مجموعه مقالات

چهارمین کنفرانس بین المللی آمار ایران

جلد اول

نظریه آمار - نظریه احتمال

۱ ۵ ۳ شهریور ۱۳۷۷

دانشگاه شهید [illegible]

عنوان‌کتاب : مجموعه‌مقالات‌چهارمین کنفرانس بین‌المللی آمار ایران، جلداول
۱-۳ شهریور ۱۳۷۷، دانشگاه شهید بهشتی، تهران

تألیف : ارائه دهندگان مقاله در کنفرانس

مترجم :

ویراستاران : دکتر محمد رضا مشکانی - دکتر محمد قاسم وحیدی اصل

ناشر : مرکز چاپ وانتشارات دانشگاه شهیدبهشتی

محل نشر : تهران

تاریخ نشر : ۱۳۷۸

تیراژ : ۱۵۰۰

نوبت چاپ : اول

شماره شابک : ۹۶۴ - ۴۵۷ / ۰۲۱ - ۹

شماره‌رده‌بندی‌دیویی : ۳۱۵/۵ ک۷۴۲

شماره‌رده‌بندی کنگره : HA ۴۵۷۰/۲/آ ک۹ ۱۳۷۷

قیمت : ۱۲۵۰۰ ریال

بسمه تعالی

پیشگفتار

چهارمین کنفرانس بین‌المللی آمار ایران در ادامه سنت برگزاری کنفرانسهای انجمن آمار ایران به همت دانشگاه شهید بهشتی از اول تا سوم شهریورماه ۱۳۷۷ برگزار گردیده است. مجموعه حاضر در بردارنده متن کامل بخشی از مقالات ارائه شده در این کنفرانس است.

برگزارکنندگان این کنفرانس با عنایت به کاربرد وسیع علم آمار کوشیده‌اند تا با ارائه یک برنامه مدون، مخاطبان بیشتری از تخصصهای گوناگون آماری را جلب کنند. به این منظور کمیته برگزاری کنفرانس تصمیم به ایجاد بخش‌های تخصصی - کاربردی زیر به عنوان اجزای اصلی تشکیل دهنده کنفرانس گرفت.

این بخشها عبارت بودند از :

نظریه احتمال، نظریه آمار، آمار در کشاورزی، آمار در اقتصاد، آموزش آمار، آمار در مدیریت، آمار در علوم تربیتی، آمار زیستی، آمار در هواشناسی، آمار بیمه (آکچواری)، آمار مهندسی و آمار رسمی بوده‌اند. کثرت شرکت کنندگان و تنوع مقالات عرضه شده در این بخش‌ها گویای این واقعیت است که خوشبختانه کنفرانس در رسیدن به هدف مذکور موفق بوده‌است.

در مجموعه حاضر مقالات مدعوین و شرکت کنندگان را به ترتیب حروف الفبا در بخشهای مربوطه آورده‌ایم. برخی از مقالات ارائه شده از سوی نویسندگان در زمان تهیه این مجموعه به صورت مناسب برای تدوین، ارسال نشده بود و به همین دلیل، در این مجلد آورده نشده‌اند. این مقالات پس از دریافت در مجلدات آینده به همین ترتیب به چاپ خواهند رسید.

در پایان لازم است از نویسندگان مقالات و مسؤلان بخشهای مختلف کنفرانس که با تلاش فراوان به غنای علمی کنفرانس افزودند، به ویژه از آقایان محمد حسن تقی و علی اکبر بروميده که در تدوین این مجموعه و ارائه انواع خدمات علمی به کنفرانس دریغ ننموده‌اند، سپاسگزاری کنیم. همچنین از آقای پارسی که در طراحی و چاپ به موقع پوستر، رنامه کنفرانس، خلاصه مقالات و مجموعه حاضر همکاری صمیمانه کرده‌اند تشکر و قدردانی می کنیم.

دکتر محمد صادق مهدوی	معاونت دانشجوئی دانشگاه
مهندس فرشید جاهدی	معاونت مالی واداری دانشگاه
دکتر بهمن هنری	معاونت آموزشی دانشگاه
دکتر احمد شمیرانی	معاونت تحصیلات تکمیلی دانشگاه
دکتر سیامک نوربلوچی	دبیرکنفرانس ورئیس دانشکده علوم ریاضی
دکتر حمید رضانیکبخت	مدیرکل روابط بین‌الملل دانشگاه

کمیته هماهنگی

آقای مسعود البرز	عضو هیات علمی دانشکده علوم ریاضی
آقای سید مهدی اعرابی	رئیس روابط عمومی دانشگاه
آقای فرشید جاهدی	معاونت مالی و اداری دانشگاه
دکتر محمد صادق مهدوی	معاونت دانشجویی و فرهنگی دانشگاه
دکتر سیامک نوربلوچی	رئیس دانشکده علوم ریاضی و دبیر کنفرانس
مهندس عباس یزدی صمدی	مدیر دفتر فنی دانشگاه

دبیر کنفرانس	دکتر سیامک نوربلوچی	استاد گروه آمار
دبیر کمیته علمی	دکتر محمد رضا مشکانی	استاد گروه آمار
دبیر کمیته اجرائی	دکتر محمد ذکائی	استادیار گروه آمار

کمیته اجرائی (اعضاء هیأت علمی دانشکده علوم ریاضی)

آقای علی آقا آزاده	دکتر عبدالرحیم شهلائی
دکتر محمدمهدی ابراهیمی	دکتر زهرا گویا
آقای مسعود البرز	دکتر محمدرضا مشکانی
دکتر سیدعلیرضا حسینیون	دکتر سیامک نوربلوچی
دکتر سیدجلال داودزاده	دکتر محمد قاسم وحیدی اصل
دکتر محمد ذکائی	

۵- دکتر محمد ذکایی دانشگاه شهید بهشتی
۶- دکتر احمد پارسیان دانشگاه صنعتی اصفهان
۷- دکتر احمدرضا سلطانی دانشگاه شیراز

بنابر ساختار کنفرانس که متشکل از بخشهای متعدد است ، کمیته علمی بخشهایی را به شرح زیر و با مسئولیت یکی از اعضای کمیته علمی و با عضویت همکاران دیگر که نام آنها در مقابل هر بخش آمده است ، تشکیل داد.

۱- بخش آمار نظری : مسئول آقای دکتر احمد پارسیان
اعضا : دکتر ناصررضا ارقامی
دکتر سیامک نوربلوچی
دکتر محمدرضا مشکانی

۲- بخش احتمال: مسئول آقای دکتر محمدقاسم وحیدی‌اصل
اعضا : دکتر احمدرضا سلطانی
دکتر بیژن ظهوری زنگنه
دکتر علی اکبر رحیم زاده ثانی

۳- بخش آمار زیستی: مسئول آقای دکتر انوشیروان کاظم‌نژاد
اعضا : دکتر محمود محمودی
دکتر سقراط فقیه زاده
دکتر محمدرضا اشراقیان
دکتر یدالله محرابی
دکتر ابراهیم حاجی زاده
دکتر سیدمحمدکاظم نائینی

۴- بخش کشاورزی : مسئول آقای مهندس علیرضا دهقان
اعضا : دکتر بهمن یزدی صمدی
دکتر پرویز وجدانی
دکتر عباس گرامی
دکتر محمدرضا قنادها
مهندس فرزان نصیری آذر
مهندس علی یوسفیان دارانی

۵- بخش اقتصاد : مسئول آقای دکتر محمد ذکائی

۷- مدیریت : مسئول آقای دکتر محمدرضا حمیدی‌زاده
اعضا : دکتر محمد ذکایی
دکتر دری
دکتر لاجوردی

۸- آمار در علوم تربیتی : مسئول آقای دکتر سیامک نوربلوچی
اعضا : دکتر رحیمی
دکتر محمد ذکایی
* دکتر حسین‌علی پورکاظمی
دکتر هاشمی پرست
دکتر سلیمی‌زاده

۹- آمار بیمه : مسئول آقای دکتر محمدرضا مشکانی
اعضا : دکتر محمدقاسم وحیدی‌اصل
آقای ضمیری
آقای شاکرین
آقای نکویی

۱۰- آمار در هواشناسی : مسئول آقای دکتر محمدرضا مشکانی
اعضا : خانم فاطمه رحیم‌زاده
دکتر غلامعلی کمالی

۱۱- آمار رسمی : مسئول آقای دکتر حمیدرضا نواب‌پور

آقای سید مهدی اعرابی رئیس روابط عمومی دانشگاه
آقای محسن باقری
آقای داریوش رباطی
آقای غضنفر حضرتی
آقای محمود بهمن نژاد

همکاران اداری دانشکده :

سیدغلامرضاسکاکی رئیس دبیرخانه کمیته هماهنگی

آقای نعمت الله رخشنده
خانم مریم فضلعلی پور
خانم مریم قادری
خانم نسرین ستاری
خانم مریم بهمنش

خانم مهین صادقی
خانم حمیرا خوشدل
آقای حجت الله پورشاد
آقای سیداسماعیل خیرآبادی
آقای سید فاضل بهبودی

مسئول دبیرخانه کمیته برگزاری : آقای محمد صدری
مسئول اداره انتظامات : آقای عزیزالله رضاییان
مسئول خرید کتاب : آقای علی‌اکبر ترابی
مدیر کمیته فرعی بودجه : آقای ناخدا
مدیر کمیته فرعی روابط بین الملل : خانم ژیلا نجفی

کمیته انتشارات :

آقای شمخال رحمتی **مدیر کمیته انتشارات**
اقای سماک امینی
آقای محمد حسین برهانی
آقای امان الله سنجری

آقای سید قاسم سید
آقای مجتبی شیر افکن
آقای ذبیح الله گنجی
آقای حسین میرزایی

نمایشگاه کامپیوتر و انفورماتیک کنفرانس :
آقای مهندس جواد فخاری

کارگاههای آموزشی و ارائه نرم‌افزارها
آقای دکتر محمدرضا فقیهی

خانم جزایری
آقای صدرالله صداقت
آقای محمد فضلعلی
آقای حسین قاسم کاشانی
آقای یحیی محمودی
آقای اسماعیل نورافشان
آقای غلامرضا یکتادوست حسینی

آقای مرتضی مصنف
آقای علیرضا مولایی
آقای حمید ناطقی
آقای لطیف یگانه

امور تغذیه :
آقای محمد روشن مدیر امور تغذیه
خانم شاهیندخت جمالی
آقای بیژن زحمتکش
آقای میر هادی طاهری
خانم مریم علیرشتی
قای سید مهدی حسین کاظمی

امور انتظامات :
هوشیار کولیوند سرپرست انتظامات
علیرضا حسین زاده
یدالله فلاحی
صیاد خدایاری
هوشنگ سلطانی
سیامک فصیحی
عزیزالله رضائیان

فتر فنی :
قای عباس یزدی ضمدی
قای کلامی
قای اصفهانی
نای کیوان
نای رحیمی
نای قربانخانی
نای علی پاشا
ای آرخ

ور مالی :
ای امامی ذیحساب مدیرکل امور مالی
ای باقر جلیلی
ای سعید سعیدی
ای داود علی پناه
انم حسنی مقدمی
ای هوشنگ منعمی

۱- بخش آمار نظری

۲- بخش احتمال

برآورد ماکسیمم درستنمائی برای توزیعهای وزنی

سلیمان خزائی[1] ، جوادبهبودیان[2]

۱-دانشگاه همدان ۲-دانشگاه شیراز

چکیده: می‌دانیم که اگر متغیر تصادفی X دارای تابع احتمال $f(x \mid \theta)$ باشد و اگر مشاهدهٔ X=x شانسی متناسب با W(x) (تابعی نا منفی) برای انتخاب شدن داشته باشد، آنگاه متغیر تصادفی X^w با تابع احتمال

$$f^w(x,\theta)=\frac{w(x)f(x/\theta)}{E[w(x)]}$$

را متغیر تصادفی وزنی با تابع وزن W(x) گویند.

منگامیکه تابع وزن بکار رفته، صعودی باشد؛ کسب اطلاعات بیشتر در مورد برآورد پارامتر مجهول جامعه و أثیر این تابع بر روی برآورد، دارای اهمیّت است.

ر این مقاله، برای خانواده‌های نمائی وزنی برآورد ماکسیمم درستنمائی (MLE) را بدست آورده، رفتار توابع زنی یکنوا را در روشهای برآورد پارامتر مثلاً روش ماکسیمم درستنمائی مطالعه می‌کنیم. بهمین منظور، رآورد ماکسیمم درستنمائی (MLE) توزیعهای وزنی را در حالتی که تابع احتمال متغیر تصادفی X به ناوادهٔ نمائی تعلق دارد را بدست آورده، تأثیر توابع وزنی یکنوا را روی برآوردهای ماکسیمم درستنمائی انواده‌های نمائی بررسی می‌کنیم.

ــبرآورد ماکسیمم درستنمائی توزیعهای نمائی وزنی:

فرض کنیم: تابع چگالی متغیر تصادفی X متعلق به خانواده نمائی بصورت زیر است :

$$f(x/\theta)=e^{\theta T(x)-c(\theta)} \qquad (۱-۱)$$

$$\frac{\partial}{\partial\theta} \ln L(\theta, x_1, \dots, x_n) = \sum_{i=1} T(x_i) - nc'(\theta)$$

$$\frac{\partial}{\partial\theta} \ln L(\theta, x_1, \dots, x_n) = 0$$

$$\Rightarrow \quad c'(\theta) = \frac{1}{n} \sum_{i=1}^{n} T(x_i) \quad (۲-۱)$$

برآورد ماکسیمم درستنمائی که آن را با $\hat{\theta}$ نشان میدهیم ، ریشه معادله (۲-۱) است.

brown(1986) ثابت کرد که اگر $c(\theta)$ تابعی اکیدا محدب نسبت به θ باشد یـ $0 < c''(\theta) = \text{var}[T(x)]$ ، آنگاه جواب این معادله منحصر بفرد است [1].پس در این مقاله فرض می کنی $c(\theta)$ تابعی اکیدا محدب نسبت به θ باشد، یعنی برآورد ماکسیمم درستنمائی θ، یکتا است . $c(\theta)$ تابع مولد تجمعی (cumullant generation function) گویند .

حال فرض کنید متغیر تصادفی X دارای تابع چگالی (۱-۱) با تابع وزن $w(x)$ باشد ، آنگاه طبق تعری ،تابع چگالی متغیر تصادفی وزنی X^w بصورت زیر است :

$$f^w(x/\theta) = \frac{f(x/\theta).w(x)}{E_\theta[w(x)]}$$

$$= \frac{w(x)}{E_\theta[w(x)]} e^{\theta T(x) - c(\theta)}$$

$$= w(x)\, e^{\theta T(x) - c(\theta) - \log E_\theta[w(x)]}$$

$$= w(x)\, e^{\theta T(x) - (c(\theta) + \log E_\theta[w(x)])}$$

بافرض $c_w(\theta) = c(\theta) + Log E_\theta[w(X)]$

$$f^w(x/\theta) = w(x) e^{\theta T(x) - c_w(\theta)} \qquad \theta \in \theta$$

با اندکی دقت ملاحظه میشود که تابع چگالی بدست آمده متعلق به خانواده نمائی با اندازه $w(x)d\mu(x)$ است . هنگامی که اندازه احتمال برابر μ است، خانواده نمائی (۱-۱) نیز به خانواده ای از توزیعهای وزنی متعلق است با تابع وزن $w(x) = e^{\theta T(x)}$ که $E[w(x)] = e^{c(\theta)}$ حال، فرض کنیم ، $\hat{\theta}^w$ برآورد ماکسیمم درستنمائی پارامتر مجهول توزیع وزنی ، یعنی θ_w باشد پس:

$$L_w(\theta, x_1, \dots, x_n) = \prod_{i=1}^{n} f^w(x_i/\theta)$$

... این ریشه منحصر بفرد میباشد .

۲-رابطه بین برآوردگرهای ماکسیمم درستنمائی وزنی و غیروزنی:

در این بخش ، این دو برآوردگر را با یکدیگر مقایسه می کنیم . چون این مقایسه ، هنگامی که فضای پارامتر خانواده های وزنی وغیروزنی یکسان نیست پیچیده می باشد و ممکن است معادله راستنمائی (۱-۲) برای یک خانواده دارای جواب باشد و برای خانواده دیگر جواب نداشته باشد ، پس فرض می کنیم فضای پارامتر در هر دو خانواده یکسان باشدو معادلات راستنمائی (۱-۳) و (۱-۲) دارای جواب منحصربفرد باشند .

قضیه (۲-۱) : فرض کنید تابع چگالی متغیر تصادفی Xمتعلق به خانواده نمائی، بصورت زیر است:

$$f(x/\theta) = e^{\theta T(x)-c(\theta)} \qquad \theta \in \theta, x \in R$$

و فرض کنید $\hat{\theta}$ و $\hat{\theta}_w$ بترتیب برآوردگرهای ماکسیمم درستنمائی پارامترهای θو θ_w هستند.

الف- اگر تابع وزن $w(x)$ ،غیرنزولی از $T(x)$ باشد آنگاه $\hat{\theta}_w \le \hat{\theta}$.

ب- اگر تابع وزن $w(x)$ ، غیرصعودی از $T(x)$ باشد آنگاه $\hat{\theta}_w \ge \hat{\theta}$.

اثبات: می دانیم که:

$$\frac{d}{d\theta}c_w(\theta) = \frac{d}{d\theta}c(\theta) + \frac{d}{d\theta}\log E[w(X)]$$

فرض می کنیم $c'(\theta) = \frac{d}{d\theta}c(\theta)$ و $c'_w(\theta) = \frac{d}{d\theta}c_w(\theta)$. چون $f(x/\theta)$ دارای خاصیت نسبت راستنمائی یکنوا (MLR) است و $w(x)$ تابعی غیرنزولی از $T(x)$ است در نتیجه $0 \le \frac{d}{d\theta}\log E_\theta[w(x)]$ و در نتیجه:

$$c'_w(\theta) \ge c'(\theta)$$

از اینکه طبق فرض بر آوردگرهای ماکسیمم درستنمائی منحصربفرد هستند، بنابراین $\hat{\theta}_w \le \hat{\theta}$.

نامساوی فوق را با برهان خلف ثابت می کنیم . پس اگر نامساوی برقرار نباشد، باید $\hat{\theta}_w > \hat{\theta}$.

چون $c(\theta)$ و $c_w(\theta)$ توابعی اکیدا محدب بر حسب θ هستند با توجه به[4] .

$$\hat{\vartheta}) + \frac{d}{d\theta}\log E_{\hat{\theta}}[w(X)] < \frac{1}{n}\sum_{i=1}^{n} T(x_i)$$

اما میدانیم که $\hat{\theta}$ ریشه معادله (۱-۲)است پس:

$$\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} T(x_i)$$

درنتیجه نامساوی بالا نمی تواند برقرار باشد لذا $\hat{\theta}_w \le \hat{\theta}$.

اثبات قسمت (ب) قضیه مشابه است.

تعریف: تابع دومتغیره $f(x, y)$ که x،y متعلق به مجموعه های یک بعدیX و Y هستندرا بطور کلی م

(Total positive)از مرتبه K گویند اگر برای تمام $x_1 < x_2 < \ldots < x_m$ و $y_1 < y_2 < \ldots < y_m$

داشته باشیم: $1 \le m \le k, y_i \in Y, x_i \in X$

$$\begin{pmatrix} y_1, y_2, \ldots, y_m \\ x_1, x_2, \ldots, x_m \end{pmatrix} = \begin{vmatrix} f(x_1, y_1), f(x_1, y_2), \ldots, f(x_1, y_m) \\ f(x_2, y_1), f(x_2, y_2), \ldots, f(x_2, y_m) \\ \\ f(x_m, y_1), f(x_m, y_2), \ldots f(x_m, y_m) \end{vmatrix} \ge 0$$

چنین تابعی را با علامت TP_k نشان می دهند.در آمار خاصیت TP_k یک تابع چگالی با فرض k=2 مه

خاصیت نسبت راستنمائی یکنوای تابع چگالی است [3]

و $\hat{\theta}$ و $\hat{\theta}_w$ بترتیب براوردگرهای ماکسیمم درستنمائی توزیع اصلی و توزیع وزنی باشند.
آنگاه:

الف- اگر تابع وزن $w(x)$ غیرنزولی از $T(x)$ باشد ، آنگاه برای هر جواب $\hat{\theta}_w$ ، $\hat{\theta}_w \leq \hat{\theta}$.

ب- اگر تابع وزن $w(x)$ غیرصعودی از $T(x)$ باشد، آنگاه برای هر $\hat{\theta}_w$ ، $\hat{\theta}_w \geq \hat{\theta}$.

اثبات: [2].

هنگامیکه تابع وزن بکار رفته ، به پارامتر بستگی داشته باشد ،براورد ماکسیمم درستنمائی پارامتر چگالی وزنی نیز به این پارامتر وابسته است بطوریکه با تغییر پارامتر تابع وزن می توان براوردهای متفاوتی را بدست آورد. مثلا اگر $w_\alpha(x)$ تابع وزنی با پارامتر α و $\hat{\theta}_\alpha$ براورد ماکسیمم درستنمائی چگالی وزنی باشد باتغییر α و $\hat{\theta}_\alpha$ نیز تغییر میکند در صورتیکه این تغییرات محسوس نباشد می توان به تابع وزن بکار رفته اطمینان نمود . برعکس ،اگر این تغییرات زیاد باشد تابع وزن بکار رفته مناسب نبوده و به اطلاعات بیشتری در مورد آن نیازمندیم .بطور ساده میتوان گفت اگر $\hat{\theta}_\alpha$، تابعی یکنوا از α باشد ،آسانتر می توان در مورد انتخاب تابع وزن تصمیم گرفت.

قضیه ۲-۳: فرض کنید، تابع چگالی $f(x/\theta)$ دارای آماره بسنده $T(x)$ است ، و $\hat{\theta}_\alpha$ براورد ماکسیمم درستنمائی توزیع وزنی با تابع وزن $w_\alpha(x)$ باشد و:

الف- بازای هر x ثابت، $\log f(x/\theta)$ ، نسبت به θ مقعر است .

ب- $f(x/\theta)$ نسبت به $(T(x) , \theta)$ ،تابع TP_2 است.

ت- $f(x/\theta)$ نسبت به θ مشتق پذیر بوده برای $w_\alpha(x) = h(T(x),\alpha)$ و $h(x) = 1$ داشته باشیم:

$$\frac{d}{d\theta}\int f(x/\theta)h(T(x),\alpha)d\mu(x) = \int \frac{d}{d\theta}f(x/\theta)h(T(x),\alpha)d\mu(x)$$

ج- براورد ماکسیمم درستنمائی $\hat{\theta}_\alpha$، برای هر α یکتا باشد . در اینصورت اگر $h(T(x),\theta)$ تابع TP2 باشد ، $\hat{\theta}_\alpha$ تابعی غیرصعودی از α است.

اثبات: [7].

قضیه ۲-۴: برای خانواده نمائی که تابع چگالی توزیع آن بصورت :

$$f(x/\theta) = e^{\theta T(x) - c(\theta)}$$

تعریف شده باشد . فرض کنید تابع $T(x)$ فرد و اندازه μ متقارن باشد $\hat{\theta}$ و $\hat{\theta}_w$ نیز بسترتیب براوردگرهای ماکسیمم درستنمائی برای توزیع اصلی و توزیع وزنی باشند.

باشدکه متعلق به خانواده نمائی است و برای هر α، برآوردماکسیمم درستنمائی وزنی یکتا باشد اگر $w(x) = h(\alpha, |T(x)|)$ و TP_2 باشد آنگاه $|\hat{\theta}_\alpha|$ تابعی غیرصعودی از αاست .

اثبات:[2].

مثال : فرض کنید دو متغیر تصادفی Z و S^2 مستقل از یکدیگر باشند بطوریکه :

$$N(0,1), KS^2 \approx X_k^2$$

و متغیر تصادفی T را بصورت $T = \frac{Z+\theta}{S}$ تعریف می کنیم ،که دارای توزیع استودنت غیرمرکزی با پا θاست.

توابع چگالی متغیرهای تصادفی Z و S را بترتیب با ϕو g_k نشان می دهیم ، داریم :

$$,s) = \phi(z) g_k(s)$$

$$s) = s\phi(st-\theta) g_k(s) \qquad s>0 \text{ ، } t \in R$$

در نتیجه :

$$\theta) = \int_0^\infty s\phi(st-\theta) g_k(s) ds \qquad t \in R$$

که$f(t,\theta)$ تابع چگالی استودنت غیرمرکزی با پارامتر θ می باشد . حال شرایط قضیه (۲-۲) را برا
بررسی می کنیم .

الف- تابع چگالی $f(t,\theta)$ تابع $TP2$ است.

اثبات: فرض کنید $\theta_1 < \theta_2$ و $t_1 < t_2$

$$\theta) = \begin{vmatrix} f(t_1,\theta_1) f(t_1,\theta_2) \\ f(t_2,\theta_1) f(t_2,\theta_2) \end{vmatrix}$$

$$(t_1,\theta_1).f(t_2,\theta_2) - f(t_1,\theta_2) f(t_2,\theta_1)$$

$$s\phi(st_1-\theta_1) g_k(s) ds \quad \int_0^\infty s\phi(st_2-\theta_2) g_k(s) ds$$

$$s\phi(st_1-\theta_2) g_k(s) ds \quad \int_0^\infty s\phi(st_2-\theta_1) g_k(s) ds$$

$$\int_0^\infty s^2 g_k^2(s) \phi(st_1-\theta_1)\phi(st_2-\theta_2)(ds)^2 -$$

$$s^2 g_k^2(s) \phi(st_1-\theta_2)\phi(st_2-\theta_1)(ds)^2$$

با علم اینکه $m(\theta) > 0$ باشد باید عبارت زیر انتگرال مثبت باشد پس :

$$[(st_1-\theta_1)^2+(st_2-\theta_2)^2]\le[(st_1-\theta_2)^2+(st_2-\theta_1)^2]$$

$(\theta_1-\theta_2)(t_1-t_2)\ge 0$ امامی دانیم که $\theta_2>\theta_1$ و $t_2>t_1$ بنابراین نامساوی فوق برقرار است . در نتیجه $m(\theta)\ge 0$. یعنی شرط (ب)قضیه(۲-۲) برقرار است.

ب- اگر $w(x)$ تابعی کران دار یا چند جمله ای باشد آنگاه:

$$\frac{d}{d\theta}\int_0^\infty sw(x)\phi(ts-\theta)g_k(s)ds=\int_0^\infty\frac{d}{d\theta}sw(x)g_k(s)\phi(ts-\theta)ds$$

یعنی شرط (ت) قضیه (۲-۲)برقرار است.

برای بررسی شرط (الف) قضیه از قضیه معروف به perekop(1971) استفاده می کنیم که بصورت زیر است.

قضیه: برای $x\in R^m$ و $y\in R^n$ فرض میکنیم $f(x,y)$ تابع لگاریتم- مقعر (تابعی که لگاریتم آن مقعر است) باشد. آنگاه :

$$G(x)=\int_{R^n}f(x,y)dy$$

نیز تابع لگاریتم مقعر است .

اثبات:[6].

با توجه به قضیه فوق لم زیرا اثبات می کنیم .

لم: اگر $f_k(t/\theta)$ تابع چگالی متغیر تصادفی استودنت غیرمرکزی با پارامتر θ و درجه آزادی k باشد. آنگاه برای هرt ، $f_k(t/\theta)$ تابع لگاریتم- مقعر از θ است .

اثبات: فرض کنید $ks^2\approx\chi^2_k$ ، بنابراین تابع چگالی آن عبارت است از:

$$h(ks^2)=\frac{1}{(\frac{k}{2})2^{\frac{k}{2}}}(ks^2)^{\frac{k}{2}}e^{-\frac{ks^2}{2}}\qquad s>0,k>0$$

وتابع چگالی متغیر تصادفی s عبارت است از:

$$g_k(s)=h(ks^2)|J|$$

$$(\frac{k}{2})^{\frac{k}{2}-1}2$$

$$f_k(t,\theta)=\int_0^{\infty} s\phi(ts-\theta)g_k(s)ds$$

پس:

$$f_k(t,\theta)=\frac{2(\frac{k}{2})^{\frac{k}{2}}}{(\frac{k}{2})}\int_0^{\infty}\exp[-Q_t(s,\theta)]ds$$

که $t\in R$

$$\phi(ts-\theta)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(ts-\theta)^2}$$

در آن $Q_t(s-\theta)=\frac{t^2k^2-2ts\theta+\theta^2}{2}+\frac{ks^2}{k}-k\log s$.

تابع (s,θ) Q_t تابع محدب از θ و s میباشد ،زیـــرا مــاتریس هســیان (Hessian-matrix) آن نیمه معین ثابت است [5] . یعنی:

$$\begin{bmatrix} \frac{\partial^2 Q_t(s,\theta)}{\partial s^2} & \frac{\partial^2 Q_t(s,\theta)}{\partial s^2} \\ \frac{\partial^2 Q_t(s,\theta)}{\partial\theta\partial s} & \frac{\partial^2 Q_t(s,\theta)}{\partial\theta^2} \end{bmatrix}\geq 0$$

در نتیجه $Q_t(s,\theta)$ - تابع مقعر و $\exp[-Q_t(s,\theta)]$ تابع لگاریتم - مقعر می باشد ،و طبق قضیه prekopa $f_k(t/\theta)$ ، لگاریتم مقعر است . بنابراین ،تابع چگالی توزیع اســتودنت غــیرمرکزی بــا پارامتر θ ، برای هر t ثابت نسبت به θ مقعر است .

