ۵/۵۳، ۱۷/۵۱)	۱۴
۰	۰	(۰/۲۴، ۰/۷۲، ۱/۱۱)	(۰/۳۹، ۰/۷۲، ۱/۱۹)	(۱/۵۲، ۵/۰۲، ۱۳/۸۸)	(۱/۵۲، ۵/۱۵، ۱۴/۱۵)	۱۵

در جدول ۱ و ۲ دیده می‌شود که با استفاده از نمونه‌های با حجم به ترتیب ۵، ۵ و با تعداد مشاهدات مورد نیاز ۸ و طرح سانسور $1, 1, 1, 1, 1, 1, 1, 1$ نتایج شبیه سازی با ۱۰۰۰۰ تکرار بدین شکل است که میانگین مربعات خطا برای θ_1 ، θ_2 و θ_3 با استفاده از سانسور پیش رونده نوع II توأم به ترتیب برابر است با ۸/۵۶، ۲۱/۲۱ و ۴۶/۹۶ و با استفاده از سانسور پیش رونده نوع II توأم تطبیقی به ترتیب برابر با ۷/۶۲، ۲۰/۹۷ و ۴۴/۵ است. همچنین مقدار اریبی تجربی برآوردگرها برای این پارامترها با استفاده از سانسور پیش رونده نوع II توأم به ترتیب برابر است با ۰/۸۳، ۱/۴۷ و ۲/۲۴ و با استفاده از سانسور پیش رونده نوع II توأم تطبیقی به ترتیب برابر با ۰/۸۲، ۱/۴۱ و ۲/۲۱ است. در سانسور پیش رونده نوع II توأم ۳۹۰ نمونه و در سانسور پیش رونده نوع II توأم تطبیقی ۳۸۵ نمونه از بین ۳۰۰۰ نمونه‌ی تولید شده فاقد هرگونه شکستی بوده‌اند و تمامی آیت‌های آنها سانسور شده است. بنابراین در این شبیه سازی سانسور پیش‌رونده نوع II توأم تطبیقی نسبت به سانسور پیش‌رونده نوع II توأم دارای برآوردهای بهتری است.

دست‌آوردهای پژوهش

در این مقاله در مورد سانسور پیش‌رونده‌ی نوع II و یکی از تعمیم‌های آن یعنی JT2PC صحبت شد. جهت رفع چند عیب بیان شده در انتهای بخش ۱ در مورد JT2PC یک طرح سانسور جدید با نام "سانسور پیش‌رونده‌ی نوع II توأم تطبیقی" معرفی گردید که تعداد سانسور در هر مرحله را یک متغیر تصادفی با تابع توزیع گسسته در نظر می‌گیرد. تابع چگالی احتمال آن به همراه چند قضیه‌ی مرتبط در بخش ۳ ارائه گردید. برای شناخت بهتر مدل جدید یک شبیه سازی به روش مونت کارلو با ۱۰۰۰۰ بار تکرار انجام شده که در آن JT2PC و AJT2PC برای داده‌های نمایی از سه جامعه‌ی متفاوت با یکدیگر مقایسه شده‌اند و نتایج آن در جداول ۱ و ۲ آورده شده است. بر اساس نتایج به دست آمده در جداول ۱ و ۲ و با توجه به ایرادات عنوان شده در پایان بخش ۱ در مورد JT2PC و همچنین با توجه به اینکه JT2PC یک زیر مدل از AJT2PC می‌باشد، می‌توان AJT2PC را یک سانسور منطقی و جایگزینی برای JT2PC در نظر گرفت.

مراجع

- [1] Balakrishnan, N. (2007), *Progressive Censoring Methodology: an Appraisal (with discussions)*, Test, 16, 211–296.
- [2] Balakrishnan, N. and Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods, and Applications*, Birkhäuser, Boston.
- [3] Balakrishnan, N., Burkschat, M. and Cramer, E. and Hofman, G. (2008), *Fisher Information Based Progressive Censoring Plans*, Computational Statistics and Data Analysis, 53, 366–380.
- [4] Balakrishnan, N. and Cramer, E. (2014), *The Art of Progressive Censoring*, Springer, New York.
- [5] Balakrishnan, N. and Feng, S. (2015), *Exact likelihood inference for k exponential populations under joint Type-II censoring*, Communications in Statistics–Simulation and Computation, 44, 591–613.
- [6] Balakrishnan, N. and Rasouli, A. (2008), *Exact likelihood inference for two exponential populations under joint Type-II censoring*, Computational Statistics and Data Analysis, 52, 2725–2738.
- [7] Balakrishnan, N., Suc, F. and Liud, K.Y. (2015), *Exact Likelihood Inference for k Exponential Populations Under Joint Progressive Type-II Censoring*, Communications in Statistics – Simulation and Computation, 44, 902–92.
- [8] Cramer, E. and Iliopoulos, G. (2009), *Adaptive Progressive Type-II Censoring*, Test, 19, 342–358.
- [9] Herd, RG. (1956), *Estimation of Parameters of a Population from a Multi-Censored Sample*. Phd thesis, Iowa State College, Ames, Iowa.
- [10] Rasouli, A. and Balakrishnan, N. (2010), *Exact likelihood inference for two exponential populations under progressive joint Type-II censoring*, Communication in Statistics–Theory and Methods, 39, 2172–2191.



بررسی نرخ لگ-بخت و ارتباط آن با نرخ شکست و نرخ شکست معکوس

جانی، ن^۱ صالحی، ا^۲ خراشادیزاده، م^۳

^{۱،۲} گروه آمار، دانشگاه صنعتی بیرجند
^۳ گروه آمار، دانشکده علوم ریاضی و آمار، دانشگاه بیرجند

چکیده

بالا بودن قابلیت اعتماد و طولانی بودن طول عمر قطعه باعث می‌شود تعیین و بررسی رفتار نرخ شکست در نقاط انتهایی مشکل‌تر شود. از آنجایی که نقاط انتهایی نرخ شکست نقش اساسی در تعیین شکل آن دارد، بدست آوردن برآوردی برای توزیع زمان خرابی در این گونه داده‌ها دچار مشکل می‌شود. همچنین استفاده از قطعاتی که تابع نرخ شکست غیریکنوا دارند (مانند توزیع‌های لگ نرمال، بر، گوسین معکوس) که برازش توزیع، توسط تابع نرخ شکست مشکل و یا بعضاً غیر ممکن می‌شود. با توجه به دو مورد ذکر شده، با مطالعه رفتار تابع نرخ شکست نمی‌توان به تنهایی توزیع زمان شکست را به آسانی بدست آورد. در نتیجه محققین معیار دیگری را به نام نرخ لگ-بخت برای حل این مشکل ارائه کردند. در این مقاله ضمن معرفی معیار لگ-بخت، به بررسی روابط بین این معیار و برخی دیگر از معیارهای قابلیت اعتماد از جمله نرخ شکست و نرخ شکست معکوس پرداخته می‌شود.

کلمات کلیدی: تابع لگ-بخت، نرخ لگ-بخت صعودی (نزولی)، نرخ شکست صعودی (نزولی)، نرخ شکست معکوس صعودی (نزولی).

^۱ narges_jani@birjand.ac.ir

^۲ salehi@birjandut.ac.ir

^۳ m.khorashadizadeh@birjand.ac.ir

۱ مقدمه

تابع نرخ شکست به دلیل رابطه یک به یک با تابع توزیع طول عمر، یکی از مهم‌ترین معیارها در مطالعات قابلیت اعتماد و مباحث طول عمر است که نقش محوری در بیشتر تحلیل‌های بقاء ایفا می‌کند. از این شاخص به طور گسترده در مسائل کاربردی همچون سالخوردگی و مشخصه سازی توزیع زمان شکست، استفاده می‌شود. بررسی رفتار تابع نرخ شکست یک روش موثر برای توصیف مدل اصلی می‌باشد.

اگر تابع نرخ شکست صعودی باشد می‌گوییم با گذشت زمان مولفه فرسوده می‌شود. اگر تابع نرخ شکست نزولی باشد بدین معناست که مولفه با گذشت زمان رو به بهبود است و اگر تابع نرخ شکست ثابت باشد به این معنی است که نرخ شکست به زمان بستگی ندارد و در طول زمان همواره برابر مقدار ثابتی است. توزیع نمایی به دلیل ویژگی فقدان حافظه، نرخ شکست آن به زمان بستگی نداشته و همواره مقداری ثابت است. این از ویژگی‌های منحصر به فرد توزیع نمایی است و در بین توزیع‌های طول عمر پیوسته، توزیع نمایی تنها مدل احتمالی است که دارای نرخ شکست ثابت است.

امروزه در مباحث قابلیت اعتماد و تحلیل بقاء، دو موضوع جدید مورد توجه قرار گرفته است. اولین موضوع مربوط به بالا بودن قابلیت اعتماد و طولانی بودن طول عمر قطعه می‌باشد که باعث می‌شود تعیین و بررسی رفتار نرخ شکست در نقاط انتهایی مشکل‌تر شود. از آنجایی که نقاط انتهایی نرخ شکست نقش اساسی در تعیین شکل آن دارد، بدست آوردن برآوردی برای توزیع زمان خرابی در این گونه داده‌ها دچار مشکل می‌شود. موضوع دوم استفاده از قطعاتی است که تابع نرخ شکست غیریکنوا دارند (مانند توزیع‌های لگ نرمال، بر، گوسین معکوس) که برازش توزیع، توسط تابع نرخ شکست مشکل و یا بعضاً غیر ممکن می‌شود. در این خصوص می‌توان به زیمر و همکاران (۱۹۹۸a) و ونگ و همکاران (۲۰۰۳، ۲۰۰۵) اشاره کرد. نکته جالب اینجاست که در این توزیع‌ها، شکل تابع نرخ شکست مطابق شکل معمول، وانی نیست. حال با توجه به دو مورد ذکر شده در بالا، با استفاده از رفتار تابع نرخ شکست نمی‌توان به تنهایی توزیع زمان شکست را به آسانی بدست آورد. برای حل این مشکل محققینی همچون زیمر و همکاران (۱۹۹۸b، ۱۹۹۸a) و ونگ و همکاران (۲۰۰۳، ۲۰۰۵) در توزیع‌های طول عمر پیوسته، مدل جدیدی برای مشخصه‌سازی زمان شکست براساس شاخص نرخ لگ-بخت (LOR) ارائه کردند که قابل مقایسه با مدل‌های نرخ شکست نیز می‌باشد. ونگ و همکاران (۲۰۰۵) نیز کرانهایی برای توزیع‌های با نرخ لگ-بخت یکنوا بدست آوردند. کلاس‌ها و ترتیب‌های تصادفی براساس تابع نرخ لگ-بخت و رابطه‌اش با برخی از کلاس‌ها و ترتیب‌های تصادفی رایج را ناوارو و همکاران (۲۰۰۸) بدست آوردند. خراشادیزاده (۲۰۱۲a) به معرفی و بررسی تابع نرخ لگ-بخت در توزیع‌های گسسته پرداخت و با معرفی تابع لگ-بخت در توزیع‌های دو متغیره و ارتباط آن با تابع مفصل، شاخص جدیدی برای محاسبه میزان تقارن و یا چولگی توزیع‌های دو متغیره ارائه کرد. همچنین خراشادیزاده و همکاران (۲۰۱۲b) نشان دادند که نرخ لگ-بخت می‌تواند توزیع را به صورت منحصر به فرد مشخص کند و همچنین نرخ لگ-بخت را برای توزیع لجستیک گسسته برحسب t و لگ لجستیک گسسته برحسب $\ln t$ را مشخصه‌سازی و یک معیار جدید از چولگی را براساس تابع نرخ لگ-بخت ارائه کردند. مشخصه‌سازی توزیع‌های طول عمر با استفاده از نرخ لگ-بخت در سالخوردگی توزیع‌های گسسته توسط خراشادیزاده و همکاران (۲۰۱۳) مورد بررسی قرار گرفت. توابع نرخ شکست، میانگین باقیمانده عمر و نرخ لگ-بخت برای برخی توابع توزیع مانند وایبل، توسط بهاتاچارجی و همکاران (۲۰۱۳) انجام شد.

در این مقاله ابتدا به معرفی تابع لگ-بخت و نرخ لگ-بخت می‌پردازیم. سپس رابطه‌ی بین نسبت لگ-بخت و معیارهای مختلف قابلیت اعتماد مانند توابع چگالی، توزیع، نرخ شکست و نرخ شکست معکوس را بررسی می‌کنیم. سپس به معرفی رده توزیع‌های نرخ لگ-بخت صعودی (ILOR) و نرخ لگ-بخت نزولی (DLOR) بر حسب t ، $\ln t$ و $\ln(\ln t)$ پرداخته و رابطه آنها را با رده توزیع‌های IFR(DFR) و IRFR(DRFR) مورد بررسی و مقایسه قرار می‌دهیم.

۲ آشنایی با مفهوم نرخ لگ-بخت

ابتدا مفهوم بخت را اینگونه تعریف می کنیم. نسبت وقوع یک پیشامد به عدم وقوع آن پیشامد، که عبارت است از

$$\text{Odds}(t) = \frac{F(t)}{1 - F(t)}.$$

اگر متغیر تصادفی T دارای تابع توزیع تجمعی $F(t)$ و تابع قابلیت اعتماد $R(t)$ باشد آنگاه لگاریتم تابع بخت که به تابع لگ-بخت نیز معروف است به صورت زیر تعریف می شود،

$$\text{LO}(t) = \ln(\text{Odds}(t)) = \ln \frac{F(t)}{R(t)} = \ln F(t) - \ln R(t). \quad (۱)$$

براحتی ملاحظه می شود که رابطه ای یک به یک بین تابع توزیع و تابع لگ-بخت وجود دارد. در واقع تابع لگ-بخت، تابع توزیع را به کمک رابطه زیر به صورت یکتا مشخص سازی می کند،

$$F(t) = \frac{\exp(\text{LO}(t))}{1 + \exp(\text{LO}(t))}. \quad (۲)$$

مشتق تابع لگ-بخت نسبت به t ، که با مفاهیم قابلیت اعتماد از جمله تابع نرخ شکست و نرخ شکست معکوس در ارتباط است، نرخ لگ-بخت نام دارد و به فرم زیر محاسبه می شود،

$$\begin{aligned} \text{LOR}_1(t) = \frac{d}{dt} \text{LO}(t) &= \frac{f(t)}{F(t)R(t)} \\ &= \frac{h(t)}{F(t)} \\ &= \frac{\lambda(t)}{R(t)} \\ &= \lambda(t) + h(t). \end{aligned} \quad (۳)$$

که در آن $\lambda(t) = \frac{f(t)}{R(t)}$ و $h(t) = \frac{f(t)}{F(t)}$ به ترتیب تابع نرخ شکست و تابع نرخ شکست معکوس نام دارند. با اندکی محاسبات جبری ملاحظه می شود که نرخ لگ-بخت، تابع چگالی احتمال متغیر تصادفی را به کمک رابطه زیر مشخص سازی می کند.

$$f(t) = \frac{\text{LOR}(t) \exp(\text{LO}(t))}{(1 + \exp(\text{LO}(t)))^2}.$$

همچنین اگر فرض کنیم $L(x) = \int_m^x \text{LOR}(u) du$ که در آن m میانه توزیع است (در نتیجه $\text{LO}(m) = 0$) آنگاه می توان توزیع طول عمر را به کمک رابطه زیر مشخص سازی نمود:

$$F(t) = \frac{e^{L(t)}}{1 + e^{L(t)}} = \frac{e^{\int_m^t \text{LOR}(u) du}}{1 + e^{\int_m^t \text{LOR}(u) du}}. \quad (۴)$$

از دیگر کمیت هایی که در ارتباط با لگ-بخت می باشند و خواص جالب و کاربردی نیز دارند، با مشتق گرفتن از تابع لگ-بخت نسبت به $\ln t$ و $\ln(\ln t)$ حاصل می شوند. اگر $x = \ln t$ و $y = \ln(\ln t)$ باشد آنگاه داریم،

$$t = e^x = e^{e^y}, \quad \frac{dt}{dx} = e^x = t, \quad \frac{dt}{dy} = e^y e^{e^y} = t \ln t.$$

لذا تابع نرخ لگ-بخت بر حسب $\ln t$ و $\ln(\ln t)$ به ترتیب به صورت زیر حاصل می شود:

$$\text{LOR}_x(t) = \frac{d\text{LO}(t)}{dx} = \frac{th(t)}{F(t)}, \quad (۵)$$

$$\text{LOR}_y(t) = \frac{d\text{LO}(t)}{dy} = \frac{t(\ln t)h(t)}{F(t)}. \quad (۶)$$

در ادامه به برخی روابط بین توابع نرخ لگ-بخت معرفی شده و رده توزیع های طول عمر می پردازیم.

۳ ویژگی‌هایی از توزیع‌ها با نرخ لگ بخت یکنوا

در این بخش ابتدا برخی رده توزیع‌های طول عمر براساس توابع نرخ لگ-بخت و نماد‌های بکار رفته را معرفی می‌کنیم و سپس روابط بین این رده توزیع‌ها را با رده توزیع‌های نرخ شکست صعودی (IFR) و نزولی (DFR) و همچنین نرخ شکست معکوس نزولی (DRFR) بررسی می‌کنیم.

تعریف ۱.۳. توزیع F دارای خاصیت نرخ لگ بخت صعودی $ILOR_i$ (نزولی، $DLOR_i$) در $\ln t, t$ و یا $\ln(\ln t)$ است اگر و فقط اگر $LOR_i(t)$ به ترتیب برای $i = 1, 2, 3$ برحسب $\ln t, t$ و یا $\ln(\ln t)$ صعودی (نزولی) باشد.

نکته ۱.۳. حاصلضرب دوتابع نامنفی و صعودی (نزولی)، صعودی (نزولی) است.

قضیه ۱.۳. عبارات زیر معادلند:

(a) توزیع F دارای خاصیت $ILOR_i$ (نزولی، $DLOR_i$) در $\ln t, t$ و یا $\ln(\ln t)$ است

(b) تابع $LO(t)$ به ترتیب برای $i = 1, 2, 3$ نسبت به $\ln t, t$ و یا $\ln(\ln t)$ محدب (مقعر) باشد.

برهان. با توجه به تعریف توابع نرخ لگ-بخت و تعریف ۱ براحتی نتیجه مطلوب حاصل می‌شود. $\square$

قضیه ۲.۳. اگر $ILOR_1, F$ باشد آنگاه IFR نیز می‌باشد. یا عبارتی،

$$ILOR_1 \Rightarrow IFR \quad (۷)$$

برهان. طبق تعریف تابع $LOR_1(t)$ داریم $h(t) = F(t)LOR_1(t)$. از طرفی با توجه به اینکه حاصلضرب دو تابع نامنفی صعودی، خود نیز نامنفی و صعودی است، لذا صعودی بودن $LOR_1(t)$ نسبت به t ، صعودی بودن نرخ شکست را نتیجه می‌دهد. $\square$

تذکر ۱.۳. $ILOR$ برحسب t شرایط قوی تری نسبت به IFR دارد و می‌توان از همه ویژگی‌های IFR استفاده کرد.

عکس قضیه فوق برقرار نیست.

به عنوان مثال توزیع وایبل را در نظر بگیرید. نرخ شکست توزیع وایبل به صورت زیر می‌باشد

$$h(t) = \frac{f(t)}{R(t)} = \lambda^\beta \beta t^{\beta-1}.$$

توزیع وایبل به ازای $\beta > 1$ دارای $h(t)$ صعودی است و دارای خاصیت IFR می‌باشد. از طرفی تابع نرخ لگ بخت آن عبارت است از

$$LOR_1(t) = \frac{h(t)}{F(t)} = \frac{\lambda^\beta \beta t^{\beta-1}}{1 - e^{-(\lambda t)^\beta}}.$$

با در نظر گرفتن $\beta = 2$ ، $LOR_1(t)$ به صورت زیر بدست می‌آید

$$LOR_1(t) = \frac{\lambda^2 2t}{1 - e^{-(\lambda t)^2}}.$$

$$\frac{d}{dt} LOR_1(t) = \frac{2\lambda^2(1 - e^{-(\lambda t)^2}) + 4\lambda^4 t^2(e^{-(\lambda t)^2})}{(1 - e^{-(\lambda t)^2})^2}$$

همانطور که مشاهده می‌کنید $LOR_1(t)$ صعودی نیست بنابراین دارای خاصیت $ILOR_1$ نمی‌باشد.

قضیه ۳.۳. اگر IFR, F باشد آنگاه $ILOR_2$ نیز می‌باشد. یا عبارتی،

$$IFR \Rightarrow ILOR_2 \quad (۸)$$

برهان. اگر IFR, F باشد یعنی $h(t)$ صعودی است. از طرفی داریم

$$\frac{d}{dt} LOR_T(t) = \frac{(h(t) + th'(t))F(t) - th(t)f(t)}{(F(t))^2} = \frac{th'(t)F(t) + h(t)(F(t) - tf(t))}{(F(t))^2}$$

زیمر و همکاران (۱۹۹۸) نشان دادند که تحت شرط $(F(t)[h(t) + th'(t)] - th(t)f(t) > 0)$ که معمولاً در قابلیت اعتماد برقرار است، مشتق $LOR_T(t)$ مثبت می‌شود بنابراین $ILOR_T$ می‌باشد. □

عکس قضیه فوق برقرار نیست.

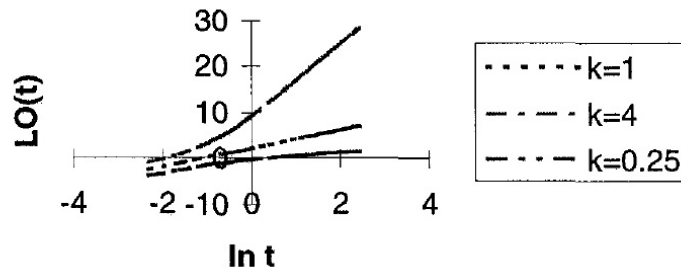
به عنوان مثال توزیع بر نوع ۱۲ را در نظر بگیرید. این توزیع، یک توزیع تک مدی است که دارای تابع نرخ شکست غیر یکنوا است که به ازای مقادیر مختلف c, k, m و m می‌تواند شکل‌های متنوعی به خود بگیرد. فرض کنید T دارای توزیع بر نوع ۱۲ با تابع چگالی احتمال و تابع توزیع به ترتیب به فرم زیر باشد

$$f(t) = \frac{mck(mt)^{c-1}}{(1 + (mt)^c)^{k+1}}, \quad F(t) = 1 - \frac{1}{(1 + (mt)^c)^k}, \quad t, c, k, m > 0$$

همچنین تابع قابلیت اعتماد و تابع نرخ شکست توزیع بر عبارت است از

$$R(t) = \frac{1}{(1 + (mt)^c)^k}, \quad h(t) = \frac{mck(mt)^{c-1}}{(1 + (mt)^c)}$$

توزیع بر، یک تعمیم از توزیع لگ لجستیک است به این معنا که اگر $k = 1$ باشد این توزیع به توزیع لگ لجستیک تبدیل می‌شود. تابع لگ-بخت برای توزیع بر، بر حسب $\ln t$ برای $k > 1$ محدب، برای $k = 1$ خطی مستقیم و برای $k < 1$ مقعر است بنابراین توزیع بر برای $k > (<) 1$ دارای خاصیت $ILOR_T (DLOR_T)$ است. اما همانطور که گفته شد خاصیت IFR در آن برقرار نیست.



شکل ۱: نمودار تابع لگ بخت توزیع بر

قضیه ۴.۳. اگر $ILOR_T, F$ باشد آنگاه $ILOR_T$ نیز می‌باشد.

برهان. طبق تعریف $ILOR_T$ داریم $ILOR_T = \ln t \cdot LOR_T$. لذا با توجه به نکته ۱.۳ و صعودی بودن LOR_T و $\ln t$ ، صعودی بودن $ILOR_T$ نتیجه می‌شود. □

نتیجه ۱.۳. در این خصوص بین رده‌های $IFR, ILOR_T, ILOR_T, ILOR_T$ ارتباطی برقرار است که به صورت زیر می‌توان نمایش داد

$$ILOR_T \Rightarrow IFR \Rightarrow ILOR_T \Rightarrow ILOR_T. \quad (9)$$

مثال ۱.۳. تابع لگ-بخت برحسب $\ln t$ برای توزیع های نرمال بریده شده، وایبل و گاما محدب است و در نتیجه همه این توزیع ها دارای خاصیت $ILOR_2$ هستند. توزیع نرمال بریده شده دارای خاصیت $ILOR_2$ است در نتیجه لگ نرمال نیز $ILOR_2$ است. توزیع گوسین معکوس نیز به دلیل اینکه تابع لگ-بخت آن در $\ln(\ln t)$ محدب است بنابراین $ILOR_2$ است. به آسانی متوجه می شویم که توزیع لگ بر (y) ، لگ بر است اگر $T = \ln y$ پر باشد، $ILOR_2$ برای $k > 1$ و $DLOR_2$ برای $k < 1$ است. اگر $k = 1$ باشد لگ بر، لگ لگ لجستیک است و $LOR_2(t)$ توزیع لگ لگ لجستیک ثابت است.

همانطور که در قبل بیان شد نرخ شکست ثابت است اگر و فقط اگر توزیع متناظر آن نمایی باشد. به همین شیوه می توان قضیه زیر را بیان کرد.

اگر و تنها اگر نرخ لگ-بخت ثابت باشد، یعنی $LOR = c$ و $LO(x) = cx + b$ ، آنگاه توزیع متناظر آن لجستیک استاندارد برحسب t و لگ لجستیک استاندارد برحسب $x = \ln t$ می باشد. بنابراین لگ لجستیک برحسب $\ln t$ از نظر نرخ لگ-بخت مشابه نمایی از نظر نرخ شکست است.

قضیه ۵.۳. اگر F دارای خاصیت $DLOR_2$ باشد آنگاه $DLOR_2$ نیز می باشد.

برهان. اگر $DLOR_2$ برقرار باشد در نتیجه $LOR_2(t) = \left(\frac{th(t)}{F(t)}\right) \ln t$ نزولی است در نتیجه طبق نکته ۱.۳، $LOR_2(t) = \frac{LOR_2}{\ln t}$ نیز نزولی است. $\square$

قضیه ۶.۳. اگر F دارای خاصیت $DLOR_2$ باشد آنگاه دارای خاصیت DFR نیز می باشد.

برهان. اگر $LOR_2(t) = \left(\frac{th(t)}{F(t)}\right)$ نزولی باشد بنابراین داریم:

$$h(t) = \frac{LOR(t)F(t)}{t}$$

$$h'(t) = \frac{(LOR(t)f(t) + LOR'(t)F(t))t - LOR(t)F(t)}{t^2}$$

در نتیجه تحت شرط $0 < LOR(t) + tf(t)LOR'(t) - F(t)(tLOR'(t) - LOR(t)) < 0$ می شود بنابراین دارای خاصیت DFR می باشد. $\square$

قضیه ۷.۳. اگر F دارای خاصیت DFR باشد آنگاه دارای خاصیت $DLOR_1$ نیز می باشد.

برهان. F دارای خاصیت DFR می باشد بنابراین $h(t)$ نزولی است در نتیجه طبق نکته (۱.۳)، $DLOR_1$ نیز برقرار است. $\square$

قضیه ۸.۳. اگر F دارای خاصیت $DLOR_1$ باشد آنگاه دارای خاصیت $DRFR$ نیز می باشد.

برهان. $DLOR_1$ ، F است بنابراین طبق رابطه (۲) و نکته (۱.۳)، $\lambda(t)$ نیز نزولی است بنابراین دارای خاصیت $DRFR$ می باشد. $\square$

نتیجه ۲.۳. در این خصوص بین رده های $DLOR_2$ ، $DLOR_1$ ، $DRFR$ ، DFR ارتباطی برقرار است که به صورت زیر می توان نمایش داد

$$DLOR_2 \Rightarrow DLOR_1 \Rightarrow DFR \Rightarrow DLOR_1 \Rightarrow DRFR. \quad (10)$$

مراجع

- [۱] م. خراشادیزاده، مفاهیمی از اندازه‌های قابلیت اعتماد و روابط بین آنها در توزیع‌های طول عمر گسسته، رساله دکتری رشته آمار ریاضی، دانشگاه فردوسی، دانشکده علوم، گروه آمار، ۲۰۱۲a.
- [2] Bhattacharjee, S., Nanda, A. K. and Misra, S. K. (2013), *Inequality involving expectations to characterize distributions*, Statistics and Probability Letters. 83, 2113–2118.
- [3] Navarro, J., Ruiz, J. M. and Aguila, Y. (2008), *Characterizations and ordering properties based on log-odds functions*, Statistics. 42(4), 313–328.
- [4] Khorashadizadeh, M., Rezaei Roknabadi, A. H. and Mohtashami Borzadaran, G. R. (2012b), *Discrete Log-odds rate function and a new measure of skewness*, The 11th Iranian Statistical Conference, Iran University of Science Technology, Tehran, Iran, August. 28–30.
- [5] Khorashadizadeh, M., Rezaei Roknabadi, A. H. and Mohtashami Borzadaran, G. R. (2013), *Characterization of life distributions using log odds rate in discrete ageing*, Communications in statistics – Theory and Methods. 42(1), 76–87.
- [6] Wang, Y., Hossain, A. M. and Zimmer, W. J. (2003), *Monotone log odds rate distributions in reliability analysis*, Communications in Statistics-Theory and Methods. 32(11), 2227–2244.
- [7] Wang, Y., Hossain, A. M. and Zimmer, W. J. (2005), *Tables of Bounds for Distributions with Monotone Log-Odds Rate*, Communications in Statistics-Simulation and Computation. 34, 1–20.
- [8] Zimmer, W. J., Wang, Y. and Pathak, P. K. (1998a), *Log-odds rate and monotone log-odds rate distributions*, Journal of Quality Technology. 30(4), 376–385.
- [9] Zimmer, W. J., Keats, J. B. and Prairie, R. R. (1998b), *Characterization of non-monotone hazards*, Proceedings Annual Reliability and Maintainability Symposium, IEEE Xplore. 176–181.



ارتباط بین بعضی از توزیع های طول عمر

جمشیدیان، ا. ر^۱

گروه آمار، دانشکده علوم کامپیوتر و ریاضی، دانشگاه غیر انتفاعی شیخ بهایی

چکیده

در این مقاله ابتدا واژه های استاندارد مورد استفاده در بحث قابلیت اعتماد را یادآوری کرده و همچنین یک تعداد از کلاس توزیع های طول عمر معرفی شده در مقالات مختلف جمع آوری می شود. و رابطه بین این توزیع های بیان شده با یکدیگر را به نمایش گذاشته می شود.

کلمات کلیدی: طول عمر، تابع نرخ شکست، طول عمر باقی مانده.

۱ مقدمه

فرض کنید X یک متغیر تصادفی نامنفی است که معرف (نشانه) یکی از شاخص های طول عمر می باشد، که اصطلاحاً به آن زمان تا شکست گفته می شود و تابع چگالی احتمال آن را با $f(t)$ نمایش می دهیم. تابع توزیع مربوطه عبارتست از:

$$F(t) = P(x \leq t) = \int_0^t f(u) du$$

که با افزایش t ، این عبارت انباشتگی احتمال شکست می باشد.

تعریف ۱.۱. تابع قابلیت اعتماد (یا تابع بقاء) را با $\bar{F}(t)$ نشان داده می شود که عبارتست از احتمال اینکه X دارای مقداری بیشتر از t باشد:

$$\bar{F}(t) = 1 - F(t) = P(X > t) = \int_t^{\infty} f(u) du$$

به عبارت دیگر تابع بقاء احتمال اینکه طول عمر از t تجاوز کند، می باشد. و این یک اندازه اصلی در قابلیت اعتماد است.

^۱ Jamshidian@shbu.ac.ir

در مطالعات بقا و مباحث طول عمر، این سوال مطرح است که اگر بدانیم طول عمر موجود مورد بررسی بزرگتر یا مساوی زمان t می باشد $(X \geq t)$ چقدر احتمال دارد که این موجود در زمانی بلافاصله بعد از t از بین برود. پاسخ به این سوال در مباحث قابلیت اعتماد و تحلیل بقا نقش عمده ای را بازی می کند که به آن تابع نرخ شکست می گویند.

تعریف ۲.۱. فرض کنید $f(t)$ و $\bar{F}(t)$ به ترتیب تابع چگالی احتمال و تابع قابلیت اعتماد آن باشد. تابع $h(t)$ را تابع نرخ شکست (تابع نرخ خطر) می نامند و بصورت زیر نشان می دهند.

$$h(t) = \frac{f(t)}{\bar{F}(t)}, t > 0, F(t) < 1$$

در واقع این مقدار ریسک شکست را در طول $(t, t+s)$ توصیف می کند وقتی که در زمان t موجود مورد بررسی زنده و سالم بوده است.

نکته ۱.۱. تابع نرخ شکست، احتمال خراب شدن نیست (یعنی لزوماً مقداری بین صفر و یک اختیار نمی کند) این کمیت نرخ خراب شدن را برای موجود زنده در زمان t نشان می دهد. طبیعی است هر چقدر مقدار این کمیت بیشتر باشد، انتظار می رود که قابلیت اعتماد آن کمتر باشد. مثلاً اگر X_1, X_2 طول عمر دو قطعه با تابع شکستهای h_1, h_2 باشند، هرگاه $h_1(t) \leq h_2(t)$ آنگاه $\bar{F}_1(t) \geq \bar{F}_2(t)$ که در آن $\bar{F}_1(t)$ و $\bar{F}_2(t)$ به ترتیب تابع قابلیت اعتماد X_1, X_2 می باشند.

نکته ۲.۱. تابع نرخ شکست $h(t)$ می تواند رفتارهای متفاوتی براساس تغییرات زمان داشته باشد. متداول ترین نوع رفتار $h(t)$ می تواند ثابت، صعودی، نزولی و وانی شکل باشد.

اکثر پدیده ها در طبیعت دارای نرخ شکست صعودی هستند. مثلاً طول عمر یک تلویزیون، طول عمر یک لامپ الکتریکی و غیره، همچنین در طبیعت می توان پدیده هایی پیدا کرد که طول عمر آنها نرخ شکست نزولی داشته باشد، مثلاً آلومینیم و بعضی از فلزات دیگر دارای نرخ خطر نزولی هستند. یک واقعیت مهم در مورد تابع نرخ شکست آن است که رابطه آن با توزیع طول عمر X رابطه ای یک به یک است.

قضیه ۱.۱. اگر X نشان دهنده طول عمر یک موجود باشد و تابع نرخ شکست آن h و تابع قابلیت اعتماد آن $\bar{F}$ باشد آنگاه:

$$\bar{F}(t) = e^{-\int_0^t h(x)dx}, t > 0$$

۲ نرخ شکست صعودی (نزولی) (DFR)IFR

همانطور که در نکته ۲.۱ بیان شد، تابع نرخ شکست در t رفتارهای متفاوتی از خود نشان می دهد (صعودی، نزولی، ثابت و...) بر این اساس تعریف زیر را داریم.

تعریف ۱.۲. فرض کنید F تابع توزیع متغیر تصادفی X باشد. F را متعلق به کلاس توزیع های با نرخ شکست صعودی (نزولی) گوئیم هرگاه $h(t)$ ، تابع نرخ شکست X ، تابعی غیر نزولی (غیر صعودی) از t باشد. و معمولاً از علامت اختصاری (DFR)IFR برای نشان دادن آنکه F دارای نرخ شکست صعودی (نزولی) است استفاده می شود. توزیع نمایی توزیعی است که هم IFR و هم DRF می باشد، به عبارت دیگر توزیع نمایی در مرز بین دو کلاس DRF، IFR قرار دارد، یعنی دارای نرخ شکست ثابت می باشد.

تعریف ۲.۲. (قابلیت اعتماد شرطی). فرض کنید X طول عمر یک موجود باشد و بدانیم $X > t$ (یعنی موجود طول عمری بیشتر از t دارد) قابلیت اعتماد شرطی X به شرط آنکه $X > t$ در زمان $s+t$ را با $\bar{F}(s/t)$ نمایش داده و برابر است با:

$$\bar{F}(s/t) = P(X > s+t | X > t) = \frac{P(X > s+t)}{P(X > t)} = \frac{\bar{F}(s+t)}{\bar{F}(t)}$$

توجه داشته باشید که به ازای هر t تابع $\bar{F}(s/t)$ یک تابع قابلیت اعتماد است. یعنی تمام شرایط تابع قابلیت اعتماد را دارد به عبارت دیگر $\bar{F}(0/t) = 1, \bar{F}(\infty/t) = 0, \bar{F}(s/t)$ تابعی نزولی از s است. بر این اساس قضیه زیر را می توان بیان کرد.

قضیه ۱.۲. شرط لازم و کافی برای آنکه F , IFR, DFR باشد آنست که تابعی نزولی (صعودی) از t باشد. در واقع این قضیه بیان می کند هرگاه تابع نرخ شکست X افزایش (کاهش) می یابد قابلیت اعتماد شرطی آن کاهش (افزایش) پیدا می کند و برعکس.

توزیع F نرخ شکست صعودی IFR دارد اگر $\bar{F}(x/t) = \frac{\bar{F}(x+t)}{\bar{F}(t)}$ در $t \geq 0$ نزولی باشد برای همه $x \geq 0$.
و توزیع F نرخ شکست نزولی DFR دارد اگر $\bar{F}(x/t) = \frac{\bar{F}(x+t)}{\bar{F}(t)}$ برای $t \geq 0$ صعودی باشد برای همه $x \geq 0$.

۳ میانگین و واریانس طول عمر باقی مانده

فرض کنید X نشان دهنده طول عمر یک موجود زنده باشد که دارای تابع چگالی احتمال $f(t)$ و تابع بقای $\bar{F}(t)$ است. طبق تعریف میانگین X که با $E(X)$ نمایش داده می شود برابر است با:

$$E(X) = \int_0^{\infty} t f(t) dt$$

و بسادگی می توان نشان داد: $E(X) = \int_0^{\infty} \bar{F}(t) dt$ یعنی $E(X)$ برابر با انتگرال از صفر تا بی نهایت روی تابع قابلیت اعتماد می باشد این حقیقت در مورد تابع قابلیت اعتماد شرطی نیز صادق است.

میانگین طول عمر باقی مانده که با نماد اختصاری MRL نشان داده می شود، مدت زمان باقی مانده برای یک موجود زنده تا زمان t می باشد. MRL یک مفهوم اصلی در قابلیت اعتماد و تجزیه بقای عمر می باشد. اگر X یک متغیر تصادفی غیر منفی با توزیع تجمعی F باشد MRL در زمان t بصورت زیر تعریف می شود.

$$\mu(t) = E(X - t | X \geq t) = \frac{1}{\bar{F}(t)} \int_t^{\infty} \bar{F}(u) du \quad t \geq 0$$

تابع $\mu(t)$ را میانگین طول عمر باقی مانده می نامند. این تابع به عنوان تابعی پویا در هر لحظه از زمان، میانگین عمر باقی مانده موجود را نشان می دهد. واضح است $\mu(0) = E(X)$ امروزه در تحلیل داده های طول عمر و بقا، میانگین طول عمر باقی مانده و تابع نرخ شکست نقش مهمی را ایفا می کنند. بین میانگین طول عمر باقی مانده و توزیع طول عمر یک رابطه یک به یک وجود دارد.

قضیه ۱.۳. اگر X دارای تابع قابلیت اعتماد $\bar{F}(t)$ و میانگین طول عمر $\mu(t)$ باشد آنگاه:

$$\bar{F}(t) = \frac{\mu(0)}{\mu(t)} e^{-\int_0^t \frac{1}{\mu(u)} du}$$

نکته ۱.۳. در مباحث قابلیت اعتماد $\mu(0) = E(X)$ را میانگین زمان شکست^۱ (MTTF) می گویند.

نکته ۲.۳. میانگین طول عمر باقی مانده $\mu(t)$ و تابع نرخ شکست $h(t)$ بصورت زیر با هم مرتبط می باشند.

$$h(t) = \frac{1 + \mu'(t)}{\mu(t)}$$

اگر میانگین طول عمر باقی مانده $\mu(x)$ نزولی (صعودی) باشد، آنگاه به متغیر تصادفی x می گویند. دارای توزیع میانگین طول عمر باقی مانده نزولی (صعودی) می باشد و آن را با نماد DMRL (IMRL) نشان می دهند.

به متغیر تصادفی سالخورده $x \geq 0$ می گویند دارای توزیع طول عمر باقی مانده نزولی DMRL است اگر معادله ی زیر در $x \geq 0$ نزولی باشد.

$$\mu(x) = E(X - \frac{x}{X} > x) = \frac{\int_x^{\infty} \bar{F}(u) du}{\bar{F}(x)}$$

^۱ Mean Time To Failure

یک طول عمر با توزیع F و تابع بقا $\bar{F}$ را می گویند دارای میانگین توزیع طول عمر باقی مانده صعودی IMRL است اگر $\mu(x)$ در $x \geq 0$ صعودی باشد. چون F ، IMRL است، پس $\mu(x) = \frac{V(x)}{\bar{F}(x)}$ در $x \geq 0$ صعودی می باشد با $V(x) = \int_x^\infty \bar{F}(u)du$ بنابراین $\mu'(x) \geq 0$ به ازای $x \geq 0$ در صورتی که مشتق وجود داشته باشد. این معادل است با $F \in IMRL$ اگر و تنها اگر

$$f(x)V(x) - \bar{F}(x)^2 \geq 0.$$

$$\bar{F}^2(x) \leq f(x)V(x) \text{ بنابراین}$$

لم ۱۰۳. DMRL (IMRL) شامل IFR (DFR) می باشد

رفتار کلاسهای اصلی توزیعهای طول عمر، بر حسب نرخ شکست شان، میانگین طول عمر باقی مانده و واریانس طول عمر باقی مانده مشخص می شوند، که در آزمون طول عمر و مدلهای قابلیت اعتماد مفید می باشند. واریانس طول عمر باقی مانده برای اندازه گیری میزان پراکندگی توزیع طول عمر استفاده می شود. اگر X یک متغیر تصادفی غیر منفی که نشان دهنده طول عمر یک مولفه با تابع توزیع $F(t)$ و تابع قابلیت اعتماد $\bar{F}(t) = 1 - F(t)$ باشد. واریانس طول عمر باقی مانده در زمان t بصورت زیر تعریف می شود.

$$\sigma^2(t) = \frac{2\bar{F}(t)G(t) - V^2(t)}{\bar{F}^2(t)}$$

که:

$$G(t) = \int_t^\infty V(u)du \quad V(t) = \int_t^\infty \bar{F}(u)du$$

رابطه واریانس طول عمر باقی مانده با نرخ شکست $h(t)$ و میانگین طول عمر باقی مانده $\mu(t)$ بصورت زیر می باشد.

$$\sigma'^2(t) = h(t)(\sigma^2(t) - \mu^2(t))$$

همچنین توزیع نمایی بطور منحصر بفرد دارای واریانس طول عمر باقی مانده ثابت می باشد. به عبارت دیگر اگر X نشان دهنده متغیر تصادفی نا منفی با تابع توزیع F باشد، آنگاه واریانس طول عمر باقی مانده در زمان t در بازه $t \geq 0$ ثابت است اگر و تنها اگر F بصورت تابع توزیع نمایی باشد.

تابع توزیع را که واریانس طول عمر باقی مانده نزولی (صعودی) داشته باشد با نماد DVRL (IVRL) نشان می دهند. اگر واریانس طول عمر باقی مانده یک تابع غیر صعودی (غیر نزولی) باشد یعنی اگر $\sigma^2(x)$ غیر صعودی یا غیر نزولی، در x باشد آنگاه به F توزیع DVRL یا IVRL برای متغیر تصادفی x می باشد.

ابویوسف (۲۰۰۴) نشان داد که یک توزیع طول عمر F ، با $F(0) = 0$ و تابع بقا $\bar{F}$ ، DVRL (IVRL) است اگر

$$\frac{1}{\bar{F}(t)} \int_t^\infty v(y)dy \leq (\geq) \mu^2(t)$$

یا

$$\bar{F}(t) \int_t^\infty v(y)dy \leq (\geq) v^2(t) \quad v(t) = \int_t^\infty \bar{F}(x)dx$$

لم ۲۰۳. DMRL $\Rightarrow$ DVRL

۴ چند توزیع دیگر طول عمر

اگر یک سیستم منسجم داشته باشیم که مولفه هایش IFR باشند لزومی ندارد که طول عمر سیستم نیز IFR باشد. بنابراین یک کلاس کوچکتر را معرفی می کنیم که بتوانیم در مورد طول عمر کل سیستم اظهار نظر کنیم، براین اساس تعریف زیر بیان می شود.

تعریف ۱.۴. توزیع F دارای خاصیت متوسط نرخ شکست صعودی است اگر

$$\frac{\int_0^t h(x)dx}{t}$$

صعودی در t می باشد. این خاصیت را با نماد IFRA نشان می دهند.

یک توزیع $F(t)$ دارای خاصیت متوسط نرخ شکست صعودی است اگر و تنها اگر $-\frac{1}{t} \log[\bar{F}(t)]$ در t صعودی باشد.

قضیه ۱.۴. فرض کنید مولفه های مستقل در یک سیستم منسجم دارای توزیع طول عمر IFRA است آنگاه توزیع طول عمر سیستم نیز دارای یک توزیع طول عمر IFRA می باشد.

یک حالت خاص از قضیه ۱.۴ این است که اگر مولفه های یک سیستم منسجم مستقل IFR باشند آنگاه توزیع طول عمر سیستم IFRA می باشد.

لم ۱.۴. $F \in IFR \Rightarrow F \in IFRA$

تعریف ۲.۴. توزیع $F(t)$ را نو بهتر از استفاده شده می نامند و با نماد NBU نشان می دهند اگر:

$$\bar{F}(x+y) \leq \bar{F}(x)\bar{F}(y) \quad \forall x, y \geq 0$$

این عبارت هم ارز حالتی است که احتمال بقا شرطی $\frac{\bar{F}(x+y)}{\bar{F}(x)}$ از یک واحد به سن X کمتر از احتمال بقا متناظر با $\bar{F}(y)$ از یک واحد جدید می باشد. یعنی

$$P(X > x+y | X > x) = \frac{\bar{F}(x+y)}{\bar{F}(x)} \leq \bar{F}(y) = p(X > y)$$

برای متغیر تصادفی X تساوی برقرار است اگر و تنها اگر F توزیع نمایی باشد.

لم ۲.۴. $F \in IFRA \Rightarrow F \in NBU$

لم ۳.۴. $F \in IFR \Rightarrow F \in IFRA \Rightarrow F \in NBU$

توزیع $F(t)$ را نو بهتر از استفاده شده در امید می نامند و با نماد NBUE نشان می دهند اگر

$$\mu(t) \leq E(X)$$

که $E(X)$ مقدار امید ریاضی متغیر تصادفی X می باشد. به عبارت دیگر یک واحد استفاده شده در سن t دارای میانگین طول عمر باقی مانده کوچکتر از یک واحد نو است اگر F نو بهتر از استفاده شده در امید NBUE باشد.

بنابراین میانگین طول عمر باقی مانده (MRL) یک مولفه استفاده شده از میانگین طول عمر یک مولفه جدید تجاوز نمی کند. همچنین می توان توجه کرد که توزیع نو بهتر از استفاده شده در امید است اگر

$$\int_t^\infty \bar{F}(x)dx = \bar{F}(t).\theta \quad \forall t > 0$$

که θ میانگین توزیع F می باشد.

لم ۴.۴. $F \in NBU \Rightarrow F \in NBUE$

لم ۵.۴. NBUE شامل IFRA، IFR می باشد

یک متغیر تصادفی سالخوردگی $x \geq 0$ ، دارای توزیع نو بهتر از استفاده شده در ترتیب محدب است و با نماد NBUC نشان می دهند اگر

$$\int_x^\infty \bar{F}(t+a)du \leq \bar{F}(t) \int_x^\infty \bar{F}(u)du \quad x, t \geq 0$$

واریانس طول عمر باقی مانده نزولی خالص را با نماد NDVRL نشان می دهند NDVRL ایده ای شبیه به NDVRL دارد. می گوئیم F توزیع طول عمر دارای NDVRL است اگر واریانس شرطی اش از واریانس اولیه تجاوز نکند یعنی

$$\sigma^2(x) \leq \sigma^2_{(\cdot)} = \sigma^2$$

$$\text{لم ۰۶۰۴} \quad DVRL \Rightarrow NDVRL$$

$$\text{لم ۰۷۰۴} \quad NBUE \Rightarrow NDVRL$$

۵ دست‌آوردهای پژوهش

بر اساس مطالب بیان شده می توان ارتباط توزیع های طول عمر را به صورت شکل ۱ به نمایش گذاشت.

$$\begin{array}{ccc} IFR & \Rightarrow & IFRA \Rightarrow NBU \Rightarrow NBUC \\ \Downarrow & & \Downarrow \\ DMRL & \Rightarrow & NBUE \\ \Downarrow & & \Downarrow \\ DVRL & \Rightarrow & NDVRL \end{array}$$

شکل ۱: رابطه ی بین توزیع های طول عمر

مراجع

- [1] I.A. Ahmad and A.R. Mugdadi (2005), *Bounds of moment generating functions of some life distributions*, Statistical Papers 46,575-585.
- [2] I.A. Ahmad and A.R. Mugdadi. (2004), *Further moments inequalities of life distributions with hypothesis testing applications: the IFRA, NBUC and DMRL classes*, Journal of Statistical planning and inference 120, 1-12. .
- [3] M.H.M. Al-Sadi (2004), *Properties of classes of life distribution based on the conditional variance*, Applied Probability Trust 41, 953-960.
- [4] R.E. Barlow and F. Proschan (1975), *Statistical theory of reliability and life testing*.
- [5] S.E. Abu-Youssef (2004), *Nonparametric test for monotone variance residual life class of life distributions with hypothesis testing applications*, Applied Mathematics and computation 158, 817-826.