نتیجه :اگر $f_k(t/\theta)$ تابع چگالی توزیع استودنت غیر مرکزی بــا پــارامتر θ و تــابع وزن مــورد

monograph series.

2. *Iyenger, S. and Zhao peng-liang (1994). Maximum likelihood estimation for weithed distributions. statistics and probability letters, 21,37-47.*

3. *Kotz, S. and Johnson, **N.L**. (1985). Encyclopedia of statistical sciences. v,9,289.*

4. *Lehmann,E.(1991).Theory of point estimation.*

5. *Mahfoud, M. and Patil ,G.P. (1982). On weigthed distributions. in statistics and probability: Essays in honor of C.R.Rao, G.Kallianpur et al Amesterdam: North Holland, 479-492.*

6. *Perkopa, A.(1971). Logarithmic concave measures with applications Acta scientiarum mathematicarum 32,302-316 .*

۷ -خزائی سلیمان . "برآوردماکسیمم درستنمائی برای توزیعها ی وزنی یک متغیره وتوزیعهای وزنی دو متغیره" دانشگاه شیراز، پایان نامه کارشنا سی ارشد آمار ۱۳۷۵ .

تابع زیان LINEX

ناهید سنجری ، نرگس عباسی

دانشگاه شیراز

چکیده در این مقاله، برآورد بیز و بهترین برآورد پایا با کمترین زیان را برای دترمینان ماتریس کواریانس ، $|\Sigma|$ ، همچنین برآوردی برای اثر ماتریس کواریانس ، $tr\Sigma$ ، بر اساس یک نمونه Nتایی جمعیت نرمال pـمتغیره با میانگین μ و ما تریس کواریانس Σ ، تحت تابع زیان Linex ارائه نموده‌ا .در حالتیکه بین مولفه های نمونه ضریب همبستگی مشترکی چون ρ وجود دارد، برآورد بیز برای ا پارامتر را محاسبه کرده‌ایم .

۱ مقدمه

بر مبنای یک نمونه تصادفی ازنرمال p-متغیره با میانگین μ وکواریانس Σ ،تحت تابع زیان Linex

$$L(\hat{\theta},\theta) = b\{e^{a(\frac{\hat{\theta}}{\theta}-1)} - a(\frac{\hat{\theta}}{\theta}-1) - 1\} \qquad (1-1)$$

$$L(\hat{\theta},\theta) = b\{e^{a(\frac{\theta}{\hat{\theta}}-1)} - a(\frac{\theta}{\hat{\theta}}-1) - 1\} \qquad (1-2)$$

که‌در آن $b > 0$ و $a \neq 0$ است ، برآوردهای نقطه‌ای برای اثر ماتریس کواریانس و دترمینان ماتریس کواریانس مورد بررسی قرار می دهیم .

در خانواده نمایی یک پارامتری اگر $\hat{\theta}$، آماره کافی برای θ باشد آنگاه بهترین برآوردکننده برای θ

$$\frac{1 - e^{\frac{}{n+2}}}{2a} \sum_{i=1}^{n}(X_i - \bar{X})^2$$

در قسمتهای فوق ملاحظه شد که ضریب آماره تابعی از a و n می باشد ، فرض می کنیم که برآوردکننده‌ها به شکل $g(e^z) \mid S \mid$ برای $\mid \Sigma \mid$ و $h(e^z) trS$ برای $tr\Sigma$ باشند که h و g توابعی مناسب از z می باشند. با این ایده، در قسمتهای ۲و۳ بهترین مضرب از trS ، برآورد MRE و برآورد بیز برای $\mid \Sigma \mid$ و درقسمت آخر، برآورد بیز را برای ضریب همبستگی مشترک در نمونه را محاسبه کرده‌ایم .

۲ برآوردیابی $tr\Sigma$

فرض کنید $X_1, \cdots, X_N$ یک نمونه‌ای تصادفی از $N_p(\mu, \Sigma)$ ، بطوری که μ و Σ هر دو نامعلومند ، $\bar{X}$ و $S = \sum_{i=1}^{N}(X_i - \bar{X})^2$ آمارهای بسنده می‌نیمال و

$$S \sim W_{p-1}(\Sigma, N-1)$$

بر مبنای trS ما می خواهیم $tr\Sigma$ را تحت تابع زیان ۱-۱ برآورد کنیم .

لم ۱.۲ برای هر ماتریس مثبت قطعی (p.d) ، $p \times p$ مانند Σ داریم :

$$\mid \Sigma \mid \leq p^{-p}(tr\Sigma)^p$$

اثبات : بنابر $\lambda_1, \ldots, \lambda_p$ ، مقادیر ویژه ماتریس Σ، داریم : $\mid \Sigma \mid = \Pi_{i=1}^{p}\lambda_i$ ، $tr\Sigma = \sum_{i=1}^{p}\lambda_i$

$$\frac{\sum_{i=1}^{p}\lambda_i}{p} \geq \sqrt[p]{\lambda_1 \ldots \lambda_p}.$$

قضیه ۲.۲ تحت تابع زیان ۱-۱ و باقرار دادن $\theta = tr\Sigma$ ، $\hat{\theta} = ctr\Sigma$ ، $n = N - 1$ به طوریکه $I - \frac{2ac}{tr\Sigma}\Sigma$ یک ماتریس مثبت قطعی باشد ، آنگاه

$$\begin{aligned} R(ctr\Sigma, tr\Sigma) &\geq b\{e^{-\frac{2ac}{np+2}}(1 + \frac{np}{2}) + \frac{np}{2} + a - 1\} \\ &= R(c^* tr\Sigma, \sigma^2 I) \end{aligned}$$

$$
\begin{aligned}
&= p^{-p}(p - 2ac)^p \\
&= (1 - \frac{2ac}{p})^p
\end{aligned}
$$

بنابر این

$$
\begin{aligned}
R(ctrS, tr\Sigma) &\geq b\{e^{-a}(1 - \frac{2ac}{p})^{\frac{-np}{2}} - anc + a - 1\} \\
&= R(ctrS, \sigma^2 I) \\
&\geq R(c^* trS, \sigma^2 I) \\
&= b\{(e^{\frac{-2ap}{np+2}} - 1)(1 + \frac{np}{2}) + a - 1\}
\end{aligned}
$$

که در آن $c^* = \frac{p}{2a}(1 - e^{\frac{-2a}{np+2}})$.

در حالت $p = 2$ فرمول تابع مخاطره به فرم زیر خواهد شد:

$$
\begin{aligned}
R(ctrS, tr\Sigma) &= b\{e^{-a}(1 - 2ac + a^2c^2k)^{\frac{-n}{2}} - anc + a - 1\} \\
&= R_{k,n}(ctrS, tr\Sigma)
\end{aligned}
$$

که در آن $0 < k < 1$ و $k = \frac{4|\Sigma|}{(tr\Sigma)^2}$.اگر نقطه می نیمم $R_{k,n}(ctrS, tr\Sigma)$ را با $c^*_{k,n}$ نشان دهیم مقادیر $c^*_0 = \frac{1}{2a}(1 - e^{\frac{-2a}{n+2}})$ و $c^*_1 = \frac{1}{a}(1 - e^{\frac{-a}{n+1}})$ به ترتیب $R_{0,n}(ctrS, tr\Sigma)$ و $R_{1,n}(ctrS, tr\Sigma)$ را می نیمم می کنند.

در جدول ۲ـ۱ برای حالت $a = -1$ ، با در نظر گرفتن برخی مقادیر k و n، مقدار $c^*_{k,n}$ را بدست آورده‌ایم . بعد از چند تبدیل مناسب بر روی $k, n, c^*_{k,n}$ و از طریق رگرسیون چندگانه برآورد کلی را برا $c^*_{k,n}$ را به شکل زیر ارائه می نماییم :

$$
\frac{1}{2 - k}(e^{\frac{1.0076(2-k)^{1.01}}{n+2-k}} - 1)
$$

در عمل نمی توان از مقدار فوق استفاده نمود زیرا به مقدار مجهول k بستگی دارد.

قضیه ۲.۳ تحت زیان Linex و $\theta = tr\Sigma$ ، $\hat{\theta} = ctrS$ $a \geq -1$ و $p = 2$ برای هر $c^*_0 <$

از انجا که برای هر n و $\circ < k_{\circ} < k_{۱} < ۱$ داریم :

$$R_{k_{\circ}}(ctrS, tr\Sigma) > R_{k_{۱}}(ctrS, \Sigma).$$

$$R_{k}(c_{k}^{*}trS, \Sigma) = \frac{۱ - ۲ac + a^{۲}c^{۲}k}{۱ - ack} - anc$$

براحتی برای $a > \circ$ میتوان رابطه (*) را ثابت کرد. برای $-۱ < a < \circ$ ، با توجه به اینکه به ازای هر k یک مقدار c_{k}^{*} وجود دارد و $\circ < c_{k}^{*} < \frac{۱}{n}$ ، میتوان نشان داد. مقدار مشتق $\frac{dk}{dc_{k}^{*}}$ مثبت است ودر نتیجه قضیه به اثبات میرسد.

جدول- ۲.۱: مقادیر بهینه $c_{k,n}^{*}$ برای حالت $a = -۱$.

k	$n = ۲$	$n = ۳$	$n = ۴$	$n = ۵$	$n = ۸$	$n = ۱۰$	$n = ۲۰$	$n = ۴۰$
۰/۰۵	۰/۳۲۸	۰/۲۴۸	۰/۱۹۹	۰/۱۶۶	۰/۱۱۰	۰/۰۰۹	۰/۰۴۷	۰/۰۲۴
۰/۱۵	۰/۳۳۶	۰/۲۵۲	۰/۲۰۱	۰/۱۶۷	۰/۱۱۱	۰/۰۹۱	۰/۰۴۷	۰/۰۲۴
۰/۲۵	۰/۳۴۴	۰/۲۵۶	۰/۲۰۴	۰/۱۶۹	۰/۱۱۲	۰/۰۹۱	۰/۰۴۷	۰/۰۲۴
۰/۳۵	۰/۳۵۱	۰/۲۶۰	۰/۲۰۶	۰/۱۷۱	۰/۱۱۳	۰/۰۹۲	۰/۰۴۸	۰/۰۲۴
۰/۴۵	۰/۳۵۹	۰/۲۶۴	۰/۲۰۹	۰/۱۷۳	۰/۱۱۳	۰/۰۹۲	۰/۰۴۸	۰/۰۲۴
۰/۵۵	۰/۳۶۶	۰/۲۶۸	۰/۲۱۱	۰/۱۷۴	۰/۱۱۴	۰/۰۹۳	۰/۰۴۸	۰/۰۲۴
۰/۶۵	۰/۳۷۳	۰/۲۷۲	۰/۲۱۳	۰/۱۷۶	۰/۱۱۵	۰/۰۹۳	۰/۰۴۸	۰/۰۲۴
۰/۷۵	۰/۳۸۰	۰/۲۷۵	۰/۲۱۶	۰/۱۷۷	۰/۱۱۵	۰/۰۹۴	۰/۰۴۸	۰/۰۲۴
۰/۸۵	۰/۳۸۶	۰/۲۷۹	۰/۲۱۸	۰/۱۷۹	۰/۱۱۶	۰/۰۹۴	۰/۰۴۸	۰/۰۲۴
۰/۹۵	۰/۳۹۲	۰/۲۸۲	۰/۲۲۰	۰/۱۸۰	۰/۱۱۷	۰/۰۹۴	۰/۰۴۸	۰/۰۲۴

۳ براوردیابی $|\Sigma|$

بنابر $\frac{|S|}{|\Sigma|} \sim \chi_{n}^{۲} \cdots \chi_{n-p+۱}^{۲}$ مقدار تابع مخاطره برابر است با [۲]:

$$\Sigma^{-1} \sim W_p(n', A^{-1})$$

که A^{-1} ماتریس مثبت قطعی و n' مقدار طبیعی معلوم، قرار میدهیم. بنابراین

$$\Sigma^{-1} \mid S \sim W_p(n+n', (S+A)^{-1})$$

و لذا

$$\frac{\mid S+A \mid}{\mid \Sigma \mid} \sim \chi^2(n+n') \dots \chi^2(n+n-p+1)$$

پس برآورد بیز برای Σ برابر است با:

$$\delta^* = c^*_{Bayes} \mid S+A \mid$$

که در آن $c^*_{Bayes} = g(n+n', a)$.

برای حالت $p = 2$ و $a = -1$ ، با تغییر دادن مقادیر n در جدول ۳ـ۱ مقادیر بهینه c^*_{MRE} و c^*_{Bayes} را بدست می آوریم و درآخر توسط رگرسیون ساده برآوردی برای $g(n,-1)$ را نمایش می دهیم :

$$\begin{aligned} g(n,-1) &= \frac{1}{2}\left(e^{-0.025+\frac{0.529}{n+2}} - 1\right) \\ &\simeq \frac{1}{2}\left(e^{\frac{1}{2(n+1)}} - 1\right). \end{aligned}$$

جدول ۳ـ۱ : مقادیر c^*_{MRE} برای برآوردیابی $\mid \Sigma \mid$

n	۲	۴	۶	۸	۱۰	۱۵	۲۰	۵۰
c^*_{MRE}	۰/۱۳۱۲	۰/۰۴۹۷	۰/۰۲۳۷	۰/۰۱۳۸	۰/۰۰۹۱	۰/۰۰۴۱	۰/۰۰۲۳	۰/۰۰۰

۴ برآورد بیز برای ρ

فرض کنید ساختار ماتریس کواریس Σ به فرم زیر باشد:

$$\begin{pmatrix} 1 & \rho & \dots & \rho \end{pmatrix}$$

زیر [illegible] [۱].

$$g(\rho) \propto \frac{(۱+(p-۱)\rho)^{\frac{(p-۱)n}{۲}}(۱-\rho)^{\frac{n}{۲}}}{((۱+(p-۱)\rho)(p-۱)(۱-r)+(۱-\rho)(۱+(p-۱)r))^{\frac{pn}{۲}}}$$

که در آن $\frac{-۱}{p-۱} < \rho < ۱$ و r ضریب همبستگی نمونه‌ای

$$r = \frac{۲\sum_{u<v} S_{uv}}{(p-۱)trS}$$

می‌باشد.

جدول ۴ـ۱ مقدار برآورد بیز برای ρ به ازای مقادیر مختلف n و r تحت تابع زیان ۱ـ۲ محاسبه شده است . در شکل ۴ـ۲ مشاهده می کنیم برای هر n ، برآورد بیز برای ρ مضربی از r نیست . مقدار این برآورد به ازای $r = ۱$ وجود ندارد.

جدول- ۴ـ۱: برآورد بیز تحت تابع زیان ۱ـ۲ برای ρ .

r	$n=۲$	$n=۳$	$n=۴$	$n=۵$	$n=۷$	$n=۱۰$	$n=۱۵$	$n=۲۰$
۰٫۹۹	۰٫۹۷	۰٫۹۹	۰٫۹۹	۰٫۹۹	۰٫۹۹	۰٫۹۹	۰٫۹۹	۰٫۹۹
۰٫۹۵	۰٫۹۰	۰٫۸۳	۰٫۸۴	۰٫۸۶	۰٫۸۹	۰٫۹۱	۰٫۹۲	۰٫۹۳
۰٫۹۰	۰٫۹۲	۰٫۸۰	۰٫۷۸	۰٫۷۹	۰٫۸۴	۰٫۸۱	۰٫۸۶	۰٫۸۷
۰٫۸۵	۰٫۹۶	۰٫۷۹	۰٫۷۵	۰٫۷۴	۰٫۷۵	۰٫۷۸	۰٫۸۰	۰٫۸۱
۰٫۷۵	۱٫۰۷	۰٫۸۲	۰٫۷۲	٫۶۹	۰٫۶۷	۰٫۶۸	۰٫۶۹	۰٫۷۰
۰٫۶۵	۱٫۲۲	۰٫۸۹	۰٫۷۵	۰٫۶۸	۰٫۶۲	۰٫۶۰	۰٫۶۰	۰٫۶۱
۰٫۵۵	۱٫۴۴	۱٫۰۲	۰٫۸۲	۰٫۷۱	۰٫۶۰	۰٫۵۵	۰٫۵۳	۰٫۵۳
۰٫۴۵	۱٫۷۶	۱٫۲۱	۰٫۹۵	۰٫۸۱	۰٫۶۴	۰٫۵۳	۰٫۴۷	۰٫۴۶
۰٫۳۵	۲٫۲۶	۱٫۵۳	۱٫۱۸	۰٫۹۷	۰٫۷۳	۰٫۵۷	۰٫۴۶	۰٫۴۱
۰٫۲۵	۳٫۱۶	۲٫۱۳	۱٫۶۱	۱٫۳۱	۰٫۹۶	۰٫۷۰	۰٫۵۱	۰٫۴۲
۰٫۱۵	۵٫۲۷	۳٫۵۲	۲٫۶۵	۲٫۱۳	۱٫۵۴	۱٫۰۹	۰٫۷۵	۰٫۵۸
۰٫۰۵	۱۵٫۸	۱۰٫۵۵	۷٫۹	۶٫۳۴	۴٫۵۵	۳٫۱۶	۲٫۱۳	۱٫۵۹

مراجع

sian, A. and Sanjari Farsipour, N.(1993).On the Admissibility [۲]
Indmissibility of Estimators of Scale Parameters Using Asymmetric
Loss Function. *Commun. Statist. Theor. Meth.* **22**,2877-2901.

در کلاس طرح بلوک‌های کامل متعادل

علی شادرخ ، عباس گرامی

دانشگاه تربیت مدرس

مقدمه فرض کنید v تیمار داریم که می‌خواهیم در یک طرح کاملا تصادفی آنها را با هم مقایسه کنیم و مایلیم برای هر تیمار n تکرار داشته باشیم در این صورت به تعداد nv واحد آزمایشی همگن نیاز داریم. بدیهی است که به دست آوردن تعداد زیاد واحد همگن همیشه کار راحتی نیست. در طرح بلوکهای کامل تصادفی، واحدهای همگن را در یک بلوک قرار می‌دهیم که تعداد واحدهای آزمایشی این بلوک باید برابر با v تیمار آزمایشی باشد و این کار همیشه به راحتی ممکن نمی‌شود.

هدف از طرح‌هایی که در بارهٔ آنها صحبت شد مقایسه چند تیمار آزمایشی با هم بود، دراین طرح‌ها مقایسه زوج تیمارها ارزش یکسان دارد. اما بعضی وقتها هدف مقایسه چند تیمار آزمایشی با یک تیمار خاص است که این تیمارهای خاص را تیمار کنترل(1) می‌نامیم. در این حالت ارزش مقایسه زوج تیمارهای آزمایشی با کنترل از زوج تیمارهای آزمایشی بیشتر است.

مثلاً در امور پزشکی برای درمان یک بیماری یک داروی جدید تولید می‌شود و می‌خواهیم این دارو را در درمان بیماری با داروهای قبلی مورد سنجش قرار دهیم.

و یا مثلاً فرض کنید برای انجام کاری معین یک روش استاندارد در حال استفاده است، و همچنین چهار روش جدید قابل قبول دیگر هم طراحی شده‌است. حال می‌خواهیم روشهای جدید که همان تیمارهای آزمایشی هستند با روش جاری، که نام آن تیمار کنترل است، را مقایسه کنیم.

نحوه تخصیص واحدهای آزمایشی به تیمارها طرح نامیده می‌شود که آن را با d نمایش می‌دهیم و مسلم است که طرح‌های متعددی برای انجام یک آزمایش می‌توان تهیه کرد. آن طرحی مطلوبتر است که مقایسه‌های مورد نظر را

همگن هستند و طرح متناظر با این مدل را کاملا تصادفی گویند. لازم به توضیح است که هر مدل آماری دارای اجزایی است که یکی از آن اجزاء بیان رابطه تابعی بین متغیر پاسخ با فاکتورها و عوامل مورد بررسی و یا مزاحم است. در این رابطه y_{ij} شاهد jام برای تیمار iام، t_i اثر تیمار iام، μ میانگین کل و ε_{ij} اشتباه آزمایشی تکرار jام تیمار iام است.

مثال ۲-۱) : در مدل با رابطه (۲-۱) $y_{ij} = \mu + t_i + b_j + \varepsilon_{ij}$ فرض بر این است که واحدهای آزمایش گروههای همگن تقسیم بندی شده‌اند که به هر گروه یک بلوک گفته می‌شود و فرض بر آن است که ناهمگنی واحدها درون بلوکها از بین رفته است، یعنی هر بلوک دارای واحدهای آزمایشی تقریبا همگن است. طرح متناظر با این مدل را طرح بلوکهای تصادفی گوییم اگر، تعداد واحد آزمایشی همگن در هر بلوک برابر با تعداد تیمارهای مورد بررسی باشد. ان را طرح بلوک های کامل تصادفی می‌نامند که b_j نشان دهنده اثر بلوک jام است.

در این رابطه yها بیانگر مشاهدات بدست آمده از تیمار iام در ردیف lام بلوک jام است و t_i اثر تیمار iام و b_j اثر بلوک jام در مدل است. در روابط مدل‌های سه گانه فوق فرض بر آن است که ε_{ijl} ها متغیرهای تصادفی مستقل میانگین صفر و واریانس σ^2 می‌باشند.

با توجه به بیان مقدمات فوق حال می‌توان نسبت به چگونگی دقت بیشتر در مقایسه‌ها صحبت کرد. در این مجموعه v تعداد تیمار و b تعداد بلوک و k را تعداد واحدهای آزمایشی هر بلوک (اندازه بلوک) را نشان می‌دهند. و اندیسهای i ، j ، l روی $1 \leq i \leq v$ ، $1 \leq j \leq b$ ، $1 \leq l \leq k$ تعریف می‌شود. و همچنین پارامترهای v ، b ، k را با سه‌تائی (v,b,k) نمایش می‌دهیم. هدف اولیه و اصلی آن است که بررسی شود که آیا در بین تیمارهای مورد آزمون تیماری وجود دارد که از تیمار کنترل بهتر باشد. در حرکت به سوی این هدف باید عبارتهای $t_0 - t_i$ $1 \leq i \leq v$) را با بیشترین دقت ممکن برآورد شوند.

تحت فرضیات مدل‌های بالا روش حداقل مربعات بهترین برآورد خطی ناأریب (2) - $\hat{t}_{d0}$ (B.L.U.E) t_{di} $(1 \leq i \leq v)$ برای مقایسه‌های $t_0 - t_i$ تحت طرح d به دست می‌دهد. باید مطمئن بود که تیمارها را به واحدهای آزمایشی طوری نسبت داده شوند که مقایسه‌های $t_0 - t_i$ قابل برآورد باشند. طرح‌هایی که دارای چنین خاصیتی هستند را طرح‌های همبند (3) می‌گویند. در این بررسی توجه خود را به چنین طرح‌هایی معطوف می‌کنیم.

۳ – هدف و تعریف مسئله

طرح های بهین می گویند ، اما در بعضی شرایط پیدا کردن چنین طرحی فوق‌العاده مشکل و یا غیرممکن است. در این مواقع ما بدنبال طرح‌های می‌گردیم که به طرح‌های بهین نزدیک باشند، و یا به عبارت دیگر طرح‌های نزدیک به طرح‌های بهین، که هدف نهایی این مقاله است.

با توجه به مطالب فوق در این جا لازم می باشد که طرح‌های بهین را معرفی کنیم.

فرض کنید که رابطه مدل بصورت :

$$1 \leq l \leq k \,,\; 1 \leq j \leq b \text{ و } 0 \leq i \leq v \,,\; y_{ij} = \mu + t_i + \beta_j + \varepsilon_{ij}$$

باشد واندازه هر بلوک $k \leq v$ باشد، حال اگر n_{dij} عبارت باشد از تعداد دفعاتی که تیمار i ام در بلوک j ام طرح d ظاهر می‌شود. آنگاه $N_d=(n_{dij})$ راماتریس تقاطع[1] می‌نامیم.اگر $(0 \leq i \leq v)$ $r_{di}=\sum_{j=1}^{b} n_{dij}$ باشد آنگاه C_d را ماتریس اطلاعات می‌نامیم. $C_d=diag(r_{d0},r_{d1},...,r_{dv})-k^{-1}N_dN'_d$ این ماتریس $n.n.d$[2] است و اگر $r(C_d)=v$ باشد تمام مقایسه های دو به دو بین اثرات تیمارها قابل برآورد هستند و این ماتریس ناویژه است.

اگر ماتریس P را بصورت :

$$P = \begin{vmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \vdots & & & & \\ 1 & 0 & 0 & \dots & -1 \end{vmatrix}$$

در نظر بگیریم آنگاه PC_dP' ماتریس واریانس کوواریانس بردار برآوردگر حداقل مربعات مقایسه‌های $t_0-t_1, t_0-t_2, ..., t_0-t_v$ می‌باشد که در آن C_d یکی ازمعکوس‌های تعمیم یافته C_d می‌باشد.

فرض کنید که $C(v,b,k)$ کلاس تمام طرح‌های ممکن برای پارامترهای k,b,v باشند و $C=PC_dP'$ آنگاه با توجه به تعاریف فوق هدف ما پیدا کردن طرحی است مثل $d \in C(v,b,k)$ که مقدار توابع خاصی که روی ماتریس C تعریف می‌شوند را مینیمم کند

مثال :

طرح بهین نوع MV: طرح $d \in C(v,b,k)$ را طرح بهین نوع MV گویند اگر در بین تمام طرح‌های $C(v,b,k)$ $\underset{1 \le i \le v}{Max}\, Var(\hat{t}_{d0} - \hat{t}_{di})$ آن کمترین باشد.

طرح بهین نوع A: طرح $d \in C(v,b,k)$ را طرح بهین نوع A گویند اگر در بین تمام طرح‌های $C(v,b,k)$ $\sum_{i=1}^{v} Var(\hat{t}_{d0} - \hat{t}_{di})$ کمترین باشد. توجه داشته باشیم اگر μ_iها مقادیر ویژه ماتریس C باشند آنگاه:

$$tr(C) = \sum_{i=1}^{v} var(\hat{t}_{d0} - \hat{t}_{di}) = \sum_{i=1}^{v} \mu_i$$

۳ - معرفی مدل و تعاریف

فرض کنید که تعداد تیمارهای ما $v+1$ تا باشد آنها را با $0 , 1 , ... , v$ نمایش می‌دهیم که 0 تیمار کنترل می‌باشد. همچنین b بلوک به اندازه k داریم، اگر y_{ijl} مشاهده بدست آمده از تیمار iام $(0 \le i \le v)$ در l امین $(1 \le l \le k)$ واحد آزمایشی از بلوک jام $(1 \le j \le b)$ باشد آنگاه ما یک مدل خطی جمع‌پذیر بدون اثر متقابل و بصورت زیر داریم:

$$y_{ijl} = \mu + t_i + b_j + \varepsilon_{ijl}$$

که در آن μ میانگین کل و b_j اثر بلوک jام $(1 \le j \le \beta)$ و t_i اثر تیمار iام $(0 \le i \le v)$ می‌باشد و ε_{ijl} خطای تصادفی مستقل با میانگین صفر و واریانس σ^2 می‌باشد.

هدف ما در این جا آن است که طرحی پیدا کنیم که برآوردهای خطی نااریب $\hat{t}_i - \hat{t}_0$ $(1 \le i \le v)$ دارای کمترین واریانس در بین تمام طرح بلوکهای با $k > v$ باشد. برای رسیدن به این هدف و با توجه به این که معیار A بهترین معیار برای این هدف می‌باشد ما تمام توجه خود را روی چنین معیاری معطوف می‌کنیم.

در سال *Ting and Notz 1988* در طرح بلوکهای کامل $(k > v)$ و در دامنه خاصی از پارامترها $2 \le v \le 50$ ، $2 \le b \le 50$ ، $v \le k \le 30$ طرح‌های بهین نوع A را بر اساس یک قضیه به دست آورند قبل از بیان این قضیه نیاز به چند تعریف است که در ذیل آورده می‌شوند.

فرض کنید که $C(v,b,k)$ کلاس تمام طرح‌های ممکن برای پارامترهای k,b,v باشد و n_{dij} تعداد دفعاتی است که تیمار i ام در بلوک j ام در طرح d ظاهر می‌شود آنگاه:

$$r_{di} = \sum^{b} n_{dij}$$

r و v تیمار است و طریقه قرار گرفتن تیمارها در واحدهای آزمایشی طوری است که :

(i) $n_{ij}=0$ یا 1 $(1 \leq j \leq b\ ,\ 1 \leq i \leq v)$ یعنی دوتائی[1] است.

(ii) هر تیمار r بار تکرار می‌شود.

(iii) هر دو تیمار در یک طرح با هم λ بار در بلوکهای مشترک ظاهر می‌شوند.

توجه داشته باشید که به ازای هر مقدار دلخواه از پارامترها ممکن است چنین طرحی وجود نداشته باشد. بین پارامترهای این طرح‌ها یک سری روابط ریاضی وجود دارد که به این صورت بیان می‌شوند.

(i) $rv=bk=n$

(ii) $\lambda(v-1)=r(k-1)$

(iii) $b \geq v$

علت ناقص بودن این طرح‌ها این است که $v \geq k$ و همچنین علت متعادل بودن آنها هم این است که هر تیمار r بار و هر زوج تیمار λ بار در بلوکهای مشترک ظاهر می‌شوند.

تعریف ۲-۳) طرح $d\in C(v\ ,\ b\ ,\ k)$ را طرح BTB[2] (طرح بلوک تیمارهای متعادل) $(k>v)$ گوئیم اگر روابط زیر برقرار باشد.

$$\lambda_{01}= \lambda_{02} ,\ldots, \lambda_{0v} = \lambda_0 \qquad (۳-۱)$$

$$\lambda_{12}= \lambda_{13},\ldots, \lambda_{(v-1)(v)} = \lambda_1 \qquad (\lambda_{ij} = \sum_{l=1}^{b} n_{il}n_{jl})$$

تعریف ۳-۳) $(Kiefer\ 1975)$ یک طرح $d\in C(v,b,k)$ را طرح BBD[3] (طرح بلوکهای متعادل) گوئیم اگر تمام r_{di} ها $(1\leq i\leq v)$ مساوی باشند و برای $i<j$، λ_{ij} ها برابر و همچنین برای هر i $|n_{ij} - \frac{k}{v}| \leq 1$ باشد.

شباهت این طرح‌ها با طرح‌های BIB در این است که در آن طرح‌ها هم تعداد تکرار هر تیمار در طرح یکسان است.

حال در موقعیتی هستیم که قضیه مورد نظر را بیان کنیم.

$[(kb-r_0)/v]$ یا $[(kb-r_0)/v)] + 1$ $\quad 1 \leq i \leq v \quad (i)$

: $[r_i/b]$ یا $[r_i/b] + 1$ $\quad 1 \leq i \leq v, \quad 1 \leq j \leq b \quad (ii)$

$= [r_0/b]$ یا $[r_0/b] + 1$ $\quad 1 \leq j \leq b \quad (iii)$

(iv) اگر r_0 یک عدد صحیح و مثبتی باشد $(1 \leq r_0 \leq \frac{bk}{2})$ که $F_A(r)$ را که در زیر تعریف می‌کنیم می
کند آنگاه d یک طرح بهین نوع A در کلاس طرح‌های $C(v,b,k)$ می‌باشد.

$$) = kv[kv-h(r)]^{-1} + akv\ [c-pr - h(r)]^{-1}$$

که

$$: bkv\ (k-1) + bv\alpha(v - 2k + v\alpha)$$

$$[(bk-r)/vb] \qquad a = (v-1)^2$$

$$v(k-1)+k-2v\alpha$$

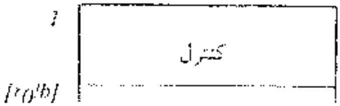
$$= b([r/b])^2 + 2([r/b](r-b[r/b]) + (r - b[r/b])$$

۴ - روش های پیدا کردن طرح بلوکهای کامل متعادل بهین نوع A

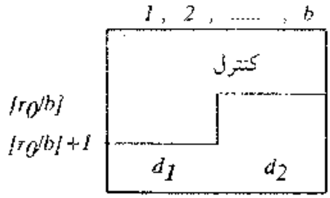
طرح‌های BTB که در شرایط قضیه (۱-۳) صدق کنند به دو نوع تقسیم می‌شوند:

I) تیپ R *(Rectangular)*

طرح‌های تیپ R (شکل ۱-۳) برای زمانی است که $r_0 = O(mod\ b)$ در این طرح‌ها هر بلوک شامل تکرار از تیمار کنترل است و قسمت باقی مانده بلوکها که آن را d_0 می‌نامیم یک طرح $BD(v,b,k-r_0/b)$ تشکیل می‌دهد.



طرح تیپ S(شکل ۲-۳) برای زمانی است که $r_0 \neq O(mod\ b)$



شکل ۲-۳)

قسمت d_1 یک طرح $BBD\ (v\ ,\ r_0 - b\ [r_0\ /\ b\ ]\ ,\ k - [\ r_0\ /\ b\ ]\ -1)$ و قسمت d_2 یک طرح $BBD\ (v,b-(r_0-b[r_0/b],\ k-[r_0/b])$ می‌باشد. توجه داشته باشید که در دو قسمت $d_2 d_1$ تیمارهای آزمایشی به تعداد یکسان تکرار می‌شوند. (توجه لزومی ندارد که هر دو طرح $d_2 d_1$ طرح‌های BBD باشند. اگر $d_1 \cup d_2$ هم طرح BBD باشد شرایط لازم برقرار است)

در سال *Notz and Ting 1988* برای پیدا کردن طرح های بهین الگوریتم زیر را پیشنهاد کردند.

گام اول :

ابتدا مقدار r_0 را چنان پیدا می‌کنیم که $F_A(r)$ را می‌نیمم کند.

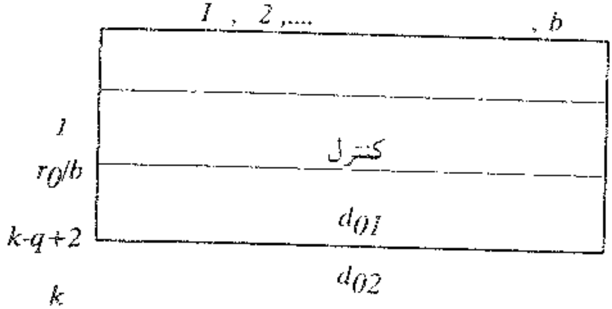
گام دوم :

حالت I) $r_0 = O(mod\ b)$ آنگاه طرح بهین نوع A از تیپ R است و باید شرایط لازم برای BBD بودن قسمت d_0 بررسی شود یعنی باید کنترل کنیم که کمیت های زیر مقادیر صحیحی باشند.

$$q_1 = (kb - r_0)/v \qquad (i)$$

$$q_2 = q_1\ (q-1)/(v-1) \qquad (ii) \qquad (۳-۲$$

(ii) اگر $(kb-r_0)/bv$ عدد صحیحی نباشد آنگاه یک طرح BBD که هر تیمار آزمایشی در هر **بلوک به تعد**
$[(kb-r_0)/bv]$ تکرار شود را به عنوان بهترین d_{01} انتخاب می‌کنیم (شکل ۳-۳) و یک $BIB(v,b,q)$ را به عنوا
d_{02} انتخاب می‌کنیم.



شکل (۳-۳)

حالت II) $r_0 \neq 0 (mod\ b)$

در این حالت طرح بهین نوع A از تیپ S است که می‌توان آن را به دو طرح تیپ R تقسیم کرد که اگر یکی $r_0^* = (r_0-b[r_0/b])([r_0/b]+1)$ و $b^* = r_0-b\ [r_0/b]$ می‌باشد و در دیگری $r_0^{**} = r_0 - r_0^*$ ، $b^{**} = b-b^*$ می‌باشد که برای بررسی این طرح‌های BBD باید فرآیند حالت (I) را تکرار کنیم.

ذیلاً دو مثال برای طرح‌های بهین نوع A تیپ R , S که با استفاده از این الگوریتم به دست **آمده‌اند را بی**
می‌کنیم.

مثال ۱-۳) اگر $v=4$, $b=12$, $k=6$ باشد آنگاه طرح زیر یک طرح بهین نوع A از تیپ R می‌باشد.

$$\begin{array}{cccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\
3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\
4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4
\end{array}$$

طرح (۱-۳)

مثال ۲-۳) اگر $v=4$, $b=18$, $k=7$ باشد طرح زیر یک طرح بهین نوع A از تیپ S می‌باشد.

$$\begin{array}{cccccccccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

در سال $Gerami\ 1991$ با استـفاده از قضـیه ۱-۳ (قضیه 4 $Ting\ and\ Notz\ 1988$) الگوریتم پیدا کردن طرح‌های بهین را توسعه داد قبـل از بیان این موضوع یک تعریف وجود دارد که در زیر بیان می‌کنیم.

تعریف ۴-۳ $(John\ and\ Mitchell\ 1977)$

برای مقادیر k,b,v معلوم یک طرح d را RGD[1] (طرح نمودار منظم) می‌نامیم اگر :

$$n_{ij} = [\frac{k}{v}] \text{ یا } [\frac{k}{v}] + 1 \qquad (i)$$

r_i ها همه با هم برابر باشند $\qquad (ii)$

$$j \neq j' \,,\ i \neq i' \text{ برای هر } |\lambda_{ii'} - \lambda_{jj'}| \leq 1 \qquad (iii)$$

در سال $Gerami\ 1991$ روش پیدا کردن طرح‌های بهین نوع A در کلاس طرح بلوکهای کامل را چـنین بیان کردند.

طرح‌های BIB که در شرایط قضیه (۲-۳) صادق هستند به دو گروه تقسیم می‌شوند

I) طرح‌های تیپ R $(Rectangular)$ شکل (۱-۳)

این طرح‌ها برای زمانی است که $r_0 = O\ mod\ (b)$ باشد. در این طرح ها در هر بلوک تیمار کنترل $\frac{r_0}{b}$ دفعه ظاهر می‌شود و قسمت مابقی طرح را که d_0 می‌نامیم یک طرح $BBD\ (v,b,\ k-\frac{r_0}{b})$

II) طرح‌های تیپ S $(Step)$ شکل (۲-۳)

این طرح‌ها برای زمانی است که $r_0 \neq O\ mod(b)$ باشد. پس در این حالت :

d_1 یک طرح $BBD(v,r_0-b[r_0/b]\ ,\ k-[r_0/b]-1)$

$d2$ یک طرح $BBD(v,b-r_0+b\ [r_0/b]\ ,\ k-[r_0/b]$

یا این که

d_1 یک طرح $RGD(v,r_0-b[r_0/b]\ ,\ k-[r_0/b]-1\ ,\ n_1\ ,\ n_2\ ,\ \lambda_1\ ,\ \lambda_2)$

d_2 یک طرح $RGD(v,b-r_0+b\ [r_0/b]\ ,\ k-[r_0/b]\ ,\ n_1\ ,\ n_2\ ,\ \lambda_3\ ,\ \lambda_4)$

گام اول :

مقدار r_0 را چنان پیدا می‌کنیم که $F_A(r)$ را روی L می‌نیمم کند.

گام دوم :

حالت $I)$ $r_0 = O\ mod(b)$ طرح بهین نوع A از تیپ R است و ما باید شرایط لازم برای BBD بود طرح d_0 را بررسی کنیم. در اصل باید بررسی کنیم که آیا کمیت های زیر مقادیر صحیحی هستند.

$)\ r = (bk\text{-}r_0)/\ v$

$i)\ \lambda = r(k_1\text{-}1)/v\text{-}1$ (۳-۲)

که در آن $_1 = k\text{-}(r_0/b)\text{-}v[(bk\text{-}r_0)/vb]$

اگر کمیت‌های (۳-۲) مقادیر صحیحی نباشند آنگاه قضیه (۲-۳) نمی‌تواند مورد استفاده قرار بگیرد اما ا
کمیت های r , λ صحیح باشند آنگاه ساختن چنین طرح BBD به دو گروه تقسیم میشود.

(i) اگر $(bk\text{-}r_0)/vb$ عدد صحیح باشد آنگاه یک طرح BBD که در هر بلوک آن تیمارهای آزمایشی $\frac{b\text{-}r_0}{vb}$ دفعه ظاهر می‌شود. بهترین طرح d_0 است.

(ii) اگر $\frac{bk\text{-}r_0}{vb}$ عدد صحیح نباشد آنگاه یک طرح BBD که تیمارهای آزمایشی $[(bk\text{-}r_0)/vb]$ دفعه هر بلوک ظاهر می‌شوند را به عنوان بهترین طرح ممکن d_{01} انتخاب می‌کنیم و یک طرح $BIB(v,b,k_1)$ را عنوان بهترین طرح d_{02} انتخاب می‌کنیم (شکل ۳-۲)

حالت $II)$ $r_0 \neq O\ mod\ (b)$

در این حالت طرح بهین نوع A از تیپ S می‌باشد. این طرح به صورت ترکیبی از دو طرح در دو حال ساخته می‌شود.

(i) از دو طرح تیپ R که در یکی $r_0^* = b^*([r_0/b]\ +1\ ),\ b_0^* = r_0\ \text{-}b\ [r_0/b]$ و دیگری $r_0^{**} = b^{**}[r_0/b],\ b^{**} = b\text{-}b^*$ می‌باشد. برای بررسی وجود چنین طرح‌های BBD باید فرایند حالت (I) تکرار کنیم که این کار سختی است.

یا اینکه :

(ii) d_1 شامل b^* بلوک است که هر تیمار آزمایشی $[(k\text{-}1\text{-}[r_0/b])/v]$ دفعه در هر بلوک تکرار می‌شو

... در اولین RGD ظاهر شوند آنگاه آنها با هم در دومین RGD λ_3 دفعه ظاهر می‌شوند و اگر آنها λ_2 دفعه در بلوکهای یکسان در اولین RGD ظاهر شوند آنگاه آنها باید در λ_4 دفعه در بلوکهای یکسان در دومین RGD ظاهر می‌شوند و همچنین شرایط زیر باید برقرار باشد.

$$r_1^* = b^* k^* / v \quad \text{باید عدد صحیحی باشد.} \qquad (i)$$

$$r_2^* = b^{**} k^{**} / v \quad \text{باید عدد صحیحی باشد.} \qquad (ii)$$

$$n + n_2 = v - 1 \qquad (iii)$$

$$r_1^* (k^* - 1) = n_1 \lambda_1 + n_2 \lambda_2 \qquad (iv)$$

$$r_1^{**} (k^{**} - 1) = n_1 \lambda_3 + n_2 \lambda_4 \qquad (v) \qquad (۳-۴)$$

$$\lambda_1 + \lambda_3 = \lambda_2 + \lambda_4 \qquad (vi)$$

برای بررسی وجود یک چنین RGD به ازای مقادیر v, b, k, r_0 باید تمام شرایط (۳-۴) برقرار باشد که یک کار سختی است.

مثال ۳-۴) برای مقادیر $v=6$, $b=7$, $k=12$ مقدار $r_0=24$ به دست می‌آید. طرحی که از این مقدار r_0 به دست می‌آید از تیپ S است که در d_1 هر تیمار آزمایشی در هر بلوک یک بار تکرار می‌شود. بعلاوه یک RGD $(6,3,2,1,4,1,0)$ که در آن هر تیمار با یک تیمار دیگر $\lambda_1=1$ دفعه دربلوکهای یکسان ظاهر می‌شود و با چهار تیمار دیگر $\lambda_2=0$ دفعه در بلوکهای یکسان ظاهر می‌شود این RGD در زیر با $S_{6,3,2}$ نمایش داده شده

$$S\ 6,3,2: \quad \begin{matrix} 1 & 2 & 5 \\ 4 & 5 & 6 \end{matrix}$$

و در d_2 هر تیمار آزمایشی در هر بلوک یکبار تکرار می‌شود. بعلاوه یک $RGD(6,4,3,1,4,0,1)$ که در آن هر تیمار با یک تیمار دیگر $\lambda_3=0$ دفعه در بلوکهای یکسان ظاهر می‌شوند. و هر تیمار با چهار تیمار دیگر $\lambda_4=1$ دفعه در بلوکهای یکسان ظاهر می‌شود. این RGD در زیر با $S_{6,4,3}$ نمایش داده شده و همچنین طرح (۳-۳) طرح کامل مثال فوق است.

$$S\ 6,4,3: \quad \begin{matrix} 1 & 1 & 2 & 3 \\ 2 & 5 & 4 & 4 \\ 3 & 6 & 6 & 5 \end{matrix}$$

0	0	0	1	1	1	1	
1	1	1	2	2	2	2	
2	2	2	3	3	3	3	طرح (۳-۳)
3	3	3	4	4	4	4	
4	4	4	5	5	5	5	
5	5	5	6	6	6	6	
6	6	6	1	1	2	3	
1	2	3	2	5	4	4	
4	5	6	3	6	6	5	

۴- طرح های نزدیک به طرح‌های بهین نوع A

در کلاس طرح های $C(v,b,k)$ برای بعضی از مقادیر پارامتر با استفاده از قضیه (۱-۳) نمی‌توانیم طرح‌های بهین پیدا کنیم. یعنی در اصل یک طرح BTB متناظر با آن پارامترها و مقدار r_0 وجود ندارد.

اما از آن جایی که $F(r_0) = min\ F(r)$ برای هر $r \in L = \{r: 1 \leq r \leq \frac{bk}{2}\}$ می‌باشد.

پس $F(r_0)$ برای هر $d \in C(v,b,k)$ یک کران پایین $\sigma^{-2} \sum Var(\hat{t}_{d0}-\hat{t}_{di})$ ارائه می‌دهد که مقدار

عبارت $AE_d = \dfrac{min\ \{\sum var\ (\hat{t}_{d0} - \hat{t}_{di})\ r_d \in D\}}{\sum Var(\hat{t}_{d0}-\hat{t}_{di})}$ به عنوان کارایی

متناظر با معیار A تعریف می‌کنیم. واضح است، طرحی که بیشترین مقدار AE_d را دارا باشد کاراترین طرح است و آن را به عنوان طرح نزدیک به طرح بهین می‌توان انتخاب کرد.

پس وقتی که BTB متناظر با r_0 وجود نداشته باشد با استفاده از قضیه (۱-۳) می‌توان طرح‌های نزدیک به طرح‌های بهین را پیدا کرد از این قضیه می‌توان نتیجه گرفت که یک طرح d^* که BTB باشد و مقدار r^*_{d0} هم به مقدار r_0 نزدیک باشد، آنگاه این طرح دارای کارائی بالایی است.

مثال ۱-۴) فرض کنید $v=4$, $k=8$, $b=18$ باشد. آنگاه برای این پارامترها مقدار $r_0=51$ می‌باشد و $F(r_0) = 0.255019$ ولی برای $r_0=51$ طرح BTB وجود ندارد، اما برای مقدار $r^*_0=48$ یک طرح d^* که BTB می‌باشد وجود دارد و $F(r^*_0) = 0.2553263$ می‌باشد پس کارآئی این طرح برابر با :

$$\frac{0.255019}{0.2553264} = 0.9987$$

[illegible] d^* [illegible] $C(4,18,8)$ انتخاب می‌کنیم.

طرح (۱-۴)

1	1	1	2	2	2	3	3	3	4	4	4	2	3	4	3	4	4
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

مثال ۲-۴) فرض کنید که $v=9$, $k=15$, $b=24$ باشد. آنگاه برای این پارامترها $r_0=94$ می‌باشد و $F(r_0)=$ 0.403433 برای $r_0=94$ طرح BTB وجود ندارد ولی برای $r_0^*=90$ یک طرح BTB وجود دارد.
اگر طرح متناظر با مقدار $r_0^*=90$ را d_0^* بنامیم آنگاه $\frac{F(r_0)}{F(r_0^*)}=0.9996$ می‌باشد پس طرح d^* را به عنوان طرح نزدیک به طرح بهین در کلاس $C(10,24,15)$ انتخاب می‌کنیم.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6.	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	3	4	5
1	1	1	1	2	2	2	2	3	3	3	4	4	5	6	6	7	7	2	6	3	4	5	8
3	4	5	8	4	5	7	8	5	7	9	6	9	6	8	9	9	8	9	7	6	8	7	9

طرح (۲-۴)

در سال *Gerami 1991* یک الگوریتم برای پیدا کردن طرح‌های نزدیک به طرح های بهین ارائه کردند.

گـــام اول : مـــقـــــدار r_1 نـزدیک بـــه r_0 را طـــــوری پـــیـــــدا مـی‌کنیم کـــه روی تـــمام مـــقادیر $r \in L_1 = \{r: 1 \leq r \leq \frac{bk}{2},\ r \neq r_0\}$ مقدار $F(r_1)$ می‌نیمم باشد.

گام دوم : گام اول را انجام می‌دهیم و باید $r \neq r_0 r_1$ اگر شرطهای مورد نیاز آن گام برقرار باشد که کار تمام

$$1 \leq i \leq v$$

آنگاه واضح است که ME_d کارایی طرح d نسبت به معیار بهینگی MV می‌باشد. و هر طرحی که ME_d آن بیشتر باشد کاراترین طرح است و ما آن را به عنوان طرح نزدیک به طرح بهین نوع MV انتخاب می‌کنیم در هر دو مثال (۷-۳) و (۸-۳) هر دو طرح به دست آمده علاوه بر آنکه طرح نزدیک **به طرح** بهین معیار A بودند. طرح نزدیک به طرح بهین نوع MV هم بودند.

برای پیدا کردن طرح های نزدیک به طرح‌های بهین الگوریتم زیر را پیشنهاد می‌کنیم.

گـــــام اول: ابتـــدا مقـــدار r_0^* راکـه بـه r_0 نـزدیک باشـــد و هـــم چنـــین $F(r)$ را روی $r \in L = \{r: 1 \leq r \leq \frac{bk}{2}\}$ می‌نیمم کند و $F(r_0)/F(r_0^*) > 0.999$ باشد را پیدا می‌کنیم.

گام دوم: شرایط لازم را در الگوریتم پیدا کردن طرح‌های بهین را بررسی می‌کنیم اگر شرایط برقرار بود که جستجو تمام می‌شود و طرح متناظر با این r_0^* طرح نزدیک به طرح بهین است اگر نه دوباره گام اول را تکرار می‌کنیم.

عادل محمدپور ، جواد بهبودیان

دانشگاه شیراز

چکیده: یکی از آزمون‌هایی که در آمار کاربرد فراوان دارد آزمون تی می‌باشد. در صورتی که نمونه تصادفی از توزیع نرمال آمده باشد استنباطها در مورد توزیع بخوبی انجام خواهد شد. اگر برخی از این فرضها را از دست بدهیم آیا تاثیری در آزمون خواهد داشت یا خیر؟ (اگر تاثیری نداشته باشد می‌گوئیم آزمون تی پرزور است.) هدف اصلی مقاله پاسخ به پرسش بالا میباشد. دراین مقاله ضمن بررسی کارهای انجام شده به محدودیت های موجود اشاره کرده و با شبیه سازی آنها را بصورت ملموس‌تری نشان خواهیم داد. برای این منظور سه فرض استقلال، همتوزیعی، یا نرمال بودن را حذف کرده مسئله را بررسی می‌کنیم.

۱ مقدمه

یکی از آزمون‌هایی که در آمار کاربرد فراوان دارد آزمون تی می‌باشد. این آزمون بر اساس سه فرض استقلال، همتوزیعی، و نرمال بودن بنا نهاده شده است. این آزمون پرتوانترین آزمون بطور یکنواخت در کلاس آزمون‌های نااریب و پایا (UMPI, UMPU) برای فرض‌های دوسوئی در آمار کلاسیک میباشد. در عمل برسی فرضهای اولیه برای این آزمون عملاً امکان پذیرنیست، چرا که آزمونی پرتوان برای انجام این مهم نداریم و آزمون‌های ناپارامتری موجود نیز تنها برای رد کردن این فرضها کارامی باشند [۱،۲،۵،۶]. با توجه به این نکته برای ما اهمیت زیادی دارد که بدانیم تحت چه شرایطی این آزمون خوب عمل میکند و تحت چه شرایطی به نتیجه آزمون نمیتوان اعتماد کرد. برای این منظور ما یک (یاچند) فرض از فرض‌های اولیه را برداشته و با شبیه‌سازی چنین داده‌هایی و با انجام آزمون تی سعی به پاسخ دادن پرسش بالا می‌کنیم.

۲ شبیه‌سازی

دراین قسمت نحوه کار به این صورت است که یک یاچند فرض از فرض های اولیه را برداشته و [illegible]

را در سطح $\alpha = ۰/۰۵$ با تابع آزمون

$$\phi(x) = \begin{cases} ۱ & |\frac{\bar{X}-۰}{S/\sqrt{n}}| > t(n-۱; ۱-\frac{۰/۰۵}{۲}) \\ ۰ & |\frac{\bar{X}-۰}{S/\sqrt{n}}| < t(n-۱; ۱-\frac{۰/۰۵}{۲}), \end{cases}$$

انجام می‌دهیم، که در آن n، $\bar{X}$، و $S^۲$ بترتیب حجم نمونه، میانگین نمونه، و واریانس نمونه می باشند. این عمل را ۱۰۰۰ بار تکرار کرده و درصد دفعات رد فرض صفر را محاسبه می‌کنیم. اگر این تعداد حدود پنج درصد دفعات بود می‌گوئیم: آزمون تی نسبت به این توزیع بخصوص و حجم نمونه پرزور می‌باشد.

جدول پیوست نتیجه شبیه سازی را نشان میدهد. ما از علامتهای $+, \pm, \mp, -, =$ بترتیب جهت پرزور بودن و در نهایت جهت حساس بودن آزمون استفاده کرده‌ایم. سه ستون اول نشاندهنده فرض‌هایی است که باآنها داده‌ها را تولید کرده‌ایم. برای مثال سطر اول بیان کننده این مطلب می باشد که داده‌ها دارای سه فرض استقلال، همتوزیعی، و نرمال بودن میباشند. بررسی همه توزیع‌ها از توان ما خارج است. ما تنها برخی از توزیع‌ها را بررسی می کنیم. جهت مقایسه بهتر، انحراف‌معیار (مقیاس) توزیع های انتخاب شده را یکسان بوده، و علاوه بر آزمون تی آزمون ناپارامتری ویلکاکسون [۳] نیز آورده شده است.

برای خلاصه شدن جدول از علائم زیر استفاده شده است.

i: فرض استقلال

id: فرض همتوزیعی

n: فرض نرمال بودن

$N(\mu, \sigma^۲)$: توزیع نرمال با میانگین μ و واریانس $\sigma^۲$

$EN(\mu, \sigma^۲, \rho)$: توزیع نرمال جابجاپذیر با میانگین μ، واریانس $\sigma^۲$، و ضریب همبستگی ρ [۷]

$C(\mu, \sigma)$: توزیع کُشی با مبدأ μ و مقیاس σ

$EC(\mu, \sigma, \rho)$: توزیع کُشی جابجاپذیر با مبدأ μ، مقیاس σ، و ضریب هموردایی ρ [۴]

$E(\lambda)$: توزیع نمایی با میانگین $۱/\lambda$

$U(a, b)$: توزیع یکنواخت در فاصله a و b

$B(p, n)$: توزیع دوجمله‌ایی با احتمال موفقیت p و تعداد n آزمایش

$G(p)$: توزیع هندسی با احتمال موفقیت p

$P(\lambda)$: توزیع پواسن با میانگین λ

$a - b$: پارامتر در فاصله $[a, b]$، با توجه به حجم نمونه تغییر می‌کند. برای

nal of the American Statistical Association, 74, 194-199.

[2] Chen, G. and Adatia A. (1997). Independence and t Distribution. *The American Statistican* , 51, 176-177. ¿

[3] Conver, W. J. (1980). *Practical Nonparametric Statistics.* 2nd ed. Wiley, New York.

[4] Modarres, R. and Nolan, J. P. (1994). A Method for Simulating Stable Random Vectors. *Computational Statistics*, 9, 11-19.

[5] Mohammadpour, A. and Behboodian J. (1996). Exchangeable and IID Normal Data. *Third Iranian Statistics Conference*, 129-143.