مطالعه‌ای بر نگهداری و تعمیر پیشگیرانه سیستم‌های یک بار مصرف و برآورد پارامترها

چرکزی، ح^۱ احمدی، ج^۲

^{۱،۲} گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

چکیده

در مقاله حاضر، ابتدا به معرفی سیستم یک بار مصرف می‌پردازیم. سپس سیاست دوره‌ای را برای نگهداری، تعمیر پیشگیرانه و جایگزینی این سیستم‌ها، برای به حداقل رساندن متوسط هزینه و متوسط در دسترس بودن این سیستم‌ها بیان می‌کنیم. نهایتاً روشی برای برآورد پارامتر توزیع طول عمر سیستم یک بار مصرف ارائه می‌کنیم.

کلمات کلیدی: برآورد ماکسیمم درستی، تعمیر مینیمال، سیاست دوره‌ای بازرسی، سیستم یک بار مصرف، فرآیند پواسن ناهمگن.

۱ مقدمه

یکی از مباحث مهم در قابلیت اعتماد، بحث نگهداری و تعمیر سیستم‌هاست. با توجه به هزینه‌ی بالای خرید ماشین آلات و سیستم‌های نو، الگوهای نگهداری و تعمیر از اهمیت ویژه‌ای برخوردارند. سیستم‌های پیچیده مانند اتومبیل‌ها، سیستم‌های ارتباطی، تجهیزات پزشکی، هواپیماها و قطارها پس از خرابی به علت اینکه هزینه‌ی تعمیر آن‌ها به مراتب کمتر از هزینه‌ی جایگزینی با سیستم نو است، معمولاً تعمیر می‌شوند. یک سیستم را قابل تعمیر گویند، هرگاه پس از خرابی طی یک فرآیند مشخص (بدون تعویض کل سیستم) دوباره به وضعیت فعال برگردد. در

^۱ h.charckazi1993@um.ac.ir

^۲ ahmadi-j@um.ac.ir

ارتباط با سیستم‌های قابل تعمیر کتاب‌های زیادی نوشته شده است که از آن جمله می‌توان به کتاب آشور و فینگلد (۱۹۸۴) اشاره نمود. از مباحث مورد مطالعه در زمینه سیستم‌های قابل تعمیر، می‌توان به محاسبه‌ی تعداد خرابی‌های مورد انتظار در طول دوره‌ی گارانتی سیستم، بررسی قابلیت اعتماد و نرخ خرابی سیستم، تعیین زمان بازرسی و جایگزینی سیستم و کمینه کردن هزینه‌ی نگهداری و تعمیر سیستم اشاره نمود. برای بررسی مشخصه‌های قابلیت اعتماد سیستم‌های قابل تعمیر معمولاً از نتایج نظری احتمال و فنون فرآیندهای تصادفی استفاده می‌شود. هنگامی که با سیستم‌هایی سروکار داریم که شکست آن‌ها ممکن است باعث صدمات جبران ناپذیری شود، الگوهای تعمیر پیشگیرانه نقش مهمی در جلوگیری از خرابی‌های اساسی سیستم‌ها ایفا می‌کنند. در نگهداری پیشگیرانه، یک سیستم قبل از آن که دچار شکست شود براساس شیوه‌هایی خاص، برخی از مولفه‌های معیوب مورد بررسی و تعمیر قرار می‌گیرد. در واقع الگوهای پیشگیرانه برای کاهش فشار و استرس روی سیستم‌ها به‌کار می‌رود تا احتمال خرابی سیستم را کاهش دهند. از طرفی در دنیای واقعی، با سیستم‌هایی مانند موشک، گلوله‌های انفجاری، کیسه‌های هوا در ماشین‌ها و دستگاه‌های انفجاری الکتریکی سروکار داریم که تنها از یک واحد یکپارچه تشکیل شده و انجام عملکرد آن تنها یک بار است. این سیستم‌ها، بعد از شروع فعالیت دیگر قابل تعمیر نیستند و آن‌ها را سیستم‌های یک بار مصرف گویند. از آن‌جا که، در آزمایش سیستم‌های یک بار مصرف تنها وضعیت دستگاه در یک زمان خاص می‌تواند ثبت شود، زمان‌های دقیق شکست را نمی‌توان از طریق آزمایش بدست آورد. در نتیجه، طول عمر واحدهای تحت آزمون در زمان بازرسی همیشه از چپ یا راست سانسور شده است. بنابراین، برای تحلیل این گونه سیستم‌ها، نیاز به بررسی‌های بیشتری نسبت به سیستم‌های دیگر است.

۲ نگهداری سیستم یک بار مصرف در حالت آماده به کار

به دلیل استفاده زیاد از سیستم‌های یک بار مصرف برای اهداف مهمی از قبیل قسمت‌های پزشکی و جنگ، برای حفظ در دسترس بودن بالای سیستم، بازرسی‌های زیاد و مکرر صورت گیرد؛ با این حال، بازرسی‌های مکرر ممکن است هزینه‌ی زیادی را متحمل شوند که برای مصرف‌کنندگان غیر قابل قبول است. علاوه بر این، بازرسی‌های مکرر، با توجه به مدت زمان نگهداری طولانی سیستم‌ها، منجر به در دسترس بودن کم سیستم‌ها می‌شود (زیرا برخی از سیستم‌ها ممکن است در این فاصله‌ی زمانی از بین روند). بنابراین، لازم است برنامه‌های بازرسی، بهینه سازی شوند. مشکلات سیاست بازرسی توسط بسیاری از محققان از جمله کیتاگاوا و همکاران (۲۰۱۴) و ناکاگاوا و کوادا (۱۹۸۳) مورد مطالعه قرار گرفته است. زمانی که یک خرابی شناسایی می‌شود، سیستم نیاز به برخی از انواع الگوهای نگهداری و تعمیر از قبیل جایگزینی یا تعمیر دارد. تقریباً تمام مطالعات از مشکلات سیاست بازرسی، پس از شناسایی یک خرابی، جایگزینی را در نظر می‌گیرند. ولی در مواردی، با سیستم‌هایی سر و کار داریم که جایگزینی آن‌ها، هزینه‌های زیادی را متحمل می‌شود در این گونه موارد، معمولاً تعمیر مینیمال مقرون به صرفه‌تر از جایگزینی سیستم است. تعمیر مینیمال توسط بارلو و هانت (۱۹۶۰) معرفی شد که در آن، پس از تعمیر سیستم از کار افتاده، سیستم دوباره شروع به کار می‌کند اما وضعیت آن از نظر سالخوردگی مشابه یک سیستم نو نیست بلکه، مشابه حالتی می‌شود که بلافاصله پیش از، از کار افتادن سیستم داشته است. هدف ما، مطالعه‌ی سیستم‌های یک بار مصرف براساس نتایج کیتاگاوا و همکاران (۲۰۱۴) است و فرض بر این است که این سیستم‌ها در موقعیت آماده به کار و یا در انبار قرار دارند.

مفروضات مدل در نظر گرفته شده در این مطالعه به شرح زیر است:

۱. سیستم متشکل از یک جزء یکپارچه و دارای خاصیت نرخ خطر صعودی است.
۲. در زمان بازرسی، تمام خرابی‌ها قابل شناسایی باشند.
۳. در طول بازرسی در صورت وجود خرابی، تعمیر مینیمال انجام می‌شود.
۴. فاصله‌ی بازرسی پس از $(m-1)$ امین تعمیر مینیمال، برای $m = 1, \dots, n$ ، برابر $T^{(m)}$ است.
۵. با شناسایی n امین خرابی، سیستم تحت بررسی دیگر قابل استفاده نیست و در نتیجه برای حفظ موجودی تعداد سیستم‌ها در انبار، سیستم نو سفارش و جایگزین می‌شود.
۶. میزان مخاطره‌ی این سیستم در طول تعمیر یا در صورت خرابی تغییر نمی‌کند.

۳ تجزیه و تحلیل مدل

در این بخش، نرخ هزینه‌های مورد انتظار و متوسط در دسترس بودن سیستم یک بار مصرف بررسی می‌شود. توزیع خرابی سیستم یک بار مصرف را می‌توان با استفاده از یک فرآیند پواسون ناهمگن به صورت زیر بیان کرد. فرض کنید $N(t)$ نشان‌دهنده تعداد خرابی‌های جزئی سیستم تا زمان t و $P_m(t)$ احتمال این که تعداد خرابی‌های شناسایی شده از سیستم تا زمان t ، برابر m باشد. در این صورت با استفاده از خواص فرآیند پواسن ناهمگن می‌توانیم بنویسیم

$$\begin{aligned} P_m(t) &= P(N(t) = m) \\ &= \frac{\exp\{-H(t)\} (H(t))^m}{m!}, \quad m = 0, 1, \dots \end{aligned} \quad (1)$$

که در آن، $H(t)$ تابع شدت تجمعی است.

اگر فرض کنیم Y_m نشان‌دهنده زمان انتظار تا وقوع m امین خرابی باشد، آنگاه با استفاده از معادله (۱). داریم

$$\begin{aligned} G^{(m)}(t) &= P(Y_m \leq t) \\ &= P(N(t) \geq m) \\ &= 1 - \sum_{i=0}^{m-1} \frac{[H(t)]^i}{i!} \exp(-H(t)). \end{aligned} \quad (2)$$

در نتیجه، متوسط زمان فعال بودن سیستم بین $(m-1)$ امین و m امین خرابی به صورت زیر خواهد بود

$$\begin{aligned} \mu^{(m)} &= \int_0^\infty \bar{G}^{(m)}(t) dt - \int_0^\infty \bar{G}^{(m-1)}(t) dt \\ &= \int_0^\infty \frac{(H(t))^{m-1}}{(m-1)!} e^{-H(t)} dt. \end{aligned} \quad (3)$$

اگر زمان خرابی سیستم دارای توزیع وایبول با پارامتر مقیاس η و پارامتر شکل β ، $(\beta > 1)$ باشد، از معادله (۳) داریم

$$\mu^{(m)} = \frac{\eta}{\beta} \frac{\Gamma(m + \frac{1}{\beta} - 1)}{\Gamma(m)}. \quad (4)$$

اگر X_m فاصله‌ی زمانی بین دو تعمیر مینیمال باشد یعنی $X_m = Y_m - Y_{m-1}$ ، در این صورت تابع قابلیت اعتماد سیستم پس از $(m-1)$ امین تعمیر مینیمال، $\bar{F}^{(m)}(x)$ ، با استفاده از معادله (۲) برای $m \geq 2$ به صورت زیر بیان می‌شود

$$\begin{aligned} \bar{F}^{(m)}(x) &= P(X_m > x) \\ &= 1 - \int_0^\infty P\{X_m \leq x | Y_{m-1} = y\} dG^{(m-1)}(y) \\ &= 1 - \int_0^\infty \frac{F(x+y) - F(y)}{\bar{F}(y)} dG^{(m-1)}(y) \\ &= 1 - \left\{ \int_0^\infty \bar{F}(y) \frac{(H(y))^{m-2}}{(m-2)!} h(y) dy - \int_0^\infty \bar{F}(x+y) \frac{(H(y))^{m-2}}{(m-2)!} h(y) dy \right\} \\ &= \int_0^\infty \bar{F}(x+y) \frac{(H(y))^{m-2}}{(m-2)!} dH(y). \end{aligned} \quad (5)$$

اگر مدت زمان بین جایگزینی پی‌درپی سیستم را، به عنوان یک دوره در نظر بگیریم، آنگاه کیتاگاوا و همکاران (۲۰۱۶) متوسط در دسترس بودن سیستم و نرخ هزینه‌های مورد انتظار را به صورت زیر بدست آورده‌اند

$$Av(T^{(1)}, \dots, T^{(n)}) = \frac{E[\text{عامل زمان در هر دوره}]}{E[\text{یک دوره}]}, \quad (6)$$

و

$$C(T^{(1)}, \dots, T^{(n)}) = \frac{E[\text{هزینه‌های متحمل شده در هر دوره}]}{E[\text{یک دوره}]} \quad (۷)$$

 از طرفی، اگر τ متوسط زمان در هر دوره باشد، آنگاه

$$\begin{aligned} \tau &= \sum_{m=1}^{n-1} \sum_{k=1}^{\infty} \int_{(k-1)T^{(m)}}^{kT^{(m)}} \{k(T^{(m)} + a_I) + a_R\} dF^{(m)}(x) \\ &+ \sum_{k=1}^{\infty} \int_{(k-1)T^{(n)}}^{kT^{(n)}} \{k(T^{(n)} + a_I) + a_P\} dF^{(n)}(x) \\ &= \sum_{m=1}^n \sum_{k=1}^{\infty} \left\{ (T^{(m)} + a_I) \bar{F}^{(m)}(kT^{(m)}) \right\} + (n-1)a_R + a_P, \end{aligned} \quad (۸)$$

که در آن a_I متوسط زمان بازرسی، a_R متوسط زمان تعمیر مینیمال و a_P متوسط زمان جایگزینی سیستم است. در نتیجه از (۶) متوسط در دسترس بودن سیستم به صورت زیر بیان می شود

$$\begin{aligned} Av(T^{(1)}, \dots, T^{(n)}) &= \frac{\sum_{m=1}^n \mu^{(m)}}{\tau} \\ &= \frac{\sum_{m=1}^n \mu^{(m)}}{\sum_{m=1}^n \sum_{k=1}^{\infty} \left\{ (T^{(m)} + a_I) \bar{F}^{(m)}(kT^{(m)}) \right\} + (n-1)a_R + a_P}. \end{aligned} \quad (۹)$$

هزینه کل مورد انتظار در هر دوره به صورت زیر داده می شود

$$\begin{aligned} C &= \sum_{m=1}^{n-1} \sum_{k=1}^{\infty} \int_{(k-1)T^{(m)}}^{kT^{(m)}} \{kC_I + C_R\} dF^{(m)}(x) \\ &+ \sum_{k=1}^{\infty} \int_{(k-1)T^{(n)}}^{kT^{(n)}} \{kC_I + C_P\} dF^{(n)}(x) \\ &= C_I \sum_{m=1}^n \sum_{k=1}^{\infty} \bar{F}^{(m)}(kT^{(m)}) + (n-1)C_R + C_P, \end{aligned} \quad (۱۰)$$

که در آن C_I هزینه بازرسی، C_R هزینه تعمیر مینیمال و C_P هزینه جایگزینی سیستم در هر زمان است. و در نتیجه، نرخ هزینه مورد انتظار برای $(0, \infty)$ برابر است با

$$\begin{aligned} C(T^{(1)}, \dots, T^{(n)}) &= \frac{C}{\tau} \\ &= \frac{C_I \sum_{m=1}^n \sum_{k=1}^{\infty} \bar{F}^{(m)}(kT^{(m)}) + (n-1)C_R + C_P}{\sum_{m=1}^n \sum_{k=1}^{\infty} \left\{ (T^{(m)} + a_I) \bar{F}^{(m)}(kT^{(m)}) \right\} + (n-1)a_R + a_P}. \end{aligned} \quad (۱۱)$$

هنگامی که n بزرگ می شود، فواصل بازرسی برای اطمینان از متوسط در دسترس بودن از پیش تعیین شده، کوتاه تر می شوند و نرخ هزینه مورد انتظار افزایش می یابد. در حالی که، وقتی n کوچک می شود، از آن جا که انتظار می رود فاصله بین جایگزینی ها کوتاه تر شود، هزینه جایگزینی سیستم افزایش می یابد. بنابراین، باید مقدار n و فواصل بازرسی را به منظور به حداقل رساندن معادله (۱۱) تعیین کنیم به قسمی که $Av \geq \alpha$ ، که در آن α متوسط در دسترس بودن هدف سیستم است.

۴ تعیین فواصل بازرسی بهینه

در این بخش، روش محاسبه برای تعیین فواصل بهینه را برای سیاست بازرسی غیردوره ای بیان می کنیم. با توجه به اینکه سیاست غیر دوره ای دارای پارامترهای زیادی است، به کارگیری روش هایی که جواب بهینه تولید می کنند بسیار دشوار و زمان بر می باشد و اغلب برای مسائلی با اندازه ی کوچک میسر است. از این رو بهتر است برای حل آن ها از روش های متاهوریستیک استفاده شود. در

این بخش، تعیین فواصل بازرسی بهینه با استفاده از نرم افزار ریاضی *Matlab* و با روش الگوریتم تبرید شبیه‌سازی شده، با تعیین پارامترهای مدل که در جدول ۱ ارائه گردیده، انجام گرفته است. (برای مشاهده‌ی اطلاعات بیشتر در این زمینه به وو و همکاران (۲۰۰۸) مراجعه کنید.) برای حل یک مسئله بهینه‌سازی، الگوریتم تبرید شبیه‌سازی شده (*SA*) ابتدا از یک جواب اولیه شروع می‌کند و سپس در یک حلقه تکرار به جواب‌های همسایه حرکت می‌کند. اگر جواب همسایه، بهتر از جواب فعلی باشد، الگوریتم آن را به عنوان جواب فعلی قرار می‌دهد، در غیراینصورت، الگوریتم آن جواب را به صورت احتمالی به عنوان جواب فعلی می‌پذیرد که این احتمال به صورت زیر محاسبه می‌شود

$$p = \exp\left(-\frac{d}{t}\right) \quad (12)$$

که در آن d تفاوت در نرخ هزینه (تفاوت بین هزینه‌ی جواب فعلی و جواب همسایه) و t یک پارامتر به نام دما است. در هر دما، چندین تکرار اجرا می‌شود و سپس دما به آرامی کاهش داده می‌شود. در گام‌های اولیه، دما خیلی بالا قرار داده می‌شود تا احتمال بیشتری برای پذیرش جواب‌های بدتر وجود داشته باشد. با کاهش تدریجی دما، در گام‌های پایانی احتمال کم‌تری برای پذیرش جواب‌های بدتر وجود خواهد داشت و بنابراین الگوریتم به سمت یک جواب خوب همگرا می‌شود.

در هر مرحله، الگوریتم تبرید شبیه‌سازی شده، چند حالت را در همسایگی جواب بهینه در نظر می‌گیرد، به طور احتمالی تصمیم می‌گیرد که سیستم را از جواب منتقل کند یا در همین حالت باقی بماند. این احتمالات در نهایت سیستم را به حالت با انرژی کم‌تر میل می‌دهد. همسایه‌های یک جواب، جواب‌های جدیدی از مسئله هستند که با تغییر در جواب کنونی و با توجه به روشی از پیش تعیین شده ایجاد می‌شود. بنابراین، الگوریتم فوق را برای یک n ثابت می‌توان به صورت زیر بیان کرد:

گام اول یک جواب اولیه $\{T^{(1)}, T^{(2)}, \dots, T^{(n)}\}$ انتخاب می‌کنیم به طوریکه با استفاده از معادله‌ی (۹.۳) داشته باشیم $Av \geq \alpha$ و نرخ هزینه‌ی آن را محاسبه می‌کنیم.

گام دوم فواصل بازرسی را به طور تصادفی به منظور تعیین جواب‌های جدید به اندازه ± 0.3 تغییر می‌دهیم.

گام سوم نرخ هزینه و در دسترس بودن جدید را محاسبه می‌کنیم.

اگر دسترسی جدید رضایت‌بخش باشد یعنی $Av \geq \alpha$ و نرخ هزینه جدید، کمتر از نرخ هزینه قدیمی باشد، جواب را تجدید می‌کنیم. اگر در دسترس بودن جدید رضایت‌بخش باشد ولی نرخ هزینه جدید، بیشتر باشد، با احتمال p جواب را تجدید می‌کنیم. با این وجود، جواب جدید، رد می‌شود.

گام چهارم اگر جواب پس از تعداد زیادی تکرار متوالی بدون تغییر بماند، آن را به عنوان جواب بهینه در نظر می‌گیریم. در غیراینصورت t را ضربدر ۰.۹۵ کرده و به گام دوم می‌رویم.

به عنوان مثال فرض کنید که خرابی واحدهای آزمایش دارای توزیع وایبول با پارامتر مقیاس η و پارامتر شکل β باشد. پارامترهای مدل تحت مطالعه در جدول ۱ ارائه شده است. فواصل بازرسی بهینه از سیاست بازرسی غیر دوره‌ای برای مقادیر مختلف n در جدول ۲ خلاصه‌سازی شده است که در آن، زمان برحسب روز و هزینه بر حسب ۱۰۰۰ دلار است.

با توجه به جدول ۲

۱. با افزایش n ، برای اطمینان از متوسط در دسترس بودن از پیش تعیین شده، فواصل بازرسی کوتاه و نرخ هزینه‌های مورد انتظار افزایش می‌یابد.

۲. هنگامی که n کوچک باشد، با توجه به این که فاصله‌ی بین جایگزینی‌ها کوتاه‌تر می‌شود، هزینه‌ی جایگزینی سیستم، نسبتاً گران می‌شود.

۳. به ازای فواصل بازرسی که برای $n = 4$ به دست آمده، سیاست تحت مطالعه دارای نرخ هزینه کمتری است یعنی این فواصل یک جواب بهینه برای مسئله است.

جدول ۱: پارامترهای مدل

تابع توزیع (CDF)	پارامتر مقیاس (η)	پارامتر شکل (β)
وایبول	۳۰۰۰	۲
هزینه	۱۰	۴۰۰
متوسط زمان	۳	۸
	بازرسی	تعمیر مینیمال
	جایگزینی	

جدول ۲: فواصل بازرسی بهینه

n	$T^{(۱)}$	$T^{(۲)}$	$T^{(۳)}$	$T^{(۴)}$	$T^{(۵)}$	$T^{(۶)}$	نرخ هزینه
۲	۴۶۵	۳۱۷	—	—	—	—	۰/۱۲۳۹
۳	۴۱۴	۲۸۸	۲۴۷	—	—	—	۰/۱۱۶۶
۴	۳۸۵	۲۶۱	۲۲۵	۲۱۱	—	—	۰/۱۱۵۱
۵	۳۶۰	۲۴۷	۲۱۰	۱۹۵	۱۷۹	—	۰/۱۱۶۰
۶	۳۳۶	۲۳۱	۲۰۴	۱۸۲	۱۶۸	۱۶۲	۰/۱۱۸۲

۵ برآورد پارامترهای توزیع طول عمر سیستم یک بار مصرف

هدف از این بخش، انجام استنباط آماری درباره‌ی پارامترهای توزیع طول عمر سیستم‌های یک بار مصرف می‌باشد. برای این منظور، با استفاده از نتایج بالاکریشنان و همکاران (۲۰۱۴) و (۲۰۱۵)، تعداد N واحد از یک سیستم یک بار مصرف را در نظر گرفته و K را تعداد دفعات بازرسی در نظر می‌گیریم. به‌طوری‌که N_i تعداد واحدهای آزمون شده در زمان بازرسی i ام وقتی که $N = \sum_{i=1}^K N_i$ است. فرض کنید $0 < t_1 < t_2 < \dots < t_i$ ، زمان‌های بازرسی و تعداد خرابی‌های مشاهده شده در زمان بازرسی i ام، برابر n_i باشند. داده‌های آزمون سیستم یک بار مصرف بدست آمده از چنین آزمایش طول عمری برای $i = 1, \dots, K$ به‌صورت (n_i, N_i, t_i) بیان می‌شود. همان‌طور که در بخش ۱ اشاره شد، زمان دقیق خرابی را نمی‌توان برای این سیستم‌ها تعیین کرد و داده‌ها همیشه از چپ یا راست سانسور شده‌اند. در نتیجه، تابع درستمایی مشاهده شده براساس داده‌ها به‌صورت زیر بیان می‌شود

$$L(\lambda) \propto \prod_{i=1}^K (F(t_i))^{n_i} (1 - F(t_i))^{N_i - n_i} \\ \propto \prod_{i=1}^K (1 - R(t_i))^{n_i} (R(t_i))^{N_i - n_i}. \quad (۱۳)$$

با فرض این‌که، طول عمر سیستم‌های یک بار مصرف دارای توزیع نمایی با میانگین λ باشند، با استفاده از (۱۳) خواهیم داشت

$$\ell(\lambda) \equiv \sum_{i=1}^K \left(n_i \ln \left(1 - e^{-\frac{t_i}{\lambda}} \right) + (N_i - n_i) \ln \left(e^{-\frac{t_i}{\lambda}} \right) \right). \quad (۱۴)$$

معادله‌ی درستنمایی براساس داده‌ها برابر است با

$$\sum_{i=1}^K \frac{t_i}{\lambda^2} \left((N_i - n_i) - \frac{n_i e^{-\frac{t_i}{\lambda}}}{1 - e^{-\frac{t_i}{\lambda}}} \right) = 0, \quad (15)$$

برای حل این معادله باید از روش‌های عددی استفاده کرد. در صورت داشتن داده‌ها، می‌توان برآوردی برای λ از معادله‌ی (۱۵) بدست آورد. به‌عنوان مثال، با فرض این‌که زمان‌های بازرسی، $0 < t_1 < \dots < t_i$ ، از پیش تعیین شده باشند، تعداد خرابی‌های مشاهده شده در زمان بازرسی i ام، دارای توزیع دوجمله‌ای با پارامترهای $(N_i, F(t_i))$ خواهد بود، که در آن $F(t_i)$ تابع توزیع طول عمر سیستم‌های یک بار مصرف در زمان بازرسی i ام برای $i = 1, \dots, K$ است. در این مقاله تعداد خرابی‌های سیستم را با ۱۰۰۰ بار تکرار از توزیع دوجمله‌ای با پارامترهای $(N_i, F(t_i))$ شبیه‌سازی کرده‌ایم. برای برخی از مقادیر K اگر داده‌ها مطابق جدول ۳ باشند، آن‌گاه برآورد λ به روش نیوتن رافسون با مقدار اولیه‌ی $\lambda_0 = 2$ به‌سادگی قابل محاسبه است.

جدول ۳: برآورد ماکسیمم درستنمایی پارامتر λ براساس داده‌ها

K	۲	۴	۶	۸	۱۰
(N_i, t_i)	(۵۰, ۰/۵)	(۳۰, ۰/۵)	(۲۵, ۰/۵)	(۲۰, ۰/۵)	(۱۵, ۰/۵)
	(۵۰, ۱)	(۳۰, ۱)	(۲۰, ۱)	(۲۰, ۱)	(۱۵, ۱)
		(۲۰, ۱/۵)	(۲۰, ۱/۵)	(۱۵, ۱/۵)	(۱۵, ۱/۵)
		(۲۰, ۲)	(۱۵, ۲)	(۱۵, ۲)	(۱۰, ۲)
			(۱۰, ۲/۵)	(۱۰, ۲/۵)	(۱۰, ۲/۵)
			(۱۰, ۳)	(۱۰, ۳)	(۱۰, ۳)
				(۵, ۳/۵)	(۱۰, ۳/۵)
				(۵, ۴)	(۵, ۴)
					(۵, ۴/۵)
					(۵, ۵)
$\hat{\lambda}$	۰/۴۹۳۳	۰/۴۳۵۹	۰/۳۸۳۹	۰/۳۴۴۴	۰/۳۴۴۵

مراجع

- [۱] م. اسدی، آشنایی با نظریه قابلیت اعتماد، مرکز نشر دانشگاهی، ۱۳۹۲.
- [2] Ascher, H. E. and Feingold, H. (1984), *Repairable System Reliability*, Dekker, NewYork.
- [3] Balakrishnan, N. and Ling, M. H. (2014), *Gamma lifetimes and one-shot device testing analysis*, Reliability Engineering and System Safety, Vol. 126, 54-64.
- [4] Balakrishnan, N., Ling, M. H. and So, H. Y. (2015), *Likelihood inference under proportional hazards model for one-shot devicetesting*, IEEE Transactions on Reliability, Vol. 65, 446-458.

- [5] Barlow, R. E. and Hunter, L. C. (1960), *Optimum preventive maintenance policies*, Journal of the Operations Research Society of America, Vol. 8, No. 1, 90-100.
- [6] Kitagawa, T., Yuge, T. and Yangi, S. (2014), *Optimum periodic inspection interval and replacement policies for a one-shot system with minimal repair*, Advanced Reliability and Maintenance Modeling VI, 247-254.
- [7] Kitagawa, T., Yuge, T. and Yangi, S. (2016), *Periodic and non-periodic inspection policies for a one-shot system with minimal repair*, Journal of Japan Industrial Management Association, Vol. 66, 387-395.
- [8] Nakagawa, T. and Kowada, M. (1983), *Analysis of a system with minimal repair and its application to replacement policy*, European Journal of Operational Research, Vol. 12, 176-182.
- [9] Wu, T. H., Chang, C. C., and Chung, S. H. (2008), *A simulated annealing algorithm for manufacturing cell formation problems*, Expert Systems with Applications, 34, 1609-1617.



مدل‌های شکنندگی و خواص توزیعی آن‌ها

حوتی، ف^۱ احمدی، ج^۲

^{۱،۲} گروه آمار، دانشکده ریاضی، دانشگاه فردوسی مشهد

چکیده

در مطالعات بقا، مدل‌های آماری متعددی معرفی شده‌اند که اکثر آن‌ها بر اساس فرض همگن بودن جامعه مورد استفاده قرار می‌گیرند. اما گاهی به دلایلی چون عوامل خطر مشاهده نشده یا ناشناخته، خطرات واحدهای آماری مورد مطالعه، متفاوت بوده و فرض همگنی جامعه برقرار نیست. در اینصورت استفاده از روش‌های دیگر نظیر مدل‌های شکنندگی می‌تواند مفید واقع شود. در حقیقت این مدل‌ها به منظور تبیین تغییرات ناشی از عوامل خطر مشاهده نشده یا ناشناخته به کار می‌روند. در این مقاله قصد داریم مروری بر انواع مدل‌های شکنندگی داشته و برخی خواص توزیعی آن‌ها را مطالعه کنیم. همچنین نوع وابستگی بین متغیرهای طول عمر و شکنندگی بر اساس هر یک از مدل‌ها بررسی می‌شوند. در نهایت مدل شکنندگی نرخ خطر معکوس متناسب را معرفی کرده و ویژگی‌های آن را بررسی خواهیم کرد.

کلمات کلیدی: مدل شکنندگی ضربی، مدل شکنندگی جمعی، مدل شکنندگی بخت متناسب، مدل شکنندگی نرخ خطر معکوس متناسب، مفاهیم وابستگی، بطور کامل مثبت از مرتبه ۲، منظم معکوس از مرتبه ۲.

۱ مقدمه

در تحلیل داده‌های بقا همواره دو هدف اصلی مدل‌بندی آن‌ها به منظور یافتن ترکیبی مناسب از متغیرهای کمکی که طول بقای اعضای مورد مطالعه را تحت تأثیر قرار می‌دهند و نیز یافتن برآوردی مناسب برای تابع مخاطره مدنظر بوده است. برای این منظور تاکنون مدل‌های بسیاری

^۱ hooti.fatemeh@gmail.com

^۲ ahmadi-j@um.ac.ir

برای برآزش به داده‌ها معرفی شده‌اند. به عنوان مثال از بین مدل‌های پارامتری و نیمه پارامتری می‌توان مدل نرخ خطر متناسب (کاکس)، مدل زمان شکست شتابیده، مدل نرخ خطر معکوس متناسب، مدل نرخ خطر جمعی، مدل بخت متناسب و مدل‌های شکنندگی را نام برد. برای مثال لی و لی (۲۰۰۸) خواص مدل‌های نرخ خطر معکوس آمیخته را مطالعه کردند. سانکاران و همکاران (۲۰۱۴) مدل‌های نرخ خطر جمعی را بررسی کردند و به برآورد نرخ خطر جامعه بر اساس آن پرداختند. اما اکثر این مدل‌ها بر اساس فرض همگنی جامعه مورد استفاده قرار می‌گیرند در حالی که گاهی بنا به دلایلی چون عوامل خطر مشاهده نشده یا ناشناخته، خطرات واحدهای مورد مطالعه، متفاوت بوده و فرض همگنی جامعه برقرار نیست. در اینصورت استفاده از روش‌های دیگر نظیر مدل‌های شکنندگی می‌تواند مفید باشد.

واپل و همکاران (۱۹۷۹) واژه شکنندگی را برای تحلیل داده‌های بقا معرفی نمودند که یک مدل اثر تصادفی برای داده‌های بقا می‌باشد. این مدل‌ها به طور گسترده در مطالعات بقا به منظور تبیین تغییرات ناشی از عوامل خطر مشاهده نشده به کار می‌روند. با استفاده از مدل‌های شکنندگی می‌توان ناهمگنی بین زمان‌های بقا در گروه‌های مختلف و همبستگی ایجاد شده بین زمان‌های بقای هر گروه را تعیین نمود. برای مثال در مطالعه طول عمر یک فرد، یک عامل ژنتیکی و یا نقص ذاتی و یا اثر یک دارو می‌توانند به عنوان اثرات پنهان در مطالعه مذکور بیان شوند. مدل‌هایی که در برگیرنده این اثرات پنهان هستند به مدل‌های شکنندگی معروف هستند. برای مثال در خصوص کارهای انجام شده در این زمینه می‌توان به ژو و لی (۲۰۰۸)، گوپتا و گوپتا (۲۰۰۹)، گوپتا و پنگ (۲۰۱۳)، سانکاران و گلیجا (۲۰۱۱) اشاره کرد. همچنین برای اطلاعات بیشتر می‌توان به کتاب‌های موجود در این زمینه از قبیل هنگال (۲۰۱۱) و وینک (۲۰۱۱) و لیست مراجع در آن‌ها مراجعه نمود. در این مقاله انواع مدل‌های شکنندگی بیان و برخی از خواص توزیعی آن‌ها ارائه می‌شود. نوع وابستگی بین متغیرهای طول عمر و شکنندگی بر اساس هر مدل را بررسی خواهیم کرد. همچنین مدل شکنندگی نرخ خطر معکوس متناسب را معرفی و به مطالعه خواص توزیعی و وابستگی بر اساس آن می‌پردازیم. برای این منظور از برخی مفاهیم و تعاریف وابستگی که در نلسن (۲۰۰۶) آمده است، استفاده خواهیم کرد.

۲ مدل‌های شکنندگی نرخ خطر ضربی، جمعی و بخت متناسب و خواص آن‌ها

مدل شکنندگی نرخ خطر متناسب

پرمکاربردترین مدل‌های شکنندگی معرفی شده، مدل شکنندگی نرخ خطر متناسب یا مدل شکنندگی ضربی می‌باشد. در این مدل متغیر شکنندگی به صورت ضرب در تابع نرخ خطر پایه، وارد مدل می‌شود. فرض کنید T متغیر تصادفی پیوسته و نشان دهنده طول عمر و V بیانگر متغیر شکنندگی باشد. مدل شکنندگی نرخ خطر متناسب به صورت

$$r(t|v) = vr \cdot (t), \quad t > 0, \quad (1)$$

در نظر گرفته می‌شود که در آن $r \cdot (t)$ تابع نرخ خطر پایه و مستقل از v است. بنابراین تابع بقای شرطی متغیر T به شرط V عبارت است از

$$\begin{aligned} \bar{F}(t|v) &= [\bar{F} \cdot (t)]^v \\ &= \exp \left(-v \int_0^t r \cdot (x) dx \right), \quad t > 0, \end{aligned} \quad (2)$$

که در آن $\bar{F} \cdot (t)$ تابع بقای پایه است. تساوی دوم در (۲) با توجه به این که تابع نرخ خطر، تابع بقای متغیر را از رابطه $\bar{F}(t) = \exp \left(- \int_0^t r(x) dx \right)$ به طور یکتا مشخص می‌کند، برقرار است. در اینصورت تابع چگالی شرطی متناظر با (۲) برابر است با

$$f(t|v) = vr \cdot (t) \exp \left(-v \int_0^t r \cdot (x) dx \right).$$

با استفاده از (۲) تابع بقای غیر شرطی (جامعه) متغیر T پس از تأثیر متغیر پنهان V به صورت زیر به دست می‌آید

$$\bar{F}(t) = M_V (\ln \bar{F} \cdot (t)), \quad t > 0,$$

که در آن $h(\cdot)$ تابع چگالی و $M_V(\cdot)$ تابع مولد گشتاور متغیر شکنندگی V هستند. همچنین تابع چگالی غیر شرطی T بر حسب $r_{\cdot}(t)$ به صورت زیر حاصل می شود

$$f(t) = r_{\cdot}(t) \int_{\cdot}^{\infty} u \exp\left(-u \int_{\cdot}^t r_{\cdot}(x) dx\right) h(u) du.$$

بنابراین تابع نرخ خطر آن به صورت

$$\begin{aligned} r(t) &= \int_{\cdot}^{\infty} u r_{\cdot}(t) \frac{(\bar{F}_{\cdot}(t))^u h(u)}{\int_{\cdot}^{\infty} (\bar{F}_{\cdot}(t))^u h(u) du} du \\ &= \int_{\cdot}^{\infty} r(t|u) h(u|T > t) du, \end{aligned} \quad (۳)$$

خواهد بود. دلیل تساوی آخر در (۷) این است که

$$\begin{aligned} h(v|T > t) &= -\frac{d}{dv} \frac{P(T > t, V > v)}{P(T > t)} \\ &= -\frac{d}{dv} \frac{\int_v^{\infty} \bar{F}(t|u) h(u) du}{\bar{F}(t)} \\ &= \frac{\bar{F}(t|v) h(v)}{\bar{F}(t)}. \end{aligned} \quad (۴)$$

گوپتا و گوپتا (۲۰۰۹) نشان دادند که اگر مدل شکنندگی نرخ خطر ضربی (۱) برقرار باشد. آنگاه الف) $r(t|v)$ نسبت به v صعودی است.

ب) متغیرهای تصادفی T و V ، وابستگی نسبت درستمایی منفی، $NLRD(T, V)$ دارند.

ج) متغیرهای تصادفی T و V ، در گوشه راست نزولی، $RCSD(T, V)$ هستند.

بنابراین نتیجه می شود $\bar{F}(t|V > v)$ تابعی نزولی از v ، یعنی $RTD(T|V)$ است و نیز $\bar{H}(v|T > t)$ تابعی نزولی از t ، یعنی $RTD(V|T)$ است. همچنین تحت مدل (۱) می توان نشان داد $E(T|V > v)$ تابعی نزولی از v و $E(V|T > t)$ تابعی نزولی از t هستند.

مدل شکنندگی نرخ خطر جمعی

یکی دیگر از مدل های شکنندگی که در آن متغیر شکنندگی به صورت جمع با تابع نرخ خطر پایه، وارد مدل می شود، مدل شکنندگی نرخ خطر جمعی نام دارد. این مدل به صورت

$$r(t|v) = r_{\cdot}(t) + v, \quad t > \cdot, \quad (۵)$$

در نظر گرفته می شود. توجه داریم که در این مدل $r(t|V = \cdot) = r_{\cdot}(t)$ ، برابر تابع نرخ خطر جامعه ای بدون هیچ گونه اثر شکنندگی است. بنابراین تابع بقای شرطی متغیر T به شرط V برابر

$$\bar{F}(t|v) = \bar{F}_{\cdot}(t) \exp(-vt), \quad t > \cdot, \quad (۶)$$

خواهد بود. تابع چگالی شرطی متناظر با (۶) عبارت است از

$$f(t|v) = f_{\cdot}(t) \exp(-vt) + \bar{F}_{\cdot}(t) v \exp(-vt), \quad t > \cdot.$$

بنابراین تابع بقای غیر شرطی متغیر T پس از تأثیر متغیر پنهان V به صورت زیر حاصل می شود

$$\bar{F}(t) = \bar{F}_{\cdot}(t) M_V(-t). \quad (۷)$$

در این صورت از (۷) تابع چگالی غیر شرطی به صورت زیر نتیجه می شود

$$f(t) = f_{\cdot}(t) \int_{\cdot}^{\infty} \exp(-ut) h(u) du + \bar{F}_{\cdot}(t) \int_{\cdot}^{\infty} u \exp(-ut) h(u) du. \quad (۸)$$

با توجه به روابط (۷) و (۳) و با استفاده از (۸)، تابع نرخ خطر غیر شرطی در این مدل به صورت

$$r(t) = r_*(t) + E(V|T > t),$$

خواهد بود.

گوپتا و گوپتا (۲۰۰۹) نشان دادند تحت مدل شکنندگی نرخ خطر جمعی (۵)،

الف) $r(t|v)$ نسبت به v صعودی است.

ب) اگر تابع نرخ خطر پایه صعودی باشد، متغیرهای تصادفی V و T ، وابستگی نسبت درستیابی منفی، $NLRD(T, V)$ دارند.

ج) متغیرهای تصادفی V و T ، در گوشه راست نزولی، $RCS D(T, V)$ هستند.

مدل شکنندگی بخت متناسب

مدل شکنندگی بخت متناسب بر اساس دو تابع بخت شرطی متغیرهای V و T و تابع بخت پایه تعریف می‌شود. در این مدل متغیر شکنندگی به صورت ضرب در تابع بخت پایه وارد مدل می‌شود. تابع بخت بر اساس هر یک از توابع توزیع شرطی و پایه به صورت زیر تعریف می‌شوند

$$\phi(t|v) = \frac{\bar{F}(t|v)}{1 - \bar{F}(t|v)}, \quad \phi_*(t) = \frac{\bar{F}_*(t)}{1 - \bar{F}_*(t)},$$

که در آن $\bar{F}(t|v)$ و $\bar{F}_*(t)$ توابع بقای شرطی و پایه هستند. بنابراین مدل شکنندگی بخت متناسب به صورت

$$\phi(t|v) = v\phi_*(t), \quad t > 0, \quad (9)$$

تعریف می‌شود. از رابطه (۹) تابع بقای شرطی بر حسب تابع بخت پایه عبارت است از

$$\bar{F}(t|v) = \frac{v\phi_*(t)}{1 + v\phi_*(t)}. \quad (10)$$

بنابراین تابع چگالی شرطی متناظر با (۱۲) برابر

$$f(t|v) = \frac{v\phi_*(t)r_*(t)}{F_*(t)[1 + v\phi_*(t)]^2}, \quad (11)$$

خواهد بود. در این صورت تابع نرخ خطر شرطی به صورت زیر حاصل می‌شود

$$r(t|v) = \frac{r_*(t)}{F_*(t)[1 + v\phi_*(t)]}. \quad (12)$$

از رابطه (۱۲) تابع بقای غیر شرطی متغیر T پس از تأثیر متغیر پنهان V به صورت زیر به دست می‌آید

$$\bar{F}(t) = \int_0^\infty \frac{u\phi_*(t)}{1 + u\phi_*(t)} h(u) du. \quad (13)$$

تابع چگالی متناظر با آن برابر

$$f(t) = \int_0^\infty \frac{u\phi_*(t)r_*(t)}{F_*(t)[1 + u\phi_*(t)]^2} h(u) du, \quad (14)$$

است. بنابراین تابع نرخ خطر غیر شرطی بر اساس روابط (۱۳) و (۱۴) به صورت زیر نتیجه می‌شود

$$r(t) = \int_0^\infty \frac{r_*(t)}{F_*(t)[1 + u\phi_*(t)]} \left(\frac{uh(u)/[1 + u\phi_*(t)]}{\int_0^\infty \{uh(u)/[1 + u\phi_*(t)]\} du} \right) du. \quad (15)$$

از طرفی داریم

$$\begin{aligned} r(t) &= \int_{\cdot}^{\infty} \frac{f(t, u)}{\bar{F}(t)} du \\ &= \int_{\cdot}^{\infty} \frac{f(t|u)}{\bar{F}(t|u)} \frac{h(u)P(T > t|u)}{\bar{F}(t)} du \\ &= \int_{\cdot}^{\infty} r(t|u)h(u|T > t) du, \end{aligned} \quad (16)$$

که تساوی آخر با توجه به (۸) نتیجه شده است. بنابراین از مقایسه (۱۵) و (۱۶) تابع چگالی شرطی متغیر $V|T > t$ به صورت زیر خواهد بود

$$h(v|T > t) = \frac{vh(v)/[1 + v\phi_{\cdot}(t)]}{\int_{\cdot}^{\infty} \{uh(u)/[1 + u\phi_{\cdot}(t)]\} du}.$$

گوپتا و پنگ (۲۰۱۳) نشان دادند اگر مدل شکنندگی بخت متناسب (۹) برقرار باشد. آنگاه

الف) $r(t|v)$ نسبت به v نزولی است.

ب) متغیرهای تصادفی T و V ، وابستگی نسبت درستنمایی مثبت، $PLRD(T, V)$ دارند.

ج) متغیرهای تصادفی T و V ، در گوشه چپ نزولی، $LCSD(T, V)$ هستند.

د) متغیرهای تصادفی T و V ، در گوشه راست صعودی، $RCSI(T, V)$ هستند.

۳ مدل شکنندگی نرخ خطر معکوس متناسب

در این بخش ابتدا به معرفی مدل شکنندگی نرخ خطر معکوس متناسب می‌پردازیم، که این مدل می‌تواند به منظور برازش به داده‌هایی با سانسور چپ مورد استفاده قرار گیرد. سپس نوع وابستگی بین متغیر طول عمر و شکنندگی تحت این مدل بررسی می‌شود. در این مدل متغیر شکنندگی به صورت ضرب در تابع نرخ خطر معکوس پایه وارد مدل می‌شود. مدل شکنندگی نرخ خطر معکوس متناسب به صورت

$$r^*(t|v) = vr^*(t), \quad t > 0, \quad (17)$$

خواهد بود، که در آن $r^*(t)$ تابع نرخ خطر معکوس پایه و مستقل از v است. با استفاده از (۱۷) تابع توزیع شرطی متغیر T به شرط V به صورت زیر حاصل می‌شود

$$\begin{aligned} F(t|v) &= [F_{\cdot}(t)]^v \\ &= \exp\left(-v \int_t^{\infty} r^*(x) dx\right), \quad t > 0, \end{aligned} \quad (18)$$

که در آن $F_{\cdot}(t)$ تابع توزیع پایه است. تساوی دوم در (۱۸) با توجه به این که تابع نرخ خطر معکوس، تابع توزیع متغیر را از رابطه $F(t) = \exp\left(-\int_t^{\infty} r^*(x) dx\right)$ ، به طور یکتا مشخص می‌کند، برقرار است. در اینصورت تابع چگالی شرطی متناظر با (۱۸) برابر است با

$$f(t|v) = vr^*(t) \exp\left(-v \int_t^{\infty} r^*(x) dx\right). \quad (19)$$

با استفاده از (۱۸) تابع توزیع غیر شرطی (جامعه) متغیر T پس از تأثیر متغیر پنهان V به صورت زیر به دست می‌آید

$$\begin{aligned} F(t) &= \int_{\cdot}^{\infty} F(t|u)h(u) du \\ &= M_V(\ln F_{\cdot}(t)), \quad t > 0. \end{aligned} \quad (20)$$

تابع چگالی غیر شرطی T بر حسب $r^*(t)$ به صورت زیر حاصل می‌شود

$$\begin{aligned} f(t) &= \int_{\cdot}^{\infty} f(t|u)h(u) du \\ &= r^*(t) \int_{\cdot}^{\infty} u \exp\left(-u \int_t^{\infty} r^*(x) dx\right) h(u) du. \end{aligned} \quad (21)$$

بنابراین تابع نرخ خطر معکوس آن به صورت زیر نتیجه می‌شود

$$\begin{aligned} r^*(t) &= \int_{\cdot}^{\infty} ur^*(t) \frac{[F_{\cdot}(t)]^u h(u)}{\int_{\cdot}^{\infty} [F_{\cdot}(t)]^u h(u) du} du \\ &= \int_{\cdot}^{\infty} r^*(t|u) h(u|T \leq t) du. \end{aligned} \quad (22)$$

دلیل برقراری آخرین تساوی در (۲۲) این است که

$$\begin{aligned} h(v|T \leq t) &= \frac{d}{dv} \frac{\int_{\cdot}^v F(t|u) h(u) du}{F(t)} \\ &= \frac{F(t|v) h(v)}{F(t)} \\ &= \frac{\exp(-v \int_t^{\infty} r^*(x) dx) h(v)}{\int_{\cdot}^{\infty} \exp(-u \int_t^{\infty} r^*(x) dx) h(u) du}. \end{aligned} \quad (23)$$

قضیه ۱.۳. فرض کنید مدل شکنندگی نرخ خطر معکوس متناسب (۱۷)، برقرار باشد. در اینصورت الف) $r^*(t|v)$ تابعی صعودی از v است.

ب) متغیرهای تصادفی T و V ، وابستگی نسبت درستنمایی مثبت، $PLRD(T, V)$ دارند.

ج) متغیرهای تصادفی T و V در گوشه چپ نزولی، $LCSD(T, V)$ هستند.

برهان. برهان قسمت الف) با توجه به رابطه (۱۷) واضح است. برای اثبات قسمت ب) بنا به تعریف این نوع وابستگی، کافی است نشان دهیم تابع چگالی توأم $TP_2(T, V)$ است. بنا به (۱۹) برای هر $v_1 \leq v_2$ داریم

$$\frac{f(t|v_2)}{f(t|v_1)} = \frac{v_2}{v_1} \exp \left\{ - (v_2 - v_1) \int_t^{\infty} r^*(x) dx \right\},$$

که ملاحظه می‌شود تابعی صعودی از t است پس TP_2 می‌باشد. به منظور اثبات درستی قسمت ج) می‌توان نشان داد تابع توزیع توأم (T, V) ، $F(t, v)$ تابعی TP_2 است. برای هر $v_1 \leq v_2$ از (۱۸) نتیجه می‌شود

$$\frac{F(t|v_2)}{F(t|v_1)} = \exp \left\{ - (v_2 - v_1) \int_t^{\infty} r^*(x) dx \right\},$$

که به وضوح تابعی صعودی از t است. $\square$

نتیجه ۱.۳. از قضیه ۱.۳ نتیجه می‌شود که $F(t|V \leq v)$ تابعی نزولی از v ، یعنی $LTD(T|V)$ ، است و نیز $H(v|T \leq t)$ تابعی نزولی از t ، یعنی $LTD(V|T)$ ، است.

نتیجه ۲.۳. می‌توان نشان داد که $E(V|T \leq t)$ تابعی صعودی از t و $E(T|V \leq v)$ تابعی صعودی از v هستند.

مراجع

- [1] Gupta, R. C. and Gupta, R. D. (2009), *General frailty model and stochastic orderings*, Journal of Statistical Planning and Inference, 139, 3277–3287.
- [2] Gupta, R. C. and Peng, C. (2013), *Proportional odds frailty model and stochastic comparisons*, The Institute of Statistical Mathematics, 66, 897–912.
- [3] Hanagal, D. D. (2011), *Modeling survival data using frailty models*, Chapman and Hall/ CRC, New York.