[6] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[7] Tong, Y. L. (1990). *The Multivariate Normal Distribution.* Springer-Verlag, New York.

ssumption(s)				$n=2$		$n=5$		$n=10$		$n=25$		$n=50$	
a	i	id	Distribution	$t-$	W	$t-$	W	$t-$	W	$t-$	W	$t-$	W
	*	*	$N(0,.01)$	+	+	+	+	+	+	+	+	+	+
	*	*	$N(0,1)$	+	+	+	+	+	+	+	+	+	+
	*	*	$N(0,100)$	+	+	+	+	+	+	+	+	+	+
	*	*	$C(0,0.1)$	+	+	+	+	+	+	+	+	+	+
	*	*	$C(0,1)$	+	+	+	+	+	+	+	+	+	+
	*	*	$C(0,10)$	+	+	+	+	+	+	+	+	+	+
	*	*	$E(100)$	±	+	∓	+	∓	±	±	∓	+	−
	*	*	$E(1)$	±	+	∓	+	∓	±	±	∓	+	−
	*	*	$E(0.01)$	±	+	∓	+	∓	±	±	∓	+	−
	*	*	$U(-.17,.17)$	+	+	+	+	+	+	+	+	+	+
	*	*	$U(-1.7,1.7)$	+	−	+	+	+	+	+	+	+	+
	*	*	$U(-17,17)$	+	+	+	+	+	+	+	+	+	+
	*	*	$B(1,0.01)$			=	=	=	=	=	=	=	=
	*	*	$B(4,0.5)$			+	+	+	+	+	+	+	+
	*	*	$B(400,0.5)$			+	+	+	+	+	+	+	+
	*	*	$G(0.099)$	=	+	=	=	=	=	=	=	=	=
	*	*	$G(0.62)$	=	+	±	±	±	±	±	−	+	=
	*	*	$G(0.95)$	±	+	±	+	±	±	±	∓	+	−
	*	*	$P(0.01)$			=	=	=	=	=	=	=	=
	*	*	$P(1)$			±	+	+	+	+	+	+	+
	*	*	$P(100)$			+	+	+	+	+	+	+	+

*		*	$EN(0,1,0.1)$	+	+	±	+	∓	∓	−	−	=	=
*		*	$EN(0,1,0.2)$	+	+	∓	+	−	−	=	=	=	=
*		*	$EN(0,1,0.3)$	+	+	−	+	−	−	=	=	=	=
*		*	$EN(0,1,0.4)$	±	+	−	+	=	=	=	=	=	=
*		*	$EN(0,1,0.5)$	±	+	−	+	=	=	=	=	=	=
*		*	$EN(0,1,0.6)$	±	+	=	+	=	=	=	=	=	=
*		*	$EN(0,1,0.7)$	∓	+	=	+	=	=	=	=	=	=
*		*	$EN(0,1,0.8)$	∓	+	=	+	=	=	=	=	=	=
*		*	$EN(0,1,0.9)$	∓	+	=	+	=	=	=	=	=	=
		*	$EC(0,1,0.1)$	+	+	±	+	±	∓	∓	−	=	=
		*	$EC(0,1,0.2)$	+	+	∓	+	∓	−	∓	=	=	=
		*	$EC(0,1,0.3)$	±	+	∓	+	−	−	−	=	=	=
		*	$EC(0,1,0.4)$	±	+	−	+	=	−	−	=	=	=
		*	$EC(0,1,0.5)$	±	+	−	+	=	−	−	=	=	=
		*	$EC(0,1,0.6)$	±	+	=	+	=	−	=	=	=	=
		*	$EC(0,1,0.7)$	±	+	=	+	=	−	=	=	=	=
		*	$EC(0,1,0.8)$	±	+	=	+	=	=	=	=	=	=
		*	$EC(0,1,0.9)$	∓	+	=	+	=	=	=	=	=	=
*	*		$N(0,10^{-4}-1)$	+	+	+	+	+	+	±	±	+	+
*	*		$N(0,1-100)$	+	+	+	+	+	+	+	+	+	+
*	*		$N(0,10-10^4)$	+	+	+	+	+	+	+	+	+	+
	*		$C(0,0.01-1)$	+	+	+	+	+	+	+	+	+	+
	*		$C(0,1-10)$	+	+	+	+	+	+	+	+	+	+
	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

لیلا محمدی

دانشگاه صنعتی اصفهان

۱ مقدمه

خانواده توزیعهای کامل در آمار از اهمیت ویژه‌ای برخوردار هستند . خانواده دیگری از توزیعه اهمیت خاصی دارد ، خانواده توزیعهای کامل کراندار[۱] است . برای خانواده توزیعهای کامل برا می‌توان کلاس $UMVUE$ ها را مشخص کرد . اما برای کلاس توزیعهای کامل کراندار لزومی ن که برای پارامترهای برآورد پذیر کلاسی از $UMVUE$ ها وجود داشته باشد . در این مقاله ما روشر برای تولید خانواده توزیعهای کامل کراندار ارائه می‌دهیم . پس از آن کلاس برآورد گرهای ناریب و وکلاس $MVUE$ ها را تعیین می‌کنیم . در پایان با ارائه یک مثال ، کاربرد مطالب گفته شده را ن می‌دهیم .

۲ خانواده توزیعهای کامل کراندار

خانواده توزیعهای X کامل کراندار[۲] گفته می‌شود اگر و فقط اگر برای هر تابع کراندار $T(X)$ دا باشیم :

$$T(X)) = 0 \quad , \quad \forall\theta \in \Theta \Longrightarrow \quad P_\theta(T(X) = 0) = 1 \quad , \forall\theta \in \Theta.$$

بنابراین هر آماره کامل یک آماره کامل کراندار خواهد بود . اما دراین مقاله نشان خواهیم داد عک این مطلب صحیح نیست . اگر یک خانواده از توزیعها نه کامل باشد و نه کامل کراندار ، آنگاه یک متعلق به دامنه X و یک تابع کراندار $T(X)$ پیدا خواهد شد که $T(x_0) \neq 0$ و برای هر $\theta \in \Theta$ دا $E_\theta(T(X)) = 0$. برای خانواده توزیعهای کامل کراندار که در قضیه ۱ معرفی می‌شوند می‌توانیم کلا برآوردگرهای ناریب صفر را تعیین کنیم . با استفاده از این کلاس و قضیه $Lehmann-Sheffeh$ می‌توان کلاس $MVUE$ ها را تعیین کرد . همچنین با تعیین این کلاس می‌توان فرم توابع $\gamma(\theta)$ را برای آنها تحت خانواده توزیعهای کامل کراندار ، $MVUE$ وجود دارد مشخص کرد . توضیح اینکه

... همچنین فرض کنید $\varphi : [a,b] \longrightarrow [a,b]$ یک تابع حقیقی اکیداً صعودی و مشتق پذیر باشد که $\varphi(a) = a$ و $\varphi(b) = b$. اگر p یک مقدار ثابت متعلق به فاصله $(\frac{1}{2}, 1)$ باشد و $f_\theta(x)$ به صورت زیر تعریف شود ،

$$f_\theta(x) = p g_\theta(x) + (1-p)\frac{\partial \varphi^{-1}(x)}{\partial x} g_\theta(\varphi^{-1}(x)) \qquad (1)$$

آنگاه خانواده چگالیهای $\{f_\theta(x) : \theta \in \Theta, x \in (a,b)\}$ کامل کراندار است .

اثبات : چون

$$h_\theta(x) = \frac{\partial \varphi^{-1}(x)}{\partial x}\, g_\theta(\varphi^{-1}(x))$$

یک تابع چگالی است و $f_\theta(x)$ ترکیب محدبی از توابع چگالی $g_\theta(x)$ و $h_\theta(x)$ است بنابراین $f_\theta(x)$ برای هر $\theta \in \Theta$ روی فاصله (a,b) یک تابع چگالی می‌باشد . حال فرض کنید $E_\theta^\xi(T(X))$ نمایانگر امید ریاضی آماره $T(X)$ نسبت به چگالی ξ با پارامتر θ باشد . اگر

$$E_\theta^f(T(X)) = 0 \ , \ \forall \theta \in \Theta$$

می‌خواهیم ثابت کنیم

$$T(x) = 0 \ , \ \forall x \in (a,b)$$

یا $T(x)$ یک تابع بی‌کران از x روی (a,b) می‌باشد . فرض کنید $x_0 \in (a,b)$ وجود دارد به طوری که $T(x_0) \neq 0$. دنباله $\{c_n\} \subset (a,b)$ را چنان پیدا خواهیم کرد که وقتی $n \to \infty$ آنگاه $|T(c_n)| \to +\infty$.
از فرض $E_\theta^f(T(X)) = 0$ نتیجه می‌شود که برای هر $\theta \in \Theta$

$$\begin{aligned}
E_\theta^f(T(X)) &= \int_a^b T(x) f_\theta(x) dx \\
&= \int_a^b T(x)(p g_\theta(x) + (1-p)\frac{\partial}{\partial x}\varphi^{-1}(x) g_\theta(\varphi^{-1}(x)))dx \\
&= p\int_a^b T(x) g_\theta(x) dx + (1-p)\int_a^b T(x)\frac{\partial}{\partial x}\varphi^{-1}(x) g_\theta(\varphi^{-1}(x))dx \\
&= p\int_a^b T(x) g_\theta(x) dx + (1-p)\int_a^b T(\varphi(x)) g_\theta(x) dx \\
&= \int_a^b (pT(x) + (1-p)T(\varphi(x)))g_\theta(x)dx
\end{aligned}$$

$$pT(x) + (۱ - p)T(\varphi(x)) = ۰ \ , \ \forall x \in (a,b)$$

یعنی :

$$T(\varphi(x)) = -\frac{p}{۱-p}T(x) \ , \ \forall x \in (a,b) \qquad (۳)$$

چون $x_۰ \in (a,b)$ ، پس (۳) نتیجه می‌دهد که

$$\begin{aligned} T(\varphi(x_۰)) &= -\frac{p}{۱-p}T(x_۰) \\ T(\varphi(\varphi(x_۰))) &= (-\frac{p}{۱-p})^۲ T(x_۰) \\ T(\varphi(\varphi(\varphi(x_۰)))) &= (-\frac{p}{۱-p})^۳ T(x_۰) \\ &\vdots \\ T(\varphi_n(x_۰)) &= (-\frac{p}{۱-p})^n T(x_۰) \end{aligned}$$

که در آن برای هر $n \in N$

$$\varphi_n(x_۰) = \underbrace{\varphi(\varphi(...(\varphi}_{n}(x_۰))...))$$

فرض کنید $\mathbf{c_n} = \varphi_n(x_۰)$ ، **آنگاه**

$$\begin{aligned} \lim_{n\to\infty} \mid T(c_n) \mid &= \lim_{n\to\infty} \mid T(\varphi_n(x_۰)) \mid \\ &= \lim_{n\to\infty} \mid (-\frac{p}{۱-p})^n T(x_۰) \mid \\ &= \mid T(x_۰) \mid \lim_{n\to\infty} (\frac{p}{۱-p})^n \\ &= +\infty \end{aligned}$$

چون $\frac{۱}{۲} < p < ۱$ و در نتیجه $\frac{p}{۱-p} > ۱$ و از طرفی $\mid T(x_۰) \mid > ۰$. بنابراین $T(x)$ یک تابع بی‌کرا بوده و قضیه ثابت است .□

در رابطه با قضیه ۱ تذکر چند نکته در اینجا ضروری است :

۲—در قضیه۱ اگر $\varphi(x) = x$ آنگاه $f_\theta(x) = g_\theta(x)$. یعنی $f_\theta(x)$ کامل خواهد بود .

۳—در قضیه۱ اگر $p \in \{0,1\}$ آنگاه خانواده چگالیهای $f_\theta(x)$ کامل خواهد بود

۴—در قضیه۱ اگر $p = \frac{1}{2}$ آنگاه خانواده چگالیهای $f_\theta(x)$ نه کامل خواهد بود و نه کامل کراندار .

۵—در قضیه۱ اگر φ یک تابع اکیدا نزولی مشتق پذیر باشد طوری که $\varphi(a) = b$ و $\varphi(b) = a$ و تعریف کنیم

$$f_\theta(x) = pg_\theta(x) - (1-p)\frac{\partial\varphi^{-1}(x)}{\partial x}g_\theta(\varphi^{-1}(x))$$

آنگاه خانواده چگالیهای $\{f_\theta(x) : \theta \in \Theta, x \in (a,b)\}$ کامل کراندار خواهد بود .

بنابراین با توجه به قضیه۱ اگر

$$U_\varphi = \{T(X) : T(\varphi(x)) = -\frac{p}{1-p}T(x) \ , \ \forall x \in (a,b)\}$$

آنگاه U_φ کلاس برآوردگرهای نااریب صفر در کلاس $\{f_\theta(x) : \theta \in \Theta, x \in (a,b)\}$ خواهد بود .

قضیه ۲ : آماره $T(X)$ متعلق به U_φ است اگر و فقط اگر $r \in (a,b)$ و تابع ψ روی $[r, \varphi(r))$ وجود داشته باشند طوری که

$$T(x) = (-\frac{p}{1-p})^n \psi(c_{-n}(x)) \ , \ x \in I_{n,r} \ , \ \forall n \in Z \qquad (۴)$$

که در آن برای هر $x \in (a,b)$:

$$c_{-n}(x) = \varphi_{-n}(x) = \underbrace{\varphi^{-1}(...(\varphi^{-1}}(x))...) \ , \ \forall n \in N$$

اثبات : ابتدا ثابت می‌کنیم برای هر $r \in (a,b)$

$$(a,b) = \bigcup_{n\in Z} I_{n,r}.$$

برای هر $n \in Z$ واضح است که $I_{n,r} \subset (a,b)$. بنابراین $\bigcup_{n\in Z} I_{n,r} \subset (a,b)$. حال اگر $x_1 \in (a,b)$ و برای هر $n \in Z$ ، $x_1 \notin I_{n,r}$ ، آنگاه برای هر $n \in Z$ ، $x_1 < c_n(r)$ و یا برای هر $n \in Z$ ، $x_1 > c_n(r)$. اگر برای هر $n \in Z$ داشته باشیم $x_1 < c_n(r)$ پس $x_1 < lim_{n\to -\infty} c_n(r)$. اگر فرض کنیم $x_2 = lim_{n\to -\infty} c_n(r)$ آنگاه $x_1 < x_2$. چون برای هر $x \in (a,b)$ ، $c_n(x)$ در n اکید صعودی است و $x_2 \in (a,b)$ پس $c_{-1}(x_2) < c_0(x_2)$ یعنی $\varphi^{-1}(x_2) < x_2$. اما این تناقض است زیر $x_2 = lim_{n\to -\infty} c_n(r)$ و لزوما باید $\varphi^{-1}(x_2) = x_2$. به همین طریق اگر برای هر $n \in Z$ داشته باشیم $x_1 > c_n(r)$ می‌توان به تناقض رسید . از این تناقض‌ها نتیجه می‌شود که $(a,b) \subset \bigcup_{n\in Z} I_{n,r}$. بنابراین $(a,b) = \bigcup_{n\in Z} I_{n,r}$.

حال قضیه را به صورت زیر اثبات می‌کنیم .

شرط لازم :

فرض کنید $T(x)$ به فرم(۴) باشد . ثابت می‌کنیم برای هر $x \in (a,b)$ داریم $T(\varphi(x)) = -\frac{p}{1-p}T(x)$

برای اثبات به طریق زیر عمل می‌کنیم :

$$T(x) = (-\frac{p}{1-p})^n \psi(c_{-n}(x)) \; , \; x \in [c_n(r), c_{n+1}(r)) \; , \; \forall n \in Z.$$

در نتیجه

$$\begin{aligned}
T(\varphi(x)) &= (-\frac{p}{1-p})^n \psi(c_{-n}(\varphi(x))) \quad , \varphi(x) \in [c_n(r), c_{n+1}(r)) \; , \quad \forall n \in Z \\
&= (-\frac{p}{1-p})^n \psi(c_{-n+1}(x)) \; , \; x \in [\varphi^{-1}(c_n(r)), \varphi^{-1}(c_{n+1}(r))) \quad , \forall n \in Z \\
&= (-\frac{p}{1-p})^n \psi(c_{-(n-1)}(x)) \; , \; x \in [c_{n-1}(r), c_n(r)), \; \forall n \in Z \\
&= (-\frac{p}{1-p})^{n+1} \psi(c_{-n}(x)) \; , \; x \in [c_n(r), c_{n+1}(r)) \; , \; \forall n \in Z
\end{aligned}$$

شرط کافی :

حال فرض کنید $T(X) \in U_p$ و $r \in (a,b)$ را ثابت در نظر بگیرید . تعریف کنید

$$\psi(x) = T(x)\ ,\ \forall x \in [r, \varphi(r))$$

فرض کنید $x \in I_{n,r}$. با استفاده از استقرا ریاضی روی n قضیه را ثابت می‌کنیم .
اگر $n = ۰$آنگاه برای $x \in [r, \varphi(r))$ داریم

$$T(x) = \psi(x) = (-\frac{p}{۱-p})^{۰}\psi(c_{۰}(x)).$$

اگر $n = ۱$ آنگاه برای هر $x \in [\varphi(r), \varphi_{۲}(r))$ داریم $\varphi^{-۱}(x) \in [r, \varphi(r))$ ، بنابراین

$$\begin{aligned} T(x) &= -\frac{p}{۱-p}T(\varphi^{-۱}(x)) \\ &= -\frac{p}{۱-p}\psi(\varphi^{-۱}(x)) \\ &= -\frac{p}{۱-p}\psi(c_{-۱}(x))\ ,\quad \forall x \in [\varphi(r), \varphi_{۲}(r)). \end{aligned}$$

بنابراین قضیه برای $n = ۰, ۱$ ثابت است .حال فرض می‌کنیم قضیه برای $n = k \in N$ صحیح باشد،آنگاه

$$\begin{aligned} T(x) &= (-\frac{p}{۱-p})^{k}\psi(c_{-k}(x))\ ,\ x \in [c_k(r), c_{۱+k}(r)) \\ &= (-\frac{p}{۱-p})^{k}\psi(c_{-k}(x))\quad ,\ \varphi(x) \in [c_{k+۱}(r), c_{k+۲}(r)) \end{aligned} \tag{۵}$$

از طرفی

$$T(x) = (-\frac{p}{۱-p})^{-۱}T(\varphi(x))\quad ,\ \varphi(x) \in [c_{k+۱}(r), c_{k+۲}(r)) \tag{۶}$$

بنابراین

$$T(x) = (-\frac{p}{1-p})^{k+1}\psi(c_{-(k+1)}(x))\ ,\ x \in I_{k+1,r}.$$

یعنی قضیه برای $n = k + 1$ ثابت است .

با استفاده از استقرا ریاضی برای n های منفی نیز می‌توان این مطلب را نشان داد .

بنابراین شرط کافی نیز ثابت است .□

با توجه به قضایای ۱ و ۲ نتیجه زیر حاصل می‌شود :

اگر $p = \frac{1}{2}$ و $\varphi(x) = 1$ را در فاصله $[r, \varphi(r))$ در نظر بگیریم آنگاه

$$T(x) = (-1)^n\ ,\ x \in I_n\ ,\ \forall n \in Z.$$

یعنی $T(X)$ یک تابع کراندار است ، در حالیکه $T(X) \in U_\varphi$. بنابراین در این حالت خانواده چگالیها $f_\theta(x)$ نه کامل است و نه کامل کراندار .

۳ کلاس $MVUE$ ها برای توزیعهای کامل کراندار

قضیه ۳ : فرض کنید

$$M_\varphi = \{S(X) : E_\theta(S(X)) < \infty, S(x) = S(\varphi(x))\ ,\ \forall x \in (a,b)\}$$

تحت فرضیات قضایای قبل M_φ کلاس برآوردگرهای نااریب با کمترین واریانس تحت خانوا چگالی‌های $f_\theta(x)$ می‌باشد .

اثبات : طبق قضیه $Lehmann - Sheffeh$ هر $S(X)$ برای امید خودش تحت چگالی f_θ برآورد نااریب با کمترین واریانس است اگر و فقط اگر برای هر برآوردگر نااریب صفر مثل $T(X)$ وهر $\theta \in$ داشته باشیم

$$Cov_\theta(T(X), S(X)) = 0$$

از طرفی

$$T(\varphi(x)) = (-\frac{p}{1-p})T(x) \ , \ \forall x \in (a,b)$$

زیرا $T(X) \in U_\varphi$. بنابراین

$$-\frac{p}{1-p}T(x)S(\varphi(x)) = -\frac{p}{1-p}T(x)S(x) \ , \ \forall x \in (a,b)$$

در نتیجه

$$S(\varphi(x)) = S(x) \ , \ \forall x \in (a,b).$$

بنابراین $S(X)$ ، $MVUE$ امید خودش است اگر و فقط اگر $S(X) \in M_\varphi$. $\square$

قضیه ۴ : هر تابع $\gamma(\theta)$ تحت تابع چگالی f_θ دارای $MVUE$ است اگر و فقط اگر به فرم زیر باشد .

$$\gamma(\theta) = \sum_{n=-\infty}^{\infty} \int_{c_n(r)}^{c_{n+1}(r)} \psi(c_{-n}(x))g_\theta(x)dx$$

به طوری که $r \in (a,b)$ و ψ یک تابع غیر ثابت روی $[r, \varphi(r))$ است .

اثبات : آماره $S(X)$ با امید متناهی متعلق به M_φ است اگر و فقط اگر $r \in (a,b)$ و تابع غیرثابت ψ روی $[r, \varphi(r))$ وجود داشته باشند طوری که

$$S(x) = \psi(c_{-n}(x)) \ , \ \forall x \in [c_n(r), c_{n+1}(r)).$$

اثبات مطلب فوق شبیه اثبات قضیه ۲ است و از آن صرفنظر می‌کنیم . طبق قضیه ۳ هر $S(X)$ برای $E_\theta^f(S(X))$ برآوردگر نااریب با کمترین واریانس است اگر و فقط اگر $S(X) \in M_\varphi$. بنابراین $\gamma(\theta) = E_\theta^f(S(X))$ به طوری که $S(X) \in M_\varphi$. اما به راحتی ثابت می‌شود که

$$E_\theta^f(S(X)) = E_\theta^g(S(X)) \ ,$$

توجه کنید در قضیه ۴ اگر $\psi(x)$ در فاصله $[r,\varphi(r))$ کراندار باشد آنگاه $\psi(c_{-n}(x))$ نیز در فاصله $[c_n(r), c_{n+۱}(r))$ کراندار خواهد بود . چون در فاصله $[c_n(r), c_{n+۱}(r))$ داریم $S(x) = \psi(c_{-n}(x))$ بنابراین $S(x)$ در این فاصله کراندار بوده و $E_\theta^f(S(X))$ نیز کراندار شده در نتیجه سری فوق همگرا خواهد بود .

مثال : خانواده چگالیهای $f_\theta(x)$ در زیر کامل کراندار است

$$f_\theta(x) = \frac{۱}{۳}\theta e^{-\theta x} + \frac{۲}{۳}۲\theta e^{-۲\theta x}\,, x > ۰, \theta > ۰.$$

(چون اگر در قضیه ۱ قرار دهیم $p = \frac{۱}{۳}, \varphi(x) = ۲x, g_\theta(x) = \theta e^{-\theta x}$ آنگاه $f_\theta(x)$ را به صورت فوق خواهیم داشت) با توجه به قضیه ۴ ،

$$\gamma(\theta) = \sum_{n\in Z}\int_{۲^n}^{۲^{n+۱}} \psi(۲^{-n}x)\theta e^{-\theta x}dx$$

هر تابع به فرم $\gamma(\theta)$ دارای $MVUE$ خواهد بود . اگر $\psi(x) = x$ را در فاصله $[۱,۲)$ در نظر بگیریم خواهیم داشت :

$$\begin{aligned}
\gamma(\theta) &= \sum_{n\in Z}\int_{۲^n}^{۲^{n+۱}} ۲^{-n}x\theta e^{-\theta x}dx \\
&= \sum_{n\in Z} ۲^{-n}\frac{۱}{\theta}[-e^{-\theta x}(\theta x + ۱)]_{۲^n}^{۲^{n+۱}} \\
&= \sum_{n\in Z} \frac{۲^{-n}}{\theta}[e^{-\theta ۲^n}(\theta ۲^n + ۱) - e^{-\theta ۲^{n+۱}}(\theta ۲^{n+۱} + ۱)] \\
&= \sum_{n\in Z}[e^{-\theta ۲^n} + \frac{۱}{\theta}۲^{-n}e^{-\theta ۲^n} - ۲e^{-\theta ۲^{n+۱}} - \frac{۱}{\theta}۲^{-n}e^{-\theta ۲^{n+۱}}].
\end{aligned}$$

و عکس تعمیم یافته به کمک شبیه سازی

بهزاد یعقوبیان[1]، حمیدرضا نواب پور[2]

۱-دانشگاه علوم پزشکی همدان ۲-دانشگاه علامه طباطبائی

چکیده : هنگامی که در ماتریس طرح همخطی وجود دارد،واریانس برآوردگر حداقل مربعات متورم می شود ، دراین حالت به کار گیری رگرسیون اریب به منظور کاهش واریانس ، مناسبتر از استفاده از رگرسیون حداقل مربعات است .

ما از بین برآوردگرهای اریب به معرفی چهار برآوردگر ریج[3] ، مؤلفه اصلی[4] ، فشرده[5] و عکس تعمیم یافته[6] پرداخته و دقت آنها را در برآورد پارامترهای رگرسیون خطی به کمک شبیه سازی مقایسه می کنیم .

واژه های کلیدی: همخطی ، رگرسیون اریب ، براوردگر ریج ، براوردگر مؤلفه اصلی ، برآوردگر فشرده وبرآوردگر عکس تعمیم یافته .

۱- مقدمه

فرض کنید y بردار متغیر وابسته و X ماتریس طرح ، شامل متغیرهای مستقل باشد. برای بسیاری از مقاصد ویا توصیف ارتباط متغیرهای y و X به برآورد پارامترهای رابطه رگرسیون $y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3$ نیاز داریم .

گاهی بین متغیرهای مستقل همبستگی خطی وجود دارد که اصطلاحا ٬ آنرا همخطی می گو ییم . همخطی بین k متغیر مبین وجود رابطه خطی بین آنها است، به عبارت دیگـر یـک متغیر را مـی تـوان از ترکیب خطی $(k-1)$ متغیر دیگر بدست آورد.

زمانی که درمتغیرهای مستقل همخطی وجودندارد، برای تعیین ضرایب ، β ، روش حداقل مربعات را مورد استفاده قرار می دهیم اما وجود همخطی بین متغیرهای مستقل ، واریانس برآوردگرهای حداقل مربعات را متورم می سازد . در این حالت بهتر است بجای این روش از روش برآوردگرهای اریب استفاده شود .

---

[3] Ridge.

[4] Prencipal Component .

۸

۴- برآوردگرهای اریب

برآوردگر ریج بر پایه افزودن یک ماتریس قطری ، با مؤلفه های مثبت k به ماتریس $X' X$ بنـــا می شود. این برآوردگر به وسیله فرمول زیر محاسبه می گردد .

$$\beta_{rr} = (X'X + K)^{-1} X' y$$

K ماتریس قطری با مؤلفه های $k > 0$ می باشد .روش کلی تر ریج دارای اریبی کمتری است، که آن را ریج سودار می نامند .در این روش مقادیر روی قطر ماتریس K متفاوتند واندازه k_i ها بستگی به مقدار ویژه متغیر متناظر آن دارد .هر چه مقدار ویژه کوچکتر باشد k_i متناظر آن بزرگتر است .

برآوردگر مؤلفه اصلی برپایه روش مؤ لفه اصلی است. در این روش متغیرهایی که دارای مقادیر ویژه کوچک هستند حذف می شوند. برآورد حاصل از این روش عبارت است از:

$$\beta_{pc} = [P'_r A_r^{-1} P_r] X' y$$

P_r ماتریس بردارهای ویژه $X'X$ است که (p-r) بردارمتناظر با مقادیر ویژه کوچکتر آن حذف شده است و A_r ماتریس قطری (r ⊗ r) است کهمؤلفه های روی قطر آن مقادیر ویژه متناظر با بردارهای P_r هستند.

اگر به رابطه این برآوردگر با برآوردگر حداقل مربعات توجه کنیم اریبی آن مشاهده خواهد شد.

$$\beta_{pc} = [P'_r A_r^{-1} P_r](X'X)(X'X)^{-1} X' y$$

$$= [P'_r A_r^{-1} P_r](X'X)\beta_{ms}$$

بنابر این

$$E(\beta_{pc}) = [P'_r A_r^{-1} P_r](X'X)\beta$$

یعنی برآوردگر مؤلفه اصلی برآوردگری اریب ودارای این ویژگی که واریانس را کاهش می دهد.

$$=\lambda E(\beta \quad) = \lambda\beta$$

۳-طرح مطالعه

چنانچه هر یک از شرایط حجم نمونه ،n، یا انحراف معیار،σ ،وحتی معیار مقایسه برآوردگرها عوض شود این امکان وجود دارد که نتایج مطالعه تغییر کند . بنابر این برای مطالعه شبیه سازی ، حجم نمونه وانحراف معیار های مختلفی را در نظر گرفته و برای هر زوج از آنها ، ۴۰۰ مجموعه داده تولید کردیم تا برآوردگرها را در شرایط مختلف مقایسه کرده واثر خطای تصادفی را کا هش داده با شیم .

برای اینکه هم حجم نمونه و هم انحراف معیار را در مقایسه برآئردگرها بررسی کنیم ، پنج اندازه نمونه ۱۶، ۳۲، ۶۴، ۱۲۸، ۲۵۶ و پنج انحراف معیار ۸ ۰/۰، ۰/۲، ۰/۶، ۱/۲ و ۲ را در نظر گرفتیم .

متغیر های مستقل X_1 ، X_2 و X_3 به ترتیب دارای توزیع X_1"N(5,1.5) X_2"N(2,1) و $X_3 = 3X_2 + a$ که در آن a"N(0,0.25) . بنابر این همبستگی نسبتا قوی بین دو متغیر X_2 و X_3 وجود دارد.