- [4] Li, X. and Li, Z. (2008), *mixture model of proportional reversed hazard rate*, Communications in Statistics-Theory and Methods, 37, 2953–2963.
- [5] Nelsen, R. B. (2006), *An introduction to copulas*, Springer, New York.
- [6] Sankaran, P. G., Asokan, M. V. and Anjana, S. (2014), *Additive reversed hazard rates models*, American Journal of Mathematical and Management Sciences, 33, 315–329.
- [7] Sankaran, P. G. and Gleeja, V. L. (2011), *On proportional reversed hazard frailty models*, Metron, LXIX, 151–173.
- [8] Vaupel, J. W., Manton, K. G. and Stanllard, E. (1979), *The impact of heterogeneity on individual frailty on the dynamic of mortality*, Demography, 16, 439–454.
- [9] Wienke, A. (2011), *Frailty models in survival analysis*, Chapman and Hall/ CRC, New York.
- [10] Xu, M. and Li, X. (2008), *Negative dependence in frailty models*, Journal of Statistical Planning and Inference, 138, 1433–1441.



بررسی عوامل موثر بر زمان شکست ایمپلنت با استفاده از دو مدل رگرسیون خطرات متناسب کاکس و مدل بقای چندسطحی با رویکرد بیزی

رحیمی فروشانی، ع^۱ اکبری شارک، ن^۲ رازی فرد، م. ج^۳

^{۱،۲} گروه اپیدمیولوژی و آمار زیستی، دانشگاه علوم پزشکی و خدمات بهداشتی و درمانی تهران
^۳ مرکز تحقیقات دندانپزشکی، دانشگاه علوم پزشکی و خدمات بهداشتی و درمانی تهران

چکیده

در مطالعات دندانپزشکی یک فرد معمولاً چندین مراجعه دارد از این رو داده‌های زمان شکست خوشه‌ای گروه‌بندی یا گسسته جمع‌آوری می‌شود که دارای ساختار همبستگی می‌باشند. در روش‌های معمول این همبستگی نادیده گرفته می‌شود که این باعث کم‌برآوردی تغییرپذیری اثرات متغیرهای کمکی یا کم‌برآوردی احتمالات بقا می‌شود. هدف این مطالعه شناسایی عوامل موثر بر شکست ایمپلنت بدون در نظر گرفتن ساختار همبستگی و استفاده از رگرسیون خطرات متناسب کاکس و با در نظر گرفتن ساختار همبستگی و استفاده از مدل بقای چندسطحی (خطرات متناسب گروه‌بندی) با رویکرد بیزی و مقایسه نتایج آن‌ها می‌باشد.

کلمات کلیدی: داده خوشه‌ای، همبستگی درون خوشه‌ای، مدل بقای خطرات متناسب گروه‌بندی، رگرسیون خطرات متناسب کاکس.

^۱ rahimifo@tums.ac.ir

^۲ noooshin_akbari@yahoo.com

^۳ mj_khf@yahoo.com

۱ مقدمه

داده‌های مربوط به زمان رخداد برای فرآیندهای زمان گسسته، متغیر وابسته را به صورت یک سری از نتایج دودویی ثبت می‌کند که نشان دهنده این است که آیا رویداد در زمان مشاهده اتفاق افتاده است یا نه. در واقع داده‌های زمان گسسته بسیار شبیه داده‌های مقطعی سری زمانی متغیر وابسته دودویی می‌باشند. فرض کنید در یک کارآزمایی بالینی تصادفی متغیر مورد علاقه زمان وقوع یک رویداد خاص در حضور یا عدم حضور یک درمان جدید می‌باشد و داده‌ها از افراد همبسته از خوشه‌های مستقل جمع‌آوری می‌شوند با این حال گاهی بررسی و نظارت مداوم افراد مورد مطالعه گران قیمت و یا حتی غیر ممکن می‌باشد، بنابراین افراد بصورت دوره‌ای طبق یک برنامه منظم از پیش برنامه ریزی شده مورد بررسی قرار می‌گیرند از این رو داده‌های زمان شکست خوشه‌ای گروه بندی یا گسسته جمع‌آوری می‌شوند [۱]. مدل بقا گروه بندی یا گسسته به راحتی به مدل‌های چند سطحی بسط داده می‌شوند.

نقش و همکاران (۱۳۹۲)، در مطالعه‌ای با عنوان “کاربرد تکنیک‌های مدل‌یابی چند سطحی در تحلیل داده‌های تیمز و مقایسه آن با تحلیل یک سطحی” به اهمیت کاربرد مدل‌های چند سطحی پرداخته‌اند. در این مقاله تکنیک مدل‌یابی چند سطحی با استفاده از نرم افزار HLM برای تحلیل داده‌های IEA با توجه به ماهیت داده‌ها استفاده شده است چون در داده‌های IEA دانش آموزان در درون کلاس‌ها، کلاس‌ها در درون مدارس و مدارس در درون کشورها آشیانه شده‌اند. نتایج متفاوت حاصل از دو تحلیل تک سطحی و چند سطحی حاکی از اهمیت کاربرد تحلیل‌های چند سطحی برای داده‌های آشیانه‌ای از جمله داده‌های تیمز دارد [۲].

وونگ و همکاران (۲۰۰۶) مطالعه با عنوان “مدل‌سازی چند سطحی داده‌های بقای خوشه‌ای با مدل رگرسیون کاکس بر روی داده‌های ترمیم ART دندان” انجام داده‌اند. در مجموع ۲۸۴ تا ART توسط ۵ دندانپزشک قرار داده شده و وضعیت بالینی ترمیم‌های ART سالانه به مدت ۶ سال مورد بررسی قرار گرفته و زمان‌های شکست مربوط به ترمیم‌ها در گروه‌های خوشه بندی ثبت شده است در این‌جا از دو مدل خطرات متناسب چند سطحی و مدل رگرسیون لجستیک چند سطحی استفاده شده است و ضریب همبستگی تقریبی درون خوشه‌ای برای لگاریتم زمان‌های شکست بدست آمده است [۳].

گیلثورف و همکاران (۲۰۰۲) مطالعه‌ای با عنوان “تجزیه و تحلیل بقا چند سطحی از ترمیم‌های آمالگام در میان پرسنل RAF” انجام داده‌اند. استفاده از مدل چند سطحی برای تجزیه و تحلیل داده‌های بقا یک راه حل مناسب و قدرتمند برای رفع مشکل عدم استقلال در میان ترمیم‌های دندان‌های فراهم می‌کند. متغیر پاسخ مدت زمان بقای ترمیم آمالگام بود و در نهایت به مقایسه مدل‌های چند سطحی و تک سطحی می‌پردازند. با توجه به نتایج بین مدل‌های چند سطحی و تک سطحی اختلاف وجود دارد [۴].

در مطالعه حاضر مدل بقای چند سطحی با رویکرد بیزی و مدل رگرسیون خطرات متناسب کاکس بر روی داده‌های ایمپلنت دندان‌های مورد بررسی قرار گرفته‌اند. هدف از مطالعه شناسایی عوامل موثر بر شکست ایمپلنت‌های دندان‌های و مقایسه نتایج دو مدل چند سطحی و تک سطحی می‌باشد.

۲ مواد و روش‌ها

تعداد ۱۰۱۱ نفر برای درمان ایمپلنت دندان‌های طی سال‌های ۱۳۸۱ تا ۱۳۹۲ به کلینیک دندانپزشکی دانشگاه علوم پزشکی تهران مراجعه کرده‌اند که در مجموع ۲۳۶۸ ایمپلنت برای این بیماران قرار داده شده است. اطلاعات شخصی از طریق پرسش از خود بیماران و اطلاعات درمانی و مشخصات مربوط به ایمپلنت توسط دندان پزشک مربوطه ثبت شد. وضعیت ایمپلنت‌های دندان‌های طی سه دوره ارزیابی شده است. که سال‌های ۸۱ تا ۸۴ بازه زمانی اول، ۸۵ تا ۸۸ بازه زمانی دوم و ۸۹ تا ۹۲ بازه زمانی سوم می‌باشد. وضعیت ایمپلنت براساس معیارهای بوزر که عدم وجود شکایت‌های ماندگار مانند درد یا احساس جسم خارجی یا دیستری، عدم وجود عفونت در کنار ایمپلنت همراه با ترشح چرک، عدم وجود لقی، عدم وجود رادیولوژی ماندگار در اطراف ایمپلنت می‌باشد، مشخص گردید. اگر ایمپلنتی دچار شکست شود کد یک و اگر دچار شکست نشود کد صفر می‌گیرد. اگر در یک بازه زمانی شکست رخ دهد کد یک برای آن بازه و بازه زمانی بعدی ثبت می‌شود و اگر شکست برای ایمپلنتی رخ نداده باشد در هر سه بازه زمانی کد صفر می‌گیرد. متغیرهای جنسیت، سن، وضعیت سیگاری بودن، بیماری سیستمیک، اووردنچر،

نوع دندان، مکان دندان، شکل مقطع ایمپلنت، مشکل حین انجام جراحی، بایومترال، بافت سطح ایمپلنت، قطر و طول ایمپلنت، سنوس لیفت، جراحی جا به جایی عصب، زمان قرارگیری ایمپلنت و بارگیری ایمپلنت از جمله عواملی می‌باشند که نتیجه ترمیم ایمپلنت را تحت تاثیر قرار می‌دهند و در مطالعه ما مورد بررسی قرار گرفتند.

۳ مدل رگرسیون خطرات متناسب کاکس

رگرسیون کاکس یک مدل نیمه پارامتریک را برای برازش تابع مخاطره فرض میکند که می‌توان متغیرهای توضیحی یا ریسک فاکتورها را به مدل آن اضافه کرد ولی هم چنان تابع مخاطره پایه را به عنوان تابعی نامشخص ولی غیرمنفی از زمان ثابت نگاه داشت. تابع مخاطره کاکس برای متغیر توضیحی x تابعی به صورت زیر است:

$$\lambda(t; x) = \lambda_0(t) \exp(x'\beta)$$

براساس مدل فوق، مخاطره پایه $\lambda_0(t)$ تابع مخاطره برای فردی است که مقدار متغیر توضیحی‌اش برابر صفر است. چون فرض نشده که این تابع مخاطره پایه فرمی پارامتریک دارد، مدل کاکس را یک مدل نیمه پارامتریک برای تابع مخاطره گویند. تابع بقای متناسب با این مدل به صورت زیر خواهد بود:

$$s(t; x) = \exp \left[-\exp(x'\beta) \int_0^t \lambda_0(u) du \right]$$

این انتگرال، تابع تجمعی مخاطره پایه نامیده میشود و روش‌های گوناگونی برای برآورد این تابع وجود دارد. مدل کاکس کاربردی‌ترین روش برای یافتن ارتباط متغیرهای توضیحی با متغیر پاسخ بقا یا هر متغیر پاسخ دیگریست که از راست سانسور شده باشد. اگرچه این مدل محدودیتهایی نیز دارد. یکی از محدودیتهای این مدل فرض خطرات متناسب است. به این معنی که میزان مخاطره بین دو یا بیش از دو گروه از متغیر توضیحی بایستی در طول زمان بقا ثابت بماند که فرض، متناسب با تابع تجمعی پایه تعریف شده است [۵].

۴ مدل بقای خطرات متناسب گروه‌بندی با رویکرد بیزی

فرض کنید T_{ij} زمان شکست مربوط به j امین ایمپلنت از i امین فرد باشد ($i = 1, 2, \dots, n_i; j = 1, 2, \dots, n_j$) که در مقیاس پیوسته اندازه گیری شده x_{ij} یک بردار p بعدی از متغیرهای کمی مشاهده شده باشد که با T_{ij} در ارتباط است. مشروط بر اثر تصادفی e_i ، T_{ij} ها متقابلاً مستقل هستند. فرض کنید محور زمان به d فاصله زمانی مستقل $I_1, I_2, \dots, I_d$ تقسیم‌بندی شود که $I_s = (a_{s-1}, a_s]$ که $a_0 = 0$ و $a_d = \infty$. همچنین فرض کنید $s_{ij} = s$ زمان رخداد مربوط به j امین ایمپلنت از i امین فرد باشد که در بازه I_s دچار شکست شده است (یعنی $a_{s-1} < T_{ij} \leq a_s$). تابع خطر شرطی گسسته، $\lambda(s|x_{ij}, e_i)$ به صورت زیر تعریف می‌شود:

$$\lambda(s | x_{ij}, e_i) = P(S_{ij} = s | S_{ij} \geq s, x_{ij}, e_i) \quad (1)$$

برای مدل‌بندی تابع خطر گسسته (۱) از رگرسیون باینری با شکل کلی $\lambda(s | x, e) = g(\eta)$ استفاده کنیم و $\eta = \gamma_s + x^T \beta + e$ که در آن g تبدیلی است که کل خط واقعی $(-\infty, +\infty)$ را بر روی فاصله $(0, 1)$ رسم میکند، γ_s ($s = 1, \dots, d$) دنباله‌ای از پارامترهای پایه است که نشان می‌دهد هر فاصله دارای عرض از مبدا خاص خود می‌باشد و بردار پارامترهای رگرسیون می‌باشد. برای مدل بندی تابع g از مدل رگرسیون خطرات متناسب گروه‌بندی استفاده می‌کنیم:

$$g(\eta) = 1 - \exp[-\exp(\eta)]$$

بردار متغیرهای کمی می‌تواند با فرض این‌که x_{ij} در هر فاصله زمانی ثابت است، وابسته به زمان باشد. برای تعیین کامل مدل، توزیع اثرات تصادفی که ساختار وابستگی داده‌ها را مشخص می‌کند مورد نیاز است بنابراین فرض می‌شود که اثر تصادفی e_i در طول زمان ثابت است و دارای

توزیع نرمال با میانگین صفر و واریانس σ^2 می‌باشد. بطور کلی، انتخاب توزیع اثر تصادفی e_i تاثیر زیادی بر روی برآورد ندارد [۶]. از آنجا که σ^2 قدرت ارتباط بین افراد درون یک خوشه را مشخص می‌کند، بطور معمول پارامتر وابستگی نامیده می‌شود.

فرض کنید زمان‌های شکست T_{ij} بعد از فاصله I_{sij}^* دچار شکست شده یا سانسور راست برای آن رخ داده باشد. نشانگر شکست y_{ijs} را به صورت زیر تعریف می‌کنیم:

$$y_{ijs} = \begin{cases} 1 & \text{if subject } j \text{ of cluster } i \text{ fails in } I_s^* \\ 0 & \text{if subject } j \text{ of cluster } i \text{ survives through } I_s^* \end{cases}$$

بنابراین تابع درستنمایی کل داده‌ها L به صورت زیر می‌باشد

$$L = \prod_{i=1}^n \prod_{j=1}^{n_i} \prod_{s=1}^{s_{ij}^*} [g(\eta_{ijs})^{y_{ijs}} \{1 - g(\eta_{ijs})\}^{(1-y_{ijs})}]$$

براساس درستنمایی داده‌های مشاهده شده و با استفاده از رویکرد بیزی و تحمیل توزیع پیشین ناآگاهی دهنده به هریک از متغیرهای مورد نظر به برآورد پارامترها می‌پردازیم. متغیرهای $\gamma_1, \dots, \gamma_d, \beta_1, \dots, \beta_p$ بطور مستقل دارای توزیع $N(0, \varphi_1)$ می‌باشند و $1/\sigma^2$ دارای توزیع گاما $1/\varphi_1$ و $1/\varphi_2$ می‌باشد که φ_1 و φ_2 در عمل بسیار بزرگ است، مثلاً ۱۰۰۰.

تحت فرض نمایی برای T ، لگاریتم ضریب همبستگی زمان‌های شکست از یک خوشه یکسان، مستقل از متغیرهای کمکی است و به صورت زیر می‌باشد [۶]:

$$\text{corr}(\log(T_{ij}), \log(T_{il})) = \frac{\widehat{\sigma^2}}{\widehat{\sigma^2} + \pi^2/3} \quad j \neq l$$

که در آن $\widehat{\sigma^2}$ واریانس اثر تصادفی e_i می‌باشد.

نتایج تجربی نشان داد که این ضرایب همبستگی وقتی که زمان‌های شکست از مدل خطرات متناسب اثرات تصادفی پیروی می‌کنند، تقریب خوبی می‌باشند و نسبت به عدم تشخیص ساختار وابستگی اثرات تصادفی مقاوم هستند [۷، ۱].

۵ یافته‌ها

برای انجام مدل خطرات متناسب کاکس ابتدا باید متغیرهایی را که از نظر آماری معنی‌دار می‌باشند شناسایی و وارد مدل کنیم. برای این منظور ابتدا از آزمون Chi-square یا Fisher's exact test سپس از آزمون Log-Rank استفاده می‌کنیم. در آزمون Chi-square متغیرهایی معنی‌دار می‌باشند که p -مقدار آن‌ها کمتر از ۰/۱ باشد. با توجه به این آزمون متغیرهای جنس، سن، شکل مقطع ایمپلنت، وجود مشکل حین جراحی، مکان ایمپلنت، نوع دندان ایمپلنت شده، بایومترال، قطر ایمپلنت و میزان خشونت سطحی بر روی شکست ایمپلنت موثر می‌باشند. در آزمون لگ رنگ معنی‌داری متغیرها با مقایسه p -مقدار آن‌ها با ۰/۰۵ بررسی می‌شود که بر این اساس متغیرهای جنس، وجود مشکل حین جراحی، مکان ایمپلنت، نوع دندان و بایومترال معنی‌دار شده‌اند ($p\text{-value} < 0.05$).

متغیرهای معنی‌دار در آزمون‌های آماری را برای آنالیز بیشتر وارد مدل آماری می‌کنیم. جدول ۲ نتایج حاصل از برازش مدل رگرسیون خطرات متناسب کاکس را نشان می‌دهد که با توجه به مقدار p -آزمون آماری والد، متغیرهای مکان ایمپلنت ($p = 0.049$) و نوع دندان ایمپلنت شده ($p = 0.011$) در مدل کاکس موثر می‌باشند.

پس از برازش مدل کاکس، فرض متناسب بودن متغیرهای موجود در مدل را از روش افزودن متغیر وابسته به زمان بررسی می‌کنیم. نتایج آزمون مربوط به آن در جدول ۲ گزارش شده که نشان می‌دهد فرض تناسب برای همه متغیرها برقرار می‌باشد.

در مدل خطرات متناسب گروه‌بندی، یک مدل دو سطحی که ایمپلنت یا دندان ایمپلنت شده به عنوان سطح اول و فرد به عنوان سطح دوم در نظر گرفته می‌شود به داده‌ها برازش داده شده است. متغیر جنسیت ($x_1 = 0$ مرد، $x_1 = 1$ زن)، وجود مشکل حین جراحی ($x_2 = 0$ مشکل وجود دارد، $x_2 = 1$ مشکل وجود ندارد)، بایومترال ($x_3 = 0$ مواد شیمیایی مصرف شده است، $x_3 = 1$ مواد شیمیایی مصرف نشده است)، مکان

جدول ۱: تجزیه و تحلیل چندمتغیره عوامل موثر بر شکست ایمپلنت با استفاده از مدل کاکس

متغیر			رگرسیون کاکس
نسبت خطر ^۱	p-مقدار	(۹۵ % CI)	
جنس (زن در مقابل مرد)	۲/۵۵	۰/۶۶	(۰/۹۴۰ - ۶/۹۴۶)
وجود مشکل حین جراحی (نداشتن در مقابل داشتن)	۲/۳۲	۰/۲۳۸	(۰/۵۷۳ - ۹/۴۲۸)
مکان ایمپلنت (فک پایین در مقابل فک بالا)	۰/۲۹۲	۰/۴۹	(۰/۰۸۶ - ۰/۹۹۶)
بایومترال (نداشتن در مقابل داشتن)	۲/۵۶۳	۰/۱۶۱	(۰/۶۸۷ - ۹/۵۶۹)
نوع دندان ایمپلنت شده	-	۰/۱۲	-
آسیاب در مقابل پیش	۰/۱۴	۰/۰۷	(۰/۰۱۴ - ۴/۹۸)
آسیاب کوچک در مقابل پیش	۰/۲۴	۰/۰۴	(۰/۰۲۴ - ۴/۹۹)
نیش در مقابل پیش	۰/۲۶	۰/۰۹	(۰/۰۲۶ - ۵/۹۳)
شیری در مقابل پیش	۰/۰۳۷	۰/۰۳	(۰/۰۰۳ - ۲/۸۴)

جدول ۲: نتایج آزمون بررسی متناسب بودن متغیرهای کمکی در مدل کاکس

متغیر	p-مقدار
جنس	۰/۷۶۲
وجود مشکل حین جراحی	۰/۳۷۳
مکان دندان	۰/۹۷۲
بایومترال	۰/۱۰۰
نوع دندان	۰/۲۲۵

دندان ($x_4 = 0$ فک بالا، $x_4 = 1$ فک پایین) و نوع دندان ($x_5 = 0$ ، $x_6 = 0$ ، $x_7 = 0$ ، $x_8 = 1$ دندان آسیاب، $x_5 = 0$ ، $x_6 = 1$ ، $x_7 = 0$ ، $x_8 = 0$ دندان آسیاب کوچک، $x_5 = 0$ ، $x_6 = 1$ ، $x_7 = 0$ ، $x_8 = 0$ دندان نیش، $x_5 = 0$ ، $x_6 = 0$ ، $x_7 = 0$ ، $x_8 = 1$ دندان شیری) به عنوان متغیرهای کمکی وارد مدل می‌شوند. در این مدل برآوردها با روش MCMC و در نرم افزار openBUGS انجام می‌شود. در برآورد پارامترها ۵۰۰۰ شبیه‌سازی اول کنار گذاشته شد و مدل براساس ۱۰۰۰۰۰۰ شبیه‌سازی دیگر برازش داده شد که نتایج برآورد در جدول ۳ گزارش شده است.

نتایج مدل خطرات متناسب گروه‌بندی نشان می‌دهد که اثر خوشه‌بندی در میان لگاریتم زمان‌های شکست ایمپلنت‌های دندانی در افراد یکسان بسیار زیاد است ($\text{corr.child} = 0/99$). متغیرهای وجود مشکل حین عمل جراحی و بایومترال هیچ اثری بر روی زمان شکست ایمپلنت ندارند و هیچ اختلافی بین زمان‌های شکست مربوط به دندان‌های آسیاب، آسیاب کوچک، نیش و شیری با دندان پیش وجود ندارد چون فاصله اطمینان ۹۵ درصدی مربوط به آن‌ها صفر را در بر می‌گیرد. فاصله اطمینان مربوط به متغیرهای جنسیت و مکان دندان صفر را در بر نمی‌گیرد بنابراین این متغیرها در مدل معنی دار می‌باشند. برآورد پارامترهای پایه (γ_1 ، γ_2 ، γ_3) در حال افزایش می‌باشند که نشان دهنده افزایش نرخ خطر از فاصله ۱ تا ۳ می‌باشد. خطای شبیه‌سازی مربوط به همه‌ی پارامترها کمتر از ۰/۰۵ انحراف استاندارد می‌باشد که این نشان‌دهنده همگرایی

جدول ۳: تجزیه و تحلیل عوامل موثر بر شکست ایمپلنت با استفاده از مدل خطرات متناسب گروه‌بندی

مدل خطرات متناسب گروه‌بندی			متغیرها	
خطای شبیه‌سازی	انحراف استاندارد برآورد	فاصله اطمینان	میان	جنس
۰/۱۹۷۲	۶/۴۷۲	(-۲۶/۲۵ , -۱/۲۹۶)	۱۰/۲۲-	جنس
۰/۲۱۸۷	۷/۰۳۹	(-۲۶/۴۶ , ۱/۶۶۱)	۱۰/۴۴-	وجود مشکل حین عمل جراحی
۰/۱۹۷	۶/۴۷۷	(-۲۲/۰۳ , ۳/۸۷۶)	۸/۳۳۷-	بایومترال
۰/۱۷۵۷	۶/۱۱۱	(-۲۹/۴۹ , -۵/۱۸)	-۱۳/۷۵	مکان دندان
۰/۱۶۶۴	۵/۶۵۳	(-۲۱/۰۳ , ۱/۲۵۱)	-۷/۵۴	دندان آسیاب در مقابل پیش
۰/۱۱۹۶	۴/۳۳۳	(-۱۲/۲۴ , ۵/۲۸۸)	-۲/۱۱۳	دندان آسیاب کوچک در مقابل پیش
۰/۳۲۴	۱۱/۰۱	(-۴۱/۸۲ , ۱/۹۵۹)	-۱۴/۱۴	دندان نیش در مقابل پیش
۰/۲۶۶۹	۸/۶۳۹	(-۱۴/۱۲ , ۲۱/۰۱)	۲/۷۷۳	دندان شیری در مقابل پیش
۰/۲۶۶۲	۱۶/۲۲	(-۹۲/۰ , -۲۸/۹۴)	-۵۳/۲۳	بازه زمانی اول
۰/۲۳۶۸	۷/۵۶۶	(-۴۵/۰۲ , -۱۴/۲۷)	-۲۶/۸۳	بازه زمانی دوم
۰/۲۳۳۶	۷/۴۶	(-۴۳/۴۹ , -۱۳/۲)	-۲۵/۵	بازه زمانی سوم
۱۰/۴۱	۳۳۸/۰	(۱۱۵/۷ , ۱۴۳۲/۰)	۴۶۷/۰	واریانس افراد
۲/۰۸۶E - ۴	۰/۰۰۶	(۰/۹۷۲۴ , ۰/۹۹۷۷)	۰/۹۹۳	همبستگی افراد

است.

۶ بحث و نتیجه‌گیری

در مدل رگرسیون کاکس دو متغیر مکان دندان و نوع دندان رابطه معنی‌داری با شکست ایمپلنت دارند طوری که بقای ایمپلنت‌های دندانی در فک پایین بیشتر از فک بالا می‌باشد و خطر شکست در دندان‌های پیش نسبت به سایر دندان‌ها (آسیاب، آسیاب کوچک، نیش و شیری) بیشتر می‌باشد. سه متغیر جنس، وجود مشکل حین جراحی و بایومترال معنی‌دار نشده‌اند با این حال نتایج مربوط به آن‌ها نشان می‌دهد که مخاطره در خانم‌ها بیشتر است و افرادی که در حین جراحی دچار مشکل نشده‌اند و بایومترال ندارند بقای کمتری نسبت به گروه دیگر دارند. در مدل خطرات متناسب گروه‌بندی متغیر جنس و مکان دندان رابطه معنی‌داری با شکست ایمپلنت دارند طوری که بقا در خانم‌ها بیشتر از مردها می‌باشد و خطر شکست در فک پایین کمتر از فک بالا است. در رابطه با متغیرهای وجود مشکل حین جراحی و بایومترال، افرادی که دچار مشکل نشده‌اند و بایومترال ندارند بقای بیشتری نسبت به گروه‌های دیگر دارند و مخاطره در همه‌ی دندان‌ها نسبت به دندان پیش کمتر است بجز دندان شیری که مخاطره آن بالاتر می‌باشد. از آنجا که همبستگی میان لگاریتم زمان‌های شکست ایمپلنت‌های دندانی در افراد یکسان بسیار زیاد است، بنابراین تحلیل چند سطحی به بیان دقیق‌تر پدیده مورد مطالعه می‌پردازد و استفاده از آن به منظور به دست آوردن اطلاعات دقیق‌تر در مورد عوامل موثر بر شکست ایمپلنت توصیه می‌شود.

مراجع

- [1] Joop J. Hox, Multilevel Analysis Techniques and Applications, Second Edition, 2010.
- [2] Wong C.M, Lam K.F, Lo C.M. Multilevel modelling of clustered grouped survival data using Cox regression model: an application to ART dental restorations. Statistics In Medicine. 2006; 25:447-457.
- [3] Gilthorpe MS, Mayhew MT, Bulman JS. Multilevel survival analysis of amalgam restorations amongst RAF personnel. Community Dental Health. 2002; 19(1):3-11.
- [4] Kleinbaum, D.G., Klein, M., Introduction to survival analysis, Survival Analysis, 2012.
- [5] Gift HC, Reisine ST, Larach DC. The social impact of dental problems and visits. American Journal of Public Health 1992;82:1663-8.
- [6] Zupnik J, wooKim S, Ravens D, Karimbux N, Guze K. Factors Associated With Dental Implant Survival: A 4-Year Retrospective Analysis. Journal of Periodontology, 2011;82:1390-1395.
- [7] Lam KF, Lee YW. Merits of modeling multivariate survival data using random effects proportional odds model. Biometrical Journal 2004; 46:331-342.
- [8] Lam KF, Ip D. REML and ML estimation for clustered grouped survival data. Statistics in Medicine 2003; 22:2025-2034.



توزیع نسبت متغیرهای مستقل کاماراسوامی و کاربرد آن در مباحث قابلیت اعتماد

رسولی، ع^۱ ایمانی، م^۲ سعدی، س^۳

۱،۲،۳ گروه آمار، دانشگاه زنجان

چکیده

در این مقاله، ابتدا توزیع کاماراسوامی معرفی شده است. سپس تابع توزیع توام و حاشیه ای برای $W = \frac{X_1}{X_2}$ و $T = \frac{X_1}{X_1 + X_2}$ که در آن X_1 و X_2 هر دو متغیرهای مستقل و دارای توزیع کاماراسوامی اند، به دست می آید. همچنین مقدار گشتاورهای متغیرهای W و T محاسبه شده است. توزیع کاماراسوامی بدلیل انعطاف پذیری با توجه به پارامترهای آن در مدل های آمار بیزی به عنوان توزیع پیشین بکار می رود و توزیع متغیرهای T و W در مدل های فشار- قدرت مربوط به قابلیت اعتماد سیستم ها کاربرد دارد.

کلمات کلیدی: توزیع کاماراسوامی، تابع فوق هندسی گوسی، تابع چگالی توام، تابع چگالی حاشیه ای، قابلیت اعتماد سیستم

۱ مقدمه

توزیع کاماراسوامی نخستین بار توسط کاماراسوامی (۱۹۸۰) به صورت زیر معرفی شد.

$$f_Z(z) = \frac{1}{b-c} pq \left(\frac{z-c}{b-c} \right)^{p-1} \left[1 - \left(\frac{z-c}{b-c} \right)^p \right]^{q-1} \quad c < z < b,$$

^۱ rasouli@znu.ac.ir

^۲

^۳ saeed.sadi@znu.ac.ir

که $q > 0$ و $p > 0$ پارامترهای توزیع و b و c کران های متغیر آن است. در حالت کلی توزیع کاماراسوآمی با نماد $Kum(p, q, c, b)$ نشان داده می شود و با استفاده از روش تبدیل متغیر، با فرض $X = \frac{Z-c}{b-c}$ ، تابع چگالی کاماراسوآمی استاندارد به صورت زیر حاصل می شود:

$$f_X(x) = pqx^{p-1}(1-x^p)^{q-1}, 0 < x < 1. \quad (1)$$

در این جا همچنین داریم:

$$Kum(p, q) \equiv Kum(p, q, \cdot, 1),$$

و تابع توزیع تجمعی توزیع کاماراسوآمی استاندارد را به صورت زیر داریم:

$$F(x) = 1 - (1 - x^p)^q \quad 0 < x < 1,$$

همچنین فرم چنک $F^{-1}(u)$ را به صورت زیر داریم:

$$x = [1 - (1 - u)^{\frac{1}{q}}]^{\frac{1}{p}} \quad 0 < u < 1,$$

و در یک حالت خاص میانه توزیع کاماراسوآمی برابر $(1 - 0.5^{\frac{1}{q}})^{\frac{1}{p}}$ $med(X) =$ خواهد بود.

اگر $X \sim Kum(p, q)$ باشد، تابع مولد گشتاور توزیع کاماراسوآمی در نقطه صفر به صورت زیر خواهد بود:

$$\mu'_r(X) = qB(1 + \frac{r}{p}, q),$$

که در آن تابع بتا و تابع گاما به ترتیب زیر است:

$$B(\alpha, \beta) = \int_0^1 s^{\alpha-1}(1-s)^{\beta-1} ds = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

توزیع کاماراسوآمی به دلیل اینکه دارای دامنه یکسان با توزیع بتا می باشد، در آمار بیز به عنوان توزیع پیشین اطلاع بخش به کار می رود. گین (۲۰۰۰) توزیع نسبت متغیرهای بتا را ارائه داد و از آن در بررسی مدل های فشار - قدرت استفاده کرد. در این مقاله ما سعی می کنیم بر اساس توزیع کاماراسوآمی به عنوان توزیعی که انعطاف پذیرتر از توزیع بتا است، توزیع نسبت متغیرها را ارائه دهیم.

۲ تابع چگالی توام و حاشیه ای

قضیه ۱.۲. فرض کنید $X_1 \sim Kum(a, b_1)$ و $X_2 \sim Kum(a, b_2)$.

تابع چگالی متغیر تصادفی $W = \frac{X_1}{X_2}$ برای $0 \leq w \leq 1$ و $w > 1$ به ترتیب به صورت زیر است:

$$f(w) = \frac{Aw^{a-1}}{a} \cdot \frac{1}{b_2(b_2+1)} {}_2F_1(1-b_1, 2, b_2+2, w^a),$$

$$f(w) = \frac{A}{aw^{a+1}} \cdot \frac{1}{b_1(b_1+1)} {}_2F_1(1-b_2, 2, b_1+2, \frac{1}{w^a}),$$

که در آن $A = (ab_1)(ab_2)$ و

$${}_2F_1(a, b, c, x) = \int_0^1 \frac{u^{b-1}(1-u)^{c-b-1}}{B(b, c-b)(1-ux)^a} du, \quad (2)$$

تابع فوق هندسی گوسی، با نماد ${}_2F_1$ نشان داده شده است، رجوع شود به مداعی و ساکسنا (۱۹۷۳).

برهان. تابع چگالی احتمال توام را به صورت زیر داریم:

$$f(x_1, x_2) = Ax_1^{a-1}(1-x_1^{a_1})^{b_1-1} \times x_2^{a-1}(1-x_2^{a_2})^{b_2-1},$$

که در آن $0 < x_1, x_2 < 1$.

با استفاده از روش تبدیل متغیر داریم:

$$w = \frac{X_1}{X_2}, v = X_2 \Rightarrow X_1 = wv, X_2 = v,$$

و بدنبال آن برای ژاکوبین تبدیل داریم:

$$|J| = v.$$

بنابراین

$$f(w, v) = A(wv)^{a-1}(1-(wv)^{a_1})^{b_1-1}v^{a-1}(1-v^{a_2})^{b_2-1}v,$$

و تابع چگالی حاشیه ای W را به صورت زیر داریم:

$$f(w) = \int_C A(wv)^{a-1}(1-(wv)^{a_1})^{b_1-1}v^{a-1}(1-v^{a_2})^{b_2-1}v dv,$$

که در آن

$$C = \{(v, w); 0 < w, 0 < v < 1, 0 < vw < 1\}.$$

برای $0 \leq w \leq 1$ داریم:

$$f(w) = Aw^{a-1} \int_0^1 v^{a-1}(1-(wv)^{a_1})^{b_1-1}(1-v^{a_2})^{b_2-1} dv.$$

و با تغییر متغیر $v^a = t \Rightarrow av^{a-1} dv = dt$ داریم:

$$f(w) = \frac{Aw^{a-1}}{a} \int_0^1 t(1-w^a t)^{b_1-1}(1-t)^{b_2-1} dt.$$

حال با استفاده از (۲) داریم:

$$f(w) = \frac{Aw^{a-1}}{a} \cdot \frac{1}{b_2(b_2+1)} {}_2F_1(1-b_1, 2, b_2+2, w^a).$$

برای $w > 1$ با انجام تغییر متغیر دیگری به صورت $wv = z \Rightarrow dv = \frac{dz}{w}$ داریم:

$$f(w) = A \int_0^1 z^{a-1}(1-z^{a_1})^{b_1-1}(\frac{z}{w})^{a-1}(1-(\frac{z}{w})^{a_2})^{b_2-1}(\frac{z}{w}) \frac{dz}{w}.$$

با اندکی ساده کردن داریم:

$$f(w) = A \int_0^1 \frac{z^{a-1}(1-z^{a_1})^{b_1-1}}{w^{a+1}}(1-(\frac{z}{w})^{a_2})^{b_2-1} dz.$$

مجدداً با تغییر متغیر $z^a = t \Rightarrow az^{a-1} dz = dt$ داریم:

$$f(w) = A \int_0^1 \frac{t(1-t)^{b_1-1}}{aw^{a+1}}(1-\frac{t}{w^a})^{b_2-1} dt,$$

و در نهایت با استفاده از (۲) داریم:

$$f(w) = \frac{A}{aw^{a+1}} \cdot \frac{1}{b_1(b_1+1)} {}_2F_1(1-b_2, 2, b_1+2, \frac{1}{w^a}).$$

□

قضیه ۲.۲. فرض کنید $X_1 \sim Kum(a, b_1)$ و $X_2 \sim Kum(a, b_2)$ تابع چگالی $T = \frac{X_1}{X_1+X_2}$ به صورت زیر است:

اگر $0 \leq t \leq \frac{1}{2}$

$$g(t) = \frac{At^{a-1}(1-t)^{-a-1}}{a} \cdot \frac{1}{b_2(b_2+1)} {}_2F_1(1-b_1, 2, b_2+2, (\frac{t}{1-t})^a),$$

اگر $\frac{1}{2} \leq t \leq 1$

$$g(t) = \frac{At^{-a-1}(1-t)^{a-1}}{a} \cdot \frac{1}{b_1(b_1+1)} {}_2F_1(1-b_2, 2, b_1+2, (\frac{1-t}{t})^a).$$

برهان. بر اساس تبدیلات، تابع چگالی توام را به صورت زیر داریم:

$$g(y_1, t) = A(t(1-t))^{a-1} y_1^{a-1} (1 - (ty_1)^a)^{b_1-1} (1 - (y_1(1-t))^a)^{b_2-1},$$

و بدنبال آن تابع چگالی حاشیه ای متغیر T عبارت است از:

$$g(t) = A(t(1-t))^{a-1} \int_C y_1^{a-1} (1 - (ty_1)^a)^{b_1-1} (1 - (y_1(1-t))^a)^{b_2-1} dy_1.$$

که در آن

$$C = \left\{ (y_1, t); 0 < y_1 < 1, 0 < y_1 < \frac{1}{t}, 0 < y_1 < \frac{1}{1-t}, 0 < t < 1 \right\}.$$

حال با توجه به C ، دو ناحیه انتگرال گیری داریم:

$$0 \leq t \leq \frac{1}{1+b_2}$$

$$g(t) = \int_{\frac{1}{1+b_2}}^{\frac{1}{1-t}} g(y_1, t) dy_1 =$$

$$A(t(1-t))^{a-1} \int_{\frac{1}{1+b_2}}^{\frac{1}{1-t}} y_1^{a-1} (1 - (ty_1)^a)^{b_1-1} (1 - (y_1(1-t))^a)^{b_2-1} dy_1,$$

که با تغییر متغیر $v = y_1(1-t) \Rightarrow dv = (1-t)dy_1$ و اندکی ساده کردن داریم:

$$g(t) = At^{a-1} (1-t)^{-a-1} \int_{\frac{1}{1+b_2}}^1 v^{a-1} (1 - (\frac{tv}{1-t})^a)^{b_1-1} (1 - v^a)^{b_2-1} dv.$$

مجدداً با فرض $v^a = u \Rightarrow av^{a-1} dv = du$ داریم:

$$g(t) = \frac{At^{a-1} (1-t)^{-a-1}}{a} \int_{\frac{1}{1+b_2}}^1 \frac{u(1 - (\frac{t}{1-t})^a)^{b_1-1}}{(1-u)^{1-b_2}} du,$$

و در نهایت از (۲) داریم:

$$g(t) = \frac{At^{a-1} (1-t)^{-a-1}}{a} \cdot \left(\frac{1}{b_2(b_2+1)} \right) {}_2F_1(1-b_1, 2, b_2+2, (\frac{t}{1-t})^a).$$

برای $\frac{1}{1+b_2} \leq t \leq 1$ داریم:

$$g(t) = \int_{\frac{1}{1+b_2}}^{\frac{1}{t}} g(y_1, t) dy_1$$

$$= A(t(1-t))^{a-1} \int_{\frac{1}{1+b_2}}^{\frac{1}{t}} y_1^{a-1} (1 - (ty_1)^a)^{b_1-1} (1 - (y_1(1-t))^a)^{b_2-1} dy_1,$$

با تغییر متغیر $u = y_1 t \Rightarrow du = t dy_1$ و اندکی ساده کردن داریم:

$$g(t) = At^{-a-1} (1-t)^{a-1} \int_{\frac{1}{1+b_2}}^1 u^{a-1} (1 - u^a)^{b_1-1} (1 - (\frac{u(1-t)}{t})^a)^{b_2-1} du.$$

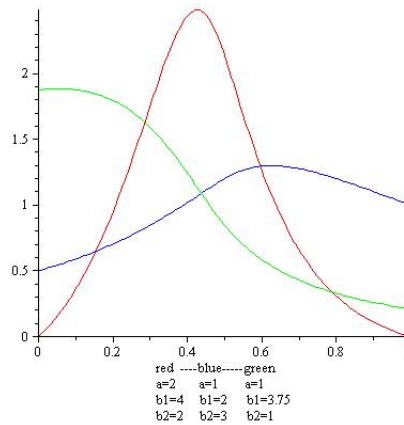
حال تغییر متغیر $u^a = v \Rightarrow au^{a-1} du = dv$ را انجام می دهیم:

$$g(t) = \frac{At^{-a-1} (1-t)^{a-1}}{a} \int_{\frac{1}{1+b_2}}^1 v(1-v)^{b_1-1} (1 - (\frac{1-t}{t})^a v)^{b_2-1} dv,$$

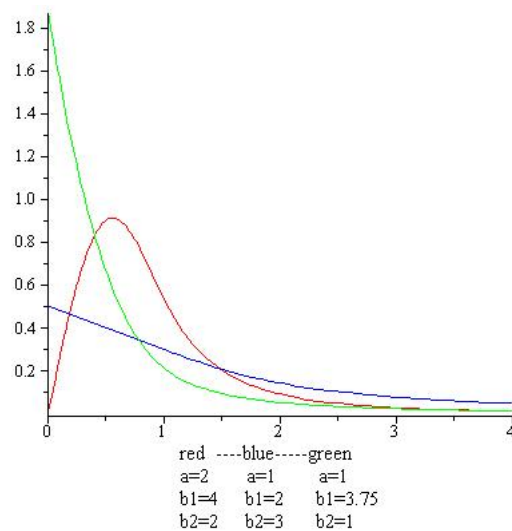
با استفاده کردن از (۲) داریم:

$$g(t) = \frac{At^{-a-1} (1-t)^{a-1}}{a} \left(\frac{1}{b_1(b_1+1)} \right) {}_2F_1(1-b_2, 2, b_1+2, (\frac{1-t}{t})^a).$$

□



شکل ۱: تابع چگالی متغیر $T = \frac{X_1}{X_1 + X_2}$



شکل ۲: تابع چگالی متغیر $W = \frac{X_1}{X_2}$

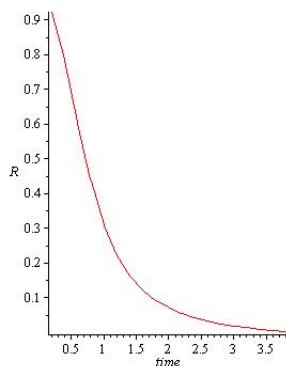
در شکل‌های ۱ و ۲ تابع چگالی متغیر T و W که در آن $X_1 \sim Kum(a, b_1)$ و $X_2 \sim Kum(a, b_2)$ به ازای مقادیر مختلف پارامترها نشان داده شده است.

حل انتگرال‌های فوق فقط با روش‌های عددی امکان‌پذیر خواهد بود و بر اساس آنها می‌توان مقادیر امید ریاضی، واریانس، میزان چولگی و کشیدگی متغیرهای T و W را محاسبه نمود. همچنین در بررسی مدل فشار-قدرت محاسبه مقدار $P(X_1 < X_2)$ و یا به طور معادل $P(W < 1)$ یعنی احتمال کمتر بودن قدرت یک سیستم در مقابل فشار وارده به آن مورد توجه در مباحث قابلیت اعتماد سیستم‌ها می‌باشد که

به عنوان مثال با این فرض که $X_1 \sim Kum(2, 4)$ و $X_2 \sim Kum(2, 2)$ ، با استفاده از روش های عددی داریم:

$$P(W > 1) = \int_1^\infty \frac{{}_4F_1(-1, 2, 6, w^{-2})}{5 \cdot w^3} = 0.3333.$$

به عنوان مثالی دیگر با فرض $X_1 \sim Kum(2, 4)$ و $X_2 \sim Kum(2, 2)$ در شکل ۳ نمودار تابع قابلیت اعتماد یک سیستم بر مبنای توزیع متغیر W یعنی $R_W(t) = P(W > t)$ ارائه گردیده است.



شکل ۳: تابع قابلیت اعتماد متغیر $W = \frac{X_1}{X_2}$

۳ نتیجه گیری

در مدلسازی فشار- قدرت بعنوان یکی از مباحث مهم در موضوع قابلیت اعتماد میزان R مهم می باشد. با این فرض که X_1 و X_2 هر دو از توزیع کاماراسوامی هستند (مثلا اگر X_1 درصد بیکاری در سالهای مختلف و X_2 درصد شغل های پایدار ایجاد شده از کل شغل های ایجاد شده در یک سال تعریف شوند) میزان R بعنوان یک شاخص عملکرد دولت می تواند مورد توجه قرار گیرد. در این مقاله در حالتی کلی تر توزیع دو متغیره تصادفی $W = \frac{X_1}{X_1 + X_2}$ و $T = \frac{X_1}{X_1 + X_2}$ بدست آمده اند که در محاسبه و بررسی احتمال فوق می تواند مورد استفاده قرار گیرد.

مراجع

- [1] Kumaraswamy, P. (1980). *A Generalized Probability Density Function for Double Bounded Random Processes*. Journal of Hydrology, 46, 79-88.
- [2] Mathai A. M. and Saxena R. K. (1973). *Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences*. Springer-Verlag, Lecture Notes No. 348, Heidelberg.
- [3] Pham-Gia, T. (2000). *Distributions of the ratios of independent beta variables and applications*, 29(12), 2693-2915.



حفظ برخی ترتیب‌های تصادفی وابسته در سیستم‌های منسجم

شریفی پیرکوهی، آ^۱ امینی، م^۲ رزمخواه، م^۳

^{۱،۲،۳} گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

چکیده

در این مقاله، حفظ ترتیب‌های تصادفی وابسته برای بردارهای تصادفی با ساختار وابستگی تغییر شکل یافته بررسی شده است. علاوه بر این به عنوان کاربردی از نتایج بدست آمده برای مقایسه تصادفی طول عمر سیستم‌های منسجم زمانی که مؤلفه‌های آن‌ها، دارای ساختار وابستگی تغییر شکل یافته هستند، مورد توجه قرار گرفته است.

کلمات کلیدی: آماره‌های مرتب، سیستم‌های منسجم، توزیع تغییر شکل یافته، مفصل.

۱ مقدمه

حفظ ترتیب‌های تصادفی در سیستم‌های منسجم با مؤلفه‌های وابسته یک موضوع مهم در نظریه قابلیت اطمینان است که مورد توجه بسیاری از محققان قرار گرفته است. ساختار وابستگی تغییر شکل یافته برای اولین بار توسط قیگگین (۱۹۸۲) و یاری (۱۹۸۷) بیان شد. پس از آن محققانی مانند: خالدی و شیکد (۲۰۱۰) و گوپتا و کومار (۲۰۱۴)، مطالعات اساسی در مورد توزیع‌های تغییر شکل یافته در مدل‌های اقتصادی، آماره‌های ترتیبی و سیستم‌های منسجم انجام دادند. ناوارو و سوردو (۲۰۱۳) و گوپتا و کومار (۲۰۱۴) در مورد شرایط برای حفظ ترتیب‌های تصادفی با ساختار وابستگی تغییر شکل یافته و حفظ سالخوردگی ترتیب‌های تصادفی تحقیقاتی انجام داده‌اند. هدف از این مقاله، حفظ برخی ترتیب‌های تصادفی در سیستم‌های منسجم وقتی مؤلفه‌ها متغیرهای تصادفی وابسته با ساختار وابستگی تغییر شکل یافته است. در بخش

^۱ h.charckazi1993@um.ac.ir

^۲ m-amini@um.ac.ir

^۳ razmkhah@um.ac.ir

۲ مفاهیم و لم و گزاره‌های مورد نیاز در بخش بعدی مطرح می‌شود و علاوه بر این حفظ نتایج در سیستم‌های منسجم و ترتیبهای تصادفی با ساختار وابستگی تغییر شکل یافته برای متغیرهای تصادفی وابسته دو بعدی بیان می‌شود و در بخش ۳ با استفاده از نتایج بدست آمده در بخش ۲، کاربردهای سیستم‌های منسجم در قالب چند مثال ارائه می‌شود.