مقادیر y را از مدل خطی زیر محاسبه می کنیم .

$$y = 5 + 4X_1 + 3X_2 + 2X_3 + e$$

که در آن e"N(0,σ^2)

پارامتر های مدل زیر بازاء هر حجم نمونه ، ۴۰۰ بار بوسیله هر چهار روش ریج ، فشرده ، مؤلفه اصلی وحداقل مربعات برآورد شده اند.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \mathfrak{M}$$

برای مقایسه برآوردها که از چهار روش فوق بدست آمدند معیارهای واریانس برآورد ، میانگین قدر

که $\hat{\beta}_{j(i)}$ برآورد مؤلفه i ام بردار β در j امین تکرار تولید داده و $\bar{\beta}_{(i)}$ برابر است با

$$\bar{\beta}_{(i)} = 1/400 \sum_{j=1}^{400} \hat{\beta}_{j(i)}$$

ب) متوسط قدر مطلق اریبی(AAB)

اریبی برای هر برآورد پارامتر به وسیله تفاضل پارامتر حقیقی از میانگین مقادیر برآورد شده پارامتر برای ۴۰۰ نمونه تولید شده محاسبه می شود .

$$\text{bias}(\hat{\beta}_{(i)}) = \bar{\beta}_{(i)} - \beta_{(i)}$$

با استفاده از اریبی،AAB بردار $\hat{\beta}$ از رابطه زیر محاسبه می شود.

$$AB(\hat{\beta}) = 1/4 \sum_{i=1}^{4} |\ \text{bias}\ (\hat{\beta}_{(i)})|$$

ج) میانگین مربع خطا (MSE)

برای محا سبه میانگین مربع خطا از رابطه زیر استفاده می کنیم .

$$\text{ISE}(\hat{\beta}) = 1/400 \sum_{j=1}^{400} (\hat{\beta}_j - \beta)' (\hat{\beta}_j - \beta)$$

همچنین می توان نوشت

$$\text{SE}(\hat{\beta}) = \sum_{i=1}^{4} \text{ETV}(\hat{\beta}_{(i)}) + \sum_{i=1}^{4} [\text{bias}(\hat{\beta}_{(i)})]^2$$
$$= \text{Total Variances} + (\text{Bias})^2$$

در این رابطه $\hat{\beta}_j$ برآورد بردار پارامتر β در j امین مرحله از تولید داده است .

۴- نتا یج شبیه سازی

از ماتریس بدست آمده برای مقادیر پارامتر β ، β ، β ، β [illegible]

(64 باشد ، آنچه از نمودارها نتیجه می شود افزایش اریبی روش حداقل مربعات نسبت به برآوردگر ریج است ، اما سایر برآوردگرها دارای اریبی بیشتری نسبت به برآوردگر حداقل مربعات و ریج می با شند .

اگر ابزار مقایسه را MSE در نظر بگیریم ، با توجه به نمودارها ی بدست آمده می توان چنین نتیجه گرفت که دقت روش حداقل مربعات ، تنها برای انحراف معیار های کوچک ، نسبت به سایر روشها بیشتر است . اما با افزایش انحراف معیار دقت روش ریج بیشتر می گردد به طوری که برای $\sigma \geq 0.6$ برآوردگر ریج دارای کمترین MSE می باشد ،همین شرایط برای برآوردگرمؤلفه اصلی نیز به طور ضعیفتر ی وجود دارد هر چند دقت آن برابر ریج نمی با شد . اما در $\sigma \geq 1.2$ از دو روش حداقل مربعات وفشرده دقیقتر عمل می کند .

Refrences :

Guilkey, D. K. and Murphy, J. L. (1975), " Directed Ridge Regression Techniques in Cases of Multicollinearity," **Journal of the American Statistical Association**, Vol. 70, No. 352, 769-775.

Gunst ,R. f. and Mason ,R. L. (1979), "Some Considerations in the Evaluation of Alternate Prediction Equations, " **Technometrics**, Vol. 21,No.1, 55-63.

Hawkins, D. M. (1973),"On the Investigation of Alternative Regression by Principal Component Analysis," **Applied Statistics**, 22, 275-286.

Hocking, R. R. (1976), "The Analysis and Selection of Variables in Linear Regression," **Biometrics**, 32, 1-49.

Hoerl , A. E.and Kennard,R.W.(1970),"Ridge Regression Biased Estimation for Nonorthogonal Problems,"**Technometrics**, 12, 55-67.

Hoerl, A. E. and Kennard, R. W. (1970), "Ridge Regression Application to Nonorthogonal Problems," **Technometrics**, 12, 69-82.

علی کریمی یکتا ، ابوالقاسم بزرگ‌نیا
دانشگاه فردوسی مشهد

چکیده در این مقاله توزیع متغیر تصادفی $X=Y+Z$ را که Y و Z مستقل ،Y متغیر تصادفی پواسون قطع شده راست ،چپ یا دو طرفه)با پارامتر θ و Z متغیر تصادفی پواسون کامل با پارامتر $\theta\theta$ می باشد ،بدست می آور و،دراینصورت X را متغیر تصادفی پواسون قطع شده (از راست ، چپ یا دو طرفه) تداخلی می نامیم.فرض می ک که تنها مشاهدات X ثبت شده ا ند و مشاهدات Y و Z بطور مجزا وجود ندارند.توزیع پواسون قطع شده تداخلی مسائلی که پیشامدها با نرخ های [۱] متفاوت در دوبازه زمانی یا دو وضعیت مختلف اتفاق می افتند ،کار برد پ می‌کند. فرض می کنیم که در طی دوره مشاهده یک مداخله [۲] همانند اعمال یک روش درمانی یا انجام یک تغ یا اعمال یک تصمیم در سیستم وجود دارد .این مداخله تاثیر خود را در تغییر پارامتر به θ به $\theta\theta$ نشان می دهد Y و Z را بترتیب، تعداد حالتها یا مشاهدات قبل وبعد از مداخله در نظر می گیریم.

۱- توزیع پواسون قطع شده در صفر تداخلی

فرض کنیم X تعداد موارد وقوع یک پیشامد نادر باشد که در دو مرحله بانرخ های متفاوت رخ می دهد یعن $X=Y+Z$ در طی زمان شمارش یک مداخله وجود دارد وهنگامیکه $Y>0$ دستگاه های شمارش گر فعال می شو Y,Z، به ترتیب موارد قبل وبعد از مداخله می باشند.

فرض کنیم Y تعداد موارد و با(یا یک بیماری واگیر دار) در یک خانواده باشد .پیشامد $Y=0$ قابل مشاهده نیست ز دستگاه های تشخیص در صورت $Y>0$ فعال می شوند .متغیر تصادفی Y یک پیشامد نادر را بوجود می‌آورد .ا مناسب است توزیع پواسون قطع شده در صفر باپارامتر θ برای متغیر تصادفی مثبت صحیح Y با تابع احتمال ز درنظر بگیریم.

$$P(Y=y)=(e^{\theta}-1)\theta^{y}/y! \qquad y=۱,۲,۳,...$$

می دهند. فرض می کنیم که چنین اثری θ را به $\rho\theta$ تغییر می دهد یعنی $E(Z)=\rho\theta$ که $\infty > \rho \geq 0$ پارامتر مداخله است و Z تعداد مواردی است که بعد از اعمال عوامل پیشگیری مشاهده می شود که مستقل از Y دارای توزیع پواسون با میانگین $\rho\theta$ است . در نظر داریم که تنها مشاهدات متغیر تصادفی $X=Y+Z$ ثبت شده است.پس

$$P_r(X=x)=\sum_{y=1}^{\infty} P_r(Y+Z=x \mid Y=y)P_r(Y=y)$$

$$=\sum_{y=1}^{\infty} P_r(Y=y)P_r(Z=x-y)$$

$$=\sum_{y=1}^{\infty}\frac{\theta^{y}}{y!(e^{\theta}-1)}\ \frac{(\rho\theta)^{x-y}e^{-\rho\theta}}{(x-y)!}$$

پس

$$P_r(X=x)=\left[e^{\rho\theta}(e^{\theta}-1)^{-1}\right]\frac{\theta^{x}}{x!}\left[(1+\rho)^{x}-\rho^{x}\right] \qquad x=1,2,3,\ldots$$

حال گشتاورهای X را می یابیم .برای این منظور ابتدا تابع مولد احتمال X را به دست می آوریم.

$$\Psi_x(S)=E(S^{X})=E(S^{Y+Z})=\left[Es^{Y}\right]\left[Es^{Z}\right]$$

$$E(S^{Y})=(e^{\theta}-1)^{-1}\sum_{y=1}^{\infty}\frac{s^{y}\theta^{y}}{y!}=(e^{\theta}-1)^{-1}(e^{s\theta}-1)$$

و همچنین داریم

$$E(S^{z})=(e^{\rho\theta})\sum_{y=1}^{\infty}\frac{(\rho\theta s)^{z}}{z!}=e^{-\rho\theta}e^{\rho\theta s}$$

پس

$$\Psi_x(s)=E(s^{X})=(e^{\theta}-1)^{-1}e^{-\rho\theta}e^{\rho\theta s}(e^{s\theta}-1)$$

با توجه به

$$E(X)=\frac{d\Psi_x(S)}{dS} \mid S=1$$

داریم

$$\mu=E(X)=\theta\left[\rho+1+(e^{\theta}-1)^{-1}\right] \qquad (1)$$

همچنین
$$\frac{d^{2}\Psi_x(s)}{ds^{2}}=(e^{\theta}-1)e^{-\rho\theta}\theta\frac{d}{ds}e^{\rho\theta s}\left[e^{s\theta}(\rho+1)-\rho\right]$$

$$X(X-1)]=\frac{d^{2}\Psi_x(s)}{ds^{2}}\mid_{s=1}=(e^{\theta}-1)^{-1}e^{-\rho\theta}\theta\left\{\rho\theta\, e^{\rho\theta}\left[e^{\theta}(\rho+1)-\rho\right]+e^{\rho\theta}\theta(\rho+1)e^{\theta}\right\}$$

و در نتیجه

$$E[X(X-1)]=\theta^{2}\{(\rho+1)^{2}+(1-2\rho)(e^{\theta}-1)^{-1}\}$$

در نتیجه واریانس X به صورت زیر به دست می آید

$$var(X)=E[X(X-1)]+E(X)-E^{2}(X)$$

$$var(X)=\mu-e^{\theta}[\theta/(e^{\theta}-1)]^{2} \qquad (2)$$

برای گشتاورهای مراتب بالاتر می توان از تابع مولد گشتاور استفاده کرد که به صورت زیر به دست می آید .

$$m_x(t)=E(e^{tX})=[Ee^{tY}][Ee^{tZ}]=m_Y(t).m_Z(t)$$

$$m_x(t)=e^{\rho\theta(e^t-1)}(e^{\theta e t}-1)/(e^{\Theta}-1)$$

۲-توزیع پواسون قطع شده دو طرفه تداخلی

ابتدا مثال زیر رادر نظر می گیریم

مثال ۲-۱ یک فروشگاه به مشتریانی که بیشتر از تعداد معین N از کالایی را با حداکثر تعداد M کالا خریداری نمایند تخفیف می دهد لذا می توانیم فرض کنیم که تعداد کالاهای فروخته شده با تخفیف دارای توزیع پواسون قطع شده می باشد .در اینصورت X تعداد کل کالاهای فروخته شده شامل آنهایی که با تخفیف بفروش رسیده (Y)و آنهایی که با قیمت معمولی بفروش رسیده (Z) ،دارای توزیع پواسون قطع شده تداخلی است .

فرض می کنیم Z دارای توزیع پواسون با پارامتر $\rho\theta$ مستقل از Y می باشد همچنین فرض می کنیم که از مشاهدات Y و Z بطور مجزا اطلاعی در دست نیست و تنها مشاهدات X ثبت شده اند.حال تابع احتمال X را می یابیم

$$p_r(X=x)=\sum_y P_r(Y=y).P_r(Y+Z=x \mid Y=y)$$

$$=\sum_y P_r(Y=y).P_r(Z=x-y)$$

$$=\sum_y \frac{\theta^y/y!\,e^{-\rho\theta}(\rho\theta)^{x-y}}{\sum_{j=N}^{M}\frac{\theta^j}{j!}\ (x-y)!}$$

$$=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]\sum_y\frac{\theta^y(\rho\theta)^{x-y}}{y!(x-y)!}$$

دو حالت در نظر می‌گیریم

الف) $N\leq x\leq M$، یا بطور معادل $N\leq y\leq x$

$$P_r(X=x)=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\frac{\theta^x}{x!}\sum_{y=N}^{x}\binom{x}{y}\rho^{x-y}$$

$$=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\left((1+\rho)^x-\sum_{j=0}^{N-1}\binom{x}{y}\rho^{x-y}\right)$$

ب) $x\underline{M}$، یا بطور معادل $N\leq y\leq x$

$$P_r(X=x)=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\frac{\theta^x}{x!}\sum_{y=N}^{M}\binom{x}{y}\rho^{x-y}$$

$$=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\frac{\theta^x}{x!}\left\{(1+\rho)^x-\sum_{y=0}^{N-1}\binom{x}{y}\rho^{x-y}-\sum_{y=M+1}^{x}\binom{x}{y}\rho^{x-y}\right\}$$

بنابراین

$$(1+\rho)^x-\sum_{j=0}^{N-1}\binom{x}{j}\rho^{x-j}\qquad\qquad x\leq M\qquad\qquad (4)$$

...ن تابع مولد احتمال X را می یابیم

$$\Psi_x(S)=E(S^X)=E\left(S^{Y+Z}\right)$$

بنا به استقلال Z,Y داریم

$$\Psi_x(S)=(Es^Y)(Es^Z)=\Psi_y(s).\Psi_z(s)$$

$$\Psi_x(S)=\left[\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\sum s^y\frac{\theta^y}{y!}=\left[\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}\sum_{y=N}^{M}\frac{(s\theta)^y}{y!}$$

$$\Psi_x(S)=\left[e^{\rho\theta}\sum_{j=N}^{M}\frac{\theta^j}{j!}\right]^{-1}e^{\theta s}\sum_{y=N}^{M}\frac{(s\theta)^y}{y!}$$

$$E(X)=\Psi_x'(S)/s=1$$

$$E(X(X-1))=\Psi_x''(S)/s=1$$

و با چند عملیات میانی بدست می آوریم

$$\left[\; M \quad j\;\right]^{-1} \left(\; \frac{M}{} \theta^j \quad \frac{M-1}{} \; j \;\right)$$

$$\sigma_x^2 = var(X) = E\left[X(X-1)\right] + \mu_x - \mu_x^2$$ و

که در آن $\mu_x^2 = (EX)^2$

پس(۷)
$$\mu_x = \theta\left\{\rho + 1 + \left[\sum_{j=N}^{M}\frac{\theta j}{j!}\right]^{-1}\left[\frac{\theta^{N-1}}{(N-1)!} - \frac{\theta^{M}}{M!}\right]\right\}$$

$$\sigma_x^2 = \theta^2\left\{(1+\rho)^2 + \frac{(2\rho+1)(\theta^{N-1}/(N-1)! - \theta^{M}/M!) + \theta^{N-2}/(N-2)! - \theta^{M-1}/(M-1)!}{\sum_{j=N}^{M}\frac{\Theta j}{j!}}\right\} + \mu_x - \mu_x^2 \quad (8)$$

۳-توزیع پواسون قطع شده از چپ تداخلی

یک مثال از این توزیع ،مینیمم تعداد مشتریان [1] لازم قبل از شروع یک برنامه است .این نوع تصمیم معمولاً ب پایه دلایل ارزش موثر [2] است .

مثال ۳-۱ در تاسیس یک دوره یا رشته ابتدا قبل از اینکه واقعاً رشته پیشنهاد شود حداقلی از دانشجویان پیش ثبت نام می شوند.این روش در دانشگاه ها یا در ادامه کلاسهای آموزشی بکار می رود.فرض کنید Y تعدا دانشجویان پیش ثبت نام شده (حداقل N)و Z تعداد دانشجویان ثبت نام شده بعد از تصمیم به تأسیس رشته باش آنگاه X ،تعداد کل دانشجویان رشته دارای توزیع پواسون قطع شده از چپ تداخلی است .

مثال ۳-۲ در سازماندهی یک گروه سیاحتی تعداد حداقلی از سیاحان (Y)باید تا قبل از زمانی معین به گرو بپیوندند در غیر اینصورت گروه منحل می شود.X تعداد نهایی سیاحان گروه است که مسئول گروه در نهایت به آ

$$P_r(X=x)=\left[\rho^{\rho\theta}\sum_{j=N}\frac{}{j!}\right]\frac{}{x!}\left\{^{(x+\rho)}\sum_{y=0}(y)^{\rho}\right\}$$

که $N0$ و $N)0, 0\leq\rho(\infty,\theta)0, x=N,N+1,\ldots$

باتوجه به (۷)داریم

$$E(X)=\left[\sum_{j=N}^{\infty}\frac{\theta^j}{j!}\right]^{-1}\theta\left\{\rho\sum_{j=N}^{\infty}\frac{\theta^j}{j!}+\sum_{j=N-1}^{\infty}\frac{\theta^j}{j!}\right\}$$

$$E[X(X-1)]=\left[\sum_{j=N}^{\infty}\frac{\theta^j}{j!}\right]^{-1}\theta^2\left\{\rho^2\sum_{j=N}^{\infty}\frac{\theta^j}{j!}+2\rho\sum_{j=N-1}^{\infty}\frac{\theta^j}{j!}+\sum_{j=N-2}^{\infty}\frac{\theta^j}{j!}\right\}$$

,

$$E[X(X-1)]=\theta^2\left\{(1+\rho)^2+\left[\sum_{j=N}^{\infty}\frac{\theta^j}{j!}\right]^{-1}\left[(2\rho+1)\frac{\theta^{N-1}}{(N-1)!}+\frac{\theta^{N-2}}{(N-2)!}\right]\right\}$$

بنابراین

$$\sigma_x^2=var(X)=\theta^2\left\{(1+\rho)^2+\left[\sum_{j=N}^{\infty}\frac{\theta^j}{j!}\right]^{-2}\left[(2\rho+1)\frac{\theta^{N-1}}{(N-1)!}+\frac{\theta^{N-2}}{(N-2)!}\right]\right\}+\mu_X-\mu_X^2$$

۴-توزیع پواسون قطع شده از راست تداخلی

این توزیع درآژانس هایی که متعهد انجام سرویس برای دو گروه از مشتریان باشند کاربرد پیدا می کند.

مثال ۴-۱ فرض کنید Y تعداد دانشجویان یک خوابگاه (حداکثر M) باشد که برای صرف غذای خاصی به یک سلف سرویس مراجعه می کنند و Zتعداد میهمانان آنها باشد .بنابراین سلف سرویس باید برای X=Y+Z مراجعه کننده ،غذا تهیه کند.اگر Y پواسون قطع شده از راست و Z مستقل از Y ،پواسون باشد X دارای توزیع پواسون قطع شده ا:

$$\left(\quad \sum_{j=0}^{\quad} j! \right) \quad x! \left[(1+\rho) \quad - \sum_{j=N+1} (y) \right] \rho$$

که $x = 0, 1, 2, \dots$, $\theta) 0$, $0 \le \rho (\infty$, $1 \le M (\infty$

با توجه به (۴) داریم

$$\mu_x = E(X) = \left[\sum_{j=0}^{M} \frac{\theta^j}{j!} \right] \theta \left\{ \rho \sum_{j=0}^{M} \frac{\theta^j}{j!} + \sum_{j=0}^{M-1} \frac{\theta^j}{j!} \right\}$$

پس

$$\mu_x = E(X) = \theta \left\{ 1 + \rho - \left[\sum_{j=0}^{M} \frac{\theta^{j-1}}{j!} \right]^{-1} \frac{\theta^M}{M!} \right\}$$

همچنین با توجه به (۶) داریم .

$$\left[X(X-1) \right] = \left[\sum_{j=0}^{M} \frac{\theta^j}{j!} \right]^{-1} \theta^2 \left\{ \rho^2 \sum_{j=0}^{N} \frac{\theta^j}{j!} + 2\rho \sum_{j=0}^{M} \frac{\theta^j}{j!} + \sum_{j=0}^{M} \frac{\theta^j}{j!} - 2\rho \frac{\theta^M}{M!} - \frac{\theta^M}{M!} - \frac{\theta^{M-1}}{(M-1)!} \right\}$$

بنابراین

$$\sigma_x^2 = \theta^2 (1+\rho)^2 - \left[\sum_{j=0}^{M} \frac{\theta^j}{j!} \right]^{-1} \left[(2\rho+1) \frac{\theta^M}{M!} + \frac{\theta^{M-1}}{(M-1)!} \right] + \mu_x - \mu_x^2$$

غلامحسین شاهکار ، حامدرضا طارقیان

دانشگاه فردوسی مشهد

چکیده به دلیل وجود حجم گستردهٔ اطلاعات و لزوم نگهداری، پردازش و انتقال سریع آنها لازم است تا اطلاعات به کدهای عددی و یا الفبا- عددی تبدیل شوند. هر یک از اجزاء این کدهای اطلاعاتی و یا ترکیب مناسبی از آنها دارای بار اطلاعاتی است. لذا اطمینان از صحیح بودن آنها، قبل از استفاده ضرورتی اجتناب ناپذیر است. برای بررسی صحت کدهای اطلاعاتی روشهای مختلفی طراحی و اعمال می شوند. هدف این مقاله نقد و بررسی برخی روشهای موجود و ارائهٔ روشی مؤثر برای بازرسی صحت کدهای ده رقمی است.

مطالب این مقاله به خصوص می تواند مورد استفاده آن دسته از مسئولین کشور ما باشد که در پی کدبندی اطلاعات سازمانی خود به صورت عددی‌اند، و لذا مایلند تا با اطلاع از نقاط ضعف و قوت روشهای مختلف کدبندی موجود، سیستمی کارآمد را برگزینند.

کلمات کلیدی : کدگذاری اطلاعات، نظریه اعداد، حساب پیمانه‌ای.

۱ - مقدمه

وجود حجم گستردهٔ اطلاعات در زمینه های مختلف علمی، اقتصادی، اجتماعی و ... و لزوم نگهداری، پردازش و انتقال سریع آنها ایجاب می کند تا اطلاعات به کدهای عددی و یا الفبا- عددی تبدیل شوند. این کدها را کدهای اطلاعاتی می نامند. کدهای اطلاعاتی در زمینه های گوناگون اطلاع رسانی مورد استفاده قرار می گیرند. این کدها می توانند در رابطه با اطلاعات مربوط به اشخاص حقیقی، مانند امور بانکی، بیمه های اجتماعی، مالیات و نظایر

حاصل کند.

برای تعیین صحت و سقم کدهای اطلاعاتی، سازمانهای مختلف از روشهای متفاوتی استفاده می کنند. به طور کلی در این روشها به منظور کشف اشتباهات کدهای اطلاعاتی، عددی را بنام **عدد بازرس** به کد اطلاعاتی اضافه می کنند. عموماً در این روشها تعیین عدد بازرس مبتنی بر حساب پیمانه‌ای است. به عبارت دیگر، عدد بازرس را باقیماندهٔ تقسیم کد اطلاعاتی یا ترکیبی از ارقام آن بر عددی به اسم پیمانه در نظر می‌گیرند.

هدف ما از تدوین این مقاله نقد و بررسی روشهای موجود در تعیین عدد بازرس و ارائهٔ روشی مؤثر برای بازرسی صحت کدهای اطلاعاتی ده رقمی است. در بخش دوم برخی از روشهای مختلف بازرسی صحت کدهای اطلاعاتی را بررسی می کنیم. همان طور که در این بخش خواهیم دید کارایی روشهای مختلف در تعیین عدد بازرس به منظور کشف اشتباهات وابسته به انتخاب عدد پیمانه است. از این رو در بخش سوم مقاله مناسبترین عدد دو رقمی به عنوان پیمانه را انتخاب می کنیم. بالاخره در بخش پایانی روشی مؤثر را معرفی می کنیم که نسبت به روشهای موجود از کارایی بیشتری در کشف اشتباهات برخوردار است.

۲ - بررسی روشهای مختلف بازرسی صحت کدهای اطلاعاتی

همان طور که گفته شد برای کشف اشتباهات کدهای اطلاعاتی، معمولاً به رشته اصلی کد اطلاعاتی عددی به عنوان عدد بازرس اضافه می کنند. این عدد را به روشی خاص، فراخور کد اطلاعاتی در نظر می‌گیرند. نظرگرفتن عدد بازرس به منظور کشف اشتباهاتی است که به دلیل دست به دست شدن کدهای اطلاعاتی مثل خواندن، نوشتن، گفتن و یا شنیدن پیش می آید. وقتی در کد اطلاعاتی j رقم اشتباه می شوند، می‌گوییم خطای به وزن j رخ داده است. روشهای مختلفی برای بازرسی کدهای اطلاعاتی در سیستم دهدهی پیشنهاد شده‌اند. در تعیین کارایی روشهای بازرسی لازم است تا توزیع فراوانی اشتباهات را بدانیم. تجربه نشان می دهد [۳] تقریباً ۹۰ درصد اشتباهات از نوع اشتباهات تک رقمی و ۱۰ درصد از نوع اشتباهات دو رقمی است که از این میان سهم اشتباهات دو رقمی در نتیجهٔ جابجایی ارقام، هفت درصد و سهم تغییر ارقام (غیر از جابجایی) سه درصد است. احتمال بروز خطاهای به وزن ۳ یا بیشتر کمتر از ۱ درصد است.

با توجه به توزیع احتمال فوق روشن است که باید اولویت را به کشف اشتباهات تک رقمی داد. بعد از آن مطلوب آن است که حتی‌المقدور اشتباهات دو رقمی کشف شوند. قابلیت روشهای مختلف در کشف اشتباهات دو رقمی بیشتر معیاری برای تشخیص کارایی آنهاست.

برای اطمینان از صحت آن کافیست ببینیم آیا مجموع ارقام به پیمانه ۹ صفر است یا خیر.

روشهای متعددی برای بازرسی صحت کدهای اطلاعاتی تدوین شده‌اند. در این جا برخی از آنها را که در حال حاضر استفاده می شوند، مورد بررسی قرار می دهیم.

در ایالات متحده کد اطلاعاتی حواله های پست مالی[1] عددی است ده رقمی به انضمام یک رقم بازرس که در آن، رقم بازرس باقیماندهٔ تقسیم مجموع ارقام کد اطلاعاتی بر عدد ۹ است [۶]. با در نظر گرفتن چنین عددی به عنوان رقم بازرس کلیهٔ اشتباهات تک رقمی کشف نمی شوند. مثلاً تغییر رقم ۹ در کد اطلاعاتی به صفر و یا بالعکس قابل کشف نیست. لازم به ذکر است که اشتباه مربوط به رقم بازرس قابل کشف است. ضمناً تنها آن دسته از اشتباهات جابجایی که در رابطه با رقم بازرس باشند، قابل کشفند. اشتباه حاصل از جابجایی دو رقم در کد اطلاعاتی مثلاً ...۵۳... بجای ...۳۵... قابل کشف نیست، در حالی که اشتباه جابجایی با احتساب رقم بازرس، مثلاً ...۵۳ بجای ۳۵... قابل کشف است. نرخ کشف اشتباهات تک رقمی در این روش ۹۸٪ و نرخ کشف تغییرات جابجایی تقریباً ۱۰٪ است [۲].

روش مشابه و با کارایی یکسان روشی است که در چکهای مسافرتی **ویزا**[2] اعمال می شود. در آنجا رقم بازرس را به این ترتیب حساب می کنند که ابتدا ارقام کد اطلاعاتی را با هم جمع کرده و سپس آن را از کوچکترین مضرب عدد ۱۰ که از مجموع ارقام بزرگتر است، کم می کنند [۲].

در ایالات متحده آژانسهای هوایی، شرکت فدرال اکسپرس[3] ، و شرکت حمل و نقل **یو.پی.اس**[4] از روشی استفاده می کنند که در کشف اشتباهات تک رقمی قابلیت کمی دارد، اما در کشف اشتباهات جابجایی از توانایی بالایی برخوردار است. مثلاً کد اطلاعاتی شرکت یو.پی.اس ۹ رقمی به انضمام یک رقم بازرس است. رقم بازرس باقیماندهٔ تقسیم کد اطلاعاتی بر عدد ۷ است. البته در این روش نیز اگر به جای رقم a رقم b بیاید که در آن $|a-b|=7$ ، در این صورت دیگر رقم بازرس قادر به کشف آن اشتباه نیست. در این روش نرخ کشف اشتباهات تک رقمی ۹۴٪، و نرخ کشف اشتباهات جابجایی ۹۴/۱٪ است [۶].

شرکت آمریکایی سی.آ.اس[5] رقم بازرسی را که به کد اطلاعاتی مواد شیمیایی نسبت می دهد، به صورت زیر تعیین می کند. این شرکت رقم بازرس کد اطلاعاتی $a_k a_{k-1} \ldots a_1$ ، $(k \leq 7)$ را باقیماندهٔ تقسیم حاصلضرب عددی $(a_k, a_{k-1}, \ldots, a_1) \times (k, k-1, \ldots, 2, 1)$ بر عدد ۱۰ در نظر می گیرد. به این ترتیب کلیه اشتباهات تک رقمی در مکانهایی با وزنهای ۱ ، ۳ یا ۷ و همچنین اشتباه مربوط به رقم بازرس قابل کشفند. اشتباهاتی به شکل $b \to a$ و در صورتی که $|a-b|=5$ و در مکانهایی که وزن آنها ۲ ، ۴ یا ۶ است، کشف نمی شوند. همچنین

مربوط به جابجایی ارقام که شامل رقم بازرس نباشند، قابل کشفند. اشتباهاتی از نوع ..ab → ba.. شامل رقم بازرس وقتی که باقیماندهٔ تقسیم $(a_6, a_5, a_4, a_3, a_2, a_1) \times (۷, ۶, ۵, ۴, ۳, ۲)$ بر عدد ۱۰، ۵ است، کشف نمی‌شوند. نرخ کشف اشتباهات جابجایی این روش ۹۸/۴٪ است [۲].

در سیستم بانکداری ایالات متحده، شماره حسابهای بانکی معمولاً ۸ رقمی به صورت $a_8 a_7 \dots a_1$ است. در این کدها رقم بازرس باقیماندهٔ تقسیم

$$C = (a_8, a_7, \dots, a_1) \times (۷, ۳, ۹, ۷, ۳, ۹, ۷, ۳)$$

بر عدد ۱۰ است. مثلاً، رقم بازرس شماره حساب ۰۹۱۹۰۲۰۴ باقیماندهٔ تقسیم $C = ۰+۰+۲۷+۹+۶۳+۰+۱۸+۰+۰+۱۲ = ۱۳۹$ بر ۱۰، یعنی ۹ است. در این روش صد درصد اشتباهات تک رقمی و ۸۸/۹٪ اشتباهات جابجایی قابل کشفند. لازم به ذکر است که تنها اشتباهات جابجایی غیر قابل کشف به صورت ab → ba وقتی است که $|a - b| = ۵$. مزیت این روش وزن دهی در مقایسه با روشهایی که تنها از دو رقم متمایز مثلاً $(۷, ۳, ۷, ۳, ۷, ۳, ۷, ۳)$ استفاده می کنند، آن است که در روش اول بیشتر اشتباهات از نوع ..abc.. → ..cba.. کشف می شوند، در حالی که در روش دوم هیچ یک از این اشتباهات قابل کشف نیستند، مگر آن که شامل رقم بازرس باشند [۲].

روش کاراتر روشی است که در اغلب کتابخانه ها از آن استفاده می کنند [۱]. در این روش کد اطلاعاتی ۱۳ رقمی به صورت $a_{13}, a_{12}, \dots, a_1$، و رقم بازرس ماندهٔ تقسیم

$$- (a_{13}, a_{12}, \dots, a_1) \times (۲, ۱, ۲, ۱, ۲, ۱, ۲, ۱, ۲, ۱, ۲, ۱, ۲) - r$$

بر عدد ۱۰ مطابق روش معمول در چکهای مسافرتی ویزاست، که در آن r تعداد ارقام بزرگتر یا مساوی ۵ در بین ارقام $a_1, a_3, a_5, a_7, a_9, a_{11}, a_{13}$ می باشد. مثلاً، رقم بازرس کد اطلاعاتی ۳۱۲۵۶۰۰۱۹۶۴۳۱ برابر است با باقیماندهٔ تقسیم :

$$- (۶+۱+۴+۵+۱۲+۰+۰+۱+۱۸+۶+۸+۳+۲) - ۲ = -۶۸$$

بر ۱۰ که عدد ۲ است. در این روش کلیهٔ اشتباهات تک رقمی و اشتباهات جابجایی به جز ۹۰ → ۰۹ کشف می

$. 0 \leq c$

یکی از مسائل مهم در تدوین روشهای بازرسی برای اطمینان از صحت کدهای اطلاعاتی، انتخاب عدد مناسب به عنوان پیمانه است. با توجه به عملکرد روشهایی که عدد بازرس آنها تک رقمی است، معلوم می شود که عدد بازرس تک رقمی قادر به کشف همهٔ اشتباهات مورد نظر نیست. از این رو برای کارایی بیشتر این روشها بهتر است تا از دو رقم بازرس استفاده شود. در بخش بعد چگونگی انتخاب پیمانه مناسب را برای کارایی بیشتر ارقام بازرس در کشف اشتباهات کدهای اطلاعاتی شرح می دهیم.

۴ - انتخاب پیمانه برای کارایی بیشتر عدد بازرس

ر این بخش هدف ما بررسی قابلیت ارقام بازرس در کشف اشتباهات کدهای اطلاعاتی اعم از اشتباهات تک رقمی، و رقمی ناشی از جابجایی یا تغییر ارقام و برخی از اشتباهات سه رقمی با توجه به عدد پیمانه است. مقدمتاً در جدول ۱ تمام اعداد دو رقمی را به عوامل اول تجزیه می کنیم.

ا استفاده از نمادهای بخش ۲، عدد بازرس کد اطلاعاتی ده رقمی N به صورت $a_9 a_8 a_7 - a_6 a_5 a_4 a_3 a_2 a_1 a_0$ ا c فرض می کنیم، به قسمی که، (به پیمانهٔ m) $N \equiv c$ ، $0 \leq c \leq m-1$ و کد اطلاعاتی مخدوش شده مراه با عدد بازرس را با

$$a'_9a'_8a'_7 - a'_6a'_5a'_4a'_3a'_2a'_1a'_0 - c'_1c'_0$$

شان می دهیم که در آن

$$N' = \Sigma\, a'_i \times 10^i \,, \qquad c' = \Sigma\, c'_p \times 10^p \,.$$

شن است که وقتی، (به پیمانهٔ m) $N - N' \equiv c - c'$ ، اشتباهات موجود غیر قابل کشف باقی می مانند. هر ه تعداد جوابهای این معادلهٔ همنهشتی کمتر باشد، پیمانهٔ m عملکرد بهتری دارد. بنابراین لازم است تا تعداد وابهای این معادلهٔ همنهشتی را در حالتهای مختلف وجود اشتباه، بررسی کنیم.

جدول ۱ – تجزیه اعداد دو رقمی به عوامل اول

۱۱	$۱۲ = ۲^{۲} \times ۳$	۱۳	$۱۴ = ۲ \times ۷$	$۱۵ = ۳ \times ۵$
$۱۶ = ۲^{۴}$	۱۷	$۱۸ = ۲ \times ۳^{۲}$	۱۹	$۲۰ = ۲^{۲} \times ۵$
$۲۱ = ۳ \times ۷$	$۲۲ = ۲ \times ۱۱$	۲۳	$۲۴ = ۲^{۳} \times ۳$	$۲۵ = ۵^{۲}$
$۲۶ = ۲ \times ۱۳$	$۲۷ = ۳^{۳}$	$۲۸ = ۲^{۲} \times ۷$	۲۹	$۳۰ = ۲ \times ۳ \times ۵$
۳۱	$۳۲ = ۲^{۵}$	$۳۳ = ۳ \times ۱۱$	$۳۴ = ۲ \times ۱۷$	$۳۵ = ۷ \times ۵$
$۳۶ = ۲^{۲} \times ۳^{۲}$	۳۷	$۳۸ = ۲ \times ۱۹$	$۳۹ = ۳ \times ۱۳$	$۴۰ = ۲^{۳} \times ۵$
۴۱	$۴۲ = ۲ \times ۳ \times ۷$	۴۳	$۴۴ = ۲^{۲} \times ۱۱$	$۴۵ = ۳^{۲} \times ۵$
$۴۶ = ۲ \times ۲۳$	۴۷	$۴۸ = ۲^{۴} \times ۳$	$۴۹ = ۷^{۲}$	$۵۰ = ۲ \times ۵^{۲}$
$۵۱ = ۳ \times ۱۷$	$۵۲ = ۲^{۲} \times ۱۳$	۵۳	$۵۴ = ۲ \times ۳^{۳}$	$۵۵ = ۵ \times ۱۱$
$۵۶ = ۲^{۳} \times ۷$	$۵۷ = ۳ \times ۱۹$	$۵۸ = ۲ \times ۲۹$	۵۹	$۶۰ = ۲^{۲} \times ۳ \times ۵$
۶۱	$۶۲ = ۲ \times ۳۱$	$۶۳ = ۳^{۲} \times ۷$	$۶۴ = ۲^{۶}$	$۶۵ = ۵ \times ۱۳$
$۶۶ = ۲ \times ۳ \times ۱۱$	۶۷	$۶۸ = ۲^{۲} \times ۱۷$	$۶۹ = ۳ \times ۲۳$	$۷۰ = ۲ \times ۵ \times ۷$
۷۱	$۷۲ = ۲^{۳} \times ۳^{۲}$	۷۳	$۷۴ = ۲ \times ۳۷$	$۷۵ = ۳ \times ۵^{۲}$
$۷۶ = ۲^{۲} \times ۱۹$	$۷۷ = ۷ \times ۱۱$	$۷۸ = ۲ \times ۳ \times ۱۳$	۷۹	$۸۰ = ۲^{۴} \times ۵$
$۸۱ = ۳^{۴}$	$۸۲ = ۲ \times ۴۱$	۸۳	$۸۴ = ۲^{۲} \times ۳ \times ۷$	$۸۵ = ۵ \times ۱۷$
$۸۶ = ۲ \times ۴۳$	$۸۷ = ۳ \times ۲۹$	$۸۸ = ۲^{۳} \times ۱۱$	۸۹	$۹۰ = ۲ \times ۳^{۲} \times ۵$
$۹۱ = ۷ \times ۱۳$	$۹۲ = ۲^{۲} \times ۲۳$	$۹۳ = ۳ \times ۳۱$	$۹۴ = ۲ \times ۴۷$	$۹۵ = ۵ \times ۱۹$
$۹۶ = ۲^{۵} \times ۳$	۹۷	$۹۸ = ۲ \times ۷^{۲}$	$۹۹ = ۳^{۲} \times ۱۱$	

۳ – ۱ – اشتباهات تک رقمی

(الف) وجود اشتباه در رقم a_i $(۰ \le i \le ۹)$ کد اطلاعاتی.

در این حالت داریم:

$$N - N' = (a_i - a_i')\, ۱۰^{i} = \alpha\, ۱۰^{i},$$

که در آن $\alpha = (a_i - a_i')$ ، و $۱ \le |\alpha| \le ۹$. وقتی $\alpha\, ۱۰^{i} = ۰$ (به پیمانه m) ، یا هم ارز آن وقتی $m \mid \alpha\, ۱۰^{i}$ (یعنی وقتی m مقسوم علیه $\alpha\, ۱۰^{i}$ است) ، یا به عبارت دیگر وقتی $m = \gamma ۲^{\beta} ۵^{\delta}$ که در آن $\gamma \mid \alpha$ ، $۱ \le \gamma \le ۹$ ، $\beta, \delta \le i$ = اشتباه تک رقمی قابل کشف نیست.

بنابراین برای کشف اشتباهات تک رقمی لازم است تا از جدول ۱ تمام اعدادی را که به صورت $\gamma ۲^{\beta} ۵^{\delta}$ هست

$35 = 7 \times 5$	$36 = 2^2 \times 3^2$	$40 = 2^3 \times 5$	$45 = 3^2 \times 5$	$48 = 2^4 \times 3$
$50 = 2 \times 5^2$	$56 = 2^3 \times 7$	$60 = 2^2 \times 3 \times 5$	$64 = 2^6$	$70 = 2 \times 5 \times 7$
$72 = 2^3 \times 3^2$	$75 = 3 \times 5^2$	$80 = 2^4 \times 5$	$90 = 2 \times 3^2 \times 5$	$96 = 2^5 \times 3$

(ب) وجود اشتباه در رقم c_p $(0 \leq p \leq 1)$ کد بازرس.

در این حالت داریم:

$$(c - c') = \beta\, 10^p ,$$

که در آن $\beta = c_p - c'_p$، $1 \leq |\beta| \leq 9$.

این اشتباه وقتی (به پیمانهٔ m) $\beta\, 10^p = 0$، کشف نمی شود. این معادلهٔ همنهشتی همان معادلهٔ همنهشتی بند (الف) است. بنابراین اعداد پیمانه‌ای که قابلیت کشف اشتباهات تک رقمی را در کد اطلاعاتی دارند، قادر به کشف اشتباهات مربوط به رقم بازرس نیز می باشند.

۳ - ۲ - اشتباهات دو رقمی ناشی از جابه جایی ارقام

(الف) جابه جایی ارقام a_i و a_j $(0 \leq i, j \leq 9 \,,\, i \neq j)$ در کد اطلاعاتی.

در این حالت داریم:

$$\begin{aligned} N - N' &= a_i 10^i + a_j 10^j - (a_j 10^i + a_i 10^j) \\ &= (a_i - a_j)(10^i - 10^j) \\ &= \alpha\, 10^j (10^k - 1) \end{aligned}$$

که در آن j را کوچکتر از i، $k = i - j$ و $\alpha = a_i - a_j$ فرض کرده‌ایم.

این نوع اشتباه وقتی (به پیمانهٔ m) $\alpha\, 10^j (10^k - 1) = 0$ که در آن $1 \leq |\alpha| \leq 9$، $0 \leq j \leq 8$ و $1 \leq k \leq 9$ کشف نمی شود. این بدان معناست که وقتی $m = \gamma 2^r 5^s \delta$ که در آن $\gamma | \alpha$، $0 \leq r, s \leq 8$، $\delta | 10^k - 1$ و $\delta < 100$، این اشتباه کشف نمی شود. در جدول ۳ به ازاء مقادیر $1 \leq k \leq 9$، اعداد $10^k - 1$ را به عوامل اول تجزیه کرده‌ایم.

$$۱۰^{۲} - ۱ = ۳^{۲} \times ۱۱$$
$$۱۰^{۳} - ۱ = ۳^{۳} \times ۳۷$$
$$۱۰^{۴} - ۱ = ۳^{۲} \times ۱۱ \times ۱۰۱$$
$$۱۰^{۵} - ۱ = ۳^{۲} \times ۴۱ \times ۲۷۱$$
$$۱۰^{۶} - ۱ = ۳^{۳} \times ۷ \times ۱۱ \times ۱۳ \times ۳۷$$
$$۱۰^{۷} - ۱ = ۳^{۲} \times ۲۳۹ \times ۴۶۴۹$$
$$۱۰^{۸} - ۱ = ۳^{۲} \times ۱۱ \times ۷۳ \times ۱۰۱ \times ۱۳۷$$
$$۱۰^{۹} - ۱ = ۳^{۴} \times ۳۷ \times ۳۳۳۶۶۷$$

با توجه به جدولهای ۱ و ۳ اعداد زیر نیز از فهرست نامزدهای مناسب برای عدد پیمانه حذف می شوند.

جدول ۴ - اعداد دو رقمی نامزد پیمانه که قابلیت کشف اشتباهات ناشی از جابجاشدن دو رقم در کد اطلاعاتی ندارند.

۱۱	۱۳	$۲۱ = ۳ \times ۷$	$۲۲ = ۲ \times ۱۱$	$۲۶ = ۲ \times ۱۳$
$۲۷ = ۳^{۳}$	$۳۳ = ۳ \times ۱۱$	۳۷	$۳۹ = ۳ \times ۱۳$	۴۱
$۴۲ = ۲ \times ۳ \times ۷$	$۴۴ = ۲^{۲} \times ۱۱$	$۴۹ = ۷^{۲}$	$۵۲ = ۲^{۲} \times ۱۳$	$۵۴ = ۲ \times ۳^{۳}$
$۵۵ = ۵ \times ۱۱$	$۶۳ = ۳^{۲} \times ۷$	$۶۵ = ۵ \times ۱۳$	$۷۷ = ۷ \times ۱۱$	$۷۸ = ۲ \times ۳ \times ۱۳$
$۸۱ = ۳^{۴}$	$۸۲ = ۲ \times ۴۱$	$۸۴ = ۲^{۲} \times ۳ \times ۷$	$۸۸ = ۲^{۳} \times ۱۱$	$۹۱ = ۷ \times ۱۳$
$۹۸ = ۲ \times ۷^{۲}$	$۹۹ = ۳^{۲} \times ۱۱$			

(ب) جابجایی ارقام کد بازرسی $c_۰$ و $c_۱$. در این حالت داریم:

$$c - c' = ۱۰ (c_۱ - c_۰) = \alpha ۱۰, \qquad ۱ \le | \alpha | \le ۹$$

چنین اشتباهی وقتی (به پیمانهٔ m) $\alpha ۱۰ = ۰$ کشف نمی شود. معادلهٔ همنهشتی فوق همان معادلهٔ همنهشتم بندهای (الف) و (ب) بخش ۳-۱ است. بنابراین پیمانه هایی که قابلیت کشف اشتباهات تک رقمی را دارند قادر به کشف اشتباهات مربوط به جابجایی ارقام کد بازرسی نیز هستند.

(ج) جابجایی ارقام a_i و c_p $(۰ \le i \le ۹ ,\ ۰ \le p \le ۱)$ در کد اطلاعاتی و بازرسی.

شوند. در جدول ۵ به ازاء $1 \le n \le 9$ ، اعداد $10^n + 1$ را به عوامل اول تجزیه کرده‌ایم.

جدول ۵ - تجزیه اعداد $10^n + 1$ به ازاء $1 \le n \le 9$ به عوامل اول.

$$10^1 + 1 = 11$$
$$10^2 + 1 = 101$$
$$10^3 + 1 = 7 \times 11 \times 13$$
$$10^4 + 1 = 73 \times 137$$
$$10^5 + 1 = 11 \times 9091$$
$$10^6 + 1 = 101 \times 9901$$
$$10^7 + 1 = 11 \times 909091$$
$$10^8 + 1 = 17 \times 5882353$$
$$10^9 + 1 = 7 \times 11 \times 13 \times 19 \times 52579$$

با توجه به مقسوم علیه های اعداد جدولهای ۱ و ۵ اعداد جدول ۶ نیز از فهرست نامزدهای عدد پیمانه حذف می شوند.

جدول ۶ - اعداد دو رقمی نامزد پیمانه که قابلیت کشف اشتباهات ناشی از جابجایی ارقام کد اطلاعاتی و کد بازرس را ندارند.

11	13	17	19	$34 = 2 \times 17$
$38 = 2 \times 19$	$42 = 2 \times 3 \times 7$	$44 = 2^2 \times 11$	$49 = 7^2$	$51 = 3 \times 17$
$52 = 2^2 \times 13$	$55 = 5 \times 11$	$56 = 2^3 \times 7$	$57 = 3 \times 19$	$68 = 2^2 \times 17$
$76 = 2^2 \times 19$	$77 = 7 \times 11$	$84 = 2^2 \times 3 \times 7$	$85 = 5 \times 17$	$95 = 5 \times 19$

با حذف اعداد جدولهای ۲ ، ۴ ، و ۶ از جدول ۱ ، مجموعهٔ S حاصل می شود که شامل تمام پیمانه های دو رقمی است که قابلیت کشف اشتباهات تک رقمی، دو رقمی ناشی از جابجایی اعم از جابجایی در کد اطلاعاتی یا عدد بازرس و همچنین جابجایی ارقام کد اطلاعاتی با ارقام عدد بازرس را دارند. روشن است که از این میان، پیمانه های مناسبتر آنهایی هستند که قابلیت کشف تغییرات دو رقمی را نیز داشته باشند.

(الف) تغییر در ارقام a_i و a_j $(0 \le i, j \le 9, i \ne j)$ کد اطلاعاتی.

در این حالت داریم:

$$N - N' = (a_i - a'_i)10^i + (a_j - a'_j)10^j$$

با قرار دادن $k = a_i - a'_i$ و $l = a_j - a'_j$ می‌بینیم که وقتی (به پیمانهٔ m) $k10^i + l10^j = 0$ ، $1 \le |k|, |l| \le 9$ ، چنین اشتباهی کشف نمی‌شود.

(پ) تغییر در ارقام c_0 و c_1 کد بازرس. در این حالت داریم:

$$c - c' = 10(c_1 - c'_1) + (c_0 - c'_0) = 10\alpha + \beta$$

که در آن، $\alpha = c_1 - c'_1$ ، $\beta = c_0 - c'_0$ ، و $1 \le |\alpha|, |\beta| \le 9$. این اشتباه وقتی (به پیمانهٔ m) $10\alpha + \beta = 0$ کشف نمی‌شود. این معادلهٔ همنهشتی حالت خاصی از معادلهٔ همنهشتی بند (الف) است. بنابراین اگر عدد بازرس قابلیت کشف اشتباهات ناشی از تغییر ارقام در کد اطلاعاتی را داشته باشد، قابلیت کشف اشتباهات مربوط به تغییر ارقام در عدد بازرس را نیز دارد.

(د) تغییر در ارقام a_i و c_p $(0 \le i \le 9, 0 \le p \le 1)$ کد اطلاعاتی و کد بازرس.

در این حالت داریم:

$$N - N' = (a_i - a'_i)10^i = \alpha 10^i$$

که در آن $\alpha = a_i - a'_i$ ، $1 \le |\alpha| \le 9$ ، و

$$c - c' = (c_p - c'_p)10^p = \beta 10^p$$

که در آن $\beta = c_p - c'_p$ ، $1 \le |\beta| \le 9$.

و کارآمدترین پیمانه ها را از بین اعداد مجموعه S استخراج می کنیم. راه حل کلی تعیین جوابهای معادلات همنهشتی خطی را در [۱] و [۴] ببینید.

اساس بررسی معادلهٔ همنهشتی بالا در تعیین پیمانهٔ m، احتمال عدم کشف اشتباهات دو رقمی، p_m است. لذا مطلوبترین عدد به عنوان پیمانه، عددی از مجموعهٔ S است که برای آن p_m مینیمم باشد. احتمال عدم کشف اشتباهات دو رقمی را به صورت زیر تعریف می کنیم.

تعریف - احتمال عدم کشف اشتباهات دو رقمی، P_m برابر است با خارج قسمت تعداد جوابهای معادلهٔ همنهشتی، $k10^i + l10^j = 0$ (به پیمانهٔ m) با $1 \le |k|, |l| \le 9$ ، $0 \le j \ne i \le 9$ ، بر حاصلضرب تعداد مقادیر ممکن k، l و زوج مرتب (i, j).

در قضیهٔ زیر ثابت می کنیم که برای هر $m \in S$ داریم:

$$P_m = \begin{cases} \dfrac{1}{m-1} , & \text{اگر ب.م.م } (m, 10) = 1 \\ \dfrac{16}{9(m-1)} , & \text{اگر } m = 2m' \text{ و ب.م.م } (m', 10) = 1 \\ \dfrac{4}{135}\left[\dfrac{100}{m-4} + \dfrac{2}{3(m-2)}\right] & \text{اگر } m = 4m'' \text{ و ب.م.م } (m'', 10) = 1 \end{cases}$$

به منظور انتخاب پیمانهٔ بهینه، در جدول ۷ با محاسبهٔ P_m، قابلیتهای هر یک از پیمانه ها را در کشف اشتباهات دو رقمی ناشی از تغییر ارقام مقایسه کرده‌ایم. ملاحظه می شود که بزرگترین عدد اول دو رقمی علاوه بر آن که قابلیت کشف کلیهٔ اشتباهات تک رقمی و دورقمی جابجایی را دارد، توان آن نیز در کشف اشتباهات ناشی از تغییر دو رقم حداکثر است.

جدول ۷ - احتمال ناتوانی اعداد مجموعهٔ S به عنوان پیمانه در کشف اشتباهات دو رقمی.

احتمال	m	احتمال	m
۰/۰۱۴۷۱	۶۹	۰/۰۴۵۴۵	۲۳
۰/۰۱۴۲۹	۷۱	۰/۰۳۵۷۱	۲۹
۰/۰۱۲۸۲	۷۹	۰/۰۳۳۳۳	۳۱
۰/۰۱۲۱۹	۸۳	۰/۰۲۳۸۱	۴۳
۰/۰۲۰۹۱	۸۶	۰/۰۳۹۵۱	۴۶
۰/۰۱۱۶۳	۸۷	۰/۰۲۱۷۴	۴۷
۰/۰۱۱۳۴	۸۹	۰/۰۱۹۲۳	۵۳
۰/۰۳۳۸۹	۹۲	۰/۰۳۱۱۹	۵۸

$\frac{\ }{m-1}$ اگر ب.م.م $(m, 10) = 1$

$P_m = \frac{16}{9(m-1)}$ اگر $m = 2m'$ و ب.م.م $(m', 10) = 1$

$\frac{4}{135}\left[\frac{100}{m-4} + \frac{2}{3(m-2)}\right]$ اگر $m = 4m''$ و ب.م.م $(m'', 10) = 1$

اثبات – لازم است تا تعداد جوابهای معادلهٔ $k10^i + l10^j = 0$ (به پیمانهٔ m) را که در آن، $|k|, |l| \le 9$ ، $0 \le i \ne j \le 9$ تعیین کنیم. بدون آن که از کلیت مطلب کاسته شود، فرض می کنیم $j < i$ و $0 < l$. اگر m نسبت به ۱۰ اول باشد، آن گاه معادله $k10^i + l10^j = 0$ (به پیمانهٔ m)، هم ارز معادلهٔ $k10^{i-j} + l = 0$ (به پیمانهٔ m) است. هر بار i، j و l را ثابت گرفته و احتمال وجود جوابی مانند k به قسمی که $|k| \le 9$ محاسبه می کنیم.

برای مقادیر ثابت i، j و $l \ne 0$ این معادله دارای یک جواب k، $0 < k < m$ است. احتمال این که $|k| \le 9$، برابر است با $\frac{18}{m-1}$. و احتمال این که این جواب نیز برابر اشتباه بوجود آمده در کد اطلاعاتی a'_i باشد، $\frac{1}{18}$ است. بنابراین $P_m = \frac{1}{m-1}$.

اگر $m = 2m'$ و m' نسبت به ۱۰ اول باشد، آن گاه، برای مقادیر $1 \le i \le 9$ و $j = 0$ (۹ حالت)، همنهشتی $10^i k + l = 0$ (به پیمانهٔ m) تنها و تنها وقتی دارای جواب است که $l \mid$ ب.م.م $(m, 10^i)$ یا $l = 2l'$، در این حالت داریم:

$$\left(\frac{10^i}{2}\right) k + l' = 0 \quad (\text{به پیمانهٔ } m')$$

که دارای یک جواب $0 < k < m'$ است. بنابراین احتمال P_m در این حالت برابر است با:

$$P_m = \frac{9}{45} \times \frac{4}{9} \times \frac{1}{m'-1}$$

برای مقادیر $1 \le j < i \le 9$ (۳۶ حالت) معادلهٔ $k10^i + l10^j = 0$ (به پیمانهٔ m)، هم ارز معادلهٔ $\frac{10^i}{2}k + \frac{10^j}{2}l = 0$ (به پیمانهٔ m') است که دارای یک جواب $0 < k < m'$ است. بنابراین $P_m = \frac{36}{45} \times \frac{1}{m'-1}$. پس در که اولاً $m = 2m'$ و ثانیاً m' نسبت به ۱۰ اول باشد، نتیجه می شود که:

$$P_m = \frac{8}{9} \times \frac{1}{m'-1} = \frac{16}{9} \times \frac{1}{m-2}$$

اگر $m = 2m' = 4m''$، و m'' نسبت به ۱۰ اول باشد، آن گاه، برای مقادیر $i = 1$ و $j = 0$ (یک حالت)، معادلهٔ $10k + l = 0$ (به پیمانهٔ m) تنها و تنها وقتی دارای جواب است که $l \mid$ ب.م.م $(m, 10)$ یا $l = 2l'$، و در این داریم

$$\frac{10}{4}k + l'' \equiv 0 \ (m'' \text{ به پیمانهٔ })$$

که دارای یک جواب $0 < k < m''$ است. بنابراین احتمال P_m برابر است با:

$$P_m = \frac{8}{45} \times \frac{2}{9} \times \frac{1}{m''-1}$$

برای مقادیر $9 \geq i \geq 2$ و $j = 1$ (۸ حالت) معادلهٔ (به پیمانهٔ m) $10^i k + 10 l \equiv 0$ تنها و تنها وقتی دارای جواب است که $10 \mid (m, 10^i)$ بهمهم یا $l = 2l'$. داریم:

$$\frac{10^i}{4}k + 5l' \equiv 0 \ (m'' \text{ به پیمانهٔ })$$

که دارای یک جواب $0 < k < m''$ است. بنابراین احتمال P_m برابر است با:

$$P_m = \frac{8}{45} \times \frac{4}{9} \times \frac{1}{m''-1}$$

بالاخره برای مقادیر $9 \geq i > j \geq 2$ (۲۸ حالت) معادلهٔ (به پیمانهٔ m) $10^i k + 10^j l \equiv 0$، هم ارز معادلهٔ (به پیمانهٔ m'') $\frac{10^i}{4}k + \frac{10^j}{4}l \equiv 0$ و دارای یک جواب $0 < k < m''$ است. پس احتمال P_m در این حالت برابر است با:

$$P_m = \frac{28}{45} \times \frac{1}{m''-1}$$

بنابراین وقتی $m = 4m''$، و m'' نسبت به ۱۰ اول است، داریم:

$$P_m = \frac{1}{45}\left[\frac{1}{m''-1}\left(28 + \frac{16}{9} + \frac{32}{9}\right) + \frac{4}{9} \times \frac{1}{m'-1}\right]$$

$$= \frac{1}{45}\left[\frac{100}{3} \times \frac{1}{m''-1} + \frac{4}{9} \times \frac{1}{m'-1}\right]$$

$$= \frac{4}{135}\left[\frac{100}{m-4} + \frac{2}{3 \times (m-2)}\right]$$

در بخش زیر روشی معرفی می کنیم که در آن عدد بازرس دو رقمی است، اما هر رقم از آن به صورتی خاص تعیین می شوند. نشان خواهیم داد که این روش علاوه بر کشف اشتباهات تک رقمی و دو رقمی ناشی از جابجایی و تغییر ارقام، قادر به کشف برخی اشتباهات سه رقمی نیز می باشد.