۲ حفظ نتایج برای ساختار وابستگی توزیع‌های تغییر شکل یافته

در اینجا به بررسی ویژگی‌های ترتیبهای تصادفی سیستم‌های منسجم با مؤلفه‌های وابسته و هم‌توزیع می‌پردازیم. نتایج به دست آمده را براساس ارائه‌ی تابع قابلیت اطمینان به عنوان تابع تغییر شکل بدست می‌آوریم و ترتیبهای تصادفی شامل ترتیب تصادفی معمولی، نرخ خطر، نرخ خطر معکوس و نسبت درستی را براساس تابع تغییر شکل به دست می‌آوریم. مزیت استفاده از این روش این است که توابع تغییر شکل به تابع توزیع سیستم بستگی ندارند و فقط وابسته به ساختار سیستم می‌باشند. توابع تغییر شکل را می‌توان در دیگر زمینه‌ها به کار برد. تابع $Q: [0, 1]^2 \rightarrow [0, 1]$ را یک تابع تغییر شکل دو بعدی می‌نامند اگر $Q(0, 0) = 0$ و $Q(1, 1) = 1$ و Q یک تابع صعودی و پیوسته باشد. فرض کنید X_1 و X_2 دو متغیر تصادفی به ترتیب با توابع توزیع F_1 و F_2 و $Q(\cdot, \cdot)$ یک تابع تغییر شکل دو متغیره باشد، در این صورت توزیع تغییر شکل یافته به صورت زیر است:

$$F_Q(t) = Q(F_1(t), F_2(t)) \quad 0 \leq t \leq 1, \quad (1)$$

با توجه به ویژگی‌های $Q(\cdot, \cdot)$ به سادگی می‌توان نشان داد که F_Q یک تابع توزیع تغییر شکل یافته است. علاوه بر این اگر Q نسبت به هر مؤلفه صعودی، F_1 و F_2 دارای تکیه‌گاه مشابه S باشند آنگاه $\bar{F}_Q$ نیز دارای تکیه‌گاه S است. با روش مشابه تابع بقاء تغییر شکل یافته به صورت

$$\bar{F}_Q(t) = \bar{Q}(\bar{F}_1(t), \bar{F}_2(t)) \quad (2)$$

بدست می‌آید، که در آن $\bar{F}_i = 1 - F_i, i = 1, 2$ و $\bar{F}_Q = 1 - F_Q$ و $\bar{Q}(u_1, u_2) = 1 - Q(1 - u_1, 1 - u_2)$ که یک تابع تغییر شکل است. معادلات (۱) و (۲) معادل هستند ولی برای سادگی معمولاً از معادله (۲) استفاده می‌شود.

تعریف ۱.۰۲. فرض کنید تابعی $C: [0, 1]^2 \rightarrow [0, 1]$ یک تابع دو متغیره باشد در این صورت $C(u_1, u_2)$ یک تابع مفصل است اگر دارای ویژگی‌های زیر باشد.

۱: برای هر u و v در $[0, 1]$

$$C(u, 0) = C(0, v) = 0 \quad \text{و} \quad C(u, 1) = C(1, v) = 1.$$

۲: برای هر u_1, u_2, v_1, v_2 در بازه‌ی $[0, 1]$ به‌طوری که $u_1 \leq u_2$ و $v_1 \leq v_2$ باشند، داشته باشیم

$$C(u_1, v_1) - C(u_2, v_1) - C(u_1, v_2) + C(u_2, v_2) \geq 0$$

فرض کنید $T = \psi(X_1, X_2)$ طول عمر یک سیستم منسجم با طول عمرهای وابسته و هم‌توزیع X_1, X_2 باشد. طول عمر یک سیستم را می‌توان با $R: R^2 \rightarrow R$ که تابعی از طول عمر مؤلفه‌ها است و فقط به ساختار سیستم وابسته است، بیان کرد. یعنی

$$\bar{F}(t_1, t_2) = \hat{C}(\bar{F}(t_1), \bar{F}(t_2)). \quad (3)$$

که در آن $\hat{C}$ ، مفصل بقاء می‌باشد.

هدف اصلی این مقاله شامل سه موضوع می‌باشد:

(i) - اگر توزیعهای پایه یکسان F_1 و F_2 نسبت به یک ترتیب تصادفی مرتب شده باشند آنگاه تحت چه شرایطی توزیع‌های تغییر شکل یافته مختلف Q_1 و Q_2 نسبت به همان ترتیب تصادفی مرتب می‌شود. به عبارت دیگر

$$F_{Q_1} \leq_{ORD} F_{Q_2}$$

(ii) - اگر توزیعهای پایه مختلف F_1, F_2, G_1, G_2 نسبت به یک ترتیب تصادفی مرتب شده باشند آنگاه تحت چه شرایطی توزیعهای تغییر شکل یافته یکسان Q نسبت به همان ترتیب تصادفی مرتب می شود. به عبارت دیگر

$$G_Q(t) = Q(G_1(t), G_2(t)).$$

(iii) - اگر توزیعهای پایه مختلف $F_i, G_i, i = 1, 2$ نسبت به یک ترتیب تصادفی مرتب شده باشند آنگاه تحت چه شرایطی توزیعهای تغییر شکل یافته مختلف Q نسبت به همان ترتیب تصادفی مرتب می شود. به عبارت دیگر

$$F_i \leq_{ORD} G_i, \quad i = 1, 2$$

تعریف ۲.۲. فرض کنید X و Y دو متغیر مطلقاً پیوسته با تکیهگاه $[0, \infty)$ و به ترتیب با توابع توزیع $F(t)$ و $G(t)$ و توابع بقای $\bar{F}(t) = 1 - F(t)$ و $\bar{G}(t) = 1 - G(t)$ باشند. همچنین فرض کنید به ترتیب $f(t)$ و $g(t)$ توابع چگالی احتمال و h_F و h_G توابع نرخ خطر $\bar{h}_F$ و $\bar{h}_G$ توابع نرخ خطر معکوس آنها باشند، در این صورت

• X کوچکتر از Y در ترتیب تصادفی معمولی است و با $X \leq_{ST} Y$ یا $F \leq_{ST} G$ نشان می دهند، هرگاه $0 \leq t \leq 1, \bar{F}(t) \leq \bar{G}(t)$.

• X کوچکتر از Y در ترتیب نرخ خطر است و با $X \leq_{HR} Y$ یا $F \leq_{HR} G$ نشان می دهند، هرگاه $\frac{\bar{F}}{\bar{G}}$ تابعی صعودی به ازای $t \geq 0$ باشد.

• X کوچکتر از Y در ترتیب نرخ خطر معکوس است و با $X \leq_{RHR} Y$ یا $F \leq_{RHR} G$ نشان می دهند، هرگاه $\frac{F}{G}$ تابعی صعودی به ازای $t \geq 0$ باشد.

• X کوچکتر از Y در ترتیب نسبت درستنمایی است و با $X \leq_{LR} Y$ یا $F \leq_{LR} G$ نشان می دهند، هرگاه $\frac{f}{g}$ تابعی صعودی به ازای $t \geq 0$ باشد.

• متغیر تصادفی X دارای تابع نرخ خطر صعودی (نزولی) $(DFR)IFR$ است، هرگاه برای هر $x > 0$ ، نسبت $\frac{\bar{F}(x+t)}{\bar{F}(t)}$ تابعی نزولی (صعودی) از t باشد که در آن $\bar{F}(t)$ تابع قابلیت اعتماد متغیر تصادفی طول عمر X است.

لم زیر به بررسی ویژگی های جدید می پردازیم.

لم ۱.۲. فرض کنید (X_1, X_2) و (Y_1, Y_2) بردارهای تصادفی با توابع توزیع حاشیه ای به ترتیب F_1 و F_2 و توابع چگالی حاشیه ای به ترتیب f_1 و f_2 و Q_1 و Q_2 دو تابع تغییر شکل دو بعدی $F_{Q_1} = Q_1(F_1, F_2)$ و $F_{Q_2} = Q_2(F_1, F_2)$ باشند بطوریکه مشتق Q موجود باشد، در این صورت

-(i)

$$f_Q(t) = \sum_{i=1}^2 f_i(t) D_i Q(F_1(t), F_2(t)) = \sum_{i=1}^2 f_i(t) D_i \bar{Q}(\bar{F}_1(t), \bar{F}_2(t)) \quad (4)$$

که در آن

$$D_i \phi(x_1, x_2) = \frac{d}{dx_i} \phi(x_1, x_2) \quad (5)$$

-(ii)

$$h_Q(t) = \sum_{i=1}^2 h_i(t) \alpha_i^Q(\bar{F}_1(t), \bar{F}_2(t)) \quad (6)$$

که در آن

$$\alpha_i^Q(u_1, u_2) = \frac{u_i D_i \phi(u_1, u_2)}{\phi(u_1, u_2)} \quad ; \quad i = 1, 2 \quad (7)$$

و $h_i = \frac{f_i}{F_i}, i = 1, 2$ ، نرخ خطر توزیع حاشیه ای $F_i, i = 1, 2$ می باشند.

-(iii)

$$\bar{h}_Q(t) = \sum_{i=1}^2 \bar{h}_i(t) \alpha_i^Q(F_1(t), F_2(t)) \quad (8)$$

که در آن $\bar{h}_i = \frac{f_i}{F_i}$ ، نرخ خطر معکوس توزیع حاشیه‌ای $F_i, i = 1, 2$ می‌باشد.

در گزاره زیر شرایطی برای حفظ ویژگی‌های توزیع تغییر شکل یافته مختلف براساس توابع توزیع حاشیه‌ای یکسان بیان می‌کنیم.

گزاره ۱.۲. فرض کنید (X_1, X_2) و (Y_1, Y_2) بردارهای تصادفی با توابع توزیع حاشیه‌ای F_1 و F_2 باشند و اگر Q_1 و Q_2 دو تابع تغییر شکل

دو بعدی $F_{Q_1} = Q_1(F_1, F_2)$ و $F_{Q_2} = Q_2(F_1, F_2)$ باشند، در این صورت

(i) $F_{Q_1} \leq_{ST} F_{Q_2}$ اگر و تنها اگر در بازه $(0, 1)^2$ ؛ $Q_1 \geq Q_2$.

(ii) $F_{Q_1} \leq_{HR} F_{Q_2}$ اگر و تنها اگر $\frac{Q_1}{Q_2}$ در بازه $(0, 1)^2$ صعودی باشد.

(iii) $F_{Q_1} \leq_{RHR} F_{Q_2}$ اگر در بازه $(0, 1)^2$ ؛ $\alpha_i^{Q_1} \geq \alpha_i^{Q_2}; i = 1, 2$.

(iv) $F_{Q_1} \leq_{RHR} F_{Q_2}$ اگر و تنها اگر $\frac{Q_1}{Q_2}$ در بازه $(0, 1)^2$ صعودی باشد.

(v) $F_{Q_1} \leq_{RHR} F_{Q_2}$ اگر در بازه $(0, 1)^2$ ؛ $\alpha_i^{Q_1} \leq \alpha_i^{Q_2}; i = 1, 2$.

گزاره زیر شرایطی برای حفظ ویژگی‌ها برای توزیع‌های تغییر شکل یافته یکسان با توابع توزیع حاشیه‌ای مختلف بدست می‌آوریم.

گزاره ۲.۲. فرض کنید (X_1, X_2) و (Y_1, Y_2) بردارهای تصادفی با توابع توزیع حاشیه‌ای به ترتیب $F_i, G_i, i = 1, 2$ باشند و اگر Q_1 و Q_2 دو

تابع تغییر شکل دو بعدی $F_{Q_1} = Q_1(F_1, F_2)$ و $F_{Q_2} = Q_2(F_1, F_2)$ باشند، در این صورت

(i) اگر $F_i \leq_{ST} G_i; i = 1, 2$ آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ، $F_Q \leq_{ST} G_Q$.

(ii) اگر $F_i \leq_{HR} G_i; i = 1, 2$ آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ، $F_Q \leq_{HR} G_Q$.

(iii) اگر $F_i \leq_{RHR} G_i; i = 1, 2$ آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ، $F_Q \leq_{RHR} G_Q$.

شرایط پیچیده‌تر برای ترتیب نرخ خطر داریم که در گزاره زیر بیان شده‌اند و همچنین ویژگی‌های مشابه برای ترتیب نرخ خطر و نرخ خطر

معکوس بدست آمده‌است.

گزاره ۳.۲. فرض کنید (X_1, X_2) و (Y_1, Y_2) بردارهای تصادفی با توابع توزیع حاشیه‌ای به ترتیب $F_i, G_i, i = 1, 2$ باشند و اگر Q_1 و Q_2 دو

تابع تغییر شکل دو بعدی $F_{Q_1} = Q_1(F_1, F_2)$ و $F_{Q_2} = Q_2(F_1, F_2)$ باشند، در این صورت

(i) اگر $F_i \leq_{HR} G_i, i = 1, 2$ آن‌گاه برای همه توابع توزیع تغییر شکل یافته

$\{F_Q \leq_{HR} G_Q, Q \in \mathcal{D}_2 = \{\phi: \beta^\phi(u_1, u_2, v_1, v_2) \searrow u_i \in (0, 1), v_i \in (1, \infty), i = 1, 2\}$

(ii) اگر $F_i \leq_{RHR} G_i, i = 1, 2$ آن‌گاه برای همه توابع توزیع تغییر شکل یافته

$\{F_Q \leq_{RHR} G_Q, Q \in \mathcal{D}_2 = \{\phi: \beta^\phi(u_1, u_2, v_1, v_2) \nearrow u_i \in (0, 1), v_i \in (0, 1), i = 1, 2\}$

(iii) اگر $F_i \leq_{LR} G_i, i = 1, 2$ باشد آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ، $F_Q \leq_{LR} G_Q$

برقرار است به‌طوری که

$$\begin{aligned} & \gamma^\phi(u_1, u_2, v_1, v_2, w_1, w_2, z_1, z_2) \\ &= \frac{w_1 z_1 u_1 D_1 \phi(u_1 v_1, u_2 v_2) + w_2 z_2 u_2 D_2 \phi(u_1 v_1, u_2 v_2)}{z_1 u_1 D_1 \phi(u_1, u_2) + z_2 u_2 D_2 \phi(u_1, u_2)} \end{aligned} \quad (9)$$

نسبت به u_1, u_2 نزولی، نسبت به v_1, v_2, w_1, w_2 صعودی و نسبت به $z_i, i = 1, 2$ صعودی (نزولی) باشد.

(iv) اگر $F_i \leq_{LR} G_i, i = 1, 2$ باشد آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ، $F_Q \leq_{LR} G_Q$

برقرار است به‌طوری که در معادله (۹) نسبت به u_1, u_2 صعودی، نسبت به v_1, v_2, w_1, w_2 نزولی و نسبت به $z_i, i = 1, 2$ نزولی (صعودی) باشد.

(v) اگر $F_i \leq_{LR} G_i, i = 1, 2$ و $F_i \in DRFR$ ، نرخ خطر معکوس نزولی باشد آن‌گاه برای همه توابع توزیع تغییر شکل یافته Q ،

$F_Q \leq_{LR} G_Q$ برقرار است به‌طوری که در معادله (۹) نسبت به $u_1, u_2, v_1, v_2, w_1, w_2$ صعودی و نسبت به $z_i, i = 1, 2$ نزولی (صعودی)

می‌باشد.

تذکر ۱۰۲. گزاره ۳.۲ (i) ویژگی بیان شده در کلاس D_2 برای بازه $u_i \in (0, \frac{1}{v_i})$ و $v_i \in (0, \infty)$ کافی است و فرضیات درباره‌ی کلاسهای سالخوردگی IFR, DFR و گزاره ۳.۲ (v) - (iii) برای حفظ ترتیب نرخ خطر بسیار پراهمیت هستند. بخاطر اینکه توابع نرخ خطر در ارئه از نسبت $\frac{g_Q}{f_Q}$ بر حسب توابع گامابه نظر می‌رسند. (متغیرهای z_1, z_2). بنابراین اگر F_1, F_2 دارای توزیع های نمایی باشند دراین صورت شرایط در باره‌ی z_1, z_2 برای حفظ ترتیب نرخ خطر نیاز نیست. (زیرا توابع نرخ خطر ثابت هستند) این حقیقت برای بدست آوردن گزاره ۱.۳ و ۲.۳ استفاده خواهد شد. برخی کاربردها از این نتایج در بخش بعدی ارائه می دهیم.

۳ کاربردهای در سیستم‌های منسجم و آماره‌های ترتیبی

در اینجا با ارائه گزاره و چند مثال از نتایج بدست آمده در بخش قبل، برای سیستم‌های منسجم با مولفه‌های مستقل و طول عمرهای مختلف مورد بررسی قرار می‌دهیم. فرض کنید X_1, X_2 مولفه‌های هم توزیع با تابع بقای $i = 1, 2$ باشد. $\bar{F}(t) = P(X_i \geq t)$ طول عمر مولفه‌ها یعنی X_1, X_2 ممکن است وابسته باشند و این وابستگی را با تابع قابلیت اطمینان توام، به صورت زیر نمایش می‌دهیم

$$\bar{F}(t_1, t_2) = P(X_1 \geq t_1, X_2 \geq t_2) \quad (10)$$

طول عمر سیستم منسجم آن بصورت زیر است

$$\mathcal{T} = \min(X_1, \max(X_2, X_3)) \quad (11)$$

بنابراین تابع قابلیت اطمینان سیستم به صورت زیر است

$$\begin{aligned} \bar{F}_T(t) &= P_r(\{X_{\{1,2\}} > t\} \cup \{X_{\{1,3\}} > t\}) \\ &= P_r(X_{\{1,2\}} > t) + P_r(X_{\{1,3\}} > t) - P(X_{\{1,2,3\}} > t) \\ &= \bar{F}(t, t, \cdot) + \bar{F}(t, \cdot, t) - \bar{F}(t, t, t) \\ &= \hat{C}(\bar{F}_1(t), \bar{F}_2(t), 1) + \hat{C}(\bar{F}_1(t), 1, \bar{F}_2(t)) - \hat{C}(\bar{F}_1(t), \bar{F}_2(t), \bar{F}_2(t)) \\ &= H(\bar{F}_1(t), \bar{F}_2(t), \bar{F}_2(t)) \end{aligned} \quad (12)$$

که H یک تابع است و $\hat{C}$ مفصل بقاء است. تابع قابلیت اطمینان هر سیستم را می‌توان با تابع تغییر شکل (H) نشان داد، که یک تابع صعودی است و به ساختار سیستم و مفصل بقای توام طول عمر مولفه‌های سیستم مستقل است. چون مولفه ها مستقل و هم توزیعند، که در آن H تابع تغییر شکل به صورت زیر داده شده است:

$$H(u_1, u_2, u_3) = \hat{C}(u_1, u_2, 1) + \hat{C}(u_1, 1, u_3) - \hat{C}(u_1, u_2, u_3).$$

که H یک تابع یه بعدب تغییر شکل یافته است. بطور خاص، اگر مفصل $\hat{C}$ مفصل حاصل ضرب باشد در این صورت ساختار تابع قابلیت اطمینان آن بصورت زیر است:

$$H(u_1, u_2, u_3) = u_1 u_2 + u_1 u_3 - u_1 u_2 u_3.$$

برای سیستم سری $X_{1:2}$ که در آن $\bar{C} = \bar{H}$ به این معنی که

$$\bar{F}_{1:2}(t) = \hat{C}(\bar{F}_1(t), \bar{F}_2(t)). \quad (13)$$

بطور مشابه برای سیستم موازی فرض کنید $X_{2:2}$ دارای $Q = C$ که در آن C مفصل است. به این معنی که

$$F_{2:2}(t) = C(F_1(t), F_2(t)) \quad (14)$$

از این رو

$$\bar{F}_{Y_1, Y_2}(t) = 1 - F_{Y_1, Y_2}(t) = 1 - C(1 - \bar{F}_1(t), \bar{F}_2(t))$$

به عبارت دیگر $H(u_1, u_2) = 1 - C(1 - u_1, 1 - u_2)$ تابع تغییر شکل است. ترتیب‌های نرخ خطر و نرخ خطر معکوس و نسبت درستیابی تحت همه‌ی ساختارهای سیستم منسجم برقرار نیستند. بنابراین از گزاره ۲.۲ (ii)، ترتیب نرخ خطر حفظ می‌شود. در گزاره زیر حفظ ویژگی‌ها برای ترتیب نرخ خطر در سیستم‌های سری با مؤلفه‌های مستقل بررسی می‌کنیم. در گزاره زیر حفظ نرخ خطر در سیستم سری با مؤلفه‌های مستقل ارائه می‌دهیم.

گزاره ۱.۳. فرض کنید $Y_{1:2} = \min(Y_1, Y_2)$ و $X_{1:2} = \min(X_1, X_2)$ طول عمر دو سیستم سری با مؤلفه‌های مستقل و به ترتیب دارای توابع توزیع حاشیه‌ای F_1, F_2 و G_1, G_2 به‌طوری که $F_1 \leq_{LR} G_1$ در این صورت
(i) اگر $F_2 \in IFR$ و $F_1 \in DFR$ باشند آن‌گاه $X_{1:2} \leq_{LR} Y_{1:2}$.
(ii) اگر $G_1 \in DFR$ و $F_2 \in IFR$ باشند آن‌گاه $X_{1:2} \leq_{LR} Y_{1:2}$.

گزاره ۲.۳. فرض کنید $Y_{1:2} = \min(Y_1, Y_2)$ و $X_{1:2} = \min(X_1, X_2)$ طول عمر دو سیستم سری با مؤلفه‌های مستقل و به ترتیب دارای توابع توزیع حاشیه‌ای F_1, F_2 و G_1, G_2 به‌طوری که $F_i \leq_{LR} G_i, i = 1, 2$ باشند در این صورت اگر F_1, F_2 یا G_1, G_2 دارای توزیع نمایی باشند آن‌گاه $X_{1:2} \leq_{LR} Y_{1:2}$.

در مثال زیر نشان می‌دهیم که ترتیب نرخ خطر لزومی ندارد برای دیگر توزیع‌ها حفظ شود.

مثال ۱.۳. دو سیستم سری با طول عمرهای $Y_{1:2} = \min(Y_1, Y_2)$ و $X_{1:2} = \min(X_1, X_2)$ را در نظر بگیرید که X_1, Y_1 مستقل و به ترتیب با تابع بقای زیر باشند

$$\bar{F}_1(t) = \exp(-2t^a) \quad \text{و} \quad \bar{G}_1(t) = \exp(-t^a)$$

و X_2, Y_2 نیز مستقل و تابع بقاء زیر می‌باشد

$$\bar{F}_2(t) = \bar{G}_2(t) = \exp(-t^b)$$

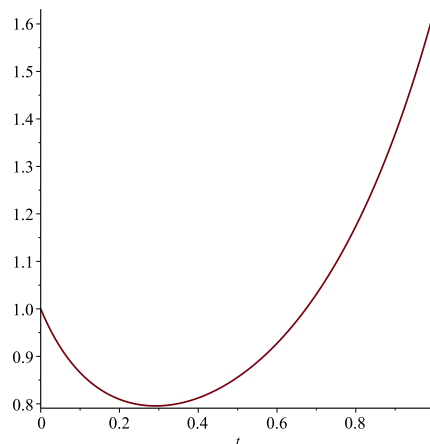
وقتی $\alpha \geq 1$ باشند در این صورت $r(t) = \frac{g_1(t)}{f_1(t)} = 2\exp(t^a)$ یک تابع صعودی نسبت به t است پس در نتیجه $F_1 \leq_{LR} G_1$. واضح است که اگر $a = 1$ و $b = 1$ به این معنی که F_1, G_1 نرخ خطر صعودی و F_2 دارای توزیع نمایی باشند، در این صورت نمودار تابع $r(t) = \frac{g_{1:2}(t)}{f_{1:2}(t)} = \frac{(2t+1)\exp(t^1)}{(2t+1)}$ که در شکل قابل مشاهده است در بازه $(0, 0.3)$ نزولی و در بازه $(0.3, \infty)$ صعودی است.

از این رو $Y_{1:2}$ و $X_{1:2}$ ترتیب نرخ خطر مناسب نیستند. اگر $a = b = 1$ به این معنی که F_1, G_1, F_2 دارای توزیع نمایی باشند در این صورت $r(t) = \frac{2\exp(-2t)}{2\exp(-2t)}$ نسبت به t صعودی و $X_{1:2} \leq_{LR} Y_{1:2}$ است. اگر $a = b = 2$ به این معنی که F_1, G_1, F_2 نرخ خطر صعودی هستند در این صورت $r(t) = \frac{2\exp(-2t^2)}{2\exp(-2t^2)}$ نسبت به t صعودی و $X_{1:2} \leq_{LR} Y_{1:2}$ است.

مثال زیر نشان می‌دهیم که گزاره ۲.۳ لزومی ندارد برای سیستم‌های منسجم (سیستم‌های موازی) درست باشد.

مثال ۲.۳. دو سیستم موازی با طول عمرهای $Y_{2:2} = \max(Y_1, Y_2)$ و $X_{2:2} = \max(X_1, X_2)$ در نظر بگیرید که X_1, Y_1 مستقل و به ترتیب با تابع بقای زیر هستند

$$\bar{F}_1(t) = \exp(-at) \quad \text{و} \quad \bar{G}_1(t) = \exp(-bt)$$


 شکل ۱: نمودار تابع $\frac{g_{1,2}(t)}{f_{1,2}(t)}$

و X_2, Y_2 نیز مستقل و تابع بقاء زیر می باشد

$$\bar{F}_2(t) = \bar{G}_2(t) = \exp(-ct)$$

برای $a, b, c > 0$ وقتی $a \geq b$ باشند در این صورت $r(t) = \frac{g_1(t)}{f_1(t)} = \frac{b \exp(-bt)}{a \exp(-at)}$ یک تابع صعودی نسبت به t است پس در نتیجه $F_1 \leq_{LR} G_1$ برقرار است. واضح است که اگر $a = 2$ و $b = c = 1$ در این صورت نسبت $r(t) = \frac{g_{1,2}(t)}{f_{1,2}(t)} = \frac{\exp(-t) - \exp(-2t)}{\exp(-t) + \exp(-2t) - \exp(-3t)}$ صعودی و نزولی است. بنابراین اگر $a = 3, b = 2, c = 1$ در این صورت $r(t) = \frac{g_{1,2}(t)}{f_{1,2}(t)}$ برای اولین بار نسبت به t صعودی و بعد نزولی است. از این رو $X_{1,2} \leq_{LR} Y_{1,2}$ برقرار نیست. بنابراین ترتیب نرخ خطر حفظ نمی شود. و گزاره ۲.۳ برای حفظ سیستم موازی درست نیست.

در مثال زیر، از نتایج داده شده در گزاره ۱.۲ برای سیستم‌ها با ساختار مختلف استفاده می‌کنیم. نشان می‌دهیم که این نتایج برای سیستم‌های منسجم با ساختار وابستگی مشابه استفاده می‌شود اما دارای مولفه‌هایی با سطح وابستگی مختلفی دارند.

مثال ۳.۳. دو سیستم منسجم با طول عمرهای $T_1 = \min(X_1, X_2, X_3)$ و $T_2 = \min(X_1, \max(X_2, X_3))$ در نظر بگیرید که مولفه‌ها مستقل و سیستم‌ها دارای توابع قابلیت اطمینان زیر هستند

$$\bar{F}_{T_1}(t) = \bar{Q}_1(\bar{F}_1(t), \bar{F}_2(t), \bar{F}_3(t)) \quad \text{و} \quad \bar{F}_{T_2}(t) = H(\bar{F}_1(t), \bar{F}_2(t), \bar{F}_3(t)) = \bar{Q}_2(\bar{F}_1(t), \bar{F}_2(t), \bar{F}_3(t))$$

قرار دهید $u_1 = \bar{F}_1(t), u_2 = \bar{F}_2(t), u_3 = \bar{F}_3(t)$. بنا به مفصل حاصل ضرب و ۱۲ داریم

$$\bar{Q}_1(u_1, u_2, u_3) = u_1 u_2 u_3 \quad \text{و} \quad \bar{Q}_2(u_1, u_2, u_3) = u_1 u_2 + u_1 u_3 - u_1 u_2 u_3$$

در این صورت بنا به گزاره ۱.۲ (ii) $T_1 \leq_{HR} T_2$ برقرار است اگر و تنها اگر

$$\frac{\bar{Q}_2(u_1, u_2, u_3)}{\bar{Q}_1(u_1, u_2, u_3)} = \frac{1}{u_3} + \frac{1}{u_2} - 1, \quad u_1, u_2, u_3 \in (0, 1)$$

که نسبت به u نزولی است. فرض کنید هر دو سیستم منسجم مستقل و دارای مفصل بقاء FGM به صورت زیر باشند

$$K_{FGM}(u_1, u_2, u_3) = u_1 u_2 u_3 (1 - \theta(1 - u_1)(1 - u_2)(1 - u_3)) \quad , \quad \theta \in [-1, 1] \quad (15)$$

وقتی $\theta = 0$ باشند آنگاه دارای مفصل حاصل ضرب است. در این صورت از معادله ۱۲ داریم

$$\bar{Q}_1^\theta(u_1, u_2, u_3) = u_1 u_2 u_3 (1 - \theta(1 - u_1)(1 - u_2)(1 - u_3))$$

و

$$\bar{Q}_2^\theta(u_1, u_2, u_3) = u_1 u_2 + u_1 u_3 - u_1 u_2 u_3 (1 - \theta(1 - u_1)(1 - u_2)(1 - u_3)) \quad \theta \in [-1, 1]$$

وقتی مفصل از معادله ۱۵ داده شده است از این رو ،

$$\frac{\bar{Q}_r^\theta(u_1, u_2, u_3)}{\bar{Q}_1^\theta(u_1, u_2, u_3)} = \frac{1}{u_3(1 - \theta(1 - u_1)(1 - u_2)(1 - u_3))} + \frac{1}{u_2(1 - \theta(1 - u_1)(1 - u_2)(1 - u_3))} - 1$$

نزولی نسبت به $(\cdot, 1)_3$ اگر $\theta \leq 0$ بنابراین برای همه توابع بقاء $\bar{F}_1, \bar{F}_2, \bar{F}_3$ داریم $T_1 \leq_{HR} T_2$.

مراجع

- [1] Gupta, N. and Kumar, S. (2014). *Stochastic comparisons of component and system redundancies with dependent components*. Oper Res Lett , 42:284–289.
- [2] Khaledi, B.E. and Shaked, M. (2010). *Stochastic comparisons of multivariate mixtures*. J Multivar Anal, 101:2486– 2498.
- [3] Navarro, J., del Aguila, Y., Sordo M. A. and Suárez-Llorens, A. (2013). *Stochastic ordering properties for systems with dependent identically distributed components*. Appl Stoch Model Bus Ind, 29:264–278.
- [4] Navarro , J., del Águila, Y., Sordo, M.A .and Suárez-Llorens, A. (2015). *Preservation of Stochastic Orders under the Formation of Generalized Distorted Distributions*. Applications to Coherent Systems .Methodol Comput Appl Probab.
- [5] Quiggin, J. (1982). *A theory of anticipated utility*. J Econ Behav Organ, 3:323–343 .
- [6] Yaari, M. E. (1987). *The dual theory of choice under risk*. Econometrica, 55:95–115.



سیستم‌های مرکب دومؤلفه‌ای و سه‌مؤلفه‌ای با مؤلفه‌های وایبول

صابرزاده، ز^۱ رزمخواه، م^۲

^{۱،۲} گروه آمار، دانشکده ریاضی، دانشگاه فردوسی مشهد

چکیده

سیستم‌های مرکبی را در نظر می‌گیریم که دارای n جزء بوده به قسمی که هر یک از اجزای آن‌ها شامل دو یا سه مؤلفه باشند. هدف اصلی این مقاله بررسی میانگین باقیمانده‌ی عمر چنین سیستم‌هایی منوط بر فعال بودن همه‌ی مؤلفه‌های آن‌ها در زمان t می‌باشد. همچنین مدل‌های دوجمله‌ای دومتغیره و مدل دوجمله‌ای سه‌متغیره را که در بدست آوردن میانگین باقیمانده عمر این سیستم‌ها از آن‌ها بهره می‌گیریم را معرفی می‌کنیم. در پایان تغییرات میانگین باقیمانده‌ی عمر این سیستم‌ها را در مدل فارلی-گامبل-مورگنشترن مورد ارزیابی قرار می‌دهیم.

کلمات کلیدی: آماره‌های مرتب، پارامتر وابستگی، سیستم‌های مرکب (r_1, r_2) از n ، سیستم‌های مرکب (r_1, r_2, r_3) از n ، مدل دوجمله‌ای دومتغیره، مدل دوجمله‌ای سه‌متغیره، مدل فارلی-گامبل-مورگنشترن.

۱ مقدمه

در سال‌های اخیر مطالعات بسیاری روی سیستم‌های منسجم با n مؤلفه انجام شده است. به منظور جزییات بیشتر در مورد ساختار این سیستم‌ها می‌توان به کتاب بارلو و پروشان (۱۹۷۵) اشاره کرد. سیستم‌های k از n حالت خاصی از سیستم‌های منسجم هستند که فعال بودن آن‌ها مستلزم فعال بودن حداقل k مؤلفه‌ی آن‌ها می‌باشد. بوضوح طول عمر چنین سیستمی برابر با $X_{n-k+1:n}$ است که $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ آماره‌های مرتب بدست آمده از n طول عمر مؤلفه‌های سیستم می‌باشد. واضح است که اگر $k = 1$ باشد آن‌گاه سیستم سری حاصل می‌شود.

^۱saberzadez@yahoo.com

^۲razmkhah-m@um.ac.ir

میانگین باقیمانده عمر (MRL) یک سیستم $(n-k+1)$ از n یکی از مفاهیم مهم در قابلیت اعتماد است که به صورت $E(X_{k:n} - t | X_{k:n} > t)$ تعریف می‌شود. بسیاری از محققین به این تابع و کاربردهای آن علاقه‌مند هستند. به عنوان مثال بایراموف و همکاران (۲۰۰۲)، تابع MRL متفاوتی برای یک سیستم به صورت $E(X_{n:n} - t | X_{1:n} > t)$ تعریف کرده‌اند که با $E(X_{k:n} - t | X_{k:n} > t)$ متفاوت بوده و نشان‌دهنده‌ی MLR یک سیستم موازی است مشروط بر اینکه همه‌ی مؤلفه‌های آن در زمان t فعال باشند. اسدی و بایراموف (۲۰۰۵، ۲۰۰۶) تابع MRL را به صورت زیر توسعه دادند

$$E(X_{k:n} - t | X_{r:n} > t); 1 \leq r < k \leq n.$$

همچنین می‌توان به بایراموف و آرنولد (۲۰۰۸)، توانگر و اسدی (۲۰۱۰) و اريلمز (۲۰۱۲) اشاره نمود. در اکثر کارهایی که تا کنون انجام شده است سیستم‌های در نظر گرفته شده شامل n جزء بوده و هر یک از اجزای آن‌ها دارای یک مؤلفه می‌باشد. اما در عمل ممکن است با سیستم‌هایی مواجه شویم که دارای n بوده و هر یک از اجزای آن‌ها شامل دو یا چند مؤلفه باشند. به عنوان مثال فرض کنید یک کارخانه لوازم الکتریکی شامل n خط تولید و هر خط تولیدی به ترتیب دارای دو دستگاه وابسته A_i و B_i ($i = 1, 2, \dots, n$) می‌باشد. همچنین فرض کنید در خط i ام، هر دستگاه A_i و B_i یکی از دو جزء اساسی محصول را تولید می‌کنند و برای رسیدن به محصول نهایی لازم است که تعداد مشخصی از اجزاء تولید شوند. اگر تعداد دستگاه‌های فعال $A_1, A_2, \dots, A_n$ کمتر از r_1 و تعداد دستگاه‌های فعال $B_1, B_2, \dots, B_n$ کمتر از r_2 باشند، فرایند تولید متوقف می‌شود، یعنی فرض می‌کنیم سیستم از کار افتاده است. بنابراین، برای این‌که چنین سیستمی فعال باشد، باید حداقل r_1 تا از دستگاه‌های A_i ، و حداقل r_2 تا از دستگاه‌های B_i فعال باشند. بایراموف (۲۰۱۳) قابلیت اعتماد و MRL سیستم‌های منسجم با دو مؤلفه‌ی وابسته در هر جزء را مورد بررسی قرار داد. در این مقاله می‌خواهیم تحقیق بایراموف را به حالتی که هر یک از اجزای سیستم دارای سه مؤلفه‌ی وابسته است، بسط دهیم. ابتدا سیستم مرکب دو مؤلفه‌ای را معرفی کرده و قابلیت اعتماد آن را با توجه به مدل دوجمله‌ای دومتغیره محاسبه می‌کنیم. سپس با فرض اینکه تمام مؤلفه‌های آن در زمان t فعال باشند، MRL این سیستم را بدست می‌آوریم. همچنین تغییرات میانگین باقیمانده عمر این سیستم‌ها را با فرض اینکه طول عمر مؤلفه‌ها دارای توزیع وایبول با تابع توزیع

$$F(x) = 1 - \exp(-\lambda x^\beta), \quad x \geq 0.$$

باشند، مورد بررسی قرار می‌دهیم. در ادامه این توزیع را با نماد $W(\lambda, \beta)$ نشان می‌دهیم. این توزیع یک توزیع انعطاف پذیر در برازش داده‌های طول عمر می‌باشد. در بخش دوم، سیستم‌های مرکب دو مؤلفه‌ای را از دیدگاه تابع قابلیت اعتماد و میانگین باقیمانده عمر مورد بررسی قرار می‌دهیم. به طور مشابه در بخش سوم به ارزیابی سیستم‌های مرکب سه مؤلفه‌ای می‌پردازیم. برای در نظر گرفتن ساختار وابستگی در این مقاله، از مفصل فارلی گامبل مورگنشرن (FGM) استفاده می‌کنیم.

۲ سیستم‌های مرکب دو مؤلفه‌ای

سیستمی را در نظر بگیرید که دارای n جزء بوده و جزء i ام آن شامل دو مؤلفه‌ی A_i و B_i ($i = 1, 2, \dots, n$) باشد. فرض کنید طول عمر متناظر مؤلفه‌ها را با X_i و Y_i نشان دهیم، به عبارتی X_i طول عمر مؤلفه‌ی A_i و Y_i طول عمر مؤلفه‌ی B_i می‌باشد. همچنین فرض کنید که مؤلفه‌های i امین جزء به هم وابسته‌اند یعنی X_i و Y_i متغیرهای تصادفی وابسته با تابع توزیع توأم $F(x, y)$ باشند، اما اجزاء سیستم مستقل از هم کار کنند، به عبارتی $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ بردارهای تصادفی مستقل از هم باشند. در این سیستم، فرض بر این است که اولین مؤلفه‌ی هر جزء با اولین مؤلفه‌ی سایر اجزاء هم‌توزیع بوده و برای دومین مؤلفه‌ی اجزای سیستم نیز چنین رابطه‌ای برقرار باشد.

تعریف ۱.۲. یک سیستم n جزئی با دو مؤلفه‌ی وابسته در هر جزء را سیستم مرکب (r_1, r_2) از n می‌نامیم هرگاه سیستم فعال باشد اگر و تنها اگر حداقل r_1 تا از مؤلفه‌های $A_1, A_2, \dots, A_n$ و حداقل r_2 تا از مؤلفه‌های $B_1, B_2, \dots, B_n$ فعال باشند.

۱.۲ قابلیت اعتماد

با توجه به تعریف یک سیستم مرکب (r_1, r_2) از n بوضوح مشخص می‌شود که طول عمر این سیستم بابر $T_{r_1, r_2; n} = (X_{n-r_1+1:n}, Y_{n-r_2+1:n})$ است و قابلیت اعتماد آن به صورت زیر تعریف می‌شود

$$P(T_{r_1, r_2; n} > t) = P(X_{n-r_1+1:n} > t, Y_{n-r_2+1:n} > t). \quad (۱)$$

مشاهده می‌شود که قابلیت اعتماد این سیستم به توزیع $(n - r_1 + 1)$ امین آماره مرتب از نمونه‌ی تصادفی $X_1, \dots, X_{n-r_1+1:n}$ و $(n - r_2 + 1)$ امین آماره مرتب بدست آمده از نمونه‌ی تصادفی $Y_1, \dots, Y_{n-r_2+1:n}$ بستگی دارد. قابلیت اعتماد این سیستم را می‌توان با استفاده از مدل دوجمله‌ای دومتغیره بدست آورد. لذا ابتدا این مدل را تعریف می‌کنیم.

برای اولین بار آیتکن و گونین (۱۹۳۵) مدلی را تحت عنوان مدل دوجمله‌ای دومتغیره معرفی کردند. فرض کنید یکی از پیشامدهای E_1 و E_2 به طور همزمان یکی از دو پیشامد F_1 و F_2 با احتمال‌های توأم $P(E_i F_j) = p_{ij}$ رخ دهند به قسمی که $\sum_{i,j} p_{ij} = 1$ و $P(E_1) = p_{11} + p_{10}$. اگر ξ_1 و ξ_2 به ترتیب نمایانگر تعداد دفعاتی باشد که پیشامدهای E_1 و F_1 در یک نمونه‌ی تصادفی n تایی با جایگذاری رخ می‌دهد، آن‌گاه داریم،

$$P(\xi_1 = i, \xi_2 = j) = \sum_{k=\max\{0, i+j-n\}}^{\min\{i, j\}} c(n, k; i, j) p_{11}^k p_{10}^{i-k} p_{21}^{j-k} p_{20}^{n-i-j+k}, \quad (۲)$$

که در آن،

$$c(n, k; i, j) = \frac{n!}{k!(i-k)!(j-k)!(n-i-j+k)!}. \quad (۳)$$

حال برای محاسبه‌ی قابلیت اعتماد یک سیستم (r_1, r_2) از n کافیست در (۲) پیشامدهای $F_1 = \{Y_i > t\}$ و $F_2 = \{Y_i \leq t\}$ را در نظر بگیرید. بنابراین با توجه به این مدل قابلیت اعتماد یک سیستم مرکب دومؤلفه‌ای به صورت زیر می‌باشد

$$\begin{aligned} P(T_{r_1, r_2; n} > t) &= P(X_{n-r_1+1:n} > t, Y_{n-r_2+1:n} > t) \\ &= \sum_{i=r_1}^n \sum_{j=r_2}^n \sum_{k=\max\{0, i+j-n\}}^{\min\{i, j\}} \frac{n!}{k!(i-k)!(j-k)!(n-i-j+k)!} \\ &\quad \times p_{11}^k p_{10}^{i-k} p_{21}^{j-k} p_{20}^{n-i-j+k}, \end{aligned} \quad (۴)$$

که در آن

$$\begin{aligned} p_{11} &= P(E_1 F_1) = P(X_i > t, Y_i > t) = \bar{F}(t, t), \\ p_{10} &= P(E_1 F_2) = P(X_i > t, Y_i \leq t) = \bar{F}_X(t) - \bar{F}(t, t), \\ p_{01} &= P(E_2 F_1) = P(X_i \leq t, Y_i > t) = \bar{F}_Y(t) - \bar{F}(t, t), \\ p_{00} &= P(E_2 F_2) = P(X_i \leq t, Y_i \leq t) = F(t, t). \end{aligned}$$

۲.۲ میانگین باقیمانده‌ی عمر

یک سیستم مرکب (r_1, r_2) از n را در نظر بگیرید به طوری که در زمان t همه‌ی مؤلفه‌های آن فعال باشد. باقیمانده‌ی عمر این سیستم مرکب به صورت زیر تعریف می‌شود

$$\begin{aligned} T_{r_1, r_2; n; t} &= \{X_{n-r_1+1:n} - t, Y_{n-r_2+1:n} - t | \text{همه‌ی مؤلفه‌ها در زمان } t \text{ باشند}\} \\ &= \{X_{n-r_1+1:n} - t, Y_{n-r_2+1:n} - t | X_{1:n} > t, Y_{1:n} > t\}. \end{aligned} \quad (۵)$$

حال اگر $r_1 = r_2 = 1$ باشد آنگاه با یک سیستم موازی دو مؤلفه‌ای مواجه هستیم که در زمان t همه‌ی مؤلفه‌های آن فعال است. بنابراین با توجه به (۵) طول عمر یک سیستم $(1, 1)$ از n با تمام مؤلفه‌های فعال در زمان t برابر است با

$$T_{1,1:n;t} = \{X_{n:n} - t, Y_{n:n} - t | X_{1:n} > t, Y_{1:n} > t\},$$

بنابراین تابع MRL آن به فرم زیر می‌باشد

$$\Phi_{1,1:n}(t) = E(T_{1,1:n;t}) = \int_0^\infty P\{T_{1,1:n;t} > x\} dx,$$

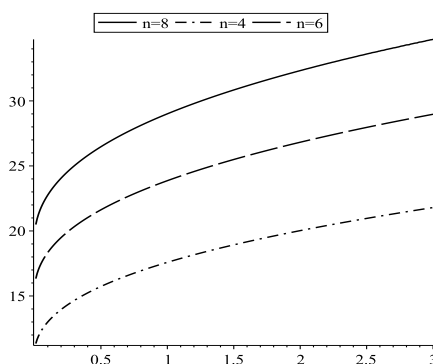
که در آن عبارت داخل انتگرال به صورت زیر می‌باشد که قبلاً توسط بایراموف (۲۰۱۳) محاسبه شده است

$$\begin{aligned} P\{T_{1,1:n;t} > x\} &= 1 - \left(1 - \frac{\bar{F}(t+x, t)}{\bar{F}(t, t)}\right)^n - \left(1 - \frac{\bar{F}(t, t+x)}{\bar{F}(t, t)}\right)^n \\ &\quad + \left(1 - \frac{\bar{F}(t+x, t)}{\bar{F}(t, t)} - \frac{\bar{F}(t, t+x)}{\bar{F}(t, t)} + \frac{\bar{F}(t+x, t+x)}{\bar{F}(t, t)}\right)^n. \end{aligned}$$

حال فرض کنید (X_i, Y_i) در i امین جزء $(i = 1, 2, \dots, n)$ دارای مدل FGM با توزیع توأم زیر باشد

$$F(x, y) = F_X(x)F_Y(y)\{1 + \alpha(1 - F_X(x))(1 - F_Y(y))\}, \alpha \in [-1, 1],$$

که در آن α پارامتر وابستگی بین متغیرهای تصادفی X و Y نامیده می‌شود. برای جزئیات بیشتر در مورد این مدل به نلسن (۲۰۰۶) مراجعه نمایید. همچنین توابع توزیع حاشیه‌ای را $W(0.5, 0.5)$ در نظر می‌گیریم. در شکل ۱ تغییرات تابع MRL نسبت به t و به ازای $\alpha = 0.5$ و $n = 4, 6, 8$ رسم شده است. مشاهده می‌شود که MRL سیستم نسبت به t صعودی می‌باشد. همچنین به ازای یک مقدار ثابت t ، میانگین باقیمانده عمر سیستم با افزایش تعداد اجزای سیستم، افزایش می‌یابد.

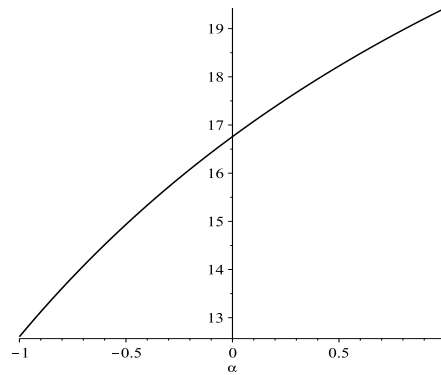


شکل ۱: رفتار تابع MRL سیستم مرکب دو مؤلفه‌ای نسبت به t تحت حاشیه‌ای‌های وایبول.

برای بررسی تغییرات MRL نسبت به پارامتر وابستگی α ، نمودار مربوطه به ازای $t = 1/2$ و $n = 4$ در شکل ۲ رسم شده است. ملاحظه می‌کنید که با افزایش وابستگی بین دو متغیر X و Y ، میانگین باقیمانده عمر نیز افزایش می‌یابد. در حقیقت MRL نسبت به پارامتر α صعودی می‌باشد.

۳ سیستم‌های مرکب سه مؤلفه‌ای

در این بخش سیستم‌هایی را در نظر داریم که دارای n جزء بوده و جزء i ام آن شامل سه مؤلفه‌ی A_i ، B_i و C_i ($i = 1, 2, \dots, n$) باشد. فرض کنید Z_i و Y_i به ترتیب طول عمر متناظر با مؤلفه‌های A_i ، B_i و C_i باشند. همچنین فرض کنید در i امین جزء X_i ، Y_i و Z_i متغیرهای تصادفی



شکل ۲: رفتار تابع MRL سیستم مرکب دو مؤلفه‌ای نسبت به α تحت حاشیه‌ای‌های وایبول.

وابسته با تابع توزیع توأم $F(x, y, z)$ باشند، اما اجزاء سیستم مستقل از هم کار کنند، به عبارتی $(X_1, Y_1, Z_1), (X_2, Y_2, Z_2), \dots, (X_n, Y_n, Z_n)$ بردارهای تصادفی مستقل از هم باشند. در این سیستم، فرض بر این است که اولین مؤلفه‌های همه‌ی اجزاء هم‌توزیع بوده و برای دومین و سومین مؤلفه‌های اجزای سیستم نیز چنین رابطه‌ای برقرار باشد.

تعریف ۱.۳. یک سیستم با n جزء و سه مؤلفه‌ی وابسته در هر جزء را سیستم مرکب (r_1, r_2, r_3) از n گوئیم هرگاه سیستم فعال باشد اگر و تنها اگر حداقل r_1 تا از مؤلفه‌های $A_1, A_2, \dots, A_n$ ، حداقل r_2 تا از مؤلفه‌های $B_1, B_2, \dots, B_n$ و حداقل r_3 تا از مؤلفه‌های $C_1, C_2, \dots, C_n$ فعال باشند.