۴ – روش پیشنهادی

روشی که ذیلاً شرح می دهیم قابلیت کشف کلیهٔ اشتباهات تک رقمی را دارد. علاوه بر آن قادر است تا اشتباهات دو رقمی و سه رقمی ناشی از تغییر یک رقم عدد بازرس و جابجایی دو رقم کد اطلاعاتی را کشف کند. کد اطلاعاتی می

ارقام اطلاعاتی را با a_i، $0 \le a_i \le 9$، $1 \le i \le 10$ و ارقام بازرس را با c_p ، که در آن $1 \le p \le 2$ نشان می دهیم ارقام بازرس را به صورت زیر تعریف می کنیم:

$$c_2 \equiv \sum i\, a_i \ (\text{به پیمانه } 11),$$
$$c_1 \equiv \sum a_i \ (\text{به پیمانه } 11).$$

در حالتهای مختلف، وجود اشتباهات تک رقمی، دو و سه رقمی که ممکن است عدد $a_{10}a_9\ldots a_1c_2c_1$ مخدوش کنند، بحث می کنیم و نشان می دهیم که روش پیشنهادی قادر به کشف کلیهٔ این اشتباهات است.

الف) اشتباهات تک رقمی

فرض کنید رقم a_k در کد اطلاعاتی، $1 \le k \le 10$ به a'_k تغییر کرده باشد. این تغییر را با $a_k \to a'_k$ نشان دهیم. به سادگی معلوم می شود که c_1 و همچنین c_2 قادر به کشف چنین اشتباهی هستند.

ب) اشتباهات ناشی از جابجایی دو رقم

فرض کنید ارقام a_k و a_h ($1 \le k < h \le 10$) جابجا شوند. این جابجایی را با $a_k \leftrightarrow a_h$ نشان می دهیم. این نوع اشتباه از طریق رقم بازرس c_1 کشف می شود، زیرا معادلهٔ همنهشتی

$$ka_k + ha_h = ka_h + ha_k \ (\text{به پیمانهٔ } 11),$$

که هم ارز معادلهٔ

$$(k - h)(a_k - a_h) = 0 \ (\text{به پیمانهٔ } 11)$$

است، و این معادله برای $1 \le |k - h|$ ، $|a_k - a_h| \le 9$ جوابی ندارد.

وقتی یکی از ارقام کد اطلاعاتی، a_k با یکی از ارقام عدد بازرسی، c_1 (یا c_2) جابجا می شود، این اشتباه را می توان اشتباه تک رقمی $a_k \to a'_k = c_1$ تلقی کرد. و بنابراین رقم c_2 (یا c_1) چنین اشتباهی را کشف می کند. وقتی c_1 و c_2 جابجا می شوند، با کنترل دو مجموع در رابطهٔ تعریفی c_1 و c_2 و ملاحظهٔ عدم موازنهٔ آنها چنین اشتباهی آشکار می شود.

ج) اشتباهات دو رقمی

وقتی ارقامی به اشتباه جای دو رقم بازرس یعنی c_1 و c_2 را می گیرند، این اشتباه به دلیل عدم برقراری

$$\alpha_k = a'_k - a_k, \quad \alpha_h = a'_h - a_h$$

که در آن $1 \le |\alpha_k|, |\alpha_h| \le 9$.

چنین اشتباهی مادام که $k\alpha_k + h\alpha_h = 0$ (به پیمانهٔ 11)، از طریق c_2 قابل کشف نیست. و مادام که $\alpha_k + \alpha_h = 0$ (به پیمانهٔ 11)، از طریق c_1 نیز قابل کشف نیست. از آنجا که بنا بر لم زیر دستگاه معادلات

$$k\alpha_k + h\alpha_h = 0 \ (\text{به پیمانهٔ } 11),$$

$$\alpha_k + \alpha_h = 0 \ (\text{به پیمانهٔ } 11),$$

$$1 \le |\alpha_k|, |\alpha_h| \le 9, \quad 1 \le k < h \le 10$$

جوابی ندارد، لذا تمام این قبیل اشتباهات دو رقمی از طریق ارقام بازرس c_1 و یا c_2 قابل کشفند.

بنابراین روش پیشنهادی قادر به کشف کلیهٔ اشتباهات یک رقمی و دو رقمی است.

لم: فرض کنید a, b, c, d چهار عدد صحیح باشند که همگی نسبت به عدد صحیح m اول‌اند. اعداد مخالف صفر x و y به پیمانهٔ m جواب دستگاه معادلات همنهشتی

$$ax + by = 0 \ (\text{به پیمانهٔ } m),$$

$$cx + dy = 0 \ (\text{به پیمانهٔ } m),$$

است، اگر و تنها اگر $ad = bc$ (به پیمانهٔ m). اثبات این لم سر راست است، و از ارائه آن صرف نظر می کنیم.

د) اشتباهات سه رقمی ناشی از تغییر یک رقم و جابه جایی دو رقم

به سادگی با توجه به بندهای (الف) و (ب) معلوم می شود که این قبیل اشتباهات سه رقمی با کنترل مجموعهای c_1 و c_2 قابل کشف‌اند.

۵ – نتیجه

در این مقاله پس از توجیه لزوم کدبندی اطلاعات و اطمینان از صحت آنها روشهای موجود بازرسی صحت کدهای اطلاعاتی را مورد نقد و بررسی قرار داده و نقاط قوت و ضعف هر یک را برشمرده‌ایم. نشان داده‌ایم که روشهای موجود قادر به کشف کلیهٔ اشتباهات دو رقمی نیستند. روشی را پیشنهاد کرده‌ایم که در آن عدد بازرس دو رقمی است اما هر رقم از آن به صورتی خاص تعیین شده‌اند. نشان داده‌ایم که این روش علاوه بر کشف کلیهٔ اشتباهات [illegible]

[1] Agnew, J. *Explorations in number theory,* Brooks/Cole, Monterey, [illegible]

[2] Gallian J.A. and Winters S. (1988) " *Modular Arithmetic in the Marketplace*". The American Mathematical Monthly, 548-551.

[3] Gumm ,H.P. (1985) " *A new class of check-digit methods for arbitrary number systems*" IEEE Transactions on Information Theory, Vol IT-31,No 1.

[4] LeVeque, W.J. *Fundamentals of number theory,* Addison Wesley, Reading,MA, 1977.

[5] Philip M. Tuchinsky, " *Internationa Standard Book Numbers*" The UMAP journal, 5 (1985) 41 - 54.

[6] Stevens , P. " *Two suggestions to improve on the efficiency of the check computations in the banking system in Belgium*" European Journal of Operations Research 42 (1989) 52-58.

[7] Stevens , P. " *A Two-error-detecting arithmetical check system for decimal identification numbers*" European Journal of Operations Research 52 (1991) 378-381.

ere

$$\alpha_{n:m}(j) = \sum_{k=1}^{\infty} g_{n:m}^{(2k)\widehat{*}}(j), \qquad \beta_{n:m}(j) = \sum_{k=1}^{\infty} g_{n:m}^{(2k-1)\widehat{*}}(j), \tag{3.4}$$

$$g_{n:m}(j) = \begin{cases} (\widehat{\tfrac{h}{h}})(j-m) & , \ j \geq n \\ 0 & , \ j < n \end{cases} \tag{3.5}$$

l for each k the sequence $g_{n:m}^{k\widehat{*}} = \{g_{n:m}^{k\widehat{*}}(j)\}_{j\geq n}$ is formed recursively by $g_{n:m}^{(k-1)\widehat{*}}$ l $g_{n:0}$ through

$$g_{n:m}^{k\widehat{*}}(j) = \sum_{i=n}^{\infty} g_{n:m}^{(k-1)\widehat{*}}(i) g_{n:0}(i+j), \ \ j \geq n, \ \ k \geq 2, \tag{3.6}$$

, $g_{n:m}^{k\widehat{*}} = g_{n:m}^{(k-1)\widehat{*}} \ \widehat{*} \ g_{n:0}$.

Refrences

] N. Aronszajn, *Theory of reproducing kernels,* Trans. Amer. Math. Soc., 68 (1950), pp. 337-404.

] H. Dym and H. P. McKean, *Gaussian Processes, Function Theory, and The Inverse Spectral Problem,* Academic Press, New York, 1976.

] H. Salehi, *On the bilateral linear predictor for minimal stationary stochastic processes,* SIAM J. Appl. Math., 26 (1974), pp. 502-507.

] H. Salehi and A. R. Soltani, *On regularity of homogeneous random fields. Prediction theory and harmonic analysis (eds V. Mandrekar and H. Salehi),* North-Holland publishing company, (1983).

where $\widehat{g}(n) = \frac{1}{2\pi}\int_0^{2\pi} e^{-inx} g(x)dx$ is the Fourier coefficient of g at n. The last equality in (2.7) follows from the fact that $\widehat{(\frac{h}{\bar{h}})}(n)$ is real. It is crucial to notice that for each n,

$$\mathcal{H}_X(-\infty, n] = \mathcal{H}_\xi(-\infty, n], \qquad \mathcal{H}_X[n, +\infty) = \mathcal{H}_\eta[n, +\infty), \tag{2.8}$$

and

$$X_n = \sum_{k=-\infty}^{n} \widehat{h}(n-k)\xi_k, \qquad X_n = \sum_{k=n}^{\infty} \widehat{h}(n-k)\eta_k, \tag{2.9}$$

and

$$\mathcal{H}_X(\,(-n,n)^c) = \mathcal{H}_\xi(-\infty, -n] \vee \mathcal{H}_\eta[n, \infty). \tag{2.10}$$

The ingredients for our interpolation procedure are (2.10) and the von Neumann's alternative projection formula (von Neumann's formula in short). The application of the von Neumann's formula in interpolation is realized by different authors, [3]. For the ease of the readers we present the von Neumann's formula, consult [1] [2].

$$Ph = P_0 h + \sum_{k=1}^{\infty}[P_1(P_2P_1)^{k-1} + P_2(P_1P_2)^{k-1} - (P_2P_1)^k - (P_1P_2)^k]h, \tag{2.11}$$

3 Recipe Formula

In this section we present the best linear interpolator formula for ξ_m, $m = -n+1, ..., n-1$, based on $\{\xi_t, \eta_s,\ t \le -n,\ s \ge n\}$. We present the main result of this section by the following theorem.

Theorem 3.1. Let $\{X_t, t \in Z\}$ be a second order stationary process, $EX_t = 0$. Let $\{\xi_t, t \in Z\}$ and $\{\eta_t, t \in Z\}$ be the (PRP) defined by (2.4), (2.6). Then the best linear interpolator of ξ_m, $m = -n+1, ..., n-1$, based on $\{\xi_t, t \le n\}$ and $\{\eta_t, t \ge n\}$ is given by

$$\widetilde{\xi}_{n:m} = -\sum_{j=-\infty}^{-n} \alpha_{n:m}(-j)\xi_j + \sum_{j=n}^{\infty} \beta_{n:m}(j)\eta_j, \tag{3.2}$$

and the error $\sigma^2 = E|\xi_m - \widetilde{\xi}_{n:m}|^2$ is given by

ctral distribution and $f = F'$ is called the spectral desity of $\mathcal{X}$. It follows
n (2.1) that

$$EX_n\overline{X}_m = \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)x}F(dx). \tag{2.2}$$

process $\mathcal{X}$ is called purely nondeterministic (PND) if $\cap_n \mathcal{H}_X(n) = \{0\}$, ere $\mathcal{H}_X(-\infty, n]$ is the span closure of the elements of $\{X_m, m \leq n\}$ in $\Omega, \mathcal{F}, P)$. Due to the Kolmogorov isomorphism, $X_n \to \exp(-inx)$, $\mathcal{H}_X =$ $(-\infty, +\infty)$ is isometrically isomorphic to $L^2(F)$, where

$$L^2(F) = \{g : \int_0^{2\pi} | g(x) |^2 F(dx) < \infty\}.$$

he process is (PND) then it has a density f for which

$$\int_0^{2\pi} log f(x)dx > -\infty. \tag{2.3}$$

der (2.3) the density f can be factored as $f = | h |^2$, where h is an outer ction of the class of Hardy functions H^2 subject to $\overline{h}(x) = h(-x)$, [2].

Based on h and Φ two distinguish sets of white noises can be constructed follows:

let

$$\beta(\Delta) = \int_\Delta \frac{1}{h} d\Phi, \quad \text{i.e.,} \qquad d\Phi = hd\beta,$$

d define

$$\xi_n = \int_0^{2\pi} e^{-inx} d\beta(x), \quad n \in Z. \tag{2.4}$$

en $\xi = \{\xi_n\}$ is a purely random process (PRP), i.e.

$$E\xi_n\overline{\xi_m} = \begin{cases} 0 & n \neq m \\ 1 & n = m, \end{cases} \tag{2.5}$$

i) let

$$\gamma(\Delta) = \int \frac{1}{\overline{h}} d\Phi, \quad \text{i.e.,} \quad d\Phi = \overline{h}d\gamma,$$

STATIONARY SEQUENCES

A.R. Soltani, F.Yaghmaee

Shiraz University

Abstract: A recipe formula in the time domain of a discrete time second order stationary process $\{X_n\}$ is derived to interpolate the unknown values of $\xi_{-m+1}, \ldots, \xi_{m-1}$ of the noise series $\{\xi_n\}$. In the case that the process is an ARMA model, the coefficients of the model and the observations $\{x_n, |n| > m\}$ are the only inputs of the recipe formula. Error terms are also evaluated.

1 Introduction

A weakly stationary second order process X_n, $n = 0, \pm 1, \pm 2, \ldots$, is the output of a certain linear system receiving a noise sequence $\{\xi_n\}$ as the input. Knowledge on the *past history* $\{X_n, n < 0\}$ is irrelevant in the estimation of ξ_0, as ξ_0 is independent of $\{X_n, n < 0\}$. This declines any extrapolation procedure in predicting ξ_0 based on the observations from the past. Instead, it is relevant to speak of the interpolation of ξ_0 based on the *past and future observations* $\{X_n, n < 0\} \cup \{X_n, n > 0\}$. Our goal is to provide, in the time domain, an interpolating formula for ξ_0 by using $\{X_n, n \neq 0\}$. In 1974 Salehi [3] presented time domain algorithm for interpolating the missing observation X_0. The J.von Neuman's alternating projection theorem is the essential tool in Salehi's procedure. The innovation processes associated with $\{X_n\}$ and $\{X_{-n}\}$ were elegantly employed to apply von Neuman's theorem. For other work of Salehi on interpolation see [4].

In this work we interpolate ξ_0, or more generally $\xi_{-n+1}, \ldots, \xi_{n-1}$. We present a rather compact and easy read formula for the interpolations and the error terms.

References

1] Chayes, J.T., Chayes, L. and Durrett, R. (1988) Connectivity properties of Mandelbrot's percolation process. *Prob. Th. Rel. Fields* **77** 307-324

2] Dekking, M. and Meester, R. (1990) On the structure of Mandelbrot's percolation process and other random Cantor sets. *J. Stat. Phys.* **58** 1109-1126

3] Grimmett, G.R. (1989) *Percolation* (Springer-Verlag, New York).

4] Mandelbrot, B. (1974) Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62** 331-358

5] Orzechowski, M. (1997) *Geometrical and topological properties of fractal percolation.* Ph.D. Thesis, University of St Andrews

6] Roy, R., Sarkar, A. and White, D.G. (1998) Backbends in directed percolation, to appear in *J. Stat. Phys.* **91**

K_n the event $\{A \cap D_n$ contains a cluster enclosing the origin$\}$. For $p < p_c$ the clusters of A are all single points a.s., so the lemma holds trivially (and anyway this case is irrelevant for proving Theorem 1); therefore assume $p \geq p_c$. It follows using standard arguments (for example by applying the previous lemma) that K_n occurs infinitely often almost surely. But on (K_n i.o.), any two unbounded connected subsets of A will both have non-empty intersection with some box B_n for which K_n occurs, and hence will be contained in the same cluster. □

It suffices now to prove:

Lemma 5 *If $p \geq p_c$ then, with probability 1, A has at least one unbounded cluster.*

Proof The proof is a standard renormalisation argument, similar to that in Section 4 of Roy *et al.* [6].

Given a percolation measure μ on the bonds of the classical square lattice L, we say that μ is *1-dependent with parameter q* if it is such that each bond $e \in L$ is open with probability q, and if this is independent of the status of any other bond that has no end-points in common with e. Using an elementary contour argument (analogous to that in the first chapter of Grimmett [3]) it can be shown that there exists a $q_0 < 1$ such that percolation occurs on the lattice μ-almost surely for every such 1-dependent measure μ with parameter $q > q_0$.

For $n \geq 0$ let Λ_n be the event

$$\Lambda_n = \mathcal{H}([0,3^{n+1}] \times [0,3^n]) \cap \mathcal{V}([0,3^n] \times [0,3^n]) \cap \mathcal{V}([2.3^n, 3^{n+1}] \times [0,3^n])$$

(see figure 1). By Lemma 3, we can choose n such that $\mathbf{P}_p(\Lambda_n) > q_0$. As shown in figure 2, we now form a criss-cross grid in $\mathbf{R}^2$ using infinitely many $3^{n+1} \times 3^n$ boxes, and say that a box is *open* if a suitable version of Λ_n occurs within it. Note that the status of two different boxes are independent if the boxes do not overlap. Thus we can think of the grid of open and closed boxes as a 1-dependent percolation model with parameter $\mathbf{P}_p(\Lambda_n)$, which percolates only if A contains at least one unbounded cluster. The lemma follows by our □

following qualitative analogue of the classical RSW Theorem is crucial.

Lemma 2 *If $\mathbb{P}_p(\mathcal{H}([0,1]\times[0,2]))>0$, then $\mathbb{P}_p(\mathcal{H}([0,1]\times[0,2/3]))>0$.*

This lemma is proved in Dekking and Meester [2], although the result they actually state there is slightly weaker than this.

We shall use a fairly well-known corollary for our proof of Theorem 1.

Lemma 3 *If $p\geq p_c$, then $\lim_{n\to\infty}\mathbb{P}_p(\mathcal{H}([0,3^{n+1}]\times[0,3^n]))=1$.*

Proof This proof is instructive as it follows in spirit the proof of Lemma 2 itself, but is considerably more straightforward in its details. Note in particular that the only non-elementary ingredient used is the FKG inequality.

Write $C_n=[0,3^{n+1}]\times[0,3^n]$. Let Γ be the event

$$\Gamma=\bigcap_{i=0}^{6}\mathcal{H}\left([i/3,1+i/3]\times[0,2/3]\right)\cap\bigcap_{i=1}^{6}\mathcal{V}\left([i/3,(i+2)/3]\times[0,1]\right),$$

and note that $\Gamma\subset\mathcal{H}(C_0)$. If $p\geq p_c$ then $\theta(p)>0$, so $\mathbb{P}_p(\mathcal{H}([0,1]\times[0,2]))>0$, and therefore $\mathbb{P}_p(\mathcal{H}([0,1]\times[0,2/3]))>0$ by Lemma 3. By the symmetries of the fractal percolation model and by the FKG inequality, it follows that

$$\mathbb{P}_p(\mathcal{H}(C_0))\geq\mathbb{P}_p(\Gamma)\geq\left(\mathbb{P}_p(\mathcal{H}([0,1]\times[0,2/3]))\right)^{13}>0.$$

To complete the proof we use a scaling argument after Chayes *et al.* [1]. Let $A^{(n)}$ denote the random Cantor set in C_0 obtained by disregarding any removals in the first n iterations of the construction of $A\cap C_0$. Thus $A^{(n)}$ is an increasing sequence of nested sets; moreover, $A^{(n)}$ has precisely the same distribution as $A\cap C_n$ rescaled by a factor 3^n. Therefore, writing $\mathcal{H}^{(n)}$ for the event that there is a horizontal crossing of C_0 in $A^{(n)}$, we have

$$\mathbb{P}_p(\mathcal{H}(C_n))=\mathbb{P}_p\left(\mathcal{H}^{(n)}\right)\uparrow\mathbb{P}_p\left(\bigcup_{n=0}^{\infty}\mathcal{H}^{(n)}\right).$$

By Kolmogorov's Zero-One Law, $\mathbb{P}_p\left(\bigcup_{n=0}^{\infty}\mathcal{H}^{(n)}\right)$ equals either 0 or 1; but this

Damien White

University of Utrecht

Abstract: We give an elementary proof of a result stated by Chayes *et al.* [1] concerning fractal percolation in $\mathbb{R}^2$.

keywords: Cantor set, fractal, percolation, renormalisation.

The process known as fractal percolation was introduced by Mandelbrot [4] in 1974. Informally, we can describe the model as follows. (For a formal description see [1] or [2].) Let $0 \leq p \leq 1$. Divide the unit square $[0,1]^2$ into 9 equal smaller squares, and in the natural way retain each of these squares with probability p, or else remove it with probability $1-p$. Iterating this procedure (suitably scaled) of subdivision and random removal on each of the retained squares yields in the limit a random Cantor set $S \subset [0,1]^2$. Indeed, we can tile the plane with i.i.d. copies of this process to obtain a random set $A \subset \mathbb{R}^2$. Of particular interest are the connected components, or *clusters*, of A.

In 1988, Chayes *et al.* [1] argued that S exhibits a first-order phase transition at some $p_c \in (0,1)$, in a sense that will be made precise below. In the same paper they also stated

Theorem 1 *If $p \geq p_c$ then, with probability 1, A has a unique unbounded cluster.*

The arguments used to support these results took recourse to complicated RSW techniques, and it remained an open problem to find rigorous and elementary proofs. This was achieved for the first-order phase transition by Dekking and Meester [2] in 1990. The purpose of this note is to show how Theorem 1 can be proved in an elementary way.

Given a rectangle $Q = [a_1, a_2] \times [b_1, b_2] \subset \mathbb{R}^2$, let $\mathcal{H}(Q)$ be the event that there is a horizontal crossing of Q in A, that is, there exists a connected set $C \subset Q \cap A$ such that C intersects both edges $\{a_1\} \times [b_1, b_2]$ and $\{a_2\} \times [b_1, b_2]$.

[12] Bickel, P. (1981). Quelques aspects de la statistique robuste. *Lect. Notes Math.* Springer, New York 876, 2–68.

[13] Hampel, F.R., Ronchetti, E.M., Rousseeuw P.J. and Stahel, W.A. (1986). Robust Statistics. The approach based on influence functions. Willey, New York

[14] Huber, D. (1981). Robust statistics. Willey, New York.

[15] Kirillov, A. and Gvishiani, A. (1979). The theorems and problems of functional analysis (Russian). Nauka, Moscow.

References

[1] Lazrieva, N., Toronjadze, T. (1997). Robust estimators in discrete time statistical models. Contiguous alternatives. *Proc. A. Razmadze Math. Inst.* 115, 58–96.

[2] Ibragimov, I. and Khasminskii, R. (1981). Statistical estimation. Asymptotic theory. New York, Springer.

[3] Le Cam, L. and Yang, G. (1990). Asymptotics in Statistics: Some basic concepts.

[4] Shiryaev, A., Spokoiny, V. (1995). Statistical experiments and decisions: Asymptotic theory. VINITI, Moscow.

[5] Chitashvili, R., Lazrieva, N., Toronjadze, T. (1990). Asymptotic theory of M-estimators in general statistical models. Report BS-R9019, 1–31. *Centre for Mathematics and Computer Sciences*, Amsterdam, Netherlands.

[6] Chitashvili, R., Lazrieva, N, .Toronjadze, T. (1990). Asymptotic theory of M-estimators in general statistical models. On asymptotic behaviour of estimators in the presence of nuisance parameters. Report BS-R9020, 1–45. *Centre for Mathematics and Computer Sciences*, Amsterdam, Netherlands.

[7] Jacod, J. (1979). Calcul stochastique et problémes de martingales. *Lect. Notes in Math.* 714, Springer, Berlin, Heidelberg, New York.

[8] Jacod, J. and Shiryaev, A. (1987). Limit theorems for stochastic processes. Springer, New York.

[9] Kabanov, Yu., Liptzer, R. and Shiryaev, A. I(1978), II(1979). Absolute continuity of locally absolutely continuous probability distributions.

$$c_n^2(\theta)\langle N_\theta^{n,\varepsilon}\rangle_T = \sum_\alpha \int (H_\alpha^{n,\varepsilon})^2\, d\widetilde{\mu}_\alpha^n \le$$

$$\le \sum_\alpha \frac{R^2}{\left(\int I_{\{\psi_\alpha^n>\eta_\alpha(1-\varepsilon)\}}d\mu_\alpha^n\right)^2} \int I_{\{u>\eta_\alpha(1-\varepsilon)\}}\widetilde{Q}_\alpha^{n,\psi}(\omega,du) \le K_n \quad P_\theta^n\text{-a.s.}, \tag{85}$$

for some constant K_n, $0<K_n<\infty$, thanks to the (62).

Thus the process $\mathcal{E}\left(c_n(\theta)N_\theta^{n,\varepsilon}\right)$ is square integrable, hence, uniformly integrable martingale. In particular, $E_\theta^n\mathcal{E}\left(c_n(\theta)N_\theta^{n,\varepsilon}\right)=1$.

Moreover, evidently, $c_n(\theta)\Delta N_\theta^{n,\varepsilon} \ge R_n > -1$, for some constant R_n.

Indeed, $H_\alpha^{n,\varepsilon}\ge 0$, $\alpha = c,\pi,b$ and

$$H_\delta^{n,\varepsilon} \ge \frac{-R\displaystyle\int I_{\{\psi_\delta^n>\eta_\delta(1-\varepsilon)\}}q_{\omega,t}^n(dx)}{\displaystyle\int I_{\{\psi_\delta^n>\eta_\delta(1-\varepsilon)\}}d\mu_\delta^n}.$$

Hence if we denote by $\lambda^n(\omega,t)$ the last expression without sign "$-$", we get that $c_n(\theta)\lambda^n(\omega,t)<1$ for some $n\ge 1$, by virtue of condition (3.151) and the fact that $\inf_n \int I_{\{\psi_\delta^n>\eta_\delta(1-\varepsilon)\}}d\mu_\delta^n > 0$. The latter follows from definition of η_δ.

Without loss of generality we assume that $c_n(\theta)\lambda^n(\omega,t)<1$ for any $n\ge 1$.

Therefore $\widetilde{P}_\theta^n := \mathcal{E}\left(c_n(\theta)N_\theta^{n,\varepsilon}\right)\cdot P_\theta^n$ is a probability measure, equivalent for each $n\ge 1$ to the measure P_θ^n, $\widetilde{P}_\theta^n \sim P_\theta^n$.

We prove now that the sequence $\{c_n^2(\theta)\langle N_\theta^{n,\varepsilon}\rangle_T\}_{n\ge 1}$ is stochastically bounded with respect to the sequence of measures $\{\widetilde{P}_\theta^n\}_{n\ge 1}$.

Denote $b_\alpha^n = \int I_{\{\psi_\alpha^n>\eta_\alpha(1-\varepsilon)\}}d\mu_\alpha^n$, $\quad \alpha = c,\pi,\delta,b$.

From (85) we get that for $\forall n\ge 1$ $P_\theta^n\left\{\omega : c_n^2(\theta)\langle N_\theta^{n,\varepsilon}\rangle_T \le K_n\right\}=1$. But, as we prove above, $\widetilde{P}_\theta^n \sim P_\theta^n$. Hence $\widetilde{P}_\theta^n\left\{\omega : c_n^2(\theta)\langle N_\theta^{n,\varepsilon}\rangle_T \le K_n\right\}=1$.

Thus $\overline{\lim}_{n\to\infty}\widetilde{P}_\theta^n\left\{c_n^2(\theta)\langle N_\theta^{n,\varepsilon}\rangle_T > d\right\} \le \overline{\lim}_{n\to\infty}\widetilde{P}_\theta^n\left\{K_n>d\right\}\to 0$ as $d\to\infty$, since the sequence of numbers $\{K_n\}_{n\ge 1}$ is bounded. Indeed, $\inf_n b_\alpha^n>0$ by the definition of η_α and the sequence $\{r_\alpha^n\}_{n\ge 1}$ from (62) is bounded for each $\alpha = c,\pi,\delta,b$.