۱.۳ قابلیت اعتماد

در این قسمت می‌خواهیم قابلیت اعتماد یک سیستم مرکب (r_1, r_2, r_3) از n را مورد مطالعه قرار می‌دهیم. با توجه به تعریف، طول عمر آن $T_{r_1, r_2, r_3; n} = (X_{n-r_1+1:n}, Y_{n-r_2+1:n}, Z_{n-r_3+1:n})$ است. بنابراین قابلیت اعتماد این سیستم به‌صورت زیر می‌باشد

$$P(T_{r_1, r_2, r_3; n} > t) = P(X_{n-r_1+1:n} > t, Y_{n-r_2+1:n} > t, Z_{n-r_3+1:n} > t). \quad (6)$$

ملاحظه می‌شود که این احتمال به توزیع توأم سه آماره مرتب $X_{n-r_1+1:n}, Y_{n-r_2+1:n}, Z_{n-r_3+1:n}$ بستگی دارد. حال می‌توان با استفاده از مدل دوجمله‌ای سه‌متغیره قابلیت اعتماد یک سیستم سه‌مؤلفه‌ای را بدست آورد. پس ابتدا به طور مختصر این مدل را بیان می‌کنیم. سه پیشامد وابسته E_1, F_1, G_1 با متمم‌های $E_1^c = E_1^c, F_1^c = F_1^c, G_1^c = G_1^c$ ، و احتمالات توأم $(i, j, k = 0, 1)$ $P(E_i F_j G_k) = p_{ijk}$ را در نظر بگیرید به قسمی که $\sum_{i,j,k} p_{ijk} = 1$. در یک نمونه‌ی تصادفی ساده و با جایگذاری، فرض کنید ξ_1, ξ_2 و ξ_3 به ترتیب تعداد رخداد‌های

F_1, E_1 و G_1 را نشان دهند. بنابراین تابع جرم احتمال (ξ_1, ξ_2, ξ_3) برابر است با

$$\begin{aligned}
 P(\xi_1 = i, \xi_2 = j, \xi_3 = k) &= \sum_{a=0}^{\min\{i,j\}} \sum_{b=0}^{\min\{j,k\}} \sum_{c=0}^{\min\{i,k\}} \sum_{d=0}^{\min\{a,b,c\}} n! \{d! (a-d)! (c-d)! \\
 &\times (i-a-c+d)! (b-d)! (j-a-b+d)! (k-b-c+d)! \\
 &\times (n-i-j-k+a+b+c-d)!\}^{-1} \\
 &\times p_{11}^d p_{11}^{a-d} p_{11}^{c-d} p_{11}^{i-a-c+d} p_{11}^{b-d} p_{11}^{j-a-b+d} p_{11}^{k-b-c+d} \\
 &\times p_{11}^{n-i-j-k+a+b+c-d}.
 \end{aligned} \quad (7)$$

برای جزئیات بیشتر به جانسون و همکاران (۱۹۹۷، صفحه‌ی ۱۴۱) مراجعه نمایید.

حال با استفاده از مدل دوجمله‌ای سه‌متغیره و با در نظر گرفتن پیشامدهای

$G = \{Z_i \leq t\}$ و $G_1 = \{Z_i > t\}$ ، $F = \{Y_i \leq t\}$ ، $F_1 = \{Y_i > t\}$ ، $E = \{X_i \leq t\}$ ، $E_1 = \{X_i > t\}$ در این مدل می‌توان قابلیت اعتماد یک سیستم مرکب سه مؤلفه‌ای را محاسبه نمود. بنابراین داریم

$$\begin{aligned}
 P\{T_{r_1, r_2, r_3:n} > t\} &= P\{X_{n-r_1+1:n} > t, Y_{n-r_2+1:n} > t, Z_{n-r_3+1:n} > t\} \\
 &= \sum_{i=r_1}^n \sum_{j=r_2}^n \sum_{k=r_3}^n P(\xi_1 = i, \xi_2 = j, \xi_3 = k),
 \end{aligned} \quad (8)$$

که تابع جرم احتمال توأم بردار (ξ_1, ξ_2, ξ_3) را می‌توان با استفاده از رابطه‌ی (۷) بدست آورد، همچنین احتمالات مربوطه عبارتند از

$$\begin{aligned}
 p_{111} &= \bar{F}(t, t, t), \quad p_{11.} = \bar{F}_{X,Y}(t, t) - \bar{F}(t, t, t), \\
 p_{1.1} &= \bar{F}_{X,Z}(t, t) - \bar{F}(t, t, t), \quad p_{1..} = F_{Y,Z}(t, t) - F(t, t, t), \\
 p_{.11} &= \bar{F}_{Y,Z}(t, t) - \bar{F}(t, t, t), \quad p_{.1.} = F_{X,Z}(t, t) - F(t, t, t), \\
 p_{..1} &= F_{X,Y}(t, t) - F(t, t, t), \quad p_{...} = F(t, t, t).
 \end{aligned}$$

۲.۳ میانگین باقیمانده‌ی عمر

یک سیستم مرکب $(1, 1, 1)$ از n را در نظر بگیرید به قسمی که تمام مؤلفه‌های آن در زمان t فعال باشند. با توجه به طول عمر این سیستم، باقیمانده عمر آن برابر است با

$$\begin{aligned}
 T_{1,1,1;n;t} &= \{X_{n:n} - t, Y_{n:n} - t, Z_{n:n} - t \mid \text{باشد } t \text{ فعال}\} \\
 &= \{X_{n:n} - t, Y_{n:n} - t, Z_{n:n} - t \mid X_{1:n} > t, Y_{1:n} > t, Z_{1:n} > t\},
 \end{aligned}$$

بنابراین تابع MRL سیستم عبارتست از

$$\Phi_{1,1,1;n}(t) = E(T_{1,1,1;n;t}) = \int_0^\infty P(T_{1,1,1;n;t} > x) dx. \quad (9)$$

رزمخواه و صابرزاده (۲۰۱۷) نشان داده‌اند که

$$\begin{aligned}
 P(T_{1,1:n;t} > x) = & \left\{ [\bar{F}(t, t, t)] \right\}^{-n} \left\{ [\bar{F}(t, t, t)]^n - [\bar{F}(t, t, t) - \bar{F}(t+x, t, t)]^n \right. \\
 & - [\bar{F}(t, t, t) - \bar{F}(t, t, t+x)]^n - [\bar{F}(t, t, t) - \bar{F}(t, t+x, t)]^n \\
 & + [\bar{F}(t, t, t) - \bar{F}(t+x, t, t) - \bar{F}(t, t+x, t) + \bar{F}(t+x, t+x, t)]^n \\
 & + [\bar{F}(t, t, t) - \bar{F}(t+x, t, t) - \bar{F}(t, t, t+x) + \bar{F}(t+x, t, t+x)]^n \\
 & + [\bar{F}(t, t, t) - \bar{F}(t, t+x, t) - \bar{F}(t, t, t+x) + \bar{F}(t, t+x, t+x)]^n \\
 & - [\bar{F}(t, t, t) - \bar{F}(t+x, t, t) - \bar{F}(t, t+x, t) + \bar{F}(t+x, t+x, t)] \\
 & - [\bar{F}(t, t, t+x) - \bar{F}(t+x, t, t+x) + \bar{F}(t, t+x, t+x) \\
 & \left. - \bar{F}(t+x, t+x, t+x)]^n \right\}. \quad (۱۰)
 \end{aligned}$$

حال فرض کنید بردار تصادفی (X_i, Y_i, Z_i) در i امین جزء $(i = 1, 2, \dots, n)$ ، از مدل FGM پیروی کند، یعنی دارای تابع توزیع احتمال توأم زیر باشد

$$\begin{aligned}
 F(x, y, z) = & F_X(x)F_Y(y)F_Z(z) \{ 1 + \alpha(1 - F_X(x))(1 - F_Y(y)) \\
 & + \beta(1 - F_X(x))(1 - F_Z(z)) + \gamma(1 - F_Y(y))(1 - F_Z(z)) \\
 & + \theta(1 - F_X(x))(1 - F_Y(y))(1 - F_Z(z)) \},
 \end{aligned}$$

که در آن $\alpha, \beta, \gamma, \theta$ به ترتیب پارامترهای وابستگی بین متغیرهای (X, Y) ، (X, Z) ، (Y, Z) و (X, Y, Z) هستند طوری که $\alpha, \beta, \gamma, \theta \in [-1, 1]$. در این مدل، روابط زیر را داریم

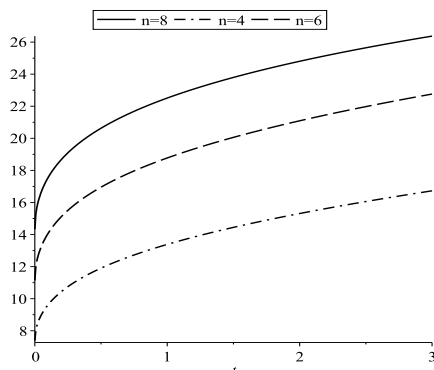
$$\begin{aligned}
 F_{X,Y}(x, y) &= F_X(x)F_Y(y) \{ 1 + \alpha(1 - F_X(x))(1 - F_Y(y)) \}, \\
 F_{X,Z}(x, z) &= F_X(x)F_Z(z) \{ 1 + \beta(1 - F_X(x))(1 - F_Z(z)) \}, \\
 F_{Y,Z}(y, z) &= F_Y(y)F_Z(z) \{ 1 + \gamma(1 - F_Y(y))(1 - F_Z(z)) \}.
 \end{aligned}$$

همچنین فرض کنید توابع توزیع حاشیه‌ای $W(0.5, 0.5)$ باشند. رفتار MRL نسبت به t و به ازای مقادیر $\alpha = \beta = \gamma = \theta = 0.5$ و $n = 4, 6, 8$ در شکل ۳ آمده است. ملاحظه می‌کنید که مشابه سیستم‌های مرکب دومؤلفه‌ای، میانگین باقیمانده عمر این سیستم نیز نسبت به t اکیدا صعودی می‌باشد. همچنین با افزایش تعداد اجزای سیستم، MRL آن افزایش می‌یابد. سیستم‌های مرکب دومؤلفه‌ای در مقایسه با سیستم‌های سه‌مؤلفه‌ای تحت توزیع‌های حاشیه‌ای و پارامترهای وابستگی یکسان، MRL بیشتری دارند.

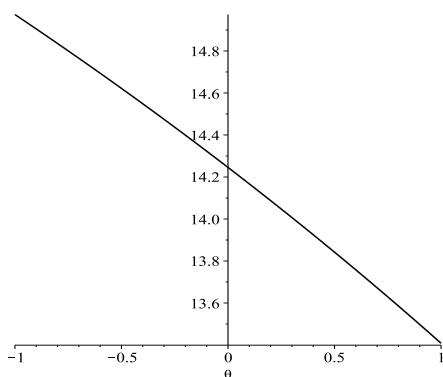
در شکل ۴ رفتار MRL سیستم را نسبت به θ و به ازای $n = 4$ و $\alpha = \beta = \gamma = 0.5$ ، $t = 1/2$ مشاهده می‌کنید. با توجه به شکل ملاحظه می‌شود که با افزایش وابستگی بین سه مؤلفه X, Y و Z میانگین باقیمانده عمر سیستم کاهش می‌یابد.

۴ بحث و نتیجه‌گیری

در این مقاله سیستم‌های مرکب n جزئی را در نظر گرفتیم که هر جزء آن‌ها شامل دو یا سه مؤلفه وابسته باشند و قابلیت اعتماد چنین سیستم‌هایی را مورد بررسی قرار دادیم. همچنین تابع MRL این سیستم‌های مرکب را با فرض اینکه همه‌ی مؤلفه‌های آن‌ها در زمان t فعال باشند را ارزیابی نمودیم. تحت مدل FGM و توابع چگالی حاشیه‌ای وایبول در هر کدام از سیستم‌های مرکب دومؤلفه‌ای و سه‌مؤلفه‌ای، به بررسی نحوه‌ی تغییرات MRL این سیستم‌ها نسبت به برخی از پارامترها پرداختیم. در حالت کلی این نتایج را می‌توان به سیستم‌های مرکب با بیش از سه مؤلفه وابسته در هر جزء تعمیم داد. هدف اصلی این مقاله محاسبه‌ی



شکل ۳: رفتار تابع MRL سیستم مرکب سه مؤلفه‌ای نسبت به t و θ تحت حاشیه‌ای‌های وایبول.



شکل ۴: رفتار تابع MRL سیستم مرکب سه مؤلفه‌ای نسبت به θ تحت حاشیه‌ای‌های وایبول.

میانگین باقیمانده‌ی عمر سیستم‌های مرکب دو یا سه مؤلفه‌ای با فرض فعال بودن همه‌ی مؤلفه‌ها در زمان t بود. همچنین در این سیستم‌ها برای بیان وابستگی بین متغیرها از مفصل FGM استفاده نمودیم و توابع توزیع حاشیه‌ای را وایبول در نظر گرفتیم. بوضوح امکان استفاده از سایر کاپولاها برای بیان ارتباط بین مؤلفه‌ها وجود دارد. همچنین می‌توان حالت‌هایی را در نظر گرفت که در زمان t برخی از مؤلفه‌های سیستم فعال باشند.

مراجع

- [1] Aitken, A. C. and Gonin, H. T.(1935), *On fourfold sampling with and without replacement*, in Proc. Roy. Soc. Edinburgh, 55, 114–125.
- [2] Asadi, M. and Bayramoglu, I. (2005), *A note on the mean residual life function of a parallel system*, Commun. Statist.-Theory Methods, 34, 475–484.

- [3] Asadi, M. and Bayramoglu, I. (2006), *The mean residual life function of a k-out-of-n structure at the system level*, IEEE Trans. Rel., 55, 314–318.
- [4] Bairamov, I., Ahsanullah, M. and Akhundov, I. (2002), *A residual life function of a system having parallel or series structures*, J. Statist. Theory Appl.
- [5] Bairamov, I. and Arnold, B. C. (2008), *On the residual lifelengths of the remaining components in an n-k+1 out of n system*, Statist. Probab. Lett., 78, 945–952.
- [6] Bairamov, I. (2013), *Reliability and mean residual life of complex systems with two dependent components per element*, IEEE Trans. Rel., 62, 276–285.
- [7] Barlow, R. E. and Proschan, F. (1975), *Statistical theory of reliability and life testing*, probability models; Holt, Rinehart and Winston, New York.
- [8] Eryilmaz, S. (2012), *On the Mean residual life of a k-out-of n: G system with a single cold standby component*, Journal of European Operational Research, 222, 273–277.
- [9] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997), *Discrete Multivariate Distributions*. New York, Wiley.
- [10] Nelsen, R. (2006), *An Introduction to Copulas*. Second Edition, Springer Verlag.
- [11] Razmkhah, M. and Saberzadeh, Z. (2017), *On the reliability of complex systems with three dependent components per element*, JIRSS, to appear.
- [12] Tavangar, M. and Asadi, M. (2010), *A study on the mean past lifetime of the components of (n-k+1)-out-of-n system at the system level*, Metrika, 72, 59–73.



خانواده‌ای از توزیع‌های دو متغیره وایبول تعمیم یافته نمایی شده- سری توانی

عابدینی، آ^۱ روزگار، ر^۲

^{۱،۲} گروه آمار، دانشکده علوم ریاضی، دانشگاه یزد

چکیده

در این مقاله، یک کلاس جدید از توزیع‌های دو متغیره تحت عنوان توزیع‌های دو متغیره وایبول تعمیم یافته نمایی شده- سری توانی را معرفی می‌کنیم. کلاس نام‌برده شامل چندین مدل از جمله: توزیع دو متغیره وایبول تعمیم یافته نمایی شده، توزیع دو متغیره وایبول تعمیم یافته نمایی شده- هندسی، -پواسن، -لگاریتمی، -دوجمله‌ای، -دوجمله‌ای منفی می‌باشد. در ادامه نحوه‌ی تشکیل و ویژگی‌های این کلاس جدید از توزیع‌های دو متغیره ارائه شده و پس از آن طریقه‌ی برآوردیابی پارامترهای مدل بر اساس تابع درست‌نمایی و الگوریتم EM ارائه شده است. این مقاله با بیان یک کاربرد از داده‌های واقعی برای برازش این مدل به پایان می‌رسد.

کلمات کلیدی: توزیع دو متغیره وایبول تعمیم یافته نمایی شده، توزیع‌های سری توانی، برآوردگر ماکسیمم درست‌نمایی، الگوریتم EM.

۱ مقدمه

در چند سال گذشته طریقه‌ی ترکیب توزیع‌های پیوسته با یک توزیع گسسته بسیار مورد توجه نویسندگان قرار گرفته است. این روش منجر می‌شود توزیع‌های جدید با انعطاف‌پذیری بالا که از لحاظ کاربرد برازش روی داده‌های واقعی مناسب‌تر می‌باشد، تولید شوند. اخیراً در حوزه‌ی مدل‌سازی

^۱ asefehabedini@stu.yazd.ac.ir

^۲ E-mail: rroozegar@yazd.ac.ir

داده‌های طول عمر که یکی از جنبه‌های اساسی استنباط آماری در زمینه‌های گوناگون علوم و تحقیقات به‌شمار می‌رود، توزیع‌های تک‌متغیره بسیار مورد مطالعه قرار گرفته است و کم‌تر نویسنده‌ای در زمینه‌ی توزیع‌های آماری دو و چندمتغیره به دلیل پیچیدگی تحلیل‌ها به مطالعه پرداخته است. در بین کارهای انجام شده در زمینه‌ی ترکیب کردن توزیع‌های پیوسته با یک توزیع گسسته خاص مانند: هندسی، پواسون، لگاریتمی، دو جمله‌ای و دو جمله‌ای منفی که حالت‌های خاص توزیع سری توانی می‌باشند، می‌توان موارد زیر را نام برد: فلورس (۲۰۱۳) [۱]، مورایس و بارتو-سوزا (۲۰۱۱) [۲] و روزگار و نادارما (۲۰۱۶) [۴]. تابع توزیع و تابع چگالی توزیع وایبول تعمیم یافته نامی شده (EEW) به ترتیب به صورت زیر هستند:

$$F_{EEW}(x; \alpha, \lambda, \xi) = (1 - e^{-\lambda H(x; \xi)})^\alpha, \quad x > 0, \alpha > 0, \lambda > 0,$$

$$f_{EEW}(x; \alpha, \lambda, \xi) = \alpha \lambda h(x; \xi) e^{-\lambda H(x; \xi)} (1 - e^{-\lambda H(x; \xi)})^{\alpha-1}.$$

اخیرا روزگار و جعفری (۲۰۱۶) [۳] به معرفی توزیع دومتغیره‌ی وایبول تعمیم یافته نامی شده (BEEW) و ارائه برخی ویژگی‌های مرتبط با آن پرداخته‌اند. فرم تابع توزیع دومتغیره به صورت زیر می‌باشد

$$F_{BEEW}(x_1, x_2) = \prod_{i=1}^2 F_{EEW}(x_i; \alpha_i, \lambda, \xi),$$

بطوری‌که

$$F_{BEEW}(x_1, x_2) = (1 - e^{-\lambda H(x_1; \xi)})^{\alpha_1} (1 - e^{-\lambda H(x_2; \xi)})^{\alpha_2} (1 - e^{-\lambda H(z; \xi)})^{\alpha_z},$$

که در عبارت بالا: $z = \min(x_1, x_2)$. این توزیع دومتغیره هم دارای بخش پیوسته مطلق و هم دارای بخش منفرد می‌باشد. در این مقاله توزیع BEEW را با کلاس توزیع سری توانی (PS) ترکیب کرده‌ایم و کلاس جدیدی از توزیع‌های دومتغیره به نام توزیع دومتغیره وایبول تعمیم یافته نامی شده-سری توانی (BEEWPS) ارائه شده است. توزیع‌های حاشیه‌ای کلاس فوق توسط طهماسبی و جعفری (۲۰۱۵) [۵] تحت عنوان وایبول تعمیم یافته نامی شده-سری توانی معرفی گردیده است.

ادامه‌ی این مقاله به شرح زیر است: توزیع BEEWPS را با ویژگی‌های گوناگون آن معرفی کرده‌ایم. برخی از حالت‌های خاص را در نظر می‌گیریم. برآورد پارامترهای مجهول را توسط الگوریتم EM بیان می‌کنیم و در نهایت کاربرد این مدل روی یک مجموعه داده‌ی واقعی را بررسی می‌کنیم.

۲ دست‌آوردهای پژوهش

۱.۲ کلاس BEEWPS

فرض می‌کنیم متغیر تصادفی گسسته‌ی N (توزیع سری توانی) با تابع جرم احتمال:

$$P(N = n) = \frac{a_n \theta^n}{A(\theta)}, \quad n = 1, 2, \dots,$$

است که در آن $a_n > 0$ وابسته به n است و $A(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$ ، $\theta \in (0, s)$ می‌تواند ∞ را نیز اختیار کند) است، به‌طوری‌که $A(\theta)$ متناهی است. برخی موارد خاص توزیع سری توانی (بریده شده در صفر) به صورت زیر می‌باشد: توزیع‌های هندسی، پواسون، لگاریتمی، دو جمله‌ای و دو جمله‌ای منفی با a_n ها به ترتیب $1, n^{-1}, \binom{k}{n}, \binom{k-1}{n-1}$ و $A(\theta)$ ها به ترتیب $e^\theta - 1, \theta(1 - \theta)^{-1}, -\log(1 - \theta), (1 + \theta)^k - 1, \frac{\theta^k}{(1 - \theta)^k}$ می‌باشد. در این‌جا منظور از $A'(\theta)$ ، $A''(\theta)$ و $A'''(\theta)$ به ترتیب مشتق اول، مشتق دوم و مشتق سوم $A(\theta)$ نسبت به پارامتر θ می‌باشد.

حال فرض می‌کنیم $\{X_{1n}, X_{2n}; n = 1, 2, 3, \dots\}$ دنباله‌ای از بردارهای تصادفی مستقل و هم‌توزیع غیرمنفی با تابع توزیع توام $F_X(\cdot, \cdot)$ می‌باشد، که $X = (X_1, X_2)$ است. در این‌جا N متغیر تصادفی سری توانی و مستقل از (X_{1i}, X_{2i}) در نظر می‌گیریم. همچنین در نظر می‌گیریم

$$Y_i = \max\{X_{i1}, \dots, X_{iN}\}, \quad i = 1, 2.$$

از این رو برای متغیر تصادفی توام $Y = (Y_1, Y_2)$ داریم:

$$P(Y_1 \leq y_1, Y_2 \leq y_2, N = n) = F_{Y_1, Y_2 | N}(y_1, y_2 | n) P(N = n) = (F_X(y_1, y_2))^n \frac{a_n \theta^n}{A(\theta)}. \quad (1)$$

بنابراین تابع توزیع تجمعی توام آن به صورت زیر می باشد:

$$F_Y(y_1, y_2) = \sum_{n=1}^{\infty} (F_X(y_1, y_2))^n \frac{a_n \theta^n}{A(\theta)} = \frac{A(\theta F_X(y_1, y_2))}{A(\theta)}.$$

در این مورد می گوئیم Y دارای یک توزیع دومتغیره F -سری توانی است. تابع توزیع حاشیه ای Y_i به صورت زیر نوشته می شود:

$$F_{Y_i}(y_i) = \frac{A(\theta F_{X_i}(y_i))}{A(\theta)}, \quad i = 1, 2.$$

گزاره ۱.۲. اگر فرض کنیم $Z_i = \min\{X_{i1}, \dots, X_{iN}\}$, $i = 1, 2$ کلاس دیگری از توزیع دومتغیره باشد، دارای تابع توزیع زیر است:

$$\bar{F}_{Z_1, Z_2}(y_1, y_2) = P(Z_1 > y_1, Z_2 > y_2) = \frac{A(\theta \bar{F}_X(y_1, y_2))}{A(\theta)},$$

که $\bar{F}_X(y_1, y_2) = P(X_1 > y_1, X_2 > y_2)$ می باشد.

در این مقاله F را توزیع دومتغیره وایبول تعمیم یافته نمایی شده که در معادله ۱ آمده در نظر می گیریم. از این رو کلاس توزیع های دومتغیره وایبول تعمیم یافته نمایی شده- سری توانی ($BEEWPS$) دارای تابع توزیع زیر است:

$$F_Y(y_1, y_2) = \begin{cases} \frac{A(\theta F_{BEEW}(y_1; \alpha_1 + \alpha_2, \lambda, \xi) F_{BEEW}(y_2; \alpha_2, \lambda, \xi))}{A(\theta)} & \text{if } y_1 \leq y_2 \\ \frac{A(\theta F_{BEEW}(y_1; \alpha_1, \lambda, \xi) F_{BEEW}(y_2; \alpha_2 + \alpha_3, \lambda, \xi))}{A(\theta)} & \text{if } y_1 > y_2 \end{cases} \quad (2)$$

که آن را با $BEEWPS(\alpha_1, \alpha_2, \alpha_3, \lambda, \xi, \theta)$ نام گذاری می کنیم. تابع چگالی Y به سادگی و با استفاده از $f_Y(y_1, y_2) = \frac{\partial^2 F_Y(y_1, y_2)}{\partial y_1 \partial y_2}$ حاصل می گردد.

گزاره ۲.۲. اگر (Y_1, Y_2) دارای توزیع $BEEWPS(\alpha_1, \alpha_2, \alpha_3, \lambda, \xi, \theta)$ باشد، آنگاه:

الف: Y_i دارای توزیع $EEWPS$ با پارامترهای $\lambda, \alpha_i + \alpha_3$ و ξ و θ است.

ب: متغیر تصادفی $U = \max(Y_1, Y_2)$ دارای توزیع $EEWPS$ با پارامترهای $\lambda, \alpha_1 + \alpha_2 + \alpha_3$ و ξ و θ است.

ج: اگر $A(\theta) = \theta$ آنگاه (Y_1, Y_2) دارای توزیع $BEEW$ با پارامترهای $\alpha_1, \alpha_2, \alpha_3, \lambda$ و ξ است.

د: $P(Y_1 < Y_2) = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3}$.

قضیه ۱.۲. اگر Y دارای توزیع $BEEWPS(\alpha_1, \alpha_2, \alpha_3, \lambda, \xi, \theta)$ باشد، آنگاه تابع چگالی توام Y به صورت زیر است:

$$f_Y(y_1, y_2) = \begin{cases} f_1(y_1, y_2) & \text{if } 0 < y_1 < y_2 \\ f_2(y_1, y_2) & \text{if } 0 < y_2 < y_1 \\ f \cdot(y) & \text{if } 0 < y_1 = y_2 = y, \end{cases} \quad (3)$$

که

$$\begin{aligned} f_1(y_1, y_2) &= \frac{\theta}{A(\theta)} f_{EEW}(y_1; \alpha_1 + \alpha_3, \lambda, \xi) f_{EEW}(y_2; \alpha_2, \lambda, \xi) [\theta f_{EEW}(y_1; \alpha_1 + \alpha_3, \lambda, \xi) \\ &\times f_{EEW}(y_2; \alpha_2, \lambda, \xi) A''(\theta f_{EEW}(y_1; \alpha_1 + \alpha_3, \lambda, \xi) f_{EEW}(y_2; \alpha_2, \lambda, \xi)) + A'''(\theta f_{EEW}(y_1; \alpha_1 + \alpha_3, \lambda, \xi) f_{EEW}(y_2; \alpha_2, \lambda, \xi))], \\ f_2(y_1, y_2) &= \frac{\theta}{A(\theta)} f_{EEW}(y_1; \alpha_1, \lambda, \xi) f_{EEW}(y_2; \alpha_2 + \alpha_3, \lambda, \xi) [\theta f_{EEW}(y_1; \alpha_1, \lambda, \xi) \\ &\times f_{EEW}(y_2; \alpha_2 + \alpha_3, \lambda, \xi) A''(\theta f_{EEW}(y_1; \alpha_1, \lambda, \xi) f_{EEW}(y_2; \alpha_2 + \alpha_3, \lambda, \xi)) + A'''(\theta f_{EEW}(y_1; \alpha_1, \lambda, \xi) f_{EEW}(y_2; \alpha_2 + \alpha_3, \lambda, \xi))], \\ f \cdot(y) &= \frac{\theta \alpha_3}{A(\theta)(\alpha_1 + \alpha_2 + \alpha_3)} f_{EEW}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda, \xi) A'(\theta f_{EEW}(y; \alpha_1 + \alpha_2 + \alpha_3, \lambda, \xi)). \end{aligned}$$

گزاره ۳.۲. با توجه به عبارات $\sum_{n=1}^{\infty} n^r a_n \theta^{n-1}$ و $\theta A''(\theta) + A'(\theta) = \sum_{n=1}^{\infty} n^r a_n \theta^{n-1}$ داریم:

$$f_1(y_1, y_2) = \sum_{n=1}^{\infty} p_n f_{EEW}(y_1; n\alpha_1 + n\alpha_r, \lambda, \xi) f_{EEW}(y_2; n\alpha_2, \lambda, \xi),$$

$$f_r(y_1, y_2) = \sum_{n=1}^{\infty} p_n f_{EEW}(y_1; n\alpha_1, \lambda, \xi) f_{EEW}(y_2; n\alpha_2 + n\alpha_r, \lambda, \xi),$$

$$f_{\cdot}(y) = \frac{\alpha_r}{\alpha_1 + \alpha_2 + \alpha_r} \sum_{n=1}^{\infty} p_n f_{EEW}(y; n\alpha_1 + n\alpha_2 + n\alpha_r, \lambda, \xi),$$

که $p_n = P(N = n) = \frac{a_n \theta^n}{A(\theta)}$ و $f_{EEW}(\cdot; n\alpha, \lambda, \xi)$ تابع چگالی توزیع EEW با پارامترهای $n\alpha$ ، λ و ξ است. توجه داشته باشید که $f_{EEW}(\cdot; n\alpha, \lambda, \xi)$ تابع چگالی متغیر تصادفی $(U_1, \dots, U_n)$ است.

نتیجه ۱.۲. تابع چگالی توام BEEWPS را به صورت زیر نیز می توان نوشت

$$f_Y(y_1, y_2) = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_r} g_a(y_1, y_2) + \frac{\alpha_r}{\alpha_1 + \alpha_2 + \alpha_r} g_s(y),$$

که

$$g_a(y_1, y_2) = \frac{\alpha_1 + \alpha_2 + \alpha_r}{\alpha_1 + \alpha_2} \begin{cases} f_1(y_1, y_2) & \text{if } y_1 < y_2 \\ f_r(y_1, y_2) & \text{if } y_1 > y_2, \end{cases}$$

$$g_s(y) = \frac{\theta}{A(\theta)} f_{EEW}(y; \alpha_1 + \alpha_2 + \alpha_r, \lambda, \xi) A'(\theta F_{EEW}(y; \alpha_1 + \alpha_2 + \alpha_r, \lambda, \xi)) \text{ if } y_1 = y_2 = y.$$

و جاهای دیگر صفر است. واضح است که $f_a(\cdot, \cdot)$ بخش پیوسته مطلق و $f_s(\cdot, \cdot)$ بخش منفرد تابع چگالی توام می باشد. بنابراین، اگر $\alpha_r = 0$ باشد، تابع چگالی توام مدل BEEWPS قسمت منفرد ندارد و تبدیل به یک توزیع پیوسته مطلق می شود.

گزاره ۴.۲. توزیع شرطی Y_1 به شرط $Y_2 = y_2$ مطلقاً پیوسته است و دارای تابع توزیع زیر می باشد

$$P(Y_1 < y_1 | Y_2 < y_2) = \begin{cases} \frac{A(\theta(1-e^{-\lambda H(y_1; \xi)})^{\alpha_1 + \alpha_r} (1-e^{-\lambda H(y_2; \xi)})^{\alpha_2})}{A(\theta(1-e^{-\lambda H(y_2; \xi)})^{\alpha_2 + \alpha_r} (1-e^{-\lambda H(y_1; \xi)})^{\alpha_1})} & \text{if } y_1 < y_2 \\ \frac{A(\theta(1-e^{-\lambda H(y_2; \xi)})^{\alpha_2 + \alpha_r} (1-e^{-\lambda H(y_1; \xi)})^{\alpha_1})}{A(\theta(1-e^{-\lambda H(y_2; \xi)})^{\alpha_2 + \alpha_r} (1-e^{-\lambda H(y_1; \xi)})^{\alpha_1})} & \text{if } y_2 < y_1. \end{cases}$$

گزاره ۵.۲. حد توزیع BEEWPS هنگامی که θ به سمت $+$ میل کند به صورت زیر نوشته می شود

$$\begin{aligned} \lim_{\theta \rightarrow +} F_Y(y_1, y_2) &= \lim_{\theta \rightarrow +} \frac{A(\theta F_X(y_1, y_2))}{A(\theta)} = \lim_{\theta \rightarrow +} \frac{\sum_{n=1}^{\infty} a_n \theta^n (F_X(y_1, y_2))^n}{\sum_{n=1}^{\infty} a_n \theta^n} \\ &= \lim_{\theta \rightarrow +} \frac{a_c (F_X(y_1, y_2))^c + \sum_{n=c+1}^{\infty} a_n \theta^{n-c} (F_X(y_1, y_2))^n}{a_c + \sum_{n=c+1}^{\infty} a_n \theta^{n-c}} = (F_X(y_1, y_2))^c \\ &= \begin{cases} (1 - e^{-\lambda H(y_1; \xi)})^{c(\alpha_1 + \alpha_r)} (1 - e^{-\lambda H(y_2; \xi)})^{c\alpha_2} & \text{if } y_1 \leq y_2 \\ (1 - e^{-\lambda H(y_1; \xi)})^{c\alpha_1} (1 - e^{-\lambda H(y_2; \xi)})^{c(\alpha_2 + \alpha_r)} & \text{if } y_1 > y_2, \end{cases} \end{aligned}$$

این رابطه تابع چگالی BEEW با پارامترهای $c\alpha_1$ ، $c\alpha_2$ ، $c\alpha_r$ ، λ و ξ است که $c = \min\{n \in N : a_n > 0\}$.

برای متغیرهای تصادفی (Y_1, Y_2, N) معادله (۱) زمانی که X دارای توزیع BEEW است را در نظر می گیریم. تا زمانی که $(Y_1, Y_2 | N = n)$ دارای توزیع $BEEW(n\alpha_1, n\alpha_2, n\alpha_r, \lambda, \xi)$ است، تابع چگالی توام (Y_1, Y_2, N) به صورت زیر است:

$$f_{Y_1, Y_2, N}(y_1, y_2, n) = \begin{cases} \frac{n^r a_n (\theta K_1(y_1, y_2))^{n-1}}{h_1(y_1, y_2)} & \text{if } y_1 < y_2 \\ \frac{n^r a_n (\theta K_2(y_1, y_2))^{n-1}}{h_2(y_1, y_2)} & \text{if } y_2 < y_1 \\ \frac{n a_n (\theta K_{\cdot}(y))^{n-1}}{h_{\cdot}(y)} & \text{if } y_1 = y_2 = y, \end{cases}$$

که

$$\begin{aligned}
h_1(y_1, y_2) &= \theta F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) F_{EEW}(y_2; \alpha_r, \lambda, \xi) A''(\theta F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) \\
&\quad \times F_{EEW}(y_2; \alpha_r, \lambda, \xi)) + A'(\theta F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) F_{EEW}(y_2; \alpha_r, \lambda, \xi)), \\
h_2(y_1, y_2) &= \theta F_{EEW}(y_1; \alpha_1, \lambda, \xi) F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi) A''(\theta F_{EEW}(y_1; \alpha_1, \lambda, \xi) \\
&\quad \times F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi)) + A'(\theta F_{EEW}(y_1; \alpha_1, \lambda, \xi) F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi)), \\
h_{\cdot}(y) &= A'(\theta F_{EEW}(y; \alpha_1 + \alpha_r + \alpha_r, \lambda, \xi)), \\
K_1(y_1, y_2) &= (1 - e^{-\lambda H(y_1; \xi)})^{\alpha_1 + \alpha_r} (1 - e^{-\lambda H(y_2; \xi)})^{\alpha_r} \\
&= F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) F_{EEW}(y_2; \alpha_r, \lambda, \xi), \\
K_2(y_1, y_2) &= (1 - e^{-\lambda H(y_1; \xi)})^{\alpha_1} (1 - e^{-\lambda H(y_2; \xi)})^{\alpha_r + \alpha_r} \\
&= F_{EEW}(y_1; \alpha_1, \lambda, \xi) F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi), \\
K_{\cdot}(y) &= (1 - e^{-\lambda H(y; \xi)})^{\alpha_1 + \alpha_r + \alpha_r} = F_{EEW}(y; \alpha_1 + \alpha_r + \alpha_r, \lambda, \xi).
\end{aligned}$$

بنابراین ما می‌توانیم امید شرطی N به شرط $Y_1 = y_1$ و $Y_2 = y_2$ را به صورت زیر بنویسیم

$$E(N|y_1, y_2) = \begin{cases} \frac{R_1(y_1, y_2)}{h_1(y_1, y_2)} & \text{if } y_1 < y_2 \\ \frac{R_2(y_1, y_2)}{h_2(y_1, y_2)} & \text{if } y_2 < y_1 \\ \frac{\theta K_{\cdot}(y) A''(\theta K_{\cdot}(y)) + A'(\theta K_{\cdot}(y))}{h_{\cdot}(y)} & \text{if } y_1 = y_2 = y, \end{cases}$$

که

$$\begin{aligned}
R_i(y_1, y_2) &= (\theta K_i(y_1, y_2))^{\tau} A'''(\theta K_i(y_1, y_2)) + \tau \theta K_i(y_1, y_2) A''(\theta K_i(y_1, y_2)) \\
&\quad + A'(\theta K_i(y_1, y_2)), \quad i = 1, 2.
\end{aligned}$$

۲.۲ حالت خاص توزیع BEEWPS

در این بخش، یک حالت خاص از چندین حالت توزیع BEEWPS را در نظر می‌گیریم که در آن $t_i = 1 - e^{-\lambda H(y_i; \xi)}$ برای $i = 1, 2$ است.

۱.۲.۲ توزیع دومتغیره وایبول تعمیم یافته نمایی شده- سری هندسی:

توزیع هندسی (بریده شده در صفر) حالت خاصی از توزیع سری توانی است با قرار دادن $a_n = 1$ و $A(\theta) = \frac{\theta}{1-\theta}$ برای $0 < \theta < 1$. بنابراین تابع توزیع توام توزیع BEEWG عبارت است از

$$F_Y(y_1, y_2) = \begin{cases} \frac{(1-\theta)t_1^{\alpha_1 + \alpha_r} t_2^{\alpha_r}}{1 - \theta t_1^{\alpha_1 + \alpha_r} t_2^{\alpha_r}} & \text{if } y_1 < y_2 \\ \frac{(1-\theta)t_1^{\alpha_1} t_2^{\alpha_r + \alpha_r}}{1 - \theta t_1^{\alpha_1} t_2^{\alpha_r + \alpha_r}} & \text{if } y_2 < y_1 \end{cases}$$

و تابع چگالی توام آن به صورت ۳ است با

$$\begin{aligned}
f_1(y_1, y_2) &= (1 - \theta) f_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) f_{EEW}(y_2; \alpha_r, \lambda, \xi) \\
&\quad \times \frac{1 + \theta F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) F_{EEW}(y_2; \alpha_r, \lambda, \xi)}{(1 - \theta F_{EEW}(y_1; \alpha_1 + \alpha_r, \lambda, \xi) F_{EEW}(y_2; \alpha_r, \lambda, \xi))^{\tau}}, \\
f_2(y_1, y_2) &= (1 - \theta) f_{EEW}(y_1; \alpha_1, \lambda, \xi) f_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi) \\
&\quad \times \frac{1 + \theta F_{EEW}(y_1; \alpha_1, \lambda, \xi) F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi)}{(1 - \theta F_{EEW}(y_1; \alpha_1, \lambda, \xi) F_{EEW}(y_2; \alpha_r + \alpha_r, \lambda, \xi))^{\tau}}, \\
f_{\cdot}(y) &= \frac{(1 - \theta) \alpha_r f_{EEW}(y; \alpha_1 + \alpha_r + \alpha_r, \lambda, \xi)}{(\alpha_1 + \alpha_r + \alpha_r)(1 - \theta F_{EEW}(y; \alpha_1 + \alpha_r + \alpha_r, \lambda, \xi))^{\tau}}.
\end{aligned}$$

تذکر ۱۰۲. هنگامی که $\theta^* = 1 - \theta$ توزیع $BEEWG$ به توزیع مارشال-الکین (۱۹۹۷) تبدیل می‌شود.

۳.۲ برآوردیابی

مراحل زیر می‌تواند برای محاسبه برآوردگرهای ماکسیمم درستنمایی پارامترها بر اساس الگوریتم EM استفاده شوند:

مرحله ۱: یک مقدار اولیه برای θ در نظر گرفته می‌شود، مثلاً $\Theta^{(0)} = (\alpha_1^{(0)}, \alpha_2^{(0)}, \alpha_3^{(0)}, \lambda^{(0)}, \xi^{(0)}, \theta^{(0)})'$.

مرحله ۲: $b_i = E(N|y_{1i}, y_{2i}; \Theta^{(0)})$ محاسبه می‌شود.

مرحله ۳: v_1, u_1, v_2, u_2 را محاسبه می‌کنیم.

مرحله ۴: تابع شبه لگاریتم درستنمایی $l_{pseudo}(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)}, \lambda, \xi, \theta^{(1)})$ را نسبت به λ و ξ ماکسیمم می‌کنیم و آن‌ها را به ترتیب $\hat{\lambda}^{(1)}$ و $\hat{\xi}^{(1)}$ می‌نامیم.

مرحله ۵: $\hat{\alpha}_i^{(1)} = \hat{\alpha}_i(\hat{\lambda}^{(1)}, \hat{\xi}^{(1)})$ را محاسبه می‌کنیم.

مرحله ۶: $\hat{\theta}$ را پیدا می‌کنیم و آن را $\theta^{(1)}$ می‌نامیم.

مرحله ۷: $\Theta^{(1)} = (\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(1)}, \lambda^{(1)}, \xi^{(1)}, \theta^{(1)})$ جایگذاری می‌شود و مراحل را تا زمانی که فرآیند همگرا شود ادامه می‌دهیم.

با استفاده از الگوریتم EM پارامترها به صورت زیر برآورد شدند

$$\hat{\alpha}_1(\lambda, \xi) = \frac{m_1 u_1 + m_2}{\sum_{i \in I_1} b_i Q_i + \sum_{i \in I_1 \cup I_2} b_i Q_i}, \quad \hat{\alpha}_2(\lambda, \xi) = \frac{m_2 v_1 + m_1}{\sum_{i \in I_1} b_i Q_i + \sum_{i \in I_1 \cup I_2} b_i Q_i},$$

$$\hat{\alpha}_3(\lambda, \xi) = \frac{m_1 + m_2 u_1 + m_2 v_2}{\sum_{i \in I_1} b_i Q_i + \sum_{i \in I_1} b_i Q_i + \sum_{i \in I_2} b_i Q_i}$$

۴.۲ مثال عددی

در این بخش توزیع مورد نظر را برای برازش به یک مجموعه داده مورد استفاده قرار می‌دهیم. در این جا شش زیرکلاس از این توزیع را در نظر گرفته ایم و نتایج را در جدول ۱ آورده‌ایم، یک مجموعه داده که مربوط به لیگ فوتبال آمریکا برای طی سه هفته متوالی مسابقات در سال ۱۹۸۶ می‌باشد و اولین بار توسط واشنگتن پست ارائه شده است را در نظر گرفته ایم، در این جا Y_1 نشان دهنده زمانی است که اولین امتیاز با ضربه به توپ حاصل شده است و Y_2 نشان دهنده زمانی است که اولین امتیاز با انتقال توپ به آخر زمین حاصل شده است. در اینجا تمام داده‌ها را بر ۱۰۰ تقسیم کرده ایم. با توجه به جدول می‌توان مشاهده کرد که تمام مدل‌ها برای این داده‌ها مناسب هستند.

جدول ۱: نتایج برای شش توزیع برازش شده

توزیع						آماره
BEEWNB	BEEWL	BEEWB	BEEWP	BEEWG	BEEW	
۰/۰۱۸۰	۰/۰۷۰۰	۰/۰۵۸۰	۰/۰۵۶۰	۰/۰۵۹۹	۰/۰۹۱۹	$\hat{\lambda}_1$
۰/۱۳۴۰	۰/۴۶۶۰	۰/۳۹۹۰	۰/۳۸۷۰	۰/۴۱۷۰	۰/۵۸۰۰	$\hat{\lambda}_2$
۰/۲۴۴۶	۰/۸۳۵۰	۰/۷۵۱۶	۰/۷۲۶۰	۰/۷۵۸۰	۱/۱۴۸۰	$\hat{\lambda}_3$
۱۱/۳۵۶۶	۱۲/۴۵۹۸	۱۱/۳۰۴۶	۱۱/۵۷۲۰	۱۲/۱۰۰	۹/۷۲۸۰	$\hat{\alpha}$
۱/۳e-۱۲	۱/۴e-۱۲	۱/۳e-۱۲	۱/۱e-۱۲	۱/۲e-۱۲	۲/۲e-۱۲	$\hat{\xi}$
۰/۷۲۵۲	۰/۸۱۸۵	۰/۳۴۲۶	۱/۱۰۰۰	۰/۶۲۳۲	—	$\hat{\theta}$
۳۸/۳۸۹۲	۳۸/۱۷۳۰	۳۸/۱۸۸۰	۳۸/۳۵۷۴	۳۸/۴۱۲۵	۳۶/۹۰۰۱	log(I)
-۶۴/۴۵۶	-۶۴/۶۷۱	-۶۴/۳۵۶	-۶۴/۶۸۹	-۶۴/۷۹۰	-۶۳/۴۴۲	AIC
-۶۲/۹۲۲	-۶۱/۴۲۵	-۶۱/۸۲۵	-۶۲/۰۶۷	-۶۲/۳۵۱	-۶۱/۶۹۲	AICC
-۵۴/۸۹۹	-۵۳/۳۴۰	-۵۳/۸۶۰	-۵۴/۰۹۰	-۵۴/۳۱۰	-۵۴/۶۷۵	BIC
۰/۱۱۲۱	۰/۰۹۷۴۱	۰/۱۱۲۵	۰/۱۰۲۰	۰/۱۰۲۹	۰/۱۲۹۹	$K - S(X_1)$
۰/۲۶۹۶	۰/۳۴۳۲	۰/۲۶۹۷	۰/۲۸۱۲	۰/۲۸۶۴	۰/۳۵۹۹	$K - S(X_2)$
۰/۴۱۷۸	۰/۳۱۹۹	۰/۳۳۲۱	۰/۳۲۹۸	۰/۳۶۷۸	۰/۳۹۳۵	$K - S(\max(X_1, X_2))$
۳/۱۰۵	۳/۴۵۸	۳/۸۷۹	۳/۲۲۱	۳/۳۹۰	—	LRT
۰/۰۵۵۷	۰/۰۶۶۸	۰/۰۷۸۹	۰/۰۵۶۹	۰/۰۷۷۵	—	مقدار $-p$

مراجع

- [1] Flores D, J., Borges, P., Cancho, V. G., and Louzada, F. (2013), *The complementary exponential power series distribution*, Brazilian Journal of Probability and Statistics. 565–584.
- [2] Morais, A. L., and Barreto-Souza, W. (2011), *A compound class of Weibull and power series distributions*, Computational Statistics and Data Analysis. 55(3), 1410–1425.
- [3] Roozegar, R., and Jafari, A. A. (2015), *On Bivariate Exponentiated Extended Weibull Family of Distributions*, arXiv preprint arXiv:1507.07535.
- [4] Roozegar, R., and Nadarajah, S. (2016), *The Quadratic Hazard Rate Power Series Distribution*, Journal of Testing and Evaluation. 45(3).
- [5] Tahmasebi, S., and Jafari, A.A. (2015), *Exponentiated extended Weibull-power series class of distributions*, arXiv preprint arXiv:1503.08653.



برخی ویژگی‌های توزیع‌هایی با نرخ خطر خطی تعمیم‌یافته

غلامی گردونک، ز^۱ صالحی، ا^۲

^{۱،۲} گروه آمار، دانشگاه صنعتی بیرجند

چکیده

در مدل‌سازی داده‌های طول عمر، بررسی و تحلیل تابع نرخ خطر برای بدست آوردن نتایج لازم عاملی بسیار مفید است. توزیع‌های معروف و شناخته شده‌ای که اغلب با آنها سروکار داریم مانند توزیع نمایی دارای نرخ خطر ثابت و توزیع رایی دارای نرخ خطر صعودی می‌باشد. در عمل نیازمند توزیع‌هایی هستیم که تابع نرخ خطر آن انعطاف پذیری بیشتری نسبت به توزیع‌های ذکر شده داشته باشد. در این مقاله خانواده توزیع نرخ خطر خطی تعمیم‌یافته را معرفی و همچنین ویژگی‌ها و خواص آماری این توزیع را مورد مطالعه و بررسی قرار می‌دهیم. این توزیع به نوعی تعمیم‌یافته توزیع‌های شناخته شده هستند. در پایان برخی نتایج مشخصه‌سازی نیز ارائه می‌کنیم.

کلمات کلیدی: تابع نرخ خطر، توزیع نمایی تعمیم‌یافته، گشتاورها، تابع نرخ خطر معکوس، مشخصه‌سازی.

۱ مقدمه

مدل‌سازی و تحلیل داده‌های طول عمر از جنبه‌های مهم علم آمار و مباحث قابلیت اعتماد محسوب می‌شود که دارای کاربردهای وسیع و گسترده‌ای در بسیاری از علوم مختلف است. از اینرو برازش یک توزیع مناسب به داده‌های واقعی بسیار مهم و اساسی می‌باشد. برای این کار توزیع‌های متنوع و شناخته شده زیادی برای تجزیه و تحلیل داده‌های طول عمر وجود دارد. در بسیاری از مطالعات نظیر مطالعات پزشکی و مهندسی داده‌های طول عمر از اهمیت بالایی برخوردار هستند. جهت آنالیز و تحلیل داده‌های طول عمر توزیع‌های متعددی برای برازش به این داده‌ها وجود دارد. از

^۱ zahra.gholami@birjandut.ac.ir

^۲ salehi@birjandut.ac.ir

جمله این توزیع‌ها توزیع نمایی، نمایی تعمیم‌یافته، رایلی و رایلی تعمیم‌یافته می‌باشد. از آنجائی که توزیع نمایی دارای نرخ خطر ثابت می‌باشد [۳] تعمیم‌یافته این توزیع را پیشنهاد دادند و خواص و ویژگی‌های این توزیع را مورد مطالعه و بررسی قرار دادند. نرخ خطر توزیع تعمیم‌یافته نمایی به ازای مقادیر مختلف پارامترهای آن به صورت صعودی یا نزولی می‌باشد. در عمل به دنبال توزیع‌هایی می‌باشیم که رفتار تابع نرخ خطر آن به صورت وانی شکل باشد و انعطاف‌پذیری بیشتری از خود نشان دهند. به همین دلیل به دنبال گسترش و تعمیم توزیع‌های مختلف هستیم. در راستای این هدف [۵]

توزیع رایلی تعمیم‌یافته را معرفی کردند. با بررسی رفتار تابع نرخ خطر آن مشاهده می‌شود که نرخ خطر آن به صورت وانی شکل یا صعودی بوده اما حالت نزولی را در بر نمی‌گیرد. مطالعات زیادی در مورد توزیع نرخ خطر خطی تعمیم‌یافته انجام شده است و نویسندگان زیادی این توزیع را گسترش دادند. [۷] توزیع نمایی خطی تعمیم‌یافته را معرفی کردند. توزیع هندسی- نرخ خطر خطی تعمیم‌یافته را [۸] مورد بررسی قرار دادند. بواسون نرخ خطر خطی تعمیم‌یافته را [۱] معرفی و ویژگی‌های آن را بررسی کردند. [۴] توزیع بتا- نرخ خطر خطی را معرفی و این توزیع را مورد مطالعه قرار دادند. [۶] توزیع نمایی تعمیم‌یافته را بسط و گسترش دادند. نرخ خطر این توزیع در مقایسه با توزیع نمایی و نمایی تعمیم‌یافته بسیار انعطاف‌پذیر می‌باشد. در این مقاله توزیع نرخ خطر خطی تعمیم‌یافته سه پارامتری که به صورت $GLFR(a, b, \theta)$ نشان داده می‌شود را معرفی می‌کنیم. این خانواده توزیع را [۳] ارائه دادند. با بررسی رفتار تابع نرخ خطر این توزیع مشاهده می‌شود که نرخ خطر این توزیع به ازای مقادیر مختلف پارامترهای آن به صورت صعودی، نزولی و وانی شکل می‌باشد. نرخ خطر این توزیع نسبت به توزیع‌های ذکر شده در فوق از انعطاف‌پذیری بیشتری برخوردار می‌باشد. در ادامه ویژگی‌های آماری این توزیع جدید از جمله گشتاورها، چندک و ... را بررسی و مورد مطالعه قرار می‌دهیم.

۲ توزیع نرخ خطر خطی تعمیم‌یافته

تابع توزیع نرخ خطر خطی با پارامترهای a و b که به صورت $LFR(a, b)$ مشخص می‌شود به صورت زیر می‌باشد:

$$F(x) = 1 - e^{-(ax - \frac{b}{\theta} x^\theta)}, \quad x, a, b > 0.$$

توزیع نرخ خطر خطی تعمیم‌یافته با معرفی یک پارامتر جدید به نام θ که انعطاف‌پذیری بیشتری در توزیع جدید ارائه می‌دهد پیشنهاد می‌شود.

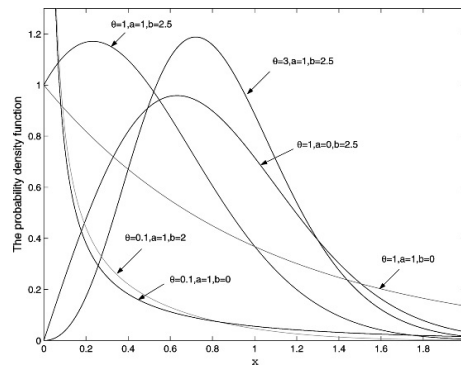
تعریف ۱.۰۲. متغیر تصادفی X دارای توزیع نرخ خطر خطی تعمیم‌یافته با پارامترهای $a, b, \theta \geq 0$ می‌باشد هرگاه تابع توزیع و تابع چگالی آن به صورت زیر باشند:

$$F(x) = [1 - e^{-(ax + \frac{b}{\theta} x^\theta)}]^\theta, \quad x \geq 0.$$

$$f(x) = \theta(a + bx) \left[1 - e^{-(ax + \frac{b}{\theta} x^\theta)} \right]^{\theta-1} e^{-(ax + \frac{b}{\theta} x^\theta)}, \quad x \geq 0.$$

که در آن θ پارامتر شکل می‌باشد. این توزیع به صورت $GLFR(a, b, \theta)$ نمایش داده می‌شود.

با اضافه شدن یک پارامتر به توزیع نرخ خطر خطی انعطاف‌پذیری بیشتری در مدل فراهم می‌شود. توزیع نرخ خطر خطی تعمیم‌یافته حالت کلی بعضی از توزیع‌های شناخته شده است. برای مثال هنگامی که $\theta = 1$ توزیع نرخ خطر خطی تعمیم‌یافته همان توزیع نرخ خطر خطی است. هنگامی که $a > 0$ و $b = 0$ توزیع نرخ خطر خطی تعمیم‌یافته به صورت توزیع نمایی تعمیم‌یافته خواهد بود. هنگامی که $a = 0$ و $b > 0$ توزیع نرخ خطر خطی تعمیم‌یافته به توزیع رایلی تعمیم‌یافته تبدیل خواهد شد. بنابراین توزیع نرخ خطر خطی تعمیم‌یافته یک مدل بسیار انعطاف‌پذیر است و زمانی که پارامترهای آن تغییر می‌کند به توزیع‌های شناخته شده تبدیل می‌شود. شکل ۱ تابع چگالی احتمال این توزیع را به ازای مقادیر مختلف نشان می‌دهد.



شکل ۱: تابع چگالی توزیع $GLFR(a, b, \theta)$

توزیع نرخ خطر خطی تعمیم یافته برای تجزیه و تحلیل داده های طول عمر از این جهت می تواند مفید باشد که رفتار تابع نرخ خطر آن انعطاف پذیری بیشتری در مقایسه با توزیع های شناخته شده مانند نمایی دارد. تابع قابلیت اعتماد و تابع نرخ خطر توزیع نرخ خطر خطی تعمیم یافته به صورت زیر می باشد:

$$\bar{F}(x) = 1 - [1 - e^{-(ax + \frac{b}{\theta} x^\theta)}]^\theta,$$

$$h(x) = \frac{\theta(a + bx)[1 - e^{-(ax + \frac{b}{\theta} x^\theta)}]^\theta - 1 e^{-(ax + \frac{b}{\theta} x^\theta)}}{1 - [1 - e^{-(ax + \frac{b}{\theta} x^\theta)}]^\theta}.$$

[۸] نشان دادند که علاوه بر تابع نرخ خطر، نرخ خطر معکوس نیز در تجزیه و تحلیل مسائل قابلیت اعتماد بسیار مهم می باشد. تابع نرخ خطر معکوس توزیع $GLFR(a, b, \theta)$ به صورت زیر محاسبه می شود:

$$r(x) = \frac{f(x)}{F(x)} = \theta \frac{(a + bx)e^{-(ax + \frac{b}{\theta} x^\theta)}}{1 - e^{-(ax + \frac{b}{\theta} x^\theta)}} = \theta \frac{f(x)}{F(x)} = \theta r_1(x),$$

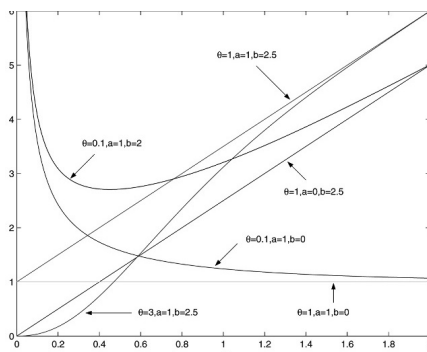
که در آن $r_1(x)$ تابع نرخ خطر معکوس توزیع $LFR(a, b)$ می باشد. حال رفتار تابع نرخ خطر توزیع نرخ خطر خطی تعمیم یافته را به ازای مقادیر مختلف پارامترهای آن بیان می کنیم.

(الف) اگر $\theta = 1$ باشد، برای $b > 0$ تابع نرخ خطر صعودی و برای $b = 0$ و $a > 0$ تابع نرخ خطر ثابت می باشد.

(ب) اگر $\theta > 1$ باشد، تابع نرخ خطر صعودی می باشد.

(ج) اگر $\theta < 1$ باشد، برای $b = 0$ تابع نرخ خطر نزولی و برای $b > 0$ تابع نرخ خطر به صورت وانی شکل می باشد.

شکل ۲ رفتار تابع نرخ خطر توزیع نرخ خطر خطی تعمیم یافته را به ازای مقادیر مختلف پارامترهای آن نشان می دهد.


 شکل ۲: تابع نرخ خطر توزیع $GLFR(a, b, \theta)$

۳ خواص آماری

یکی از ویژگی‌های مهم هر توزیعی گشتاورهای آن توزیع می‌باشد. لم زیر گشتاور k ام توزیع $GLFR(a, b, \theta)$ زمانی که $\theta \geq 1$ باشد را بیان می‌کند.

لم ۱.۳. اگر X دارای توزیع $GLFR(a, b, \theta)$ باشد، آنگاه گشتاور k ام X ، که با $\mu^{(k)}$ نشان می‌دهند به صورت زیر محاسبه می‌شود:

برای $a = 0$ و $b > 0$ ،

$$\mu^{(k)} = \frac{\theta \Gamma(\frac{k}{\theta} + 1)}{(\frac{b}{\theta})^k} \sum_{i=0}^{\infty} \frac{(-1)^{-i} \binom{\theta-1}{i}}{(i+1)^{\frac{k}{\theta}+1}};$$

برای $a > 0$ و $b > 0$ ،

$$\mu^{(k)} = \theta \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{-i} \binom{\theta-1}{i} \Gamma(k+l+1) g_i^{(l)}(\cdot)}{l! [a(i+1)]^{k+l+1}} \left[a + \frac{(k+l+1)b}{(i+1)a} \right],$$

که در آن

$$g_i^{(l)}(\cdot) = \frac{d^l}{dx^l} \exp\left(-\frac{1}{\theta}(i+1)bx^{\theta}\right) \Big|_{x=\cdot}.$$

یکی دیگر از ویژگی‌ها و خواص هر توزیعی چندک‌های آن توزیع می‌باشد که چندک‌های توزیع نرخ خطر خطی تعمیم‌یافته به صورت زیر محاسبه می‌شود:

$$x_q = \frac{1}{b} \left(-a + \sqrt{a^2 - 2b \ln(1 - q^{\frac{1}{\theta}})} \right)$$

قضیه ۱.۳. فرض X_i ها متغیرهای تصادفی مستقل که دارای توزیع $GLFR(a, b, \theta_i)$ می‌باشد. آنگاه $X_{(n)}$ دارای توزیع $GLFR(a, b, \sum_{i=1}^n \theta_i)$ می‌باشد.

برهان. X_i متغیرهای تصادفی مستقل می‌باشد از اینرو داریم:

$$\begin{aligned} P(X_{(n)} < x) &= P(X_1 < x, X_2 < x, \dots, X_n < x) = P(X_1 < x) \dots P(X_n < x) \\ &= \prod_{i=1}^n P(X_i < x) = \prod_{i=1}^n [1 - e^{-(ax + \frac{b}{\gamma} x^\gamma)}]^{\theta_i} = \left[1 - e^{-(ax + \frac{b}{\gamma} x^\gamma)} \right]^{\sum_{i=1}^n \theta_i} \end{aligned}$$

□

قضیه ۲.۳. اگر X_i ها متغیرهای تصادفی مستقل و دارای توزیع $GLFR(a, b, \theta)$ باشد، آنگاه برای $X_{(1)}$ داریم:

$$\lim_{n \rightarrow \infty} P\left(\frac{X_{(1)}}{C_n} \leq x\right) = 1 - e^{-(ax)^\theta}, \quad x > 0.$$

که در آن $C_n = n^{-\frac{1}{\theta}}$ می باشد.

فرض کنید $X_1, \dots, X_n$ یک نمونه تصادفی از توزیع $GLFR(a, b, \theta)$ باشد و اگر $X_{(1)} \leq \dots \leq X_{(n)}$ آماره های ترتیبی این نمونه تصادفی باشد آنگاه تابع چگالی $X_{(r)}$ به صورت زیر محاسبه می شود.

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} [F(x)]^{r-1} [1-F(x)]^{n-r} f(x),$$

که با ساده کردن رابطه فوق به صورت زیر خلاصه می شود:

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} f(x) \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^j [F(x)]^{r+j-1},$$

با جایگذاری مقدار تابع توزیع در رابطه فوق داریم:

$$f_{r:n}(x) = \sum_{j=0}^{n-r} K_j(n, r) f^*(x),$$

که در آن

$f^*(x)$ تابع چگالی توزیع $GLFR(a, b, \theta_{r+j})$ می باشد.

$$\theta_i = i\theta, \quad K_j(n, r) = \frac{n(-1)^j \binom{n-r}{j} \binom{n-1}{r-1}}{(r+j)}.$$

قضیه ۳.۳. گشتاور k ام آماره ترتیبی X_r به صورت زیر محاسبه می شود:

$$\text{الف: } a = 0, b > 0$$

$$\mu_{r:n}^{(k)} = \theta \left(\frac{\gamma}{b}\right)^{\frac{k}{\gamma}+1} \Gamma\left(\frac{k}{\gamma} + 1\right) \sum_{j=0}^{n-r} \sum_{i=0}^{\infty} w_j^*(n, r) \frac{(-1)^i \binom{\theta_{r+j}-1}{i}}{(i+1)^{\frac{k}{\gamma}+1}};$$

$$\text{ب: } a > 0, b \geq 0$$

$$\mu_{r:n}^{(k)} = \theta \sum_{j=0}^{n-r} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} w_j^*(n, r) \frac{(-1)^i \binom{\theta_{r+j}-1}{i} g_i^{(l)} \Gamma(k+l+1)}{l! [(i+1)a]^{k+l+1}} \left[a + \frac{(k+l+1)b}{(i+1)a} \right],$$

که در آن

$$w_j^*(n, r) = (r+j) K_j(n, r)$$

۴ مشخصه سازی

فرض کنید Y یک متغیر تصادفی به طور مطلق پیوسته که تکیه گاه آن فقط روی محور اعداد حقیقی مثبت می باشد. برای $a, b, t > 0$ و همه n های صحیح نامنفی تعاریف زیر را داریم:

$$U(t) = -\ln(1 - e^{-(at + \frac{b}{\theta}t)}) \quad , \quad \alpha^{(n)}(t) = E(U^n(Y)|Y < t).$$

با توجه به این تعاریف، نتایج زیر را داریم:

قضیه ۱.۴. متغیر تصادفی Y دارای توزیع $GLFR(a, b, \theta)$ برای $a, b, \theta > 0$ می باشد اگر و تنها اگر برای هر $t > 0$ و همه n های صحیح نامنفی رابطه زیر برقرار باشد:

$$\alpha^{(n)}(t) = U^n(t) + \frac{n}{\theta} \alpha^{(n-1)}(t) \quad (۱)$$

برهان. ابتدا فرض می کنیم که Y دارای توزیع $GLFR$ باشد، برای اثبات رابطه (۱) از تعریف $\alpha^{(n)}(t)$ استفاده می کنیم و با انتگرال گیری جزء به جزء از آن رابطه (۱) اثبات می شود.

حال فرض کنید که رابطه (۱) برقرار باشد می خواهیم ثابت کنیم که Y دارای توزیع $GLFR$ می باشد. برای این کار $f_Y(\cdot)$ و $F_Y(\cdot)$ به ترتیب تابع چگالی و تابع توزیع متغیر تصادفی Y تعریف می کنیم. با استفاده از رابطه (۱) داریم:

$$\int_0^t U^n(t) f_Y(y) dy = U^n(t) F_Y(t) + \frac{n}{\theta} \int_0^t U^{n-1}(t) f_Y(t) dy, \quad (۲)$$

از رابطه (۲) نسبت به t مشتق می گیریم و رابطه زیر بدست می آید:

$$U^n(t) f_Y(t) = U^n(t) f_Y(t) - n U^{n-1}(t) F_Y(t) \frac{f_{LFR}(t)}{F_{LFR}(t)} + \frac{n}{\theta} U^{n-1}(t) f_Y(t),$$

که در آن f_{LFR} و F_{LFR} به ترتیب تابع توزیع و تابع چگالی توزیع $LFR(a, b)$ می باشد. با توجه به اینکه $U^{n-1}(t) \neq 0$ ، داریم:

$$\frac{f_Y(t)}{F_Y(t)} = \theta \frac{f_{LFR}(t)}{F_{LFR}(t)} \quad (۳)$$

رابطه (۴) برابر با تابع نرخ خطر معکوس توزیع نرخ خطر خطی تعمیم یافته می باشد و با توجه به یکتایی تابع نرخ خطر معکوس نتیجه می شود که Y دارای توزیع نرخ خطر خطی تعمیم یافته می باشد. $\square$

قضیه ۲.۴. متغیر تصادفی Y دارای توزیع $GLFR(a, b, \theta)$ برای $a, b, \theta > 0$ می باشد، اگر و تنها اگر برای هر $t > 0$ و همه n های صحیح نامنفی رابطه زیر برقرار باشد

$$Var(U(Y)|Y < t) = \frac{1}{\theta^2} \quad (۴)$$

برهان. ابتدا فرض می کنیم که Y دارای توزیع $GLFR$ باشد، آنگاه با استفاده از قضیه ۱ داریم:

$$\alpha^{(1)}(t) = U(t) + \frac{1}{\theta}, \quad (۵)$$

و

$$\alpha^{(r)}(t) = U^r(t) + \frac{r}{\theta} \alpha^{(1)}(t) = U^r(t) + \frac{r}{\theta} \left[U(t) + \frac{1}{\theta} \right] = (\alpha^{(1)}(t))^r + \frac{1}{\theta^r}, \quad (۶)$$

بنابراین، با ادامه این روند اثبات نتیجه می‌شود.

فرض کنید رابطه (۵) برای همه $t > 0$ برقرار باشد. و $f_Y(\cdot)$ و $F_Y(\cdot)$ به ترتیب تابع چگالی و تابع توزیع متغیر تصادفی Y باشد. رابطه (۵) را می‌توان به صورت زیر بازنویسی کرد.

$$F_Y(t) \int_0^t U^\gamma(y) f_Y(y) dy - \left[\int_0^t U(y) F_Y(y) dy \right]^\gamma = \frac{1}{\theta^\gamma} F^\gamma(y) \quad (7)$$

از دو طرف رابطه (۷) نسبت به t مشتق گرفته و داریم:

$$\int_0^t U^\gamma(y) f_Y(y) dy + F_Y(t) U^\gamma(t) - \gamma U(t) \int_0^t U(y) f_Y(y) dy = \gamma \frac{F_Y(t)}{\theta^\gamma}$$

از رابطه فوق نسبت به t مشتق گرفته و با ساده کردن آن داریم:

$$F_Y(t) U(t) - \int_0^t U(y) f_Y(y) dy = - \frac{f_Y(t)}{r_{LFR}(t) \theta^\gamma} \quad (8)$$

در آخر از رابطه (۸) نیز نسبت به t مشتق گرفته و با ساده کردن آن به معادله دیفرانسیل زیر می‌رسیم:

$$\theta^\gamma F_Y(t) = \frac{1}{r_{LFR}(t)} \times \frac{d}{dt} \left(\frac{f_Y(t)}{r_{LFR}(t)} \right)$$

جواب کلی معادله دیفرانسیل فوق به صورت زیر می‌باشد:

$$F_Y(t) = c_1 e^{\alpha U(t)} + c_2 e^{-\alpha U(t)}$$

که در آن c_1 و c_2 مقادیر ثابت و دلخواه می‌باشد. از آنجاییکه

$$\lim_{t \downarrow 0} U(t) = \infty, \quad \lim_{t \rightarrow \infty} U(t) = 0$$

و $F_Y(t)$ تابع توزیع می‌باشد، اگر $c_1 = 0$ و $c_2 = 1$ در نظر می‌گیریم. با جایگذاری در رابطه بالا اثبات کامل می‌شود. $\square$

قضیه ۳.۴. فرض کنید $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و هم توزیع باشند. آنگاه X_i ها دارای توزیع $GLFR$ هستند اگر و تنها اگر $X_{(n)}$ دارای توزیع $GLFR$ باشد.

برهان. X_i متغیرهای تصادفی مستقل می‌باشد از اینرو داریم:

$$\begin{aligned} P(X_{(n)} < x) &= P(X_1 < x, \dots, X_n < x) = P(X_1 < x) \dots P(X_n < x) \\ &= \prod_{i=1}^n P(X_i < x) = \prod_{i=1}^n \left[1 - e^{-(ax_i + \frac{b}{\gamma} x_i)^\gamma} \right]^\theta = \left[1 - e^{-(ax_i + \frac{b}{\gamma} x_i)^\gamma} \right]^{n\theta} \end{aligned}$$

$\square$

مراجع

- [1] Cordeiro, G. M., Ortega, E. M. M., and Lemonte, A. J. (2015). The Poisson generalized linear failure rate model. *Communications in Statistics-Theory and Methods*, 44(10):2037–2058.
- [2] Gupta, R. C., Gupta, R. D. (2007). Proportional reversed hazard model and its applications. To appear in *J. Statist. Plann. Infer.*
- [3] Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian New Zealand Journal of Statistics*, 41(2):173–188.

- [4] Jafari, A. A. and Mahmoudi, E. (2015). Beta-linear failure rate distribution and its applications. Journal of the Iranian Statistical Society, 14(1):89-105
- [5] Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. Computational Statistics Data Analysis, 49(1):187-200.
- [6] Kundu, D. and Gupta, R. D. (2011). An extension of the generalized exponential distribution. Statistical Methodology, 8(6):485-496.
- [7] Mahmoud, M. A. W. and Alam, F. M. A. (2010). The generalized linear exponential distribution. Statistics Probability Letters, 80(1112):1005-1014.
- [8] Nadarajah, S., Shahsanaei, F., and Rezaei, S. (2014). A new four-parameter lifetime distribution. Journal of Statistical Computation and Simulation, 84(2):248-263.+9
- [9] Sarhan, A. M. and Kundu, D. (2009). Generalized linear failure rate distribution. Communications in Statistics-Theory and Methods, 38(5):642-660.

اندازه نادرستی باقیمانده آماره‌های ترتیبی

فهم شریفی، م^۱ عباس نژاد، م^۲

^{۱،۲} گروه آمار، دانشکده ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری

چکیده

در این مقاله اندازه نادرستی باقیمانده برای آماره‌های ترتیبی را مورد بررسی قرار داده و یک کران پایین برای آن ارائه می‌نماییم. سپس توزیع‌های نمایی، پاراتو و دامنه متناهی بر اساس اندازه نادرستی باقیمانده آماره‌های ترتیبی مشخصه سازی می‌شوند.

کلمات کلیدی: آماره‌های ترتیبی، آنتروپی باقیمانده، ترتیب‌های تصادفی، نرخ خطر و تابع بقا.

۱ مقدمه

فرض کنید X و Y دو متغیر تصادفی پیوسته نامنفی با تابع چگالی احتمال به ترتیب f و g باشد، اندازه عدم حتمیت یا آنتروپی شانون X به صورت زیر تعریف می‌شود:

$$H(X) = H(f) = - \int_{\cdot}^{\infty} f(x) \ln f(x) dx. \quad (1)$$

همچنین فاصله کولبک-لیبلر X از Y به صورت زیر می‌باشد،

$$H(f|g) = - \int_{\cdot}^{\infty} f(x) \ln \left(\frac{f(x)}{g(x)} \right) dx. \quad (2)$$

با جمع کردن (۱) و (۲) داریم،

$$I(f, g) = H(f) + H(f|g) = - \int_{\cdot}^{\infty} f(x) \ln g(x) dx. \quad (3)$$

^۱M.fahim.Sharifi67@gmail.com

^۲M.abbasnejad@hsu.ac.ir

این اندازه، اولین بار توسط کریج (۱۹۶۱) با عنوان اندازه نادرستی معرفی شد. بعد از آن مطالعات بسیاری بر روی این اندازه صورت گرفت. برای مثال تپلیال و تانجا (۲۰۱۳) اندازه نادرستی آماره‌های ترتیبی را مورد بررسی قرار دادند. فرض کنید X طول عمر یک سیستم باشد. بدیهی است اگر بدانیم سیستم تا زمان t فعال بوده، اندازه‌های معرفی شده در (۱)، (۲) و (۳) معیارهای مناسبی برای اندازه گیری عدم حتمیت و نادرستی نیستند. به همین دلیل ابراهیمی (۱۹۹۶) عدم حتمیت متغیر تصادفی $X_t = (X - t) | X > t; t \geq 0$ را که به عنوان متغیر تصادفی طول عمر باقیمانده شناخته می‌شود، به صورت زیر در نظر گرفت،

$$H(f; t) = - \int_t^\infty \frac{f(x)}{\bar{F}(t)} \ln\left(\frac{f(x)}{\bar{F}(t)}\right) dx. \quad (4)$$

که در آن $\bar{F}(t) = 1 - F(t)$ تابع بقای X می‌باشد.

مشابه با کار ابراهیمی اندازه نادرستی باقیمانده توسط تانجا و همکاران (۲۰۰۹) به صورت،

$$I(f, g; t) = - \int_t^\infty \frac{f(x)}{\bar{F}(t)} \ln\left(\frac{g(x)}{\bar{G}(t)}\right) dx. \quad (5)$$

ارایه شد. واضح است که اگر $t = 0$ باشد، معادلات (۴) و (۵) به ترتیب به معادلات (۱) و (۳) تبدیل می‌شود.

در بخش دوم اندازه نادرستی برای آماره‌های ترتیبی در نظر گرفته و برای آن یک کران پایین ارایه می‌کنیم. در بخش سوم، نیز نشان می‌دهیم که تحت شرایطی توزیع‌های نمایی، پاراتو و دامنه متناهی می‌توانند بر اساس اندازه نادرستی باقیمانده آماره‌های ترتیبی مشخصه سازی شوند.

۲ اندازه نادرستی باقیمانده برای آماره ترتیبی i ام

تپلیال و تانجا (۲۰۱۳) اندازه نادرستی بین چگالی آماره ترتیبی i ام و چگالی جامعه را این گونه تعریف کردند.

$$I_n(f_{i:n}, f) = - \int_0^\infty f_{i:n}(x) \ln f(x) dx, \quad (6)$$

که در آن

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} (F(x))^{i-1} (1-F(x))^{n-i} f(x).$$

تابع چگالی i امین آماره ترتیبی است و

$$B(i, n-i+1) = \frac{\Gamma(i)\Gamma(n-i+1)}{\Gamma(n+1)}$$

تابع بتا با پارامترهای i و $n-i+1$ است.

با توجه به (۶) اندازه نادرستی باقیمانده طول عمر i امین آماره ترتیبی و طول عمر جامعه را می‌توان به صورت زیر بیان نمود.

$$I_n(f_{i:n}, f; t) = - \int_t^\infty \frac{f_{i:n}(x)}{\bar{F}_{i:n}(t)} \ln\left(\frac{f(x)}{\bar{F}(t)}\right) dx. \quad (7)$$

که در آن $\bar{F}_{i:n}(t) = 1 - F_{i:n}(t)$ تابع بقای i امین آماره ترتیبی است، می‌توان نشان داد که

$$\bar{F}_{i:n}(t) = \frac{\bar{B}_{F(t)}(i, n-i+1)}{B(i, n-i+1)}$$

که در آن

$$\bar{B}_x(a, b) = \int_x^1 u^{a-1} (1-u)^{b-1} du \quad 0 < x < 1 \quad a, b > 0.$$

تابع بتای ناقص می‌باشد.

اگر $t = 0$ باشد معادله (۱) به معادله (۶) تبدیل می‌شود.

حال اندازه نادرستی باقیمانده را برای آماره ترتیبی اول برای برخی توزیع‌های طول عمر که در قابلیت اعتماد به کار می‌رود را محاسبه کرده و نتایج آن، در جدول ۱ آورده شده است.

جدول ۱: اندازه نادرستی باقیمانده برای آماره ترتیبی اول

توزیع	تابع چگالی	اندازه نادرستی باقیمانده
توزیع یکنواخت در $[0, b]$	$\frac{1}{b}$	$\ln\left(\frac{b-t}{b}\right) + \ln b$
توزیع نمایی	$\theta \exp^{-\theta x}$	$\frac{1}{n} - \ln \theta$
توزیع پارتو	$\frac{ab^a}{(x+b)^{a+1}}, a > 1, b > 0$	$-\ln\left(\frac{a}{b+t}\right) + \frac{a+1}{na}$
توزیع دامنه متناهی	$a(1-x)^{a-1}$	$-\ln\left(\frac{a}{1-t}\right) + \frac{a-1}{na}$

گزاره ۱۰۲. فرض کنید $M = f(m) < \infty$ نمای توزیع باشد یعنی $\{m = \sup\{x; f(x) \leq M\}\}$. در این صورت، یک کران پایین برای اندازه نادرستی به صورت زیر می‌باشد،

$$I_n(f_{i:n}, f; t) \geq \ln \bar{F}(t) - \ln M.$$

برهان. طبق معادله (۱) داریم:

$$\begin{aligned} I_n(f_{i:n}, f; t) &= - \int_t^\infty \frac{f_{i:n}(x)}{\bar{F}_{i:n}(t)} \ln\left(\frac{f(x)}{\bar{F}(t)}\right) dx \\ &= \ln \bar{F}(t) - \frac{1}{\bar{F}_{i:n}} \int_t^\infty f_{i:n}(x) \ln(f(x)) dx \end{aligned}$$

از آنجایی که m نمای توزیع می‌باشد پس $\ln f(x) \leq \ln M$ و بنابراین داریم:

$$I_n(f_{i:n}, f; t) \geq \ln \bar{F}(t) - \ln M.$$

□

تعریف ۱۰۲. متغیر تصادفی X را در ترتیب تصادفی نسبت درستمایی کمتر از Y گوئیم و با نماد $X \leq^{lr} Y$ نشان می‌دهیم اگر $\frac{f_X(x)}{g_Y(x)}$ نسبت به x نزولی باشد.

تعریف ۲۰۲. متغیر تصادفی X را به طور تصادفی کمتر از Y گوئیم و با نماد $X \leq^{st} Y$ نشان می‌دهیم اگر به ازای هر x ، $\bar{F}(x) < \bar{G}(x)$ باشد.

قضیه ۱۰۲. فرض کنید $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل و هم‌توزیع با تابع چگالی f و تابع توزیع F ، نشان دهنده طول عمر مؤلفه‌های یک سیستم سری باشد. در این صورت طول عمر سیستم، اولین آماره ترتیبی است. اگر f روی دامنه خود تابعی نزولی باشد، آنگاه اندازه نادرستی متناظر نیز تابعی نزولی از n است.

برهان. با قرار دادن $i = 1$ در رابطه (۱) و تغییر متغیر $u = F(x)$ داریم:

$$\begin{aligned} I_n(f_{1:n}, f; t) &= - \frac{1}{\bar{B}_{F(t)}(1, n)} \int_{F(t)}^\infty (1-u)^{n-1} \ln(f(F^{-1}(u))) du + \ln \bar{F}(t) \\ &= -E_{g_{1:n}}[\ln(f(F^{-1}(U_n)))] + \ln \bar{F}(t). \end{aligned}$$

که در آن

$$g_{1:n}(u) = \frac{(1-u)^{n-1}}{\bar{B}_{F(t)}(1, n)} \quad F(t) \leq u \leq 1$$

پس $I_n(f_{1:n}, f; t) - I_n(f_{1:n+1}, f; t)$ و

$$E_{g_{1:n+1}}[\ln(f(F^{-1}(U_{n+1}))) - E_{g_{1:n}}[\ln(f(F^{-1}(U_n)))].$$

از طرفی با توجه به نزولی بودن تابع f تابع $\ln(f(F^{-1}(u)))$ نیز تابعی نزولی از u خواهد بود. بنابراین با توجه به شیکد و شانتيکومار (۲۰۰۷) کافی است نشان دهیم که $U_{n+1} \stackrel{st}{\leq} U_n$ همچنین می‌دانیم که ترتیب در نسبت درستی، ترتیب تصادفی را نتیجه می‌دهد، پس ثابت می‌کنیم $U_{n+1} \stackrel{lr}{\leq} U_n$ یا به طور معادل نشان می‌دهیم، $\frac{g_{1:n+1}(u)}{g_{1:n}(u)}$ تابعی نزولی از u است. داریم

$$\frac{g_{1:n+1}(u)}{g_{1:n}(u)} = \frac{\bar{B}_{F(t)}(1, n)}{\bar{B}_{F(t)}(1, n+1)}(1-u)$$

که تابع نزولی از u است، بنابراین

$$I_n(f_{1:n}, f; t) - I_n(f_{1:n+1}, f; t) \geq 0.$$

□

۳ مشخصه سازی

مشخص سازی عبارت است از توصیف توزیع‌های آماری با توجه به خصوصیات معینی که منحصرًا مربوط به آن توزیع‌ها هستند.

قضیه ۱.۳. فرض کنید X یک مقدار تصادفی پیوسته نامنفی با تابع توزیع $F(\cdot)$ و تابع نرخ خطر $\lambda_F(\cdot)$ باشد و فرض کنید $I_n(f_{i:n}, f; t) < \infty, t \geq 0$.

$$I_n(f_{i:n}, f; t) = c - \ln \lambda_F(t) \quad (۸)$$

که در آن c یک مقدار ثابت است، در این صورت X دارای توزیع

$$\text{الف) } c = \frac{1}{n} \text{ اگر و تنها اگر}$$

$$\text{ب) پاراتو است اگر تنها اگر } c > \frac{1}{n}$$

$$\text{ج) دامنه متناهی است اگر و تنها اگر } c < \frac{1}{n}$$

برهان. به راحتی می‌توان نشان داد که اگر X دارای توزیع نمایی، پاراتو و دامنه متناهی باشد رابطه (۸) برقرار است. پس با مشتق گرفتن از دو طرف معادله (۸) نسبت به t داریم،

$$\frac{d}{dt}[I_n(f_{i:n}, f; t)] = -\lambda_F(t) + n\lambda_F(t)I_n(f_{i:n}, f; t)n \ln \lambda_F(t). \quad (۹)$$

با قرار دادن مقدار $I_n(f_{i:n}, f; t)$ در معادله (۹) رابطه زیر بدست می‌آید،

$$\lambda_F'(t) = (nc - 1)\lambda_F(t).$$

که یک معادله دیفرانسیل است و با حل آن داریم،

$$\lambda_F(t) = \frac{1}{at + b}. \quad (۱۰)$$

$$\text{که در آن } a = nc - 1 \text{ و } b = \frac{1}{\lambda_F(0)} \quad (۱۰)$$

الف) اگر $c = \frac{1}{n}$ آنگاه $a = 0$ با استفاده از (10) ، $\lambda_F(t)$ به یک مقدار ثابت تبدیل می‌شود که فقط برای توزیع نمایی امکان پذیر است.

ب) اگر $c > \frac{1}{n}$ آنگاه $a > 0$ و از رابطه (10) ، تابع نرخ خطر توزیع پاراتو بدست می‌آید.

ج) اگر $c < \frac{1}{n}$ آنگاه $a < 0$ و از رابطه (10) تابع نرخ خطر توزیع دامنه متناهی را نتیجه می‌دهد.

□

مراجع

- [1] Ebrahimi, N., 1996. How to measure uncertainty in the residual lifetime distributions. Sankhyā Ser. A 58, 48–57.
- [2] Kerridge, D.F., 1961. Inaccuracy and inference. J. R. Stat. Soc. Ser. B 23, 184–194.
- [3] Shaked, M., Shanthikumar, J.G., 2007. Stochastic Orders. Springer Verlag
- [4] Taneja, H.C., Kumar, V., Srivastava, R., 2009. A dynamic measure of inaccuracy between two residual lifetime distributions. Int. Math. Forum 4 (25), 1213–1220.
- [5] Thapliyal, R., Taneja, H.C., 2013. A measure of inaccuracy in order statistics. J. Stat. Theory Appl. 12 (2), 200–207.



توزیعی جدید برگرفته از خانواده وایبول

قاسمیان، ف^۱ هاشمی، ر^۲

^{۱،۲} گروه آمار، دانشکده ریاضی، دانشگاه رازی

چکیده

در این مقاله، به معرفی یک توزیع جدید برگرفته از خانواده وایبول می‌پردازیم. از آنجا که توزیع وایبول از محبوب‌ترین و پرکاربردترین توزیع‌ها در قابلیت اعتماد و تجزیه و تحلیل داده‌های طول عمر است. یک مورد خاص از این خانواده بررسی می‌شود. برخی ویژگی‌های این توزیع جدید از جمله تابع چگالی احتمال، تابع توزیع، تابع نرخ خطر، تابع بقا، تابع چگالی آماره ترتیبی و چندک‌های آن را به دست آورده و معرفی می‌کنیم.

کلمات کلیدی: قابلیت اعتماد، تابع نرخ شکست، تابع بقا، آماره ترتیبی.

۱ مقدمه

توزیع وایبول ۲ پارامتری در مدلسازی داده‌های طول عمر بسیار مفید است. با این حال این مدل تنها برای داده‌های با تابع نرخ شکست یکنواخت مناسب است. اما در بسیاری از واحد‌های مکانیکی و الکتریکی، تابع نرخ شکست آنها به صورت تک مدی و وان شکل نمایش داده می‌شود. و برای این واحد‌ها توزیع وایبول مناسب نیست بنابراین برای مدل‌سازی این داده‌ها، چندین مدل از توزیع وایبول کلی مطرح شده است. مانند توزیع وایبول نمایی [۳]، توزیع وایبول افزایشی [۵]، و توزیع وایبول انعطاف پذیر [۱]. اخیراً تعمیم جدیدی از توزیع وایبول توسط کورای^۱ [۲] ارائه شده است که توجه زیادی را به خود جلب کرده است. این توزیع دارای هر دو توزیع وایبول و وایبول معکوس در حالت خاص است. یکی دیگر از مزیت‌های این توزیع این است که دارای تابع نرخ شکست یکنواخت، وان شکل و تک مدی است. یکی از راه‌های ایجاد

^۱ fatemehghasemian67@gmail.com

^۲ r.hashemi@razi.ac.ir

^۱ Cooray

انعطاف پذیری بیشتر در توزیع‌ها اضافه نمودن پارامتر به توزیع است که می‌تواند به روش‌های مختلفی انجام شود در این مقاله هدف تولید یک توزیع جدید بر مبنای توزیعی از خانواده وایبول است.

۲ توزیع وایبول بخت

تابع توزیع تجمعی CDF این توزیع به صورت زیر است:

$$F(t) = 1 - \left[1 + \left(e^{\left(\frac{t}{\theta}\right)^\alpha} - 1 \right)^\beta \right]^{-1}, \quad 0 < t < \infty \quad (1)$$

که $\theta > 0$ یک پارامتر مقیاس است و $\beta, \alpha > 0$ پارامترهای شکل هستند. زمانی که $\beta = 1$ ، $F(t)$ تابع CDF توزیع وایبول است و زمانی که $\beta = -1$ ، تابع توزیع، توزیع وایبول معکوس است. تابع چنک این توزیع بصورت زیر است:

$$Q(u) = F^{-1}(u) = \theta \left\{ \ln \left[1 + \left(\frac{u}{1-u} \right)^{\frac{1}{\beta}} \right] \right\}^{\frac{1}{\alpha}}, \quad 0 < u < 1 \quad (2)$$

با شروع از توزیع وایبول $F_w(t)$ (هنگامی که پارامترهای شکل α, β مثبت هستند) CDF توزیع وایبول بخت به صورت زیر است:

$$F(t) = \frac{F_w^\beta(t)}{F_w^\beta(t) + [1 - F_w(t)]^\beta} \quad (3)$$

درحالی که اگر از توزیع وایبول معکوس شروع کنیم برای (هنگامی که پارامترهای شکل α, β منفی هستند) CDF توزیع وایبول بخت به صورت زیر است:

$$F(t) = \frac{F_v^{-\beta}(t)}{F_v^{-\beta}(t) + [1 - F_v(t)]^{-\beta}} \quad (4)$$

با گرفتن مشتق از CDF بر حسب t تابع چگالی احتمال به صورت زیر به دست می‌آید:

$$f(t) = \left(\frac{\alpha\beta}{t} \right) \left(\frac{t}{\theta} \right)^\alpha e^{\left(\frac{t}{\theta}\right)^\alpha} \left(e^{\left(\frac{t}{\theta}\right)^\alpha} - 1 \right)^{\beta-1} \left[1 + \left(e^{\left(\frac{t}{\theta}\right)^\alpha} - 1 \right)^\beta \right]^{-2} \quad (5)$$

تابع نرخ شکست FRF ^۲ توزیع وایبول بخت بصورت زیر است:

$$h(t) = \left(\frac{\alpha\beta}{t} \right) \left(\frac{t}{\theta} \right)^\alpha e^{\left(\frac{t}{\theta}\right)^\alpha} \left(e^{\left(\frac{t}{\theta}\right)^\alpha} - 1 \right)^{\beta-1} \left[1 + \left(e^{\left(\frac{t}{\theta}\right)^\alpha} - 1 \right)^\beta \right]^{-1} \quad (6)$$

۳ خانواده توزیع‌های تابعی از کسینوس هایپربولیک

در این قسمت به معرفی یک کلاس جدید از توزیع‌ها به نام توزیع تابعی از کسینوس هایپربولیک (HCF) می‌پردازیم. این خانواده جدید از توزیع‌ها از ترکیب توزیع پایه F با تابع کسینوس هایپربولیک بدست می‌آید. این روش باعث تولید یک پارامتر اضافی برای یک خانواده از توزیع‌ها و انعطاف پذیری بیشتر آنها می‌شود. [۴]

ابتدا تعریف تابع کسینوس هایپربولیک را یادآوری می‌کنیم. تابع کسینوس هایپربولیک براساس تابع نمایی به صورت زیر است:

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

^۲ Failure rate function

تابع کسینوس هایپربولیک تابعی زوج است و دارای بسط سری تیلور با فقط توان های زوج از x است که به صورت زیر است:

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \quad (۷)$$

تعریف ۱.۳. فرض کنید X یک متغیر تصادفی پیوسته با تابع توزیع تجمعی $F(x)$ (CDF) باشد در این صورت مدل (۸) توزیع تابعی از کسینوس هایپربولیک (HCF) نامیده می شود و تابع چگالی احتمال (PDF) آن به صورت زیر است:

$$g(x; a) = \frac{2ae^a}{e^{2a} - 1} f(x) \cosh(aF(x)) \quad (۸)$$

که $a > 0$ و $x > 0$. کاملاً واضح است که $g(x, a)$ زمانی که $a \rightarrow 0$ تابع $f(x)$ را تولید می کند.

با استفاده از بسط سری (۷)، توزیع HCF به صورت زیر است:

$$g(x; a) = \frac{2ae^a}{e^{2a} - 1} f(x) \cosh(aF(x)) = \sum_{n=0}^{\infty} w(a, n) f_U(x)$$

که U دارای توزیع کلی F است. و

$$f_U(x) = (2n + 1) f(x) F^{2n}(x)$$

و

$$w(a, n) = \frac{2ae^a}{e^{2a} - 1} \frac{a^{2n}}{(2n + 1)!}$$

تابع توزیع (CDF) توزیع مزبور به صورت زیر است:

$$G(x, a) = \frac{2e^a}{e^{2a} - 1} \sinh(aF(x)) \quad (۹)$$

تابع قابلیت اطمینان بقا $\bar{G}(x)$ و تابع نرخ شکست $h(x)$ ، (FRF) برای توزیع HCF به ترتیب به صورت زیر هستند:

$$\bar{G}(x, a) = 1 - \frac{2e^a}{e^{2a} - 1} \sinh(aF(x)) \quad (۱۰)$$

و

$$h(x, a) = \frac{\frac{2ae^a}{e^{2a} - 1} f(x) \cosh(aF(x))}{1 - \frac{2e^a}{e^{2a} - 1} \sinh(aF(x))} \quad (۱۱)$$

چندک مرتبه p -ام یعنی x_p توزیع HCF از رابطه زیر بدست می آید:

$$x_p = F^{-1} \left(\frac{1}{a} \operatorname{arcsinh} \left(\frac{e^{2a} - 1}{2e^a} p \right) \right), \quad 0 \leq p \leq 1 \quad (۱۲)$$

که

$$\operatorname{arcsinh}(x) = \ln(x + \sqrt{x^2 + 1})$$

بنابراین داریم

$$x_p = F^{-1} \left(\frac{1}{a} \ln \left(\frac{e^{2a} - 1}{2e^a} p + \sqrt{\left(\frac{e^{2a} - 1}{2e^a} p \right)^2 + 1} \right) \right) \quad (۱۳)$$

۴ توزیع جدید

همانطور که قبلاً گفته شد توزیع هایی که براساس هسته ی وایبول تولید شده اند کاربردهای زیادی در قابلیت اعتماد و تجزیه و تحلیل داده های طول عمر دارند. در این قسمت به معرفی توزیع جدید می پردازیم. توزیع جدید را با انتخاب F در رابطه (۹) به عنوان توزیع وایبول بخت معرفی می کنیم. گوییم X دارای توزیع کسینوس هایپربولیک وایبول بخت است و آن را با نماد $HCOW(\alpha, \beta, \theta, a)$ نشان می دهیم. هرگاه تابع توزیع آن به صورت

$$G(x, \alpha, \beta, \theta, a) = \frac{2ae^a \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right]}{-1 + e^{2a}} \quad (14)$$

باشد. تابع چگالی، تابع بقا و تابع نرخ شکست با توجه به رابطه (۱۴) به ترتیب به صورت زیر به دست می آیند:

$$g(x, \alpha, \beta, \theta, a) = \frac{2ae^{a+(\frac{x}{\theta})^\alpha} (-1 + e^{(\frac{x}{\theta})^\alpha})^{\beta-1} \alpha \beta (\frac{x}{\theta})^\alpha \cosh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right]}{(-1 + e^{2a})(1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta)^2 x} \quad (15)$$

$$\bar{G}(x, \alpha, \beta, \theta, a) = 1 - \frac{2ae^a \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right]}{-1 + e^{2a}} \quad (16)$$

$$h(x, \alpha, \beta, \theta, a) = \quad (17)$$

$$\frac{2ae^{a+(\frac{x}{\theta})^\alpha} (-1 + e^{(\frac{x}{\theta})^\alpha})^{\beta-1} \alpha \beta (\frac{x}{\theta})^\alpha \cosh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right]}{(1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta)^2 x (-1 + e^{2a} - 2ae^a \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right])}$$

با توجه به تعریف ۱.۳ رابطه (۱۵) زمانیکه $a \rightarrow 0$ ، $g(x, \alpha, \beta, \theta, a)$ همان تابع چگالی توزیع وایبول بخت را تولید می کند. شکل ۱ برخی از توابع چگالی توزیع $HCOW$ را برای مقادیر مختلف α, β, θ, a نشان می دهند. با توجه به این شکل، انعطاف پذیری بالای توزیع از نظر چاولگی و کشیدگی و ساختار آن مشاهده می شود. شکل ۲ توابع نرخ شکست را برای مقادیر مختلف α, β, θ, a نشان می دهد. با توجه به این نمودارها می توان ادعا نمود که توزیع معرفی شده قادر است به انواع داده های بقا و قابلیت اعتماد که دارای تابع نرخ شکست صعودی، نزولی، وانی شکل و حتی زینی شکل باشند با انعطاف پذیری مناسبی برازنده شود.

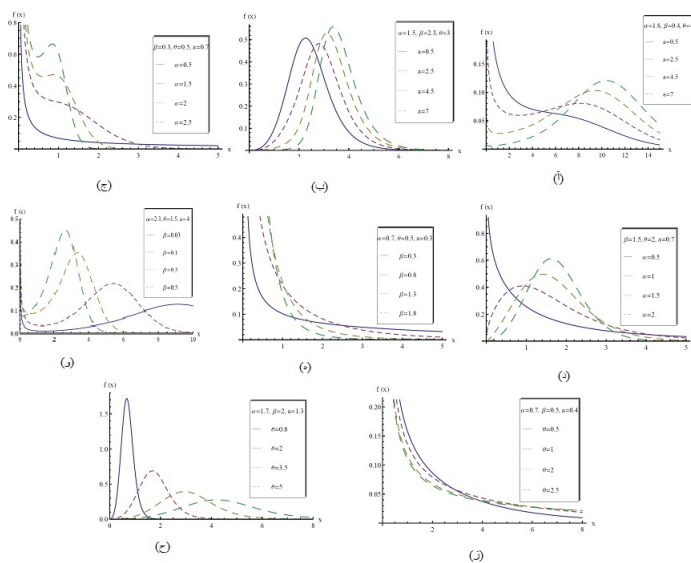
۱.۴ گشتاورها

در حالت کلی برای گشتاورهای توزیع کسینوس هایپربولیک وایبول بخت نمی توان فرم بسته ای به دست آورد. مقادیر عددی واریانس، میانگین، چاولگی و کشیدگی این توزیع در جدول ۱ آمده است.

۲.۴ چندک مرتبه P ام

از معادله (۱۴) داریم

$$\frac{2ae^a \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^\alpha})^\beta} \right) \right]}{-1 + e^{2a}} = p$$



شکل ۱: تابع چگالی برای مقادیر مختلف α ، β ، θ ، a

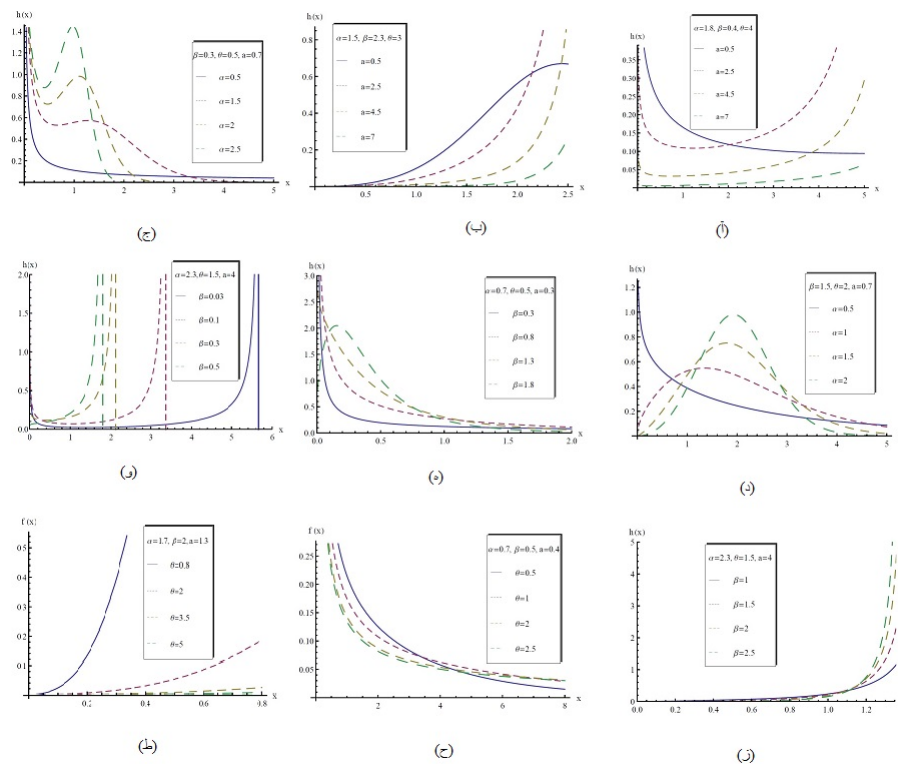
بنابراین چندک مرتبه p ام این توزیع برابر است با

$$x_p = \theta \log \left[1 + \left(\frac{1}{-1 + \frac{a}{\text{Arcsinh}[p \sinh[a]]}} \right)^{\frac{1}{\beta}} \right]^{\frac{1}{\alpha}}$$

$$= \theta \log \left[1 + \left[-\frac{\sqrt{4 - 2p^2} + 2p^2 \cosh[2a] + 2p \sinh[a]}{-2a + \sqrt{4 - 2p^2} + 2p^2 \cosh[2a] + 2p \sinh[a]} \right]^{\frac{1}{\beta}} \right]^{\frac{1}{\alpha}}$$

به ازای $P = 0.5$ رابطه بالا به صورت

$$x_{0.5} = \theta \log \left[1 + \left[-\frac{\sqrt{3/5} + 0.5 \cosh[2a] + \sinh[a]}{-2a + \sqrt{3/5} + 0.5 \cosh[2a] + \sinh[a]} \right]^{\frac{1}{\beta}} \right]^{\frac{1}{\alpha}}$$



شکل ۲: تابع نرخ شکست برای مقادیر مختلف α, β, θ, a

۳.۴ تابع چگالی آماره ترتیبی

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع $HCOW(\alpha, \beta, \theta, a)$ باشند و فرض کنید $X_{i:n}$ ، i امین آماره ترتیبی باشد در این صورت تابع چگالی $X_{i:n}$ به صورت زیر است:

$$f_{X_{i:n}}(x) = \frac{n!}{(i-1)!(n-i)!} \left[\frac{\alpha e^{\alpha} \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^{\alpha}})^{\beta}} \right) \right]}{-1 + e^{\alpha}} \right]^{i-1} \times$$

$$\frac{\alpha e^{a + (\frac{x}{\theta})^{\alpha}} (-1 + e^{(\frac{x}{\theta})^{\alpha}})^{\beta-1} \alpha \beta (\frac{x}{\theta})^{\alpha} \cosh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^{\alpha}})^{\beta}} \right) \right]}{(-1 + e^{\alpha})(1 + (-1 + e^{(\frac{x}{\theta})^{\alpha}})^{\beta})^{\alpha} x} \times$$

$$\left[1 - \frac{\alpha e^{\alpha} \sinh \left[a \left(1 - \frac{1}{1 + (-1 + e^{(\frac{x}{\theta})^{\alpha}})^{\beta}} \right) \right]}{-1 + e^{\alpha}} \right]^{n-i}$$

مراجع

- [1] Bebbington, M., Lai, C.D., Zitikis, R.(2007), *A flexible Weibull extension*, Reliab Eng Syst Safety ;92:719–726.
- [2] Cooray, K. (2006), *Generalization of the Weibull distribution: the odd Weibull family*, Statist. Modelling, 6(3) , 265–277.
- [3] Mudholkar, G. S. and Srivastava, D. K. (1993), *Exponentiated Weibull family for analyzing bathtub failure rate data*, IEEE Trans. Reliability, 42(2), 299–302.
- [4] Omid, Kh., Ali, S. (2016), *Hyperbolic cosine - F family Of distributions with an application to Exponential distribution*, Gazi University Journal of Science , 29(4), 811–829.
- [5] Xie, M. and Lai, C. D. (1995), *Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function*, Reliability Engng System Safety, 52(1) 87–93.

کشیدگی	چاولگی	واریانس	میانگین	a	θ	β	α
۱۲/۸۱۱	۵/۵۵۳	۱۶/۷۴۸	۴/۵۳۱	۰/۵	۴	۰/۴	۱/۸
۲۰/۹۰۶	۸/۵۲۰	۱۹/۳۸۸	۶/۷۶۵	۲/۵			
۵۰/۴۰۹	۱۶/۷۸۳	۱۶/۳۶۲	۸/۷۹۳	۴/۵			
۱۲۷/۲۷۲	۳۳/۴۴۴	۲/۳۸۱	۴/۴۶۷	۱/۳	۵	۲	۱/۷
۱۲۷/۲۷۲	۳۳/۴۴۴	۱/۱۶۶	۳/۱۲۵		۳/۵		
۳۰/۳۰۷	۵/۴۲۱	۶/۵۷۲	۱/۴۳	۰/۴	۰/۵	۰/۵	۰/۷
۲۸/۲۷۱	۵/۲۶۲	۲۸/۰۹۲	۲/۷۸۷	۰/۳	۰/۵	۰/۳	۰/۷
۳۲/۰۰۱	۵/۶۶۹	۱/۱۶۶	۰/۸۰۶			۰/۸	
۳۳/۵۶۲	۶/۱۸۷	۰/۴۱۴	۰/۵۱۴			۱/۳	
۵۹/۷۴۱	۷/۱۳۳	۱۷/۹۷۳	۲/۴۷۶	۰/۷	۲	۱/۵	۰/۵
۲۷/۷۰۵	۸/۰۴۱	۱/۷۹۴	۱/۷۸۰				۱
۴۷/۷۹۱	۱۴/۷۴۸	۰/۷۶۸	۱/۷۴۹				۱/۵

جدول ۱: میانگین، واریانس، چاولگی و کشیدگی توزیع $HCOW(\alpha, \beta, \theta, a)$ برای مقادیر متفاوتی از α, β, θ و a .



اندازه ی نادرستی باقیمانده تجمعی پویا

کشمیری، ز^۱ عباس نژاد، م^۲

^{۱،۲} گروه آمار، دانشکده ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری

چکیده

آنتروپی شانون معیاری از عدم حتمیت در نظریه اطلاع است که در حالت پیوسته دارای مشکلاتی از جمله منفی شدن می باشد، برای رفع این مشکلات اندازه ی جدیدی از آنتروپی بر اساس تابع بقای متغیر تصادفی به نام آنتروپی باقیمانده ی تجمعی (CRE) معرفی شد. این مفهوم برای زمانی که اطلاعی در مورد سن سیستم داریم مناسب نیست، لذا به حالت پویا تعمیم یافت. مفهوم (CRE) به اندازه ی نادرستی باقیمانده ی تجمعی (CRI) گسترش یافت و سپس به حالت پویای آن تعمیم داده شد. همچنین مسأله مشخصه سازی بر اساس اندازه ی نادرستی باقیمانده ی تجمعی پویا تحت مدل مخاطره متناسب مطالعه شد و سه توزیع نمایی، پاراتو و دامنه متناهی طبق این اندازه مشخصه سازی شدند.

کلمات کلیدی: آنتروپی شانون، طول عمر باقیمانده، نرخ مخاطره متناسب، مشخصه سازی.

۱ مقدمه

مهمترین معیار عدم حتمیت در نظریه اطلاع آنتروپی شانون (۱۹۴۸) [۱] است که برای متغیر تصادفی گسسته X با تابع احتمال $p(x) = P(X = x)$ به صورت زیر تعریف می شود:

$$H(X) = - \sum p(x) \log p(x).$$

^۱ zahra.keshmiri7021@gmail.com

^۲ m.abbasnejad@hsu.ac.ir

تعمیم این اندازه به حالت پیوسته برای متغیر تصادفی مطلقاً پیوسته نامنفی X با تابع چگالی احتمال $f(x)$ به صورت زیر می باشد:

$$H(f) = - \int_0^{\infty} f(x) \log f(x) dx.$$

در حالت پیوسته چالش هایی از جمله امکان منفی شدن آنتروپی که در عمل قابل درک نیست به وجود می آید. اولین تلاش برای برطرف کردن این چالش توسط راثو (۲۰۰۴) [۲] تحت عنوان آنتروپی باقیمانده ی تجمعی (CRE) با جایگزینی تابع بقا $(\bar{F}(x) = 1 - F(x))$ به جای تابع چگالی در دستور آنتروپی شانون آغاز شد که برای متغیر تصادفی نامنفی X به صورت زیر تعریف می شود:

$$\xi(F) = - \int_0^{\infty} \bar{F}(x) \log \bar{F}(x) dx.$$

فرض کنید X متغیر تصادفی مطلقاً پیوسته نامنفی با تابع چگالی واقعی $f(x)$ باشد. همچنین فرض کنید مدل در نظر گرفته شده برای توزیع X ، دارای تابع چگالی $g(x)$ باشد. فاصله کولبک-لیبلر بین دو توزیع به صورت زیر تعریف می شود:

$$D(f, g) = - \int_0^{\infty} f(x) \log \frac{f(x)}{g(x)} dx.$$

مفهوم اندازه ی نادرستی که رابطه ای تنگاتنگ با آنتروپی شانون و فاصله کولبک-لیبلر دارد، توسط کریج (۱۹۶۱) [۳] برای متغیر تصادفی مطلقاً پیوسته و نامنفی X به صورت زیر ارائه شد:

$$H(f, g) = - \int_0^{\infty} f(x) \log g(x) dx = H(f) + D(f, g). \quad (۱)$$

که در آن $f(x)$ تابع چگالی مدل واقعی و $g(x)$ تابع چگالی مدل فرض شده است. اندازه ی نادرستی باقیمانده ی تجمعی که با جایگذاری تابع بقا به جای تابع چگالی در رابطه (۱) به صورت:

$$\xi(F, G) = - \int_0^{\infty} \bar{F}(x) \log \bar{G}(x) dx.$$

بیان می شود که در آن $\bar{F}(x)$ و $\bar{G}(x)$ به ترتیب توابع بقای مدل واقعی و مدل فرضی است. این اندازه توسط تانجا و کومار (۲۰۱۲) [۴] ارائه شد. وقتی می دانیم یک سیستم تا زمان t کار کرده است، متغیر مورد نظر ما متغیر طول عمر باقیمانده یعنی $X_t = [X - t | X > t]$ می باشد. بدیهی است که در چنین موقعیتی هیچ کدام از اندازه های $\xi(F, G)$ ، $H(f, g)$ ، $D(f, g)$ و $H(f)$ ، $\xi(F)$ کارایی لازم را ندارند. در این راستا مفاهیمی چون اندازه آنتروپی پویا، اطلاع کولبک-لیبلر پویا، آنتروپی باقیمانده ی تجمعی پویا و اندازه نادرستی پویا معرفی شده اند. در این مقاله در بخش دوم با در نظر گرفتن $\xi(F, G)$ برای متغیر تصادفی X_t اندازه ای را به عنوان اندازه ی نادرستی باقیمانده ی تجمعی پویا ($DCRI$) معرفی می کنیم. در بخش سوم نشان می دهیم $DCRI$ تابع بقای توزیع طول عمر را به طور یکتا تعیین می کند. همچنین نشان می دهیم می توان به کمک $DCRI$ توزیع های نمایی، پاراتو و دامنه متاهی را مشخصه سازی کرد، در پایان با ذکر مثال و نتیجه ای در این راستا بحث را کامل می کنیم.

۲ اندازه ی نادرستی باقیمانده تجمعی پویا

در آزمایش طول عمر آزمونگر اطلاعاتی در مورد سن سیستم در اختیار دارد، در این صورت اندازه ی نادرستی باقیمانده تجمعی کمیت مناسبی برای این وضعیت نیست، لذا این اندازه را تحت عنوان اندازه نادرستی باقیمانده ی تجمعی پویا اصلاح می کنیم. اگر X طول عمر یک سیستم باشد به شرطی که سیستم تا زمان t عمر کند، آنگاه X_t را متغیر تصادفی طول عمر باقیمانده ی می نامیم و به صورت زیر تعریف می کنیم:

$$X_t = [X - t | X > t]$$

که بیانگر اختلاف زمان بین سن سیستم و زمان t است در صورتی که می دانیم سیستم تا زمان حداقل t عمر می کند. تابع بقای X_t به صورت

$$\bar{F}_t(x) = \begin{cases} \frac{\bar{F}(x+t)}{\bar{F}(t)} & \forall x > t \\ 1 & o.w \end{cases}$$

است.

تعریف ۱.۲. اندازه ی نادرستی باقیمانده ی تجمعی برای متغیر تصادفی X_t به صورت زیر تعریف می شود:

$$\xi(F, G; t) = - \int_t^\infty \bar{F}_t(x) \log \bar{G}_t(x) dx = - \int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)} \log \frac{\bar{G}(x)}{\bar{G}(t)} dx. \quad (2)$$

که در آن $\bar{F}_t(x)$ و $\bar{G}_t(x)$ به ترتیب تابع بقای متغیر X_t برای مدل واقعی و مدل فرضی هستند.

نکته ۱.۲. زمانی که t به صفر میل کند اندازه ی $\xi(F, G; t)$ به $\xi(F, G)$ تبدیل می شود یعنی:

$$\lim_{t \rightarrow 0} \xi(F, G; t) = \xi(F, G).$$

مثال ۱.۲. فرض کنید X متغیر تصادفی نامنفی با تابع چگالی

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{o.w} \end{cases}$$

و تابع بقای

$$\bar{F}(x) = 1 - x^2 \quad 0 \leq x \leq 1$$

باشد و همچنین متغیر تصادفی نامنفی Y از توزیع $U(0, 1)$ است که $0 \leq y \leq 1$ ، $\bar{G}(y) = 1 - y$ با جایگذاری این مقادیر در رابطه ی (۲) و ساده سازی داریم:

$$\xi(F, G; t) = \begin{cases} \frac{9(1-t) - 2(1-t)^2}{18(1+t)} & 0 \leq x < 1 \\ 0 & \text{o.w} \end{cases}$$

$$\lim_{t \rightarrow 0} \xi(F, G; t) = \frac{7}{18}$$

۳ مشخصه سازی

هدف از مشخصه سازی توزیع، شناسایی توزیع بوسیله ی ویژگی های آن است. در اینجا با استفاده از $DCRI$ تابع توزیع طول عمر را به طور یکتا تعیین می کنیم. گوئیم دو متغیر تصادفی X و Y در فرمول مخاطره ی متناسب (PHM) صدق می کنند اگر:

$$\bar{G}(x) = [\bar{F}(x)]^\beta.$$

یا به طور معادل $\lambda_G(x) = \beta \lambda_F(x)$ که در آن $\lambda_G(x) = \frac{g(x)}{\bar{G}(x)}$ و $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$ به ترتیب توابع نرخ خطر متغیرهای X و Y هستند. در اینجا لازم است رابطه بین تابع میانگین باقیمانده ی طول عمر و تابع نرخ خطر را بیان کنیم. می دانیم تابع میانگین باقیمانده ی طول عمر به صورت

$$\delta_F(t) = \frac{\int_t^\infty \bar{F}(x) dx}{\bar{F}(t)}.$$

تعریف می شود و رابطه ی بین تابع نرخ خطر و میانگین باقیمانده ی طول عمر به صورت زیر است:

$$\lambda_F(t) = \frac{1 + \delta'_F(t)}{\delta_F(t)}. \quad (3)$$

قضیه ۱.۳. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع بقای به ترتیب $\bar{F}(\cdot)$ و $\bar{G}(\cdot)$ باشند که در مدل مخاطره ی متناسب (PHM) صدق می کنند و فرض کنید $\xi(F, G; t) < \infty \quad \forall t \geq 0$ باشد آنگاه $\xi(F, G; t)$ تابع بقای $\bar{F}(\cdot)$ متغیر X را به طور یکتا تعیین می کند.

برهان. اندازه ی نادرستی باقیمانده ی تجمعی پویا را می توان به صورت زیر نوشت

$$\xi(F, G; t) = - \int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)} \log \frac{\bar{G}(x)}{\bar{G}(t)} dx = - \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(x) \log \bar{G}(x) dx + \delta_F(t) \log \bar{G}(t) \quad (4)$$

حال با توجه به این که X و Y در مدل مخاطره ی متناسب صدق می کنند داریم:

$$\xi(F, G; t) = - \frac{\beta}{\bar{F}(t)} \int_t^\infty \bar{F}(x) \log \bar{F}(x) + \beta \delta_F(t) \log \bar{F}(t).$$

حال با مشتق گیری از $\xi(F, G; t)$ نسبت به t می توان نوشت:

$$\xi'(F, G; t) = \beta \log F(t) [1 + \delta'_F(t)] - \beta \lambda_F(t) \int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)} \log \bar{F}(x) - \beta \lambda_F(t) \delta_F(t).$$

با جایگذاری (۳) و (۴) در فرمول $\xi'(F, G; t)$ داریم

$$\xi'(F, G; t) = \lambda_F(t) [\xi(F, G; t) - \beta \delta_F(t)].$$

فرض کنید G_1, F_1 و G_2, F_2 دو مجموعه از تابع توزیع احتمال که در PHM صدق می کنند و

$$\xi(F_1, G_1; t) = \xi(F_2, G_2; t) \quad \forall t \geq 0.$$

باشد

با مشتق گیری از طرفین نسبت به t داریم

$$\lambda_{F_1}(t) [\xi(F_1, G_1; t) - \beta \delta_{F_1}(t)] = \lambda_{F_2}(t) [\xi(F_2, G_2; t) - \beta \delta_{F_2}(t)]. \quad (5)$$

اگر $\lambda_{F_1}(t) = \lambda_{F_2}(t), \forall t \geq 0$ آنگاه: $\bar{F}_1(t) = \bar{F}_2(t)$ و اثبات کامل است. در غیر این صورت فرض کنید

$$A = \{t : t \geq 0 \quad \& \quad \lambda_{F_1}(t) \neq \lambda_{F_2}(t)\}$$

مجموعه ای ناتهی می باشد. بنابراین

$$\exists t_0 \in A, \lambda_{F_1}(t_0) \neq \lambda_{F_2}(t_0)$$

بدون از دست کلیت، فرض می کنیم $\lambda_{F_2}(t_0) > \lambda_{F_1}(t_0)$. بنابراین با این فرض در رابطه (۵) برای حفظ تساوی باید

$$\xi(F_1, G_1; t_0) - \beta \delta_{F_1}(t_0) > \xi(F_2, G_2; t_0) - \beta \delta_{F_2}(t_0).$$

□

در نتیجه $\delta_{F_1}(t_0) < \delta_{F_2}(t_0)$ که به تناقض می رسیم، بنابراین مجموعه A تهی است و اثبات کامل می شود.

در قضیه ی زیر توزیع های نمایی، پاراتو و دامنه متناهی را با استفاده از DCRI مشخصه سازی می کنیم:

قضیه ۲.۳. فرض کنید X و Y دو متغیر تصادفی پیوسته نامنفی باشند که در PHM صدق می کند. همچنین فرض کنید X دارای میانگین باقیمانده طول عمر $\delta_F(t)$ باشد آنگاه

$$\xi(F, G; t) = c \delta_F(t) \quad c > 0 \quad (6)$$

اگر و تنها اگر:

- i- برای $c = \beta$ توزیع نمایی باشد.
- ii- برای $c > \beta$ توزیع پاراتو باشد.
- iii- برای $c < \beta$ توزیع دامنه ی متناهی باشد.

برهان. (لزوم) i- اگر X دارای توزیع نمایی با تابع بقای $\theta > 0$ ، $\bar{F}(x) = \exp(-\theta x)$ ، آنگاه: $\delta_F(t) = \frac{1}{\theta}$ با جایگذاری در $DCRI$ و با استفاده از PHM داریم:

$$\xi(F, G; t) = \frac{\beta}{\theta} = c\delta_F(t),$$

برای $c = \beta$

ii- اگر X دارای توزیع پاراتو با تابع چگالی احتمال زیر باشد:

$$f(x) = \frac{ab^a}{(x+b)^{a+1}} \quad a > 1 \quad b > 0.$$

آنگاه تابع بقا به صورت $\bar{F}(x) = 1 - F(x) = \frac{b^a}{(a+b)^a}$ و تابع میانگین طول عمر باقیمانده به صورت $\delta_F(t) = \frac{t+b}{a-1}$ است، بنابراین $DCRI$ تحت PHM عبارت است از

$$\xi(F, G; t) = \frac{\beta a(t+b)}{(a-1)^2} = c\delta_F(t),$$

برای $c = \frac{\beta a}{a-1} > \beta$

iii- در حالتی که X دارای توزیع دامنه متناهی با تابع چگالی

$$f(x) = a(1-x)^{a-1} \quad a > 1 \quad 0 \leq x \leq 1$$

و تابع بقای $\bar{F}(t) = (1-x)^a$ و تابع میانگین طول عمر باقیمانده $\delta_F(t) = \frac{1-t}{a+1}$ باشد، آنگاه $DCRI$ تحت PHM به صورت زیر می باشد:

$$\xi(F, G; t) = \frac{\beta a(1-t)}{(a+1)^2} = c\delta_F(t),$$

برای $c = \frac{\beta a}{a+1}$

(کفایت) فرض می کنیم (۶) برقرار باشد. در این صورت با استفاده از رابطه (۴) داریم:

$$\frac{\beta}{\bar{F}(t)} \int_0^\infty \bar{F}(x) \log \bar{F}(x) dx + \beta \delta_F(t) \log \bar{F}(t) = c\delta_F(t).$$

با مشتق گیری از طرفین نسبت به t می توان نوشت

$$\begin{aligned} \frac{c}{\beta} \delta'_F(t) &= \delta'_F(t) \log \bar{F}(t) \\ &- \lambda_F(t) \delta_F(t) + \log \bar{F}(t) - \lambda_F(t) \frac{1}{\bar{F}(t)} \int_0^\infty \bar{F}(x) \log \bar{F}(x) dx \\ &= \delta'_F(t) \log \bar{F}(t) - \lambda_F(t) \delta_F(t) + \log \bar{F}(t) + \lambda_F(t) \left[\frac{c}{\beta} \delta_F(t) - \delta_F(t) \log \bar{F}(t) \right]. \end{aligned}$$

با استفاده از (۳) در رابطه بالا به دست می آوریم:

$$\lambda_F(t) \delta_F(t) = \frac{c}{\beta} \Rightarrow \delta'_F(t) = \frac{c}{\beta} - 1.$$

با انتگرال گیری از طرفین نسبت به t روی بازه $(0, x)$ داریم:

$$\delta_F(x) = \left(\frac{c}{\beta} - 1\right)x + \delta_F(0).$$

بنابراین تابع میانگین طول عمر باقیمانده $\delta_F(x)$ تابعی خطی از متغیر تصادفی پیوسته Y نامنفی X است اگر و تنها اگر: برای $c = \beta$ توزیع نمایی باشد و برای $c > \beta$ توزیع پاراتو و برای $c < \beta$ توزیع دامنه متناهی باشد. به حال و ولتر [۵] مراجعه کنید. $\square$

در قضیه ی بعد حالت کلی تر قضیه قبل را بیان می کنیم به این صورت که c را به صورت تابعی از t در نظر می گیریم:

قضیه ۳.۳. فرض کنید X و Y دو متغیر تصادفی پیوسته و نامنفی باشند که در PHM صدق کنند و

$$\xi(F, G; t) = c(t)\delta_F(t) \quad \text{if } t \geq 0 \quad (7)$$

$$\delta_F(t) = \left[k + \int_0^\infty \left\{ \frac{c(x) - \beta}{\beta} \right\} e^{\frac{c(x)}{\beta}} dx \right] e^{\frac{c(t)}{\beta}} \quad (8)$$

$$k = \delta_F(0) e^{\frac{c(0)}{\beta}}.$$

برهان. با جایگذاری (۷) در $\xi'(F, G; t)$ داریم:

$$\xi'(F, G; t) = \lambda_F(t)\delta_F(t)\{c(t) - \beta\}.$$

با مشتق گیری از (۷) نسبت به t (۶) و جایگذاری $\xi'(F, G; t)$ داریم:

$$c'(t)\delta_F(t) + c(t)\delta_F'(t) = \lambda_F(t)\delta_F(t)\{c(t) - \beta\}.$$

اکنون با استفاده از (۳) می توان نوشت

$$\delta_F'(t) + \frac{c'(t)}{\beta}\delta_F(t) = \frac{c(t) - \beta}{\beta},$$

□

که یک معادله دیفرانسیل خطی مرتبه ی اول از $\delta_F(t)$ است که با حل آن به (۸) می رسیم.

مثال ۱.۳. فرض کنید $c(t) = at + b$ و $a > 0$ با جایگذاری در فرمول (۸) به دست می آوریم:

$$\delta_F(t) = ke^{\frac{-(at+b)}{\beta}} + \frac{at - 2\beta + b}{a} - \frac{(b - 2\beta)e^{\frac{-at}{\beta}}}{a}$$

در حالت $a = 0$ همان نتیجه قضیه ۲.۳ به دست می آید.

نتیجه ۱.۳. در مثال ۱.۳ اگر $\beta = 1$ باشد $\delta_F(t)$ به صورت زیر تبدیل می شود:

$$\delta_F(t) = ke^{-at-b} + \frac{b - 2 + at}{a} - \frac{(b - 2)e^{-at}}{a}$$

مراجع

- [1] Shanon, C.E. (1948), *A mathematical theory of communication*, Bell System Technical Journal. PP, 279–423 and 623–656.
- [2] Rao, M. and Chen, Y. and Vemuri, B.C. and Wang, F. (2004), *Cumulative residual entropy: a new measure of information*, IEEE Transaction on Information Theory 50, 1220-1228.
- [3] Kerridge, D.F. (1961), *Inaccuracy and inference*, Roy. Statist. Soc. Ser. B 23, 184–194.
- [4] Tanja, H.C. and Kumar, V. (2012), *On Dynamic Cumulative Residual Inaccuracy Measure*, In Proceedings of the World Congress on Engineering 2012 Vol I.
- [5] Hall W.J. and Wellner, J.A. (1981), *Mean residual life*, North-Holland. Amsterdam. 169–184.



توابع محدب، مقعر و ارتباط آن با مفاهیم قابلیت اعتماد و مدل های اقتصاد

محمدی پور، م^۱ فشنادی، م^۲ احمدی، ج^۳

^{۱،۲،۳} گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

چکیده

یکی از تعمیم های معروف محدب (مقعر)، لگ-محدب (لگ-مقعر) می باشد که اگر تابع چگالی احتمال دارای خاصیت لگ-مقعر باشد آنگاه نرخ خطر دارای خاصیت صعودی است. در این مقاله نامساوی های دیگری از نوع محدب (مقعر) را در نظر گرفته، خواص آنها بررسی و ارتباط آنها با مفاهیم قابلیت اعتماد مطالعه می شود.

کلمات کلیدی: تابع محدب (مقعر)، نامساوی از نوع هندسی-هندسی، لگ-محدب، نامساوی از نوع هندسی-حسابی، تابع بقاء، تابع نرخ خطر.

۱ مقدمه

معمولا خانواده ای از توابع که دارای ویژگی خاصی باشد، همواره مورد توجه پژوهشگران ریاضی و علاقمندان به نظریه توزیع ها در آمار و احتمال و مطالعات نظریه قابلیت اعتماد است. توابع محدب و مقعر یکی از این موارد هستند که مقالات زیادی در خصوص بررسی ویژگیهای آنها به چاپ رسیده است. ارتباط نزدیکی بین خاصیت تقعر و تحدب تابع چگالی احتمال و تابع نرخ خطر وجود دارد که برای بررسی خواص صعودی

^۱ mo.mohammadipour@um.ac.ir

^۲ fashnadi@um.ac.ir

^۳ ahmadi-j@um.ac.ir

و نزولی بودن تابع نرخ خطر مربوط به طول عمر سیستم ها در مباحث قابلیت اعتماد مفید است. ابتدا مروری بر تابع محدب (مقعر) معمولی داشته و آنگاه تعمیم های مختلف آن را بیان می کنیم.

تابع $f: R \rightarrow R$ را محدب (مقعر) گویند، اگر نامساوی زیر برای هر $x, y \in R$ و $0 \leq \alpha \leq 1$ برقرار باشد

$$f(\alpha x + (1 - \alpha)y) \leq (\geq) \alpha f(x) + (1 - \alpha)f(y). \quad (1)$$

اگر نامساوی در (۱) اکید باشد آنگاه f را اکیدا محدب (مقعر) می گویند. در مقالات رابطه (۱) را محدب (مقعر) حسابی-حسابی (A-A) نیز می گویند. موارد زیر معادل با نامساوی (۱) هستند:

الف) در صورت وجود مشتق اول، تابع $f'(x)$ صعودی (نزولی) باشد.

ب) در صورت وجود مشتق دوم، $f''(x)$ نامنفی (نامثبت) باشد.

ویژگیهای زیادی برای توابع محدب (مقعر) معمولی بدست آمده، برخی از آنها عبارتند از:

۱. جمع دو تابع محدب (مقعر) تابعی محدب (مقعر) است. اگر یکی از آنها اکیدا محدب (مقعر) باشد آنگاه مجموع نیز اکیدا محدب (مقعر) است.

۲. ضرب یک تابع (اکیدا) محدب (مقعر) با یک اسکالر مثبت، تابعی (اکیدا) محدب (مقعر) بدست می آید.

۳. اگر $f: R \rightarrow R$ یک تابع محدب و $g: R \rightarrow R$ تابعی محدب صعودی باشد آنگاه ترکیب آنها محدب است.

تعمیم های مختلفی توسط پژوهشگران براساس تعریف محدب و مقعر صورت گرفته است که در ادامه به آنها اشاره خواهد شد. در این مقاله هدف بررسی صورت های مختلف توابع محدب و مقعر است که در مطالعات مربوط به نظریه توزیع ها، قابلیت اعتماد و مدل های اقتصاد نقش ایفا می کنند.

۲ برخی از شاخص های قابلیت اعتماد

در اینجا برخی از مفاهیم قابلیت اعتماد را یادآوری می کنیم. فرض کنید X متغیر تصادفی نامنفی پیوسته نشان دهنده ی طول عمر یک سیستم باشد که دارای تابع چگالی احتمال $f(x)$ و تابع توزیع $F(x)$ باشد، دراینصورت انتگرال چپ تابع توزیع تجمعی عبارتست از

$$F_l(x) = \int_a^x F(t)dt.$$

بگنولی و برگستر (۲۰۰۵) ضمن بررسی انتقال خاصیت محدب و مقعر بودن از $f(x)$ و $F(x)$ به $F_l(x)$ کاربردهایی از آن را در مدل های اقتصادی مطالعه کردند. $\bar{F}(x) = 1 - F(x)$ را تابع قابلیت اعتماد سیستم در زمان x گویند، فلین و هکمن (۱۹۸۳) انتگرال راست تابع قابلیت اعتماد یعنی $\bar{F}_r(x) = \int_x^b \bar{F}(t)dt$ را در نظر گرفته و استفاده از آن را در مدل های بازاریابی و اشتغال یابی بررسی کردند.

در مطالعات طول عمر سیستم ها، معمولاً دو تابع نرخ خطر و نرخ خطر معکوس را در نظر می گیرند که به ترتیب به صورت $r(x) = \frac{f(x)}{F(x)}$ و $\bar{r}(x) = \frac{f(x)}{\bar{F}(x)}$ تعریف می شوند. ارتباط محدب و مقعر بودن $f(x)$ ، $F(x)$ و $\bar{F}(x)$ با خواص صعودی و نزولی بودن $r(x)$ و $\bar{r}(x)$ مورد توجه علاقمندان به قابلیت اعتماد است. اگر سیستم تا زمان x فعال بوده، متغیر باقیمانده طول عمر آن را با $(T_x = T - x \mid T > x)$ نشان داده و امید ریاضی T_x را میانگین باقیمانده طول عمر می گویند که بصورت زیر داده می شود

$$E(T_x) = M(x) = \int_x^b \frac{tf(t)}{\bar{F}(x)} dt - x.$$

همچنین اگر سیستم قبل از زمان x از کار افتاده باشد، متغیر $T_x^* = x - T \mid T < x$ مدت زمان از کار افتادگی آن را تا زمان x نشان می

دهد. امید ریاضی T_x^* را میانگین باقیمانده عمر گذشته می گویند که بصورت زیر بدست می آید

$$E(T_x^*) = \delta(x) = x - \int_a^x \frac{tf(t)}{F(x)} dt.$$

لازم به ذکر است از تابع $\delta(x)$ در مدل های اقتصادی برای تعیین عادلانه بودن حق قیمت گذاری در معاملات خرید و فروش کالاها و ماشین های دست دوم استفاده می شود. خاصیت صعودی و نزولی بودن $\delta(x)$ برحسب x مورد توجه قیمت گذاران است.

۳ تعمیم های مختلف از توابع محدب و مقعر

یکی از تعمیم های نامساوی (۱) بصورت زیر می باشد

$$f(\alpha x + (1-\alpha)y) \leq [f(x)]^\alpha [f(y)]^{1-\alpha}, \quad 0 \leq \alpha \leq 1, \quad (2)$$

که در آن $f: [a, b] \subseteq \mathbb{R} \rightarrow (0, +\infty)$. تابعی که در نامساوی (۲) صدق کند را لگ-محدب گویند. اگر جهت نامساوی فوق عوض شود در اینصورت f را تابعی لگ-مقعر گویند. f دارای خاصیت لگ-محدب (لگ-مقعر) است اگر $\log f$ تابعی محدب (مقعر) باشد. در مقالات رابطه فوق به محدب حسابی-هندسی $(A-G)$ نیز معروف است.

برخی از ویژگیهای توابع لگ-مقعر و لگ-محدب به شرح زیر است:

۱. حاصلضرب توابع لگ-مقعر (لگ-محدب)، تابعی لگ-مقعر (لگ-محدب) می باشد.
۲. مجموع دو تابع لگ-محدب، تابعی لگ-محدب است.
۳. مجموع دو تابع لگ-مقعر، لزومی ندارد لگ-مقعر باشد.
۴. توابع لگ-مقعر و لگ-محدب تحت انتگرالگیری حفظ می شوند.
۵. یک دنباله همگرا از توابع لگ-محدب (لگ-مقعر) دارای تابع حدی لگ-محدب (لگ-مقعر) است.

تابع حقیقی $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ را محدب (مقعر) هندسی گویند هرگاه نامساوی زیر برقرار باشد

$$f(x^\alpha y^{1-\alpha}) \leq (\geq) [f(x)]^\alpha [f(y)]^{1-\alpha}, \quad 0 \leq \alpha \leq 1. \quad (3)$$

درواقع (۳) یکی از تعمیم های توابع محدب و مقعر است. تابع f محدب هندسی است یعنی $\log f$ تابعی محدب از $\log x$ است. به رابطه

(۳) محدب (مقعر) هندسی-هندسی $(G-G)$ نیز می گویند. بریچ (۲۰۱۰) نشان داده است که توزیع های آماری زیادی دارای خاصیت محدب هندسی-هندسی هستند.

یکی دیگر از تعمیم های نامساوی (۱) محدب (مقعر) هندسی-حسابی $(G-A)$ می باشد که تابع حقیقی $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ را محدب (مقعر) هندسی-حسابی $(G-A)$ گویند اگر نامساوی زیر صدق کند

$$f(x^\alpha y^{1-\alpha}) \leq (\geq) \alpha f(x) + (1-\alpha)f(y), \quad 0 \leq \alpha \leq 1. \quad (4)$$

تعمیم نامساوی هندسی-حسابی بصورت زیر می باشد

$$f\left(\prod_{i=1}^m x_i^{\alpha_i}\right) \leq \sum_{i=1}^m \alpha_i f(x_i), \quad (5)$$

که در آن $0 \leq \alpha_i \leq 1$ و $\sum_{i=1}^m \alpha_i = 1$.

بعنوان مثال اگر $f(t) = t$ و $\alpha_1 = \alpha_2 = \dots = \alpha_m = \frac{1}{m}$ (۵) نامساوی زیر نتیجه می شود

$$\left(\prod_{i=1}^m x_i \right)^{\frac{1}{m}} \leq \frac{1}{m} \sum_{i=1}^m x_i,$$

که به عنوان نامساوی میانگین حسابی-هندسی در آمار و احتمال معروف است. برای اطلاعات بیشتر درباره ی تعمیم های مختلف دیگر از محدب و مقعر می توان به نور و همکاران (۲۰۱۴) مراجعه نمود.

۴ کاربرد توابع تعمیم یافته از محدب و مقعر در قابلیت اعتماد

ابتدا لم زیر را از بنگولی و برگستر (۲۰۰۵) یاد آوری می کنیم.

لم ۱.۴. تابع $f(x)$ در بازه ی (a, b) لگ-مقعر است هرگاه یکی از شرط های زیر برقرار باشد:

الف) در صورت وجود مشتق، $\frac{f'(x)}{f(x)}$ یکنوازی نزولی در (a, b) باشد.

ب) در صورت وجود مشتق دوم، $(\log f(x))'' < 0$ باشد.

موارد زیر را در خصوص استفاده از توابع محدب (مقعر) حسابی-هندسی در مباحث قابلیت اعتماد و مدل های اقتصاد داریم

– اگر تابع چگالی احتمال، $f(x)$ ، مقعر حسابی-هندسی باشد آنگاه طول عمر سیستم دارای خاصیت نرخ خطر صعودی است.

– انتگرال چپ تابع توزیع، $F_L(x)$ ، دارای خاصیت مقعر حسابی-هندسی است اگر و تنها اگر میانگین گذشته عمر سیستم، $\delta(x)$ ، تابع صعودی باشد.

– انتگرال راست تابع بقاء، $\bar{F}_r(x)$ ، دارای خاصیت مقعر حسابی-هندسی است اگر و تنها اگر میانگین باقیمانده طول عمر، $M(x)$ ، تابع نزولی است.

– اگر تابع چگالی احتمال، $f(x)$ ، صعودی باشد آنگاه تابع بقاء، $\bar{F}(x)$ ، مقعر حسابی-هندسی و تابع نرخ خطر، $r(x)$ ، صعودی هستند.

اثبات موارد فوق با استفاده از لم ۱.۴ قابل انجام است. برای بررسی ارتباط خاصیت محدب (مقعر) هندسی-هندسی با مفاهیم قابلیت اعتماد و تحلیل بقاء لم زیر را از بریج (۲۰۱۰) بیان می کنیم

لم ۲.۴. در صورت وجود مشتق اول، تابع $f(x)$ محدب (مقعر) هندسی-هندسی است اگر و تنها اگر $\frac{xf'(x)}{f(x)}$ صعودی (نزولی) باشد.

فرض کنید X نشان دهنده متغیر تصادفی طول عمر سیستم با تابع نرخ خطر $r(x)$ باشد، در اینصورت نتایج زیر را داریم.

– اگر تابع $f(x)$ مقعر هندسی-هندسی باشد آنگاه $F(x)$ و $F_L(x)$ مقعر هندسی-هندسی می باشند.

– اگر تابع $f(x)$ محدب هندسی-هندسی باشد و $af(a) = 0$ آنگاه $F(x)$ و $F_L(x)$ محدب هندسی-هندسی می باشند.

– اگر تابع $f(x)$ مقعر هندسی-هندسی باشد آنگاه $xr(x)$ صعودی است.

– اگر تابع $f(x)$ محدب هندسی-هندسی باشد و $af(a) = 0$ آنگاه $xr(x)$ نزولی است.

در ادامه شرط کافی برای خاصیت محدب (مقعر) حسابی-هندسی را برای تابع توزیع و تابع بقاء بررسی و برای این منظور لم زیر را از آندرسن و همکاران (۲۰۰۷) بیان می کنیم

لم ۳.۴. در صورت وجود مشتق اول، $f(x)$ محدب (مقعر) هندسی-هندسی است اگر و تنها اگر $xf'(x)$ صعودی (نزولی) باشد.

فرض کنید f دارای ویژگی محدب هندسی-هندسی باشد، آنگاه

$$f(x^\alpha y^{1-\alpha}) \leq [f(x)]^\alpha [f(y)]^{1-\alpha}. \quad (۶)$$

نامساوی زیر از (۶) نتیجه می شود

$$\log \{f(x^\alpha y^{1-\alpha})\} \leq \alpha \log \{f(x)\} + (1-\alpha) \log \{f(y)\}. \quad (۷)$$

چون لگاریتم تابعی مقعر است، نامساوی

$$\alpha \log \{f(x)\} + (1 - \alpha) \log \{f(y)\} \leq \log \{\alpha f(x) + (1 - \alpha)f(y)\} \quad (۸)$$

را داریم. آنگاه از روابط (۷) و (۸) نامساوی هندسی-حسابی نتیجه می شود، یعنی

$$f(x^\alpha y^{1-\alpha}) \leq \alpha f(x) + (1 - \alpha)f(y).$$

بنابراین خاصیت محدب هندسی-هندسی، خاصیت محدب هندسی-حسابی را نتیجه می دهد.

در خصوص محدب (مقعر) هندسی-حسابی با توجه به موارد فوق نتایج زیر را درباره ی خاصیت محدب (مقعر) هندسی-هندسی از تابع چگالی به تابع بقاء و تابع توزیع داریم:

- اگر تابع $f(x)$ مقعر هندسی-هندسی باشد آنگاه $F_l(x)$ و $F(x)$ مقعر هندسی-حسابی می باشند.

- اگر تابع $f(x)$ مقعر هندسی-هندسی باشد آنگاه $\bar{F}_r(x)$ و $\bar{F}(x)$ مقعر هندسی-حسابی می باشند.

- اگر تابع $f(x)$ محدب هندسی-هندسی باشد و $af(a) = 0$ آنگاه $F_l(x)$ و $F(x)$ محدب هندسی-حسابی هستند.

- اگر تابع $f(x)$ محدب هندسی-هندسی باشد و $bf(b) = 0$ آنگاه $\bar{F}_r(x)$ و $\bar{F}(x)$ محدب هندسی-حسابی هستند.

با استفاده از نتایج بنگولی و برگستر (۲۰۰۵) و بریچ (۲۰۱۰) خواص محدب (مقعر) تعمیم یافته برخی از توزیع های معروف آماری و شاخص های قابلیت اعتماد مربوطه در جداول ۱ و ۲ خلاصه شده اند.

جدول ۱: ویژگی های مقعر (محدب) حسابی-هندسی برخی از توزیع ها

توزیع	تابع چگالی احتمال	f	F	F_l	$\bar{F}$	$\bar{F}_r$	$r(x)$	$\tilde{r}(x)$	$M(x)$	$\delta(x)$
نرمال	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	صعودی	نزولی	نزولی	صعودی
لجستیک	$\frac{e^{-x}}{(1+e^{-x})^2}$	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	صعودی	نزولی	نزولی	صعودی
پارتو	$\beta x^{-\beta-1}$	محدب حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	محدب حسابی-هندسی	محدب حسابی-هندسی	نزولی	نزولی	صعودی	صعودی
وایل	$c x^{c-1} e^{-x^c} (c \geq 1)$	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	صعودی	نزولی	نزولی	صعودی
گاما	$\frac{x^{c-1} e^{-x}}{\Gamma(c)} (c \geq 1)$	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	صعودی	نزولی	نزولی	صعودی
وایل	$c x^{c-1} e^{-x^c} (c < 1)$	محدب حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	محدب حسابی-هندسی	محدب حسابی-هندسی	نزولی	نزولی	صعودی	صعودی
گاما	$\frac{x^{c-1} e^{-x}}{\Gamma(c)} (c < 1)$	محدب حسابی-هندسی	مقعر حسابی-هندسی	مقعر حسابی-هندسی	محدب حسابی-هندسی	محدب حسابی-هندسی	نزولی	نزولی	صعودی	صعودی

جدول ۲: ویژگی های مقعر (محدب) هندسی-هندسی برخی از توزیع ها

توزیع	تابع چگالی احتمال	f	F	F_l	$\bar{F}$	$\bar{F}_r$	$xr(x)$
نرمال	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	صعودی
پارتو	$\beta x^{-\beta-1}$	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	صعودی
وایل	$c x^{c-1} e^{-x^c}$	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	صعودی
گاما	$\frac{x^{c-1} e^{-x}}{\Gamma(c)}$	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	صعودی
توانی	$c x^{c-1} (c \geq 1)$	محدب هندسی-هندسی	محدب هندسی-هندسی	محدب هندسی-هندسی	مقعر هندسی-هندسی	مقعر هندسی-هندسی	صعودی

مراجع

- [1] Bagnoli, M. and Bergstrom, T. (2005), *Log-concave probability and its applications*. Economic Theory, **26**, 445-469.
- [2] Anderson, G. D., Vamanamurthy, M. K. and Vuorinen, M. (2007), *Generalized convexity and inequalities*. The Journal of Mathematical Analysis and Applications, **335**, 1294-1308.
- [3] Baricz, A. (2010), *Geometrically concave univariate distributions*. The Journal of Mathematical Analysis and Applications, **363**, 182-196.
- [4] Noor, M. A., Noor, K. I. and Awan, M. U. (2014), Geometrically relative convex functions. Applied Mathematics Information Science, **8**, 607-616.
- [5] Flinn, C. and Heckman, J. (1983), *Are unemployment and out of the labor force behaviorally distinct labor force states*. Journal of Labor Economics, **1**, 28-43.



برآورد توزیع لگ-گاما تحت داده‌های سانسور شده فزاینده نوع II تطبیقی

مرادیان الوار، س^۱ شرفی، م^۲

^{۱،۲} گروه آمار، دانشکده علوم پایه، دانشگاه رازی

چکیده

در آزمایش‌های مربوط به طول عمر و پایداری می‌توان طرح سانسور شده فزاینده نوع II تطبیقی را برای تعادل بین مدت زمان کل آزمایش، تعداد واحدهای استفاده شده در آزمایش و کارایی استنباط آماری استفاده نمود. در این مقاله ابتدا به معرفی توزیع لگ-گاما پرداخته و سپس طرح سانسور فزاینده نوع II تطبیقی بیان می‌شود. در ادامه تحت این طرح، با استفاده از روش شبیه‌سازی مونت کارلو، برآورد ماکسیمم درستنمایی پارامترهای مکان و مقیاس، طول بازه اطمینان و احتمال پوشش توزیع مذکور محاسبه شده است. در پایان به منظور بهبود احتمال پوشش مربوط به پارامتر مکان، تبدیل باکس-کاکس صورت گرفته است.

کلمات کلیدی: احتمال پوشش، برآورد ماکسیمم درستنمایی، سانسور فزاینده نوع II تطبیقی، فاصله اطمینان، ماتریس اطلاع مشاهده شده.

۱ مقدمه

با توجه به [۵] و [۸]، توزیع لگ-گاما تبدیل لگاریتمی باز پارامتریته‌ی متغیر گامای تعمیم یافته است. از آنجایی که این توزیع یک مدل انعطاف‌پذیر برای داده‌های قابلیت اعتماد است؛ انتخاب آن به عنوان یک مدل احتمال مفید، در آزمون طول عمر و تحلیل قابلیت اعتماد شایسته

^۱moradians30@gmail.com

^۲maryamsharafi@gmail.com

است. اگر W متغیر تصادفی لگ-گاما باشد؛ با تبدیل $X = \sqrt{k}(W - \ln k)$ تابع چگالی احتمال و تابع توزیع تجمعی متناظر با آن به صورت

$$f(y) = \frac{k^{k-1/2}}{\Gamma(k)} \exp(\sqrt{k}x - k \exp(\frac{x}{\sqrt{k}})), \quad F(y) = I_{k \exp(\frac{x}{\sqrt{k}})}(k).$$

است، به طوری که $-\infty < x < \infty$ ، $k > 0$ و $I_t(k)$ تابع گامای ناقص است. با انتخاب $k = 1$ ، متغیر تصادفی X دارای توزیع مقدار فرین خواهد بود. حال زمانی که $k \rightarrow \infty$ متغیر تصادفی X دارای توزیع نرمال استاندارد می باشد. اگر $Y = \mu + \sigma X$ باشد، در این صورت تابع چگالی احتمال به این صورت تعریف می شود:

$$f(y) = \frac{1}{\sigma} \frac{k^{k-1/2}}{\Gamma(k)} \exp(\sqrt{k}(\frac{y-\mu}{\sigma}) - k \exp(\frac{y-\mu}{\sigma\sqrt{k}}))$$

$$-\infty < y < \infty, \quad -\infty < x < \infty, \quad \sigma > 0, \quad k > 0.$$

به طوری که μ پارامتر مکان، σ پارامتر مقیاس و k پارامتر شکل هستند.

۲ طرح سانسور شده فزاینده نوع II تطبیقی

از مباحث مهم در آزمایش های طول عمر و قابلیت اعتماد، مسأله سانسور است. در انجام آزمایش طول عمر ممکن است دو حالت اتفاق بیفتد؛ ۱. اگر m امین شکست قبل از زمان از پیش تعیین شده T رخ دهد، آزمایش با طرح سانسور فزاینده نوع II (از پیش تثبیت شده $R = (R_1, \dots, R_m)$) برابر است. ۲. اگر m امین شکست بعد از زمان T رخ دهد، آن گاه J شکست قبل از T رخ می دهد که، متغیر تصادفی J به صورت $J = \max(j^* : X_{j^*,m:n} < T)$ تعریف می شود. تا مرحله J بردار حذف واحدهای سالم به صورت $R = (R_1, \dots, R_J)$ است و پس از مشاهده J امین شکست تا مشاهده m امین شکست هیچ واحد سالمی از آزمایش حذف نمی شود. به محض مشاهده m امین شکست همه واحدهای باقیمانده حذف می شوند و طرح سانسور به صورت $R = (R_1, \dots, R_J, \cdot^{m-J-1}, n - m - \sum_{i=1}^J R_i)$ خواهد بود. اگر متغیر تصادفی $X = (X_{1:m:n}, \dots, X_{m:m:n})$ نمونه سانسور شده فزاینده نوع II تطبیقی با تابع چگالی احتمال $f(x)$ و تابع بقای $\bar{F}(x)$ باشد، تابع چگالی توأم X و J به صورت زیر بیان می شود:

$$f_{X,J}(x_{1:m:n}, \dots, x_{m:m:n}, j) = c_m I_{[x_{j:m:n}, x_{j+1:m:n}]}(T) \prod_{i=1}^m [\bar{F}^{r_i}(x_{i:m:n}) f(x_{i:m:n})] \quad (1)$$

به طوری که r_i و c_i به صورت روابط زیر تعریف می شوند:

$$r_i = \begin{cases} R_i & i \leq j \\ \cdot & j < i < m \\ n - \sum_{i=1}^j R_i & i = m \end{cases}, \quad c_i = \prod_{s=1}^i \lambda_s, \quad \lambda_i = \sum_{s=i}^m (r_s + 1).$$

برای مطالعه ی بیشتر به [۲]، [۴]، [۱]، [۶] و [۹] مراجعه شود.

تحت توزیع مقدار فرین با فرض نمونه های سانسور فزاینده نوع II تطبیقی توسط یه و همکاران (۲۰۱۴) استنباط آماری انجام شد. با توجه به این که توزیع لگ-گاما شکل کلی تری از توزیع مقدار فرین است، در این مقاله به استنباط آماری توزیع مذکور تحت داده های سانسور فزاینده نوع II تطبیقی پرداخته می شود.

۳ برآورد نقطه‌ای

با توجه به رابطه (۱)، تابع درستنمایی نمونه تصادفی $Y_{1:m:n}, \dots, Y_{m:m:n}$ تحت توزیع لگ-گاما به صورت

$$L = c_m \prod_{i=1}^m f(y_{i:m:n}; \mu, \sigma, k) [1 - F(y_{i:m:n}; \mu, \sigma, k)]^{r_i}$$

یا به طور معادل برای نمونه $X_{1:m:n}, \dots, X_{m:m:n}$ به این صورت خواهد بود:

$$L = \frac{c_m}{\sigma^m} \prod_{i=1}^m f(x_{i:m:n}; k) [1 - F(x_{i:m:n}; k)]^{r_i}$$

برای سادگی نماد x_i و y_i استفاده می‌شود. معادلات درستنمایی μ و σ به صورت زیر می‌باشند:

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} &= \frac{-1}{\sigma} \left[m\sqrt{k} - \sqrt{k} \sum_{i=1}^m \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m \frac{r_i f(x_i)}{1 - F(x_i)} \right] = 0 \\ \frac{\partial \ln L}{\partial \sigma} &= \frac{-1}{\sigma} \left[m + \sqrt{k} \sum_{i=1}^m x_i - \sqrt{k} \sum_{i=1}^m x_i \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m \frac{r_i x_i f(x_i)}{1 - F(x_i)} \right] = 0 \end{aligned}$$

همان‌طور که ملاحظه می‌شود معادلات درستنمایی شکل بسته ندارند؛ از این رو برآورد ماکسیمم درستنمایی پارامترهای μ و σ با استفاده از روش‌های عددی به دست خواهد آمد که با ثابت نگه داشتن پارامتر k این کار صورت می‌گیرد.

۴ برآورد فاصله‌ای

با استفاده از [۵] و [۹]، ماتریس واریانس-کوواریانس مجانبی برآوردگرهای ماکسیمم درستنمایی با معکوس کردن ماتریس اطلاع مشاهده شده (OIM)، I به دست خواهد آمد:

$$I = \begin{bmatrix} -\frac{\partial^2 L}{\partial \mu^2} |_{\hat{\mu}, \hat{\sigma}} & -\frac{\partial^2 L}{\partial \mu \partial \sigma} |_{\hat{\mu}, \hat{\sigma}} \\ -\frac{\partial^2 L}{\partial \mu \partial \sigma} |_{\hat{\mu}, \hat{\sigma}} & -\frac{\partial^2 L}{\partial \sigma^2} |_{\hat{\mu}, \hat{\sigma}} \end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}$$

عناصر ماتریس فوق مشتقات دوم نسبت به پارامترهای μ و σ است که به ترتیب به صورت زیر محاسبه می‌شوند:

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial \mu^2} &= \frac{1}{\sigma^2} \left\{ -\sum_{i=1}^m \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m \frac{r_i f(x_i)}{1 - F(x_i)} \right. \\ &\quad \times \left. \left[\sqrt{k} - \sqrt{k} \exp\left(\frac{x_i}{\sqrt{k}}\right) + \frac{f(x_i)}{1 - F(x_i)} \right] \right\} \\ \frac{\partial^2 \ln L}{\partial \sigma^2} &= \frac{1}{\sigma^2} \left\{ m + 2\sqrt{k} \sum_{i=1}^m x_i - 2\sqrt{k} \sum_{i=1}^m x_i \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m x_i^2 \exp\left(\frac{x_i}{\sqrt{k}}\right) \right. \\ &\quad \left. - 2 \sum_{i=1}^m \frac{r_i x_i f(x_i)}{1 - F(x_i)} - \sum_{i=1}^m \frac{r_i x_i^2 f(x_i)}{1 - F(x_i)} \left[\sqrt{k} - \sqrt{k} \exp\left(\frac{x_i}{\sqrt{k}}\right) + \frac{f(x_i)}{1 - F(x_i)} \right] \right\} \\ \frac{\partial^2 \ln L}{\partial \mu \partial \sigma} &= \frac{1}{\sigma^2} \left\{ m\sqrt{k} - \sqrt{k} \sum_{i=1}^m \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m x_i \exp\left(\frac{x_i}{\sqrt{k}}\right) - \sum_{i=1}^m \frac{r_i f(x_i)}{1 - F(x_i)} \right. \\ &\quad \left. - \sum_{i=1}^m \frac{r_i x_i f(x_i)}{1 - F(x_i)} \left[\sqrt{k} - \sqrt{k} \exp\left(\frac{x_i}{\sqrt{k}}\right) + \frac{f(x_i)}{1 - F(x_i)} \right] \right\} \end{aligned}$$

با توجه به کمیت‌های محوری $P_1 = \frac{\hat{\mu} - \mu}{\hat{\sigma} \sqrt{I_{11}^{-1}}}$ و $P_2 = \frac{\hat{\sigma} - \sigma}{\hat{\sigma} \sqrt{I_{22}^{-1}}}$ فاصله اطمینان تقریبی $100(1 - \alpha)\%$ پارامترهای μ و σ به این صورت خواهد بود:

$$\hat{\mu} \pm z_{\frac{\alpha}{2}} \hat{\sigma} \sqrt{I_{11}^{-1}}, \quad \hat{\sigma} \pm z_{\frac{\alpha}{2}} \hat{\sigma} \sqrt{I_{22}^{-1}}.$$

که $z_{\frac{\alpha}{2}}$ چندک $100 \frac{\alpha}{2}\%$ از توزیع نرمال استاندارد، و I_{ij}^{-1} ، (i, j) امین درایه از I^{-1} است.

۵ شبیه‌سازی

به منظور تعیین برآوردهای ماکسیمم درستنمایی پارامترهای μ و σ و عملکرد روش OIM در ساخت بازه اطمینان، از شبیه‌سازی مونت کارلو استفاده می‌شود. با ملاحظه [۷]، [۳] و [۹] ابتدا داده‌هایی از توزیع لگ-گاما با $\mu = 0$ و $\sigma = 1$ به ازای مقادیر $k = 1, 1/5, 5, 20$ با طرح‌های زیر

$$1 \text{ طرح: } R_1 = R_2 = \dots = R_{\lfloor m/2 \rfloor} = 2, R_{\lfloor m/2 \rfloor + 1} = \dots = R_{m-1} = 0;$$

$$2 \text{ طرح: } R_1 = R_2 = \dots = R_{\lfloor m/2 \rfloor} = 0, R_{\lfloor m/2 \rfloor + 1} = \dots = R_{m-1} = 2.$$

تولید می‌شود. دو مقدار مختلف برای T در نظر گرفته می‌شود. برآورد ماکسیمم درستنمایی، احتمال پوشش و طول بازه اطمینان پارامترهای μ و σ در جداول ۱، ۲، ۳ و ۴ آورده شده است. این نتایج حاصل ۱۰۰۰ شبیه‌سازی مونت کارلو است. در این جداول عدد خارج پرانتز احتمال پوشش و عدد داخل پرانتز عرض بازه می‌باشد:

جدول ۱: برآورد ماکسیمم درستنمایی، احتمال پوشش و طول بازه اطمینان ۹۵٪ برای μ و σ برای طرح ۱

OIM						
σ	μ	$\hat{\sigma}$	$\hat{\mu}$	k	T	Scheme
۸۵/۶(۰/۹۱)	۸۶/۸(۱/۲۲)	۰/۹۲۶۷	-۰/۱۰۰۴	۱	۱	(۲۵, ۱۰)
۸۶/۴(۰/۸۸)	۸۸/۰(۱/۱۶)	۰/۹۲۳۲	-۰/۰۹۴۹	۱/۵		
۸۳/۰(۰/۸۳)	۸۶/۷(۱/۰۷)	۰/۹۱۸۵	-۰/۰۷۶۶	۵		
۸۳/۳۶(۰/۸۲)	۸۸/۲(۱/۰۳)	۰/۹۲۳۱	-۰/۰۶۳۶	۲۰		

با مقایسه جداول مشاهده می‌شود که برآورد پارامتر μ با افزایش حجم نمونه بهتر است. در نمونه‌های کوچک طول بازه اطمینان پارامترهای μ و σ بزرگ بود و با افزایش حجم نمونه طول بازه کوتاه‌تر شد. احتمال پوشش برای نمونه‌هایی با حجم کوچک پایین، ولی برای نمونه‌هایی با حجم بزرگ رضایت‌بخش است. احتمال پوشش‌ها در طرح ۲ نسبت به احتمال پوشش‌ها در طرح ۱ قابل قبول‌تر هستند. همچنین طول بازه اطمینان در طرح ۲ نسبت به طرح ۱ کوتاه‌تر است. در نمونه‌های کوچک به منظور بهبود احتمال پوشش پارامتر μ و کوتاه‌تر شدن طول بازه اطمینان از تبدیل باکس-کاکس با $\tilde{\mu} = \mu^\nu / \nu$ استفاده می‌شود که $\nu = 7/5$ می‌باشد. نتایج شبیه‌سازی در جدول ۵ آمده است:

۶ دست‌آوردهای پژوهش

در این مقاله برآوردی از توزیع لگ-گاما تحت داده‌های سانسور شده فزاینده نوع II تطبیقی محاسبه شد. طبق نتایج شبیه‌سازی، طول بازه اطمینان نمونه‌های بزرگ نسبت به نمونه‌های کوچک، کوتاه‌تر و احتمال پوشش این نمونه‌ها قابل قبول‌تر بود. برای نمونه‌های کوچک از طریق تبدیل باکس-کاکس احتمال پوشش پارامتر مکان رضایت‌بخش و طول بازه اطمینان کوتاه‌تر شد.

مراجع

- [1] Al-Sobhi, M.M. and Soliman, A.A. (2015), *Estimation for the exponentiated weibull model with adaptive Type-II progressive censored schemes*. Applied Mathematical Modelling.
- [2] Balakrishnan, N. and Cramer, E. (2014), *The Art of Progressive Censoring*. Springer, New York.
- [3] Balakrishnan, N. and Sandhu, R.A. (1995), *A simple simulational algorithm for generating progressive Type II censored samples*. The American Statistician, 49, 229–230.
- [4] Cramer, E., Iliopoulos, G. (2010), *Adaptive progressive type II censoring*. Test, 19, 342–358.
- [5] Lin, C.T., Wu, S.J.S. and Balakrishnan, N. (2006), *Inference for log-gamma distribution on progressively Type II censored data*. Communications in Statistics Theory and Methods, 35, 1271–1292.
- [6] Mahmoud, M.A.W., Soliman, A.A. and El-Sagheer, R.M. (2013), *Estimation of generalized parato under an adaptive Type II progressive censoring*. Intelligent Information Management, 5, 73.
- [7] Ng, H.K.T., Kundu, D. and Chan, P.S. (2009), *Statistical analysis of exponential lifetimes under an adaptive Type II progressive censoring scheme*. Naval Research Logistic (NRL), 65, 687–698.
- [8] Sreekumar, N.V. and Thomas, P.Y. (2007), *Estimation of the parameters of log-gamma distribution using order statistics*. Metrika, 66, 115–127.
- [9] Ye, Z.S., Chan, P.S., Xie, M. and Ng, H.K.T. (2014), *Statistical inference for the extreme value distribution under adaptive Type II progressive censoring schemes*. Journal of Statistical Computation and Simulation, 84, 1099–1114.

جدول ۲: ادامه برآورد ماکسیمم درستنمایی، احتمال پوشش و طول بازه اطمینان ۹۵٪ برای μ و σ برای طرح ۱

OIM		$\hat{\sigma}$	$\hat{\mu}$	k	T	Scheme
σ	μ					
۸۵/۸(۰/۹۲)	۸۶/۸(۱/۲۳)	۰/۹۳۴۳	-۰/۱۲۷۵	۱	۱/۵	(۲۵, ۱۰)
۸۶/۱(۰/۸۸)	۸۷/۹(۱/۱۶)	۰/۹۲۱۹	-۰/۰۹۲۷	۱/۵		
۸۵/۳(۰/۸۴)	۸۹/۳(۱/۰۷)	۰/۹۲۳۱	-۰/۰۸۶۹	۵		
۸۶/۵(۰/۸۲)	۸۹/۶(۱/۰۴)	۰/۹۲۵۹	-۰/۰۶۲۴	۲۰		
۸۲/۶(۱/۰۲)	۸۳/۶(۱/۸۴)	۰/۹۰۰۶	-۰/۱۹۹۰	۱	۱	(۵۰, ۱۰)
۸۴/۲(۱/۰۲)	۸۴/۶(۱/۷۰)	۰/۹۲۲۵	-۰/۱۵۲۴	۱/۵		
۸۴/۰(۰/۹۵)	۸۶/۴(۱/۳۸)	۰/۹۰۵۹	-۰/۱۲۸۶	۵		
۸۳/۸(۰/۹۲)	۸۵/۹(۱/۲۵)	۰/۹۰۷۴	-۰/۱۰۴۰	۲۰		
۸۲/۲(۱/۰۴)	۸۵/۷(۱/۸۷)	۰/۹۱۵۴	-۰/۱۷۳۴	۱	۱/۵	
۸۴/۱(۱/۰۱)	۸۷/۱(۱/۶۹)	۰/۹۱۵۳	-۰/۱۴۷۱	۱/۵		
۸۴/۱(۰/۹۵)	۸۴/۰(۱/۳۸)	۰/۹۰۹۸	-۰/۱۳۴۴	۵		
۸۰/۷(۰/۹۱)	۸۲/۴(۱/۲۴)	۰/۸۹۶۸	-۰/۱۳۹۱	۲۰		
۹۱/۲(۰/۶۷)	۹۱/۷(۰/۹۰)	۰/۹۶۶۱	-۰/۰۵۳۸	۱	۱	(۵۰, ۲۰)
۹۰/۸(۰/۶۵)	۹۱/۰(۰/۸۵)	۰/۹۵۷۱	-۰/۰۳۹۸	۱/۵		
۸۷/۳(۰/۶۱)	۹۱/۵(۰/۷۸)	۰/۹۵۵۱	-۰/۰۴۴۶	۵		
۸۹/۹(۰/۶۱)	۹۲/۹(۰/۷۶)	۰/۹۶۵۰	-۰/۰۲۹۹	۲۰		
۸۹/۶(۰/۶۷)	۹۱/۰(۰/۸۹)	۰/۹۶۰۲	-۰/۰۳۷۰	۱	۱/۵	
۹۱/۰(۰/۶۵)	۹۱/۵(۰/۸۵)	۰/۹۵۸۲	-۰/۰۴۸۶	۱/۵		
۹۰/۴(۰/۶۲)	۹۲/۵(۰/۷۹)	۰/۹۶۳۷	-۰/۰۴۲۳	۵		
۸۹/۴(۰/۶۰)	۹۳/۴(۰/۷۶)	۰/۹۶۱۲	-۰/۰۳۵۶	۲۰		
۸۴/۴(۱/۰۷)	۸۳/۲(۲/۶۳)	۰/۹۰۰۱	-۰/۲۶۹۴	۱	۱	(۱۰۰, ۱۰)
۸۳/۱(۱/۰۵)	۸۲/۰(۲/۳۳)	۰/۹۰۳۴	-۰/۲۴۴۲	۱/۵		
۸۶/۸(۱/۰۰)	۸۵/۸(۱/۸۵)	۰/۹۰۵۵	-۰/۲۰۱۴	۵		
۸۳/۳(۰/۹۸)	۸۴/۵(۱/۶۴)	۰/۹۱۴۵	-۰/۱۴۶۰	۲۰		
۸۳/۵(۱/۱۰)	۸۲/۳(۲/۷۰)	۰/۹۲۳۵	-۰/۲۴۹۴	۱	۱/۵	
۸۲/۴(۱/۰۶)	۸۲/۷(۲/۳۴)	۰/۹۰۷۸	-۰/۲۲۰۹	۱/۵		
۸۲/۶(۱/۰۰)	۸۳/۲(۱/۸۳)	۰/۸۹۸۲	-۰/۱۹۲۵	۵		
۸۴/۰(۰/۹۸)	۸۴/۲(۱/۶۳)	۰/۹۱۰۵	-۰/۱۵۱۹	۲۰		
۹۰/۶(۰/۶۰)	۹۱/۸(۰/۸۹)	۰/۹۷۶۳	-۰/۰۳۹۷	۱	۱	(۱۰۰, ۳۰)
۹۲/۶(۰/۵۹)	۹۱/۸(۰/۸۲)	۰/۹۷۶۸	-۰/۰۳۲۳	۱/۵		
۹۰/۵(۰/۵۵)	۹۲/۵(۰/۷۲)	۰/۹۷۷۷	-۰/۰۲۶۴	۵		
۹۲/۱(۰/۵۴)	۹۰/۷(۰/۶۷)	۰/۹۸۲۵	-۰/۰۲۴۰	۲۰		
۹۳/۰(۰/۶۱)	۹۲/۲(۰/۸۹)	۰/۹۸۲۲	-۰/۰۳۶۰	۱	۱/۵	
۹۰/۳(۰/۵۸)	۹۳/۰(۰/۸۱)	۰/۹۶۷۷	-۰/۰۳۶۴	۱/۵		
۹۰/۰(۰/۵۵)	۹۲/۳(۰/۷۱)	۰/۹۷۲۵	-۰/۰۳۶۳	۵		
۹۲/۲(۰/۵۴)	۹۱/۷(۰/۶۷)	۰/۹۷۲۲	-۰/۰۱۹۰	۲۰		

جدول ۳: برآورد ماکسیمم درستنمایی، احتمال پوشش و طول بازه اطمینان ۹۵٪ برای μ و σ برای طرح ۲

OIM		$\hat{\sigma}$	$\hat{\mu}$	k	T	Scheme
σ	μ					
۸۴/۱(۱/۰۲)	۸۴/۸(۱/۳۶)	۰/۹۱۰۸	-۰/۱۴۰۸	۱	۱	(۲۵, ۱۰)
۸۴/۷(۰/۹۹)	۸۵/۰(۱/۲۶)	۰/۹۰۸۲	-۰/۱۱۵۱	۱/۵		
۸۴/۰(۰/۹۴)	۸۶/۹(۱/۱۲)	۰/۹۱۴۳	-۰/۰۸۰۱	۵		
۸۲/۸(۰/۸۹)	۸۵/۸(۱/۰۴)	۰/۹۰۰۶	-۰/۰۷۲۱	۲۰		
۸۲/۸(۱/۰۱)	۸۶/۰(۱/۳۵)	۰/۹۰۰۹	-۰/۱۲۳۶	۱	۱/۵	
۸۶/۶(۰/۹۹)	۸۷/۶(۱/۲۶)	۰/۹۱۲۶	-۰/۱۰۹۵	۱/۵		
۸۵/۳(۰/۹۴)	۸۶/۲(۱/۱۳)	۰/۹۲۲۲	-۰/۰۶۴۸	۵		
۸۲/۱(۰/۹۰)	۸۶/۳(۱/۰۴)	۰/۹۰۳۸	-۰/۰۸۲۳	۲۰		
۸۳/۹(۱/۰۸)	۸۴/۹(۲/۰۰)	۰/۹۰۸	-۰/۱۹۶۲	۱	۱	(۵۰, ۱۰)
۸۵/۱(۱/۰۹)	۸۴/۶(۱/۸۵)	۰/۹۳۶۳	-۰/۱۴۴۳	۱/۵		
۸۰/۸(۰/۹۶)	۸۱/۵(۱/۴۱)	۰/۸۷۷۷	-۰/۱۷۷۰	۵		
۸۳/۸(۰/۹۷)	۸۴/۲(۱/۳۳)	۰/۹۱۹۱	-۰/۱۱۸۳	۲۰		
۸۵/۲(۱/۰۷)	۸۴/۱(۲/۰۰)	۰/۹۰۵۲	-۰/۱۹۱۲	۱	۱/۵	
۸۱/۷(۱/۰۴)	۸۲/۹(۱/۷۸)	۰/۸۹۹۹	-۰/۱۶۷۴	۱/۵		
۸۴/۱(۰/۹۹)	۸۵/۰(۱/۴۶)	۰/۹۰۴۹	-۰/۱۳۶۱	۵		
۸۳/۱(۰/۹۶)	۸۴/۰(۱/۳۲)	۰/۹۱۱۷	-۰/۱۲۲۸	۲۰		
۸۹/۰(۰/۷۵)	۹۱/۲(۱/۰۰)	۰/۹۴۷۳	-۰/۰۵۳۳	۱	۱	(۵۰, ۲۰)
۸۷/۸(۰/۷۳)	۸۹/۳(۰/۹۳)	۰/۹۵۲۵	-۰/۰۶۴۷	۱/۵		
۸۷/۷(۰/۷۰)	۸۸/۸(۰/۸۳)	۰/۹۶۱۴	-۰/۰۴۲۰	۵		
۸۸/۵(۰/۶۷)	۹۱/۶(۰/۷۸)	۰/۹۵۵۰	-۰/۰۴۱۱	۲۰		
۸۷/۴(۰/۷۴)	۸۷/۸(۰/۹۹)	۰/۹۴۱۴	-۰/۰۸۵۰	۱	۱/۵	
۸۷/۹(۰/۷۳)	۸۸/۳(۰/۹۳)	۰/۹۴۷۲	-۰/۰۶۷۷	۱/۵		
۸۹/۳(۰/۶۸)	۸۹/۹(۰/۸۲)	۰/۹۴۴۵	-۰/۰۴۴۱	۵		
۹۰/۵(۰/۶۷)	۹۱/۷(۰/۷۸)	۰/۹۵۲۶	-۰/۰۴۱۶	۲۰		
۸۳/۶(۱/۱۱)	۸۳/۹(۲/۷۶)	۰/۹۱۳۸	-۰/۲۳۶۴	۱	۱	(۱۰۰, ۱۰)
۸۱/۳(۱/۰۸)	۸۲/۲(۲/۴۱)	۰/۹۰۴۶	-۰/۲۳۱۶	۱/۵		
۸۲/۴(۱/۰۴)	۸۱/۴(۱/۹۲)	۰/۹۱۶۹	-۰/۱۷۴۱	۵		
۸۲/۲(۰/۹۸)	۸۱/۵(۱/۶۵)	۰/۸۹۴۷	-۰/۱۷۳۴	۲۰		
۸۲/۳(۱/۰۹)	۸۲/۷(۲/۷۲)	۰/۹۰۱۲	-۰/۲۶۴۶	۱	۱/۵	
۸۲/۳(۱/۰۸)	۸۰/۹(۲/۴۱)	۰/۹۰۵۳	-۰/۲۴۲۸	۱/۵		
۸۴/۸(۱/۰۳)	۸۵/۷(۱/۹۱)	۰/۹۱۲۸	-۰/۱۶۲۴	۵		
۸۳/۸(۰/۹۹)	۸۴/۳(۱/۶۷)	۰/۹۰۷۰	-۰/۱۷۲۷	۲۰		
۹۱/۳(۰/۶۵)	۹۱/۶(۱/۰۰)	۰/۹۸۲۷	-۰/۰۲۵۸	۱	۱	(۱۰۰, ۳۰)
۹۰/۷(۰/۶۳)	۹۱/۹(۰/۹۱)	۰/۹۷۸۸	-۰/۰۳۹۳	۱/۵		
۹۰/۱(۰/۵۹)	۹۱/۷(۰/۷۶)	۰/۹۷۲۲	-۰/۰۳۷۵	۵		
۹۰/۰(۰/۵۷)	۹۰/۳(۰/۷۰)	۰/۹۷۰۰	-۰/۰۳۲۱	۲۰		

جدول ۴: ادامه برآورد ماکسیمم درستنمایی، احتمال پوشش و طول بازه اطمینان ۹۵٪ برای μ و σ برای طرح ۲

OIM		$\hat{\sigma}$	$\hat{\mu}$	k	T	Scheme
σ	μ					
۹۰/۰(۰/۶۴)	۸۹/۱(۰/۹۸)	۰/۹۶۰۵	-۰/۰۶۱۳	۱	۱/۵	(۱۰۰, ۳۰)
۹۱/۸(۰/۶۳)	۹۳/۲(۰/۹۰)	۰/۹۷۴۹	-۰/۰۳۳۳	۱/۵		
۸۹/۷(۰/۵۹)	۹۰/۳(۰/۷۶)	۰/۹۶۵۰	-۰/۰۳۲۶	۵		
۸۹/۴(۰/۵۷)	۹۱/۳(۰/۶۹)	۰/۹۵۹۱	-۰/۰۳۹۳	۲۰		

جدول ۵: احتمال پوشش و طول بازه اطمینان ۹۵٪ با استفاده از تبدیل باکس-کاکس برای μ

(n,m)			k	T	Scheme
(۱۰۰, ۱۰)	(۵۰, ۱۰)	(۲۵, ۱۰)			
۹۰/۱(۲/۳۷)	۹۲/۱(۱/۶۹)	۹۵/۷(۱/۰۸)	۱	۱	۱
۹۲/۳(۲/۰۶)	۹۰/۹(۱/۵۱)	۹۳/۳(۱/۰۵)	۱/۵		
۹۱/۰(۱/۶۴)	۹۲/۷(۱/۲۳)	۹۳/۸(۰/۹۶)	۵		
۹۱/۸(۱/۴۶)	۹۳/۷(۱/۱۱)	۹۵/۲(۰/۹۱)	۲۰		
۹۰/۸(۲/۳۴)	۹۱/۶(۱/۶۹)	۹۵/۲(۱/۱۳)	۱	۱/۵	
۹۰/۷(۲/۰۷)	۹۰/۴(۱/۵۰)	۹۵/۵(۱/۰۶)	۱/۵		
۸۹/۶(۱/۶۵)	۹۲/۲(۱/۲۵)	۹۳/۶(۰/۹۸)	۵		
۹۲/۹(۱/۴۶)	۹۰/۸(۱/۱۱)	۹۴/۱(۰/۹۲)	۲۰		
۸۹/۸(۲/۳۹)	۹۱/۹(۱/۷۸)	۹۲/۵(۱/۲۲)	۱	۱	۲
۹۱/۶(۲/۱۴)	۹۱/۰(۱/۶۰)	۹۱/۷(۱/۱۳)	۱/۵		
۸۹/۹(۱/۷۲)	۹۲/۹(۱/۳۲)	۹۳/۹(۱/۰۱)	۵		
۹۰/۹(۱/۵۴)	۹۱/۹(۱/۱۸)	۹۴/۰(۰/۹۴)	۲۰		
۸۹/۶(۲/۴۲)	۸۹/۹(۱/۷۵)	۹۱/۹(۱/۲۳)	۱	۱/۵	
۹۲/۱(۲/۱۹)	۹۲/۸(۱/۶۴)	۹۲/۷(۱/۱۲)	۱/۵		
۸۹/۹(۱/۶۹)	۹۰/۸(۱/۲۹)	۹۲/۸(۰/۹۹)	۵		
۹۲/۱(۱/۴۹)	۹۲/۷(۱/۱۶)	۹۳/۹(۰/۹۵)	۲۰		



تعیین حجم نمونه بهینه با استفاده از حدود تحمل برای طول عمر سیستم‌های k از n

نقی زاده، م^۱ وحیدیان، م^۲

^{۱،۲} گروه آمار، دانشگاه مازندران

چکیده

فاصله تحمل یک فاصله تصادفی است که با یک سطح اطمینان مشخص، نسبتی از جامعه مورد بررسی را دربر می‌گیرد و به طور گسترده‌ای در صنعت به کار می‌رود. تعیین حدود تحمل آماری برای طول عمر سیستم‌های k از n و همچنین محاسبه دقت حدود تحمل ارائه شده بر اساس داده‌های سانسور شده نوع دوم در ادبیات تحقیق مورد بررسی قرار گرفته است. یکی از موضوعات مورد توجه با در نظر گرفتن هزینه آزمایش‌ها، مساله تعیین حجم نمونه بهینه می‌باشد. در این مقاله، با استفاده از یک مساله بهینه سازی و مینیمم کردن یک تابع هزینه مناسب، حجم نمونه بهینه برای ساختن حدود تحمل آماری به دست می‌آید.

کلمات کلیدی: توزیع نمایی، حجم نمونه بهینه، حدود تحمل.

۱ مقدمه

یک F - سیستم k از n سیستمی است که زمانی از کار می‌افتد که حداقل k مولفه آن از کار بیفتد. این سیستم‌ها حالت‌های کلی‌تری از سیستم‌های موازی ($k = n$) و سیستم‌های سری ($k = 1$) هستند. برای کاربردی از این سیستم، یک سازمان اقتصادی را در نظر بگیرید. کمیته‌ای با n

^۱ m.naghizadeh@umz.ac.ir

^۲ vahidian_m@yahoo.com

عضو، پروژه ای را رد می کند اگر حداقل k عضو کمیته آن را رد کنند. اولین بار [۴] اصطلاح سیستم های k از n را به کار بردند و به دلیل سادگی ساختار، قابلیت اعتماد بالا و به صرفه بودن اقتصادی، در دهه های اخیر، مورد توجه بسیاری از پژوهشگران قرار گرفته است. مراجع [۲] - [۳] را ملاحظه کنید.

فرض کنید T طول عمر یک سیستم k از n با مولفه های توزیع شده بصورت نمایی و X زمان شکست یک مولفه دارای توزیع نمایی با تابع چگالی

$$g(x|\theta) = \frac{1}{\theta} \exp(-\frac{x}{\theta}), \quad x > 0, \quad (1)$$

باشد. تابع چگالی طول عمر سیستم، T ، به صورت زیر

$$f(t) = \frac{k \binom{n}{k} \{1 - \exp(-\frac{t}{\theta})\}^{k-1}}{\theta \exp\{\frac{(n-k+1)t}{\theta}\}}, \quad t > 0, \quad (2)$$

و قابلیت سیستم در زمان t یعنی $R(t) = Pr(T \geq t)$ به صورت

$$R(t) = \sum_{i=0}^{k-1} \frac{\binom{n}{i} \{1 - \exp(-\frac{t}{\theta})\}^i}{\exp\{\frac{(n-i)t}{\theta}\}},$$

به دست می آید. ([۱])

فاصله تحمل یک فاصله تصادفی است که با یک ضریب اطمینان مشخص، نسبتی از جامعه مورد بررسی را پوشش می دهد. کی از مباحث مورد علاقه در مورد فاصله های تحمل این است که در آزمون قابلیت تولیدات جدید، با ضریب از پیش تعیین شده γ ، حداقل نسبت β از واحدهای تولیدی، قبل از پایان دوره گارانتی از بین نروند. در این حالت، یک حد تحمل با میزان پوشش β و سطح اطمینان γ براساس داده ها می تواند برای برآورد دوره ی گارانتی به کار رود. از جدیدترین تحقیقات صورت گرفته در زمینه فواصل تحمل می توان به [۶]، [۷]، [۸] و [۹] اشاره کرد. برای ساختن حدود تحمل برای طول عمر سیستم، فرض کنید N مولفه در آزمایش قرار گیرند و $X_{1:N}, \dots, X_{r:N}$ مشاهدات سانسور شده نوع دوم از راست باشند ($r \leq N$). می توان نشان داد که برآوردگر ماکزیمم درستنمایی θ برابر $\hat{\theta} = \frac{W}{r}$ است، که در آن

$$W = \sum_{i=1}^r X_{i:N} - (N-r)X_{r:N}$$

همچنین دارای توزیع کی دو با $2r$ درجه آزادی است.

[۵] در مقاله ای به بررسی مساله تعیین حدود تحمل آماری برای طول عمر سیستم های k از n پرداخت. دقت حدود تحمل ارائه شده براساس داده های سانسور شده نوع دوم توسط [۱] محاسبه و در جداولی ارائه شدند. در این مقاله، با استفاده از یک مساله بهینه سازی، حجم نمونه بهینه را برای رسیدن به یک سطح دقت مطلوب محاسبه می کنیم. به این منظور، در بخش ۲ قابلیت اعتماد سیستم های k از n با مولفه های نمایی را بیان می کنیم. در بخش ۳ به بیان حدود تحمل و بررسی سطح دقت آنها می پردازیم. در نهایت، در بخش ۴ با استفاده از یک مساله بهینه سازی، حجم نمونه بهینه را به دست می آوریم.

۲ حدود تحمل برای طول عمر سیستم و تعیین سطح دقت آنها

فرض کنید $\mathbf{X} = (X_{1:N}, \dots, X_{r:N})$ داده های مشاهده شده و $\gamma, \beta \in (0, 1)$ باشند. آماره $L_{\beta, \gamma} \equiv L_{k; n; \beta, \gamma}(X)$ را یک حد پایین تحمل با سطح پوشش β و سطح اطمینان γ برای طول عمر سیستم گویند اگر

$$Pr \left\{ \int_{L_{\beta, \gamma}}^{\infty} f(t) dt \geq \beta \right\} = Pr\{R(L_{\beta, \gamma}) \geq \beta\} = \gamma, \quad \theta > 0. \quad (3)$$

جدول ۱: سطوح دقت حد پایین تحمل به ازای $\beta = 0.9$ ، $\gamma = 0.95$ و $\beta' = 0.95$ و مقادیر انتخابی k و n (منبع: [۱])

k=۴			k=۲			
$n=6$	$n=5$	$n=4$	$n=4$	$n=3$	$n=2$	r
0.1155	0.1167	0.1207	0.2318	0.2323	0.2343	5
0.1787	0.1808	0.1872	0.3668	0.3676	0.3707	10
0.2350	0.2377	0.2463	0.4794	0.4803	0.4842	15
0.2870	0.2904	0.3009	0.5740	0.5751	0.5793	20
0.3816	0.3860	0.3998	0.7193	0.7203	0.7246	30
0.4655	0.4707	0.4866	0.8183	0.8193	0.8231	40
0.5398	0.5454	0.5679	0.8843	0.8851	0.8882	50
0.6051	0.6110	0.6291	0.9273	0.9279	0.9302	60

با توجه به (۳) با اطمینان γ می توان تضمین کرد که حداقل 100β درصد از طول عمرهای سیستم از $L_{\beta,\gamma}$ بیشتر خواهد بود. به عبارت دیگر، با اطمینان γ ، احتمال این که مشاهده آینده T از $L_{\beta,\gamma}$ بیشتر شود، حداقل برابر β است. [۵] نشان داد حد پایین تحمل با سطح پوشش β و سطح اطمینان γ برای طول عمر سیستم برابر است با، $L_{k;n;\beta,\gamma} = \frac{2C_{k;n;1-\beta}}{\chi^2_{2r;\gamma}} W$ که در آن

$$C_{k;n;1-\beta} = \ln\left\{1 + \frac{kF_{2k,2(n-k+1);1-\beta}}{n-k+1}\right\}, \quad (4)$$

$\chi^2_{2r;\gamma}$ چندک توزیع کی دو با $2r$ درجه آزادی و $F_{2k,2(n-k+1);1-\beta}$ چندک $(1-\beta)$ توزیع F با $2k$ و $2(n-k+1)$ درجه آزادی می باشد. [۶] معیار دقت را برای فاصله های تحمل در توزیع نمایی به کار برد. سطح دقت حد پایین تحمل در $(\beta, 1)$ به صورت زیر تعریف می شود:

$$A_{\beta,\gamma}[\beta'] = \Pr\left[\beta \leq R(L_{\beta,\gamma}) < \beta' \mid R(L_{\beta,\gamma}) \geq \beta\right] \quad (5)$$

[۱] سطح دقت حد پایین تحمل برای طول عمر سیستم های k از n را با استفاده از رابطه (۵) به صورت

$$1 - \frac{\Pr\left[\chi^2_{2r} < \frac{\chi^2_{2r;\gamma} C_{k;n;1-\beta'}}{C_{k;n;1-\beta}}\right]}{\gamma} \quad (6)$$

به دست آوردند. با توجه به این که معیار دقت معرفی شده تابعی صعودی نسبت به r است، بنابراین می توان از آن برای تعیین مقدار r بهینه استفاده کرد. جدول (۱) سطوح دقت حد پایین تحمل داده شده در (۶) را به ازای $\gamma = 0.95$ ، $\beta' = 0.95$ و $\beta = 0.90$ و مقادیر انتخابی k و n نشان می دهد. یک سیستم ۲ از ۳ را در نظر بگیرید، یعنی سیستمی با ۳ مولفه که زمانی از کار می افتد که حداقل ۲ مولفه آن از کار بیفتند. برای مثال، به ازای $k=2$ و $n=3$ ، برای رسیدن به سطح دقتی به اندازه حداقل ۸۰ درصد، نیاز به مشاهده حداقل $r=40$ شکست است.

۳ تعیین حجم نمونه بهینه

در بخش ۳ مقدار مناسب r با توجه به سطح دقت محاسبه شده به دست آمد. یکی از موضوعات مورد توجه با در نظر گرفتن هزینه آزمایش ها، مساله تعیین حجم نمونه بهینه N ($r \leq N$) می باشد. یک روش مناسب برای تعیین حجم نمونه بهینه، حداقل کردن تابع هزینه به صورت $c_1 X_{r:N} + c_2 N$

می‌باشد، که در آن $c_1 > 0$ و $c_2 > 0$ به ترتیب هزینه مربوط به هر مولفه در آزمایش و هزینه مربوط به هر واحد زمان آزمایش هستند. تابع هزینه مورد انتظار آزمایش به صورت زیر تعریف می‌شود:

$$EC(N, r, \theta) = c_1 E(X_{r:N}) + c_2 N, \quad N \geq r, \quad (7)$$

که به پارامتر نامعلوم θ بستگی دارد. می‌توان نشان داد که $(N - i + 1)(X_{i:N} - X_{i-1:N})$ دارای توزیع نمایی با پارامتر θ است. با فرض $X_{.:N} \equiv \cdot$ داریم

$$E(X_{i:N} - X_{i-1:N}) = \frac{\theta}{N - i + 1}.$$

در نتیجه

$$\begin{aligned} E(X_{r:N}) &= E[X_{1:N} - (X_{2:N} - X_{1:N}) + \dots + (X_{r:N} - X_{r-1:N})] \\ &= \frac{\theta}{N} + \frac{\theta}{N-1} + \dots + \frac{\theta}{N-r+1} \\ &= \theta \sum_{i=1}^{r-1} \frac{1}{N-i}. \end{aligned} \quad (8)$$

از رابطه‌های (۷) و (۸)، تابع هزینه مورد انتظار آزمایش برابر است با

$$EC(N, r, \theta) = c_1 \theta \sum_{i=1}^{r-1} \frac{1}{N-i} + c_2 N, \quad N \geq r, \quad (9)$$

که در آن θ برآوردی برای θ است. با ثابت فرض کردن r می‌خواهیم مقدار بهینه N را با مینیمم کردن تابع هزینه (۹) به فرم زیر

$$\min EC(N, r, \theta), \quad (10)$$

$$N \geq r$$

$$N \in Z^+$$

محاسبه کنیم که در آن Z^+ مجموعه اعداد صحیح مثبت است. مساله بهینه‌سازی (۱۰) معادل با رابطه زیر است

$$\min\{EC(N, r, \theta) : N \in A\},$$

که در آن $A = \{N \in Z^+ : N \geq r\}$. در واقع حجم نمونه بهینه که آن را با N_{opt} نشان می‌دهیم، کوچکترین مقدار صحیح N ای است که برای آن داریم $EC(N+1, r, \theta) \geq EC(N, r, \theta)$. به عبارت دیگر

$$N_{opt} = \min\{N \in A : EC(N+1, r, \theta) \geq EC(N, r, \theta)\}$$

یا به طور معادل

$$N_{opt} = \min\left\{N \in A : \sum_{i=1}^{r-1} \frac{1}{N-i} - \sum_{i=1}^{r-1} \frac{1}{N+1-i} \leq \frac{1}{a}\right\}, \quad (11)$$

که در آن $a = \frac{c_1 \theta}{c_2}$. با توجه به این که

$$\sum_{i=1}^{r-1} \frac{1}{N-i} - \sum_{i=1}^{r-1} \frac{1}{N+1-i} = \frac{1}{N-r+1} - \frac{1}{N+1}$$

مقدار N^* از رابطه (۱۱) معادل با رابطه زیر است

$$N_{opt} = \min\{N \in A : (N+1)(N-r+1) \geq ra\}.$$

جدول ۲: حجم نمونه بهینه برای مقادیر انتخابی r و a

a										
۱۰۰	۸۰	۶۰	۴۰	۲۰	۱۵	۱۰	۵	۱	r	
۲۴	۲۲	۱۸	۱۶	۱۲	۱۱	۹	۸	۵	۵	
۳۷	۳۳	۲۷	۲۵	۱۹	۱۷	۱۶	۱۳	۱۰	۱۰	
۴۶	۴۲	۳۵	۳۳	۲۶	۲۳	۲۱	۱۸	۱۵	۱۵	
۵۵	۵۱	۴۳	۳۹	۳۲	۲۹	۲۷	۲۴	۲۰	۲۰	
۷۱	۶۶	۵۶	۵۲	۴۳	۴۰	۳۷	۳۴	۳۰	۳۰	
۸۶	۷۹	۶۸	۶۴	۵۴	۵۱	۴۸	۴۴	۴۰	۴۰	
۹۹	۹۳	۸۰	۷۶	۶۵	۶۱	۵۸	۵۴	۵۰	۵۰	
۱۰۱	۱۰۱	۹۲	۸۷	۷۵	۷۱	۶۸	۶۴	۶۰	۶۰	

جدول ۲ مقادیر حجم نمونه بهینه N_{opt} را به ازای مقادیر انتخابی r و a را نشان می‌دهد. از جدول ۱ دیدیم که به ازای $k = 2$ و $n = 3$ برای رسیدن به سطح دقتی به اندازه ۸۰ درصد نیاز به مشاهده حداقل $r = 40$ شکست داریم. فرض کنید $c_1 = 10$ ، $c_2 = 2$ و $\theta_0 = 4$ ، در نتیجه $a = 20$. در این صورت، حجم نمونه بهینه از جدول ۲ به ازای $r = 40$ و $a = 20$ برابر $N_{opt} = 54$ خواهد بود یعنی نیاز به حداقل ۵۴ نمونه برای داشتن کمترین هزینه داریم. به ازای $N_{opt} = 54$ و $r = 40$ داریم

$$\sum_{i=0}^{r-1} \frac{1}{N_{opt}-i} = \sum_{i=0}^{r-1} \frac{1}{54-i} = 1/3239.$$

بنابراین به ازای $\theta_0 = 4$ و از رابطه (۹) مقدار هزینه مطلوب برابر است با

$$EC(N_{opt}, r, \theta_0) = EC(54, 40, 4) = 10(4)(1/3239) + 2(54) = 160/96.$$

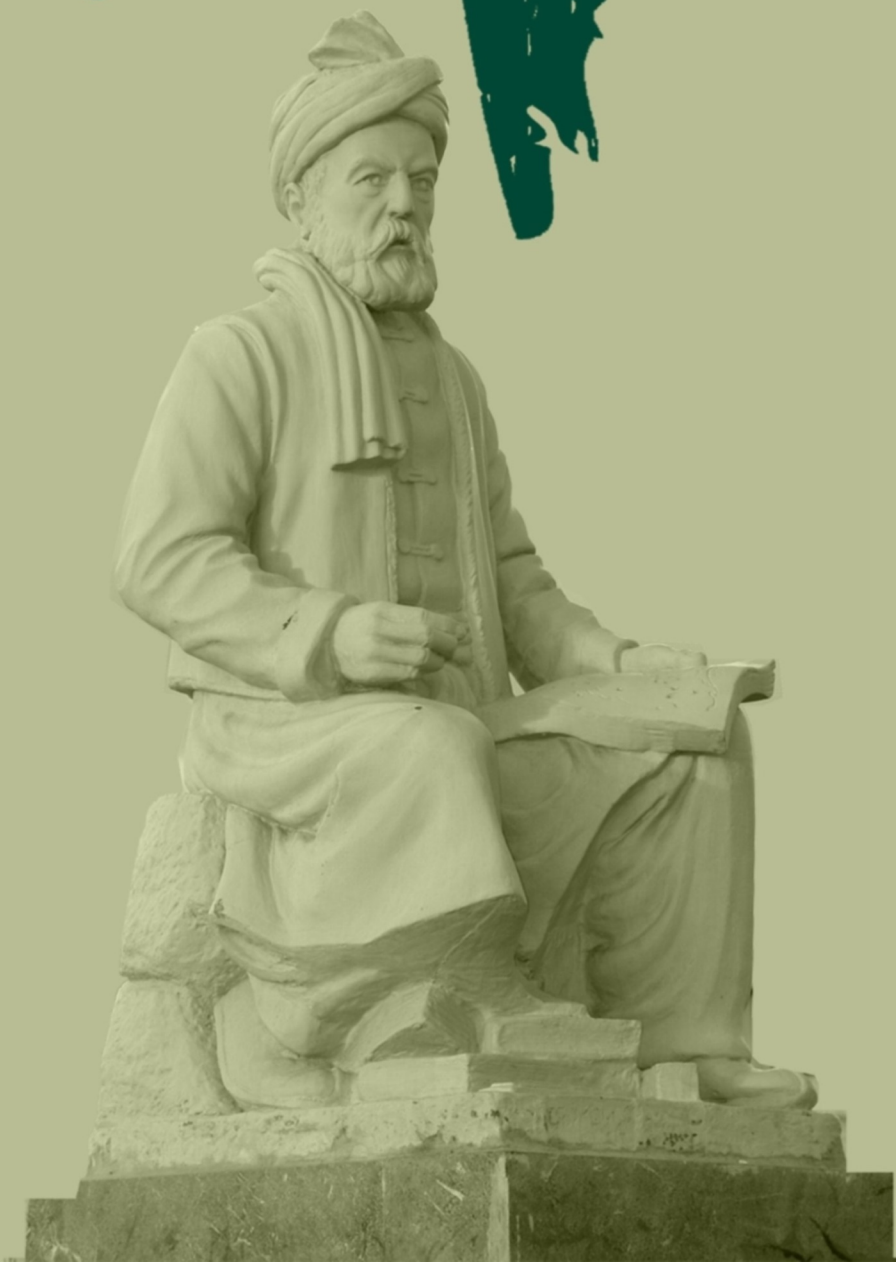
۴ بحث و نتیجه گیری

در این مقاله، ابتدا به مرور محاسبه سطح دقت حدود تحمل سیستم‌های k از n با مولفه‌های نمایی بر اساس داده‌های سانسور شده نوع دوم پرداخته شد. با توجه به سطح دقت از پیش تعیین شده می‌توان حداقل تعداد شکست‌ها را به دست آورد. با استفاده از این مقدار و به کمک بهینه‌سازی، حجم نمونه بهینه را برای رسیدن به یک سطح دقت مطلوب و حداقل تعداد شکست‌ها به دست آوردیم.

مراجع

- [۱] م. وحیدیان و م. نقی زاده. محاسبه سطح دقت حدود تحمل برای طول عمر سیستم‌های k از n ، پذیرفته شده برای چاپ در مجله علوم آماری، ۱۳۹۵.

- [2] Agarwal, M., Sen, K. and P. Mohan, (2007). GERT analysis of m-consecutive-k-out-of-n systems, *IEEE Trans. Reliability*, 56, 26-34.
- [3] Amari, V., Zuo, M. J., and G. Dill, G. (2009). A fast and robust reliability evaluation algorithm for generalized multi-state k-out-of-n systems, *IEEE Trans. Reliability*, 58, 88-97.
- [4] Chiang, D., and Niu, S. (1981). Reliability of a consecutive k-out-of-n: F system. *IEEE Trans Reliability*, 30, 87-89.
- [5] Fernández, A. J. (2010). Tolerance limits for k-out-of-n systems with exponentially distributed component lifetimes, *IEEE Trans. Reliab.* 59, 331-337.
- [6] Fernández, A. J. (2014). Computing tolerance limits for the lifetime of a k -out-of- n : F system based on prior information and censored data, *Appl. Math. Model* 38, 548-561.
- [7] Kiapour, A., and Naghizadeh Qomi, M. (2017). Equal-tailed and shortest Bayesian tolerance intervals based on exponential k-records, *Commun.Stat. Theor. and Meth.*, 46, 3949-3956.
- [8] MirMostafaei, S.M.T.K., Naghizadeh Qomi, M., and Fernandez, A. J. (2016). Tolerance limits for minimal repair times of a series system with Rayleigh distributed component lifetimes, *Applied Mathematical modeling*, 40, 3153-3163.
- [9] Naghizadeh Qomi, M., and Kiapour, A. (2017). Shortest tolerance intervals controlling both tails of the exponential distribution based on record values, *Commun.Stat. Theor. and Meth.*, 46, 271-279.



Proceeding of The 3rd Seminar on Reliability Theory and its Applications

May 17-16 2017

Department of Statistics, Ferdowsi University of Mashhad