Now, according to Remark 3, if $\eta_j = \max_\alpha\{\eta_\alpha\}$, as $\overline{H}^{n,\varepsilon}$ we take the func-

Assertion (2) follows. Assertion (3) is an easy corollary of assertion (2). □

Proof of Proposition 3. First note that $1+\chi-\widehat{\chi}>0$, since M is a martingale determining density and $\widetilde{P}\sim P$. Further, from the definition of the Girsanov transform (2) we easily get

$$\Delta_t L^d(m,M)=\frac{\psi(t,\beta_t)I_D-\widehat{\psi}(t)}{1+\chi(t,\beta_t)I_D-\widehat{\chi}(t)}, \tag{82}$$

where $D=\{(\omega,t):\mu(\{t\},E)=1\}$ with the Dirac measure $\mu(\{t\},dx)=\delta_{\beta_t}(dx)$.

On the other hand,

$$\Delta_t\Phi*(\mu-\widetilde{\nu})=\Phi(t,\beta_t)I_D-\widehat{\widetilde{\Phi}}(t), \tag{83}$$

where

$$\widehat{\widetilde{\Phi}}(t)=\int_E\Phi(t,x)\widetilde{\nu}\left(\{t\},dx\right)=\frac{\widehat{\psi}(t)}{1-\widehat{\chi}(t)}\cdot I_{\{\widehat{\chi}(t)<1\}}. \tag{84}$$

Now, substituting (26) and (84) into (83), we obtain $\Delta_t L^d(m,M)=\Delta_t\Phi*(\mu-\nu)$ and, hence, the purely discontinuous part of the $\widetilde{P}$-martingale $L(m,M)$ is equal to $\Phi*(\mu-\widetilde{\nu})$. The continuous part $L^c(m,M)=m^c-\langle m^c,M^c\rangle=\beta\cdot n-\beta\gamma\cdot\langle n\rangle$. □

Proof of Theorem 2. The method of proof of Theorem 2 developed in [1], pp. 89–95 is the general and may be used with small changes in the considered case.

We perform this fact, taking as an example the proof of key Lemma 9 of [1], pp. 94–95.

Introduce the functions

$$H_\alpha^{n,\varepsilon}=\frac{RI_{\{\psi_\alpha^n>\eta_\alpha(1-\varepsilon)\}}}{\int I_{\{\psi_\alpha^n>\eta_\alpha(1-\varepsilon)\}}d\mu_\alpha^n},\quad \alpha=c,\pi,b;$$

$$\widetilde{P}^n_\theta\left(\sup_{t\le T}\alpha_n(t)\geq\eta\right)=\widetilde{P}^n_\theta\left(\sup_{t\le T}(1+c_n(\theta)\Delta N^n_\theta(t))\geq\eta\right)\le$$

$$\le\widetilde{P}^n_\theta\left(\sup_{t\le T}c_n^2(\theta)(\Delta N^n_\theta(t))^2\geq\eta-1\right)\le\widetilde{P}^n_\theta\left(\sum_{t\le T}c_n^2(\theta)(\Delta N^n_\theta(t))^2\geq\eta-1\right). \tag{78}$$

By virtue of the Lenglart inequality, for any $\delta>0$ we have

$$\widetilde{P}^n_\theta\left(\sum_{t\le T}c_n^2(\theta)(\Delta N^n_\theta(t))^2\geq\eta\right)\le\frac{d}{\eta}+\widetilde{P}^n_\theta\left(\left(\sum_{t\le T}c_n^2(\theta)(\Delta N^n_\theta(t))^2\right)^{p,\widetilde{P}^n_\theta}\geq d\right). \tag{79}$$

ow assertion (1) follows from (74)–(79) and (12).

(2) Now we will show that

$$\left|\widetilde{B}^n_T-c_n^2(\theta)\langle L^n_\theta,N^n_\theta\rangle_T\right|\xrightarrow{\widetilde{P}^n_\theta}0\quad\text{as}\quad n\to\infty. \tag{80}$$

Denote $\overline{X}^n=X^n-\sum\Delta X^nI_{\{|\Delta X^n|\geq1\}}$. Since $\overline{X}^n$ is a special semimartingale, e unique decomposition $\overline{X}^n=\overline{M}^n+\overline{A}^n$ takes place with a predictable $\overline{A}^n$.

Rewrite $\overline{X}^n$ as follows: $\overline{X}^n=X^n-xI_{\{|x|\geq1\}}*(\mu^n-\nu^n)-xI_{\{|x|\geq1\}}*\nu^n$. We ave $\overline{M}^n=X^n-xI_{\{|x|\geq1\}}*(\mu^n-\nu^n)$.

Further, applying the triplet transformation formulas under absolutely ontinuous change of measure, we get $\widetilde{B}^n=B^n+c_n(\theta)\langle\overline{M}^n,N^n_\theta\rangle$. Hence $^n-c_n^2(\theta)\langle L^n_\theta,N^n_\theta\rangle=B^n-c_n(\theta)\langle X^n-\overline{M}^n,N^n_\theta\rangle$.

By the Lindeberg condition and the contiguity $\{\widetilde{P}^n_\theta\}\lhd\{P^n_\theta\}$, we have $^n_T\xrightarrow{\widetilde{P}^n_\theta}0$. Further,

$$\langle X^n-\overline{M}^n\rangle=x^2I_{\{|x|\geq1\}}*\nu^n-\sum_{t\le s}\left(\int xI_{\{|x|\geq1\}}\cdot\nu^n(\{t\},dx)\right)^2.$$

gain, the Lindeberg condition and the contiguity give and $C_T^{1,n} \xrightarrow{P_\theta} C_T = \Gamma_L(\theta)$.

Hence (see [8]) $X_T^n - \widetilde{B}_T^n \xrightarrow{\mathcal{L}(T|P_\theta^{n,N})} M_T$, where $\widetilde{B}_T^n = -\int_0^T \int_{|x|\geq 1} x\widetilde{\nu}^n(dt,dx)$ is a first characteristic of the process X^n with respect to the measure $P_\theta^{n,N}$.

Now the desirable follows from conditions (7) and (5). □

Proof of Proposition 2. Denote $X^n = c_n(\theta)L_\theta^n$. For each $n \geq 1$ the process X^n is a semimartingale with a triplet of predictable characteristics $(-xI_{\{|x|\geq 1\}} \cdot \nu^n, c_n^2\langle L_\theta^{n,c}\rangle, \nu^n)$ (with respect to the measure P_θ^n).

(1) The following necessary and sufficient condition for $\{\widetilde{P}_\theta^n\} \lhd \{P_\theta^n\}$ is well-known (see, e.g., [8]):

$$\lim_{\eta\to\infty}\lim_{n\to\infty} \widetilde{P}_\theta^n\left(h_T\left(\frac{1}{2};\widetilde{P}_\theta^n,P_\theta^n\right)\geq\eta\right)=0, \tag{74}$$

$$\lim_{\eta\to\infty}\lim_{n\to\infty} \widetilde{P}_\theta^n\left(\sup_{t\leq T}\alpha_n(t)\geq\eta\right)=0, \tag{75}$$

where $\left(h_t\left(\frac{1}{2};\widetilde{P}_\theta^n,P_\theta^n\right)\right)$, $0\leq t\leq T$, is the Hellinger process of order $\frac{1}{2}$, and $\alpha_n(t)=\widetilde{\rho}_\theta^n(t)/\rho_\theta^n(t-)$, where

$$\widetilde{\rho}_\theta^n=\frac{d\widetilde{P}_\theta^n}{dP_\theta^n}=\mathcal{E}(c_n(\theta)N_\theta^n).$$

It can be easily seen that

$$h\left(\frac{1}{2};\widetilde{P}_\theta^n,P_\theta^n\right)\leq\frac{1}{2}c_n^2(\theta)\langle N_\theta^n\rangle. \tag{76}$$

Indeed,

$$h\left(\frac{1}{2};\widetilde{P}_\theta^n,P_\theta^n\right)=\frac{1}{8}c_n^2(\theta)\langle N_\theta^{n,c}\rangle+\frac{1}{2}\left(\sum\left(1-\sqrt{1+c_n(\theta)\Delta N_\theta^n}\right)^2\right)^{p,P_\theta^n}$$

But, since $(1-\sqrt{1+x})^2\leq\frac{x^2}{(1+\sqrt{1+x})^2}\leq x^2$, for $x\geq-1$, we have

The necessity of condition (64) is obvious. Condition (65) which is based formula (63) strengthens, on the one hand, the ergodicity properties (4), (5) l, on the other hand, ensures asymptotic "homogeneity" of score functions the class Ψ.

This implies the necessity of the requirement of convergence of joint distri-ions both in (63) and in (67).

Condition (66) ("compliteness" of the class Ψ), makes class of score func-ns richer.

Proceed to the class $\mathcal{H}_\Psi$, namely to formula (72). Begin with the end of formula. Belonging of the sequence H to the set $\mathcal{H}(\Psi) \cap B$ guarantees, on one hand, ergodic coordination of the classes Ψ and $\mathcal{H}_\Psi$ (see (70,2)) and, the other hand, imposes boundedness type conditions on this sequence (see ditions (70,1 and (12)). In particular, the belonging to the set B implies contiguity of the sequence of alternative measures $\{\widetilde{P}^n_\theta\}$ with respect to the uence of basic measures $\{P^n_\theta\}$. Further, the belonging of each term H^n of sequence H to the set $\mathcal{K}^n$ implies a uniform integrability of the exponential rtingale $\mathcal{E}\left(c_n(\theta) N^n_\theta\right)$ and also the fact that $\inf_t \mathcal{E}_t\left(c_n(\theta) N^n_\theta\right) > 0$. From these ts follows the property $\widetilde{P}^n_\theta \sim P^n_\theta$.

The non-triviality of the optimization problem (18) follows from the rela-n $H^n \in \mathcal{H}^n_R$.

Note, finally, that the above definitions of the classes Ψ and $\mathcal{H}_\Psi$ give sta-ical sense to the risk functional $D(\psi, H, \theta)$, to the optimization problem, l as a result, to the whole problem of robust estimation considered in this per. $\square$

Let $\psi^* = \{\psi^{*,n}\}_{n \geq 1}$ be the sequence of score functions constructed in The-m 1 by the formula (48).

Theorem 2. *The sequence ψ^* is $(\Psi, \mathcal{H}_\Psi)$-optimal.*

9. Proofs.

Proof of Proposition 1. Let $W = (W_t)$, $0 \leq t \leq T$, be a standard Wiener ocess defined on a stochastic basis $(\Omega,\ \mathcal{F},\ F,\ P)$. Denote $M := \sqrt{\frac{\Gamma_1(\theta)}{T}} \cdot W$. en $C_T = E^P(M_T^2) = \Gamma_1(\theta)$. Denote, further, $X^n = c_n(\theta) L^n_\theta$, and where the martingale N_β^n $(= N_\beta^n(H^n))$ is defined by the formula (36), K_n and R_n are real numbers. Recall that in (36) index n is omitted.

Define

$$\begin{aligned}\mathcal{H}_\Psi = \ & \big\{H = \{H^n\}_{n\geq 1} : H^n \in \mathcal{H}_R^n \cap \mathcal{K}^n,\ n \geq 1; \\ & H \in \mathcal{H}(\Psi) \cap B\big\},\end{aligned} \tag{72}$$

where $\mathcal{H}_R^n$ is defined by (41).

Finally, suppose that

$$\mathcal{L}^{Q'_\delta}\left(\{\nu :\ \nu \ \text{is degenerate}\} \cup \{\nu :\ \nu(\operatorname{med}\nu) > 0\}\right) = 0. \tag{73}$$

Remark 6. Note that in the discrete time case the property $H^n \in \mathcal{K}^n$ is reduced to the property $\sup_{n,\omega,t} \lambda^n(\omega, t) < \infty$. $\square$

Remark 7. Consider, at the qualitative level, the assumptions used in the definition of classes Ψ and $\mathcal{H}_\Psi$. We begin with the class Ψ. Assumption (61) reflects the fact that in this work we do not come out of the framework of the L_2-theory. Assumption (62) strengthens the admissibility property, which statistically means that the "unbounded" score functions cannot be optimal, and they apriori are excluded from our consideration. Next, we see that for $\alpha = c, \pi, b$ and for $\alpha = \delta$ the conditions differ from each other and therefore they are given separately.

An additional assumption in the case when $\alpha = \delta$ is $\widehat{\psi}_\delta = 0$, which corresponds to the conditionally centering property of the score function $\psi_\delta^n(z, \theta|x)$ in discrete time models. This results in a necessity to consider the parameter $\beta^n = \beta(Q_\delta^{n,\prime})$, which is a nonlinear functional defined on the space of probability distributions. Note that the technical condition (69) and the analogous condition (73) ensure the continuity of the functional β^n with respect to the topology of a weak convergence of distributions.

The presence of the parameter β^n in formula (48) for the optimal score function $\psi^{*,n}$ complicates the checking of the ergodic properties and results in a necessity to consider a generalized weak convergence of distributions. This convergence reduces to a weak convergence of distributions on the metric space of probability measures or, roughly speaking, to a weak convergence of

Let $\Psi_\delta \subset \Psi_\delta^0$ be the set of sequences $\psi_\delta = \{\psi_\delta^n\}_{n\geq 1}$ with properties 1), 2), , 4) with $\alpha = \delta$ and, in addition,

5)

$$\widehat{\psi}_\delta^n = 0 \text{ for each } n \geq 1, \tag{68}$$

6)

$$\begin{aligned} &\left(Q_\delta^{n,\psi} \Longrightarrow Q_\delta^\psi\right) \Longrightarrow \mathcal{L}^{Q_\delta^\psi}\big(\{\nu : \ \nu \text{ has not a unique median}\}\cup \\ &\cup\{\nu : \ \nu \text{ degenerate in } 0\} \cup \{\nu : \int y\nu(dy) \neq 0\}\big) = 0. \end{aligned} \tag{69}$$

Remark 5. Note that from 2) follows that $\psi_\alpha \in \Phi^{\psi_\alpha}$ (take $\psi_\alpha^1 = \psi_\alpha^2 = \psi_\alpha$) nd thus $\overline{Q}^{n,\psi_\alpha} \xrightarrow{W} \overline{Q}^{\psi_\alpha}$, $\alpha = c, \pi, b$, $Q^{n,\psi_\delta} \Longrightarrow Q^{\psi^\delta}$.

Now $\Psi = (\Psi_\alpha, \ \alpha = c, \pi, \delta, b)$.

(ii) Definition of class $\mathcal{H}_\Psi$.

Let $A_\alpha \subset \Psi_\alpha^0$, $\alpha = c, \pi, \delta, b$. Define for each $\alpha = c, \pi, \delta, b$

$$\mathcal{H}(A_\alpha) = \left\{ \{H_\alpha^n\}_{n\geq 1} : \begin{array}{l} 1) \ \ \sup\limits_n \int (H_\alpha^n)^2 d\mu_\alpha^n < \infty \\ \\ 2) \ \ \int \psi_\alpha^n H_\alpha^n d\mu_\alpha^n \to, \ \forall \psi_\alpha \in A_\alpha \end{array} \right\} \tag{70}$$

The symbol "$a_n \to$" means that the sequence $\{a_n\}_{n\geq 1}$ has a finite limit.

Let $\mathcal{H}(A) = (\mathcal{H}(A_\alpha), \ \alpha = c, \pi, \delta, b)$, where $A = (A_\alpha, \ \alpha = c, \pi, \delta, b)$.

We say that the sequence $H = \{H^n\}_{n\geq 1} = \{H_\alpha^n, \ \alpha = c, \pi, \delta, b\}_{n\geq 1} \in \boldsymbol{B}$ for each sequence of martingales $\{N_\theta^n\}_{n\geq 1}$ defined by the formula (36) with $I_\alpha^n \in L_2(\mu_\alpha^n)$, $n \geq 1$, $\alpha = c, \pi, \delta, b$ the condition (12) of Proposition 2 is atisfied.

$$P_\theta^n \left(\int I_{\{|u|>\eta\}} \widetilde{Q}_\alpha^{n,\psi}(\omega, du) \leq r_\alpha^n \right) = 1, \tag{62}$$

and the sequence $\{r_\alpha^n\}_{n\geq 1}$ is bounded (for each fixed η).

Here $\widetilde{Q}_\alpha^{n,\psi}(\omega, \circ) = \int I_{\{\psi_\alpha^n \in \circ\}} d\widetilde{\mu}_{\alpha,\omega}^n$, where $\widetilde{\mu}_{c,\omega}^n(dt) = c_n^2(\theta) dC_t^n$, $\widetilde{\mu}_{\pi,\omega}^n(dt, dx) = c_n^2(\theta) I_{\{a_\theta^n = 0\}} \nu_\theta^n(dt, dx)$, $\widetilde{\mu}_{\delta,\omega}^n(dt, dx) = c_n^2(\theta) I_{\{a_\theta^n > 0\}} \nu_\theta^n(dt, dx)$, $\widetilde{\mu}_{b,\omega}^n(dt) = c_n^2(\theta) p_\theta^n(\omega, dt)$.

First we define the classes Ψ_α for $\alpha = c, \pi, b$.

For each $n \geq 1$, we denote

$$\overline{Q}_\alpha^{n,\psi}(\circ) = \int I_{\{\psi_\alpha^n \in \circ\}} d\mu_\alpha^n, \quad \overline{Q}_\alpha^{n,\psi^1,\psi^2}(\circ) = \int I_{\{(\psi_\alpha^{1,n}, \psi_\alpha^{2,n}) \in \circ\}} d\mu_\alpha^n,$$

where the measures μ_α^n are defined by (37).

Fix the sequence $\psi_\alpha^0 \in \Psi_\alpha^0$ and introduce the set

$$\Phi^{\psi_\alpha^0} = \left\{ \psi \in \Psi_\alpha^0 : \overline{Q}_\alpha^{n,\psi,\psi^0} \xrightarrow{W} \overline{Q}_\alpha^{\psi,\psi^0} \right\}. \tag{63}$$

Let $\Psi_\alpha \subset \Psi_\alpha^0$ be a set of sequences with properties:

1)

$$l_\alpha \in \Psi_\alpha, \tag{64}$$

2)

$$\left((\psi_\alpha^1, \psi_\alpha^2) \in \Psi_\alpha \right) \Longrightarrow \left(\psi_\alpha^1 \in \Phi^{\psi_\alpha^2} \right), \tag{65}$$

3)

$$\left(\widetilde{\psi}_\alpha \in \cap_{\psi_\alpha \in \Psi_\alpha} \Phi^{\psi_\alpha} \right) \Longrightarrow \left(\widetilde{\psi}_\alpha \in \Psi_\alpha \right), \tag{66}$$

4) the sequence $\{(\psi_\alpha^n)^2\}_{n\geq 1}$ is uniformly integrable with respect to the sequence of measures $\{\mu_\alpha^n\}_{n\geq 1}$.

Let $\alpha = \delta$.

For each $n \geq 1$, we denote

if $\overline{\nu}(\bullet)$ is a distribution degenerated at 0, then $F(\overline{\nu}) = 0$;

$$\left.\frac{F(\nu)}{\overline{\overline{F}}(\nu)} \text{ is bounded}\right\}, \quad d = 1, 2.$$

We say that the sequence of random measures $\{Q^n\}_{n\geq 1}$ generalized weakly nverges to the random measure Q and write $Q^n \Longrightarrow Q$ as $n \to \infty$, if ${}^{n,Q} \xrightarrow{W} \mathcal{L}^Q$ as $n \to \infty$.

The latter convergence means that conditions (57), (58) and (59) are satied, and

$$\int F(\nu)\mathcal{L}^{n,Q}(d\nu) \to \int F(\nu)\mathcal{L}^Q(d\nu) \text{ as } n \to \infty, \ \forall F \in C_d, \ d = 1, 2.$$

From definitions easily follows that $(Q^n \Longrightarrow Q) \Longrightarrow \left(\overline{Q}^n \xrightarrow{W} \overline{Q}\right)$. Indeed, is sufficient to take

$$F(\nu) = \int f(x)\nu(dx), \quad f \in C_d, \quad x \in R_d. \tag{60}$$

We consider the case described in 5. As it follows from (39), for each $n \geq 1$ ${}_n(L^n, N^n; \theta) = D_n(\psi^n, H^n; \theta)$.

We define now the classes of sequences of functions Ψ and $\mathcal{H}_\Psi$ such that $\psi = \{\psi^n\}_{n\geq 1} \in \Psi$ and $H = \{H^n\}_{n\geq 1} \in \mathcal{H}_\Psi$, then $D(\psi, H; \theta)$ $\big(= (L, N; \theta)\big) = \lim_{n\to\infty} D_n(\psi^n, H^n; \theta)$, (see (15), (16), (20), (21)).

Then we construct the $(\Psi, \mathcal{H}_\Psi)$-optimal in minimax sense sequence $\psi^* = \{\psi^{*,n}\}_{n\geq 1}$ of score functions, i.e., the sequence $\psi^* \in \Psi$ such that

$$\sup_{H\in\mathcal{H}_\Psi} D(\psi^*, H; \theta) = \inf_{\psi\in\Psi} \sup_{H\in\mathcal{H}_\Psi} D(\psi, H; \theta),$$

r each $\theta \in \Theta$.

Below the parameter θ is fixed and omitted.

(i) Definition of Class Ψ

$$\mathcal{L}^{n,Q}(B) = \int I_{\{Q^n_{\omega,t}(\bullet)\in B\}} \nu^n(d\omega, dt), \qquad \mathcal{L}^Q(B) = \int I_{\{Q_{\omega,t}(\bullet)\in B\}} \nu(d\omega, dt),$$

be σ-finite distributions on $(\mathcal{M}_d, \mathcal{B}(\mathcal{M}_d))$ induced by some σ-finite measures $\nu^n(d\omega, dt)$ and $\nu(d\omega, dt)$ defined on $(\Omega \times [0,T], \mathcal{F} \times \mathcal{B}([0,T]))$.

Here $(\mathcal{M}_d, \mathcal{B}(\mathcal{M}_d))$, $d = 1, 2$, are measurable spaces of the probability measures on $(R_d, \mathcal{B}(R_d))$, $d = 1, 2$, $\mathcal{B}(\mathcal{M}_d)$ is a Borel σ-algebra generated by open (with respect to the Prokhorov metric ρ) sets.

Introduce on $(R_d, \mathcal{B}(R_d))$, $d = 1, 2$, the measures

$$\overline{Q}^n(\bullet) = \int \nu(\bullet)\mathcal{L}^{n,Q}(d\nu), \quad n \geq 1, \tag{55}$$

and

$$\overline{Q}(\bullet) = \int \nu(\bullet)\mathcal{L}^Q(d\nu), \tag{56}$$

where $\nu(\bullet) \in \mathcal{M}_d$, and suppose that these measures satisfy conditions (52) and (53).

Introduce the functional $\overline{F}(\nu) = \int |x|^2 \nu(dx)$ defined on the space $(\mathcal{M}_d, \rho)$, $d = 1, 2$.

Then by conditions (52) and (53) we have

$$\int \overline{F}(\nu)\mathcal{L}^{n,Q}(d\nu) = \int |x|^2 \nu(dx)\mathcal{L}^{n,Q}(d\nu) = \int |x|^2 \overline{Q}^n(dx) < \infty, \tag{57}$$

$$\int \overline{F}(\nu)\mathcal{L}^Q(d\nu) = \int |x|^2 \overline{Q}(dx) < \infty, \tag{58}$$

and

$$\int \overline{F}(\nu)\mathcal{L}^{n,Q}(d\nu) \to \int \overline{F}(\nu)\mathcal{L}^Q(d\nu) \quad \text{as} \quad n \to \infty. \tag{59}$$

Denote

sator $\widetilde{\nu}^n_\theta$ possessing the same properties.

Then (51) has the form $\frac{\widetilde{a}^n_\theta - a^n_\theta}{a^n_\theta(1-a^n_\theta)} I_{\{0<a^n_\theta<1\}} = c_n(\theta) H^{\theta,n}_b I_{\{0<a^n_\theta<1\}}$. Hence $N^n_\theta(t)$ $= \int_0^t H^{\theta,n}_b I_{\{0<a^n_\theta<1\}} \left(dx_s - \nu^n_\theta(ds)\right), \quad 0 \le t \le T.$

The simplest case of this is the binomial model with a random probability of success, i.e., when the observation is a sequence of indicators $I^n_1, \dots, I^n_n$, $a^n_\theta(i) = P^n_\theta\{I^n_i = 1 | \mathcal{F}^n_{i-1}\}$, $1 \le i \le n$, $\mathcal{F}^n_i = \sigma(I^n_j,\ j \le i)$.

The latter fact explains the meaning of the index "b".

8. Construction of the Sequence of Optimal Score Functions

We need some auxiliary notions concerning the weak convergence of σ-finite distributions.

1. Let $\overline{Q}^n(\bullet), \quad n \ge 1$ and $\overline{Q}(\bullet)$ be σ-finite distributions on $(R_d, \mathcal{B}(R_d))$, $d = 1, 2$, satisfying the conditions

$$\int |x|^2 \overline{Q}^n(dx) < \infty, \quad n \ge 1, \qquad \int |x|^2 \overline{Q}(dx) < \infty, \tag{52}$$

where $x \in R_d$, $|\bullet|$ is the usual norm in R_d, $d = 1, 2$.

Suppose

$$\int |x|^2 \overline{Q}^n(dx) \to \int |x|^2 \overline{Q}(dx) \quad \text{as} \quad n \to \infty. \tag{53}$$

Define the sets

$$\overline{C}_d := \{f : f \text{ is a continuous function on } R_d,$$

$$f(0) = 0 \text{ and } \frac{f(x)}{|x|^2} \text{ is bounded}\}, \quad d = 1, 2.$$

We say that the sequence of distributions $\{\overline{Q}^n(\bullet)\}_{n \ge 1}$ weakly converges to the distribution $\overline{Q}(\bullet)$ and write

$$\overline{Q}^n \xrightarrow{W} \overline{Q} \text{ as } n \to \infty, \tag{54}$$

if relations (52) and (53) are satisfied, and

$$-\int_{C_{[0,T]}}\int_0^T\left([\dot\beta_n(t,x,\theta)]_{-m}^{m}\right)^2 dt P_\theta^n(dx). \quad \square$$

In the following special cases we briefly describe only main objects.

Poisson Type Point Process

For each $n \geq 1$, let P_θ^n be a distribution of the point process $\xi_n = (\xi_n(t))$, $0 \leq t \leq T$, with a compensator $\nu_\theta^n(t) = \int_0^t a_\theta^n(s)d\alpha_s^n$, and let P^n be a distribution of the point process with a compensator α^n.

Then

$$L_t(\dot M_\theta^n, M_\theta^n) = \int_0^t \frac{\dot a_\theta^n(s)}{a_\theta^n(s)}(dx_s - a_\theta^n(s)d\alpha_s^n), \quad L_\theta^n(t) = \int_0^t \psi_{\pi_2}^{\theta,n}(dx_s - a_\theta^n(s)d\alpha_s^n),$$

$$I_T^n(\theta) = E_\theta^n \int_0^T \left(\frac{\dot a_\theta^n(t)}{a_\theta^n(t)}\right)^2 a_\theta^n(t)d\alpha_t^n, \quad c_n^{-2}(\theta) = I_T^n(\theta).$$

The "contaminated" measure $\widetilde P_\theta^n$ in this case is the distribution of the point process with the compensator $\widetilde\nu_\theta^n(t) = \int_0^t \left(a_\theta^n(s) + c_n(\theta)H_{\pi_2}^{\theta,n}(s)\right)d\alpha_s^n$. Thus $N_\theta^n(t) = \int_0^t H_{\pi_2}^{\theta,n}(s)(dx_s - a_\theta^n(s)d\alpha_s^n)$.

Processes with Jumps at the Predictable Moments

Consider the case where $X^c = 0$, $\nu^c = 0$. Then from equalities (34) and (36) we find that

$$\begin{aligned}\left(\frac{\widetilde a_\theta^n \widetilde f_\theta^n}{a_\theta^n f_\theta^n} - 1\right) I_{\{a_\theta^n > 0\}} + \frac{\widetilde a_\theta^n - a_\theta^n}{1 - a_\theta^n} I_{\{0 < a_\theta^n < 1\}} = \\ = c_n(\theta)\left(H_\delta^{\theta,n} I_{\{a_\theta^n > 0\}} + H_b^{\theta,n} I_{\{0 < a_\theta^n < 1\}}\right).\end{aligned} \tag{51}$$

To illustrate a role of terms with indices "δ" and "b", let us consider two special cases. Let for each $n \geq 1$:

$\Big\{ \; 1 \quad$ when $t = \frac{i}{n} \cdot T, \; 1 \leq i \leq n,$

$$\int_0^T \left(\dot\beta_n(t,x,\theta)\right)^2 dt \cdot \left(E_\theta^n \int_0^T \left(\dot\beta_n(t,x,\theta)\right)^2 dt\right)^{-1} \xrightarrow{P_\theta^n} 1 \text{ as } n\to\infty.$$

The score martingale $L_\theta^n \in \mathcal{M}^2(P_\theta^n)$ is given by the formula

$$L_\theta^n(t) = \int_0^t \psi_c^{\theta,n}(s,x)(dx_s - \beta_n(s,x,\theta)ds), \quad 0\le t\le T.$$

The "contamination model" means that $\widetilde{P}_\theta^n$ is a distribution of the process $\widetilde{\xi}_n = \left(\widetilde{\xi}_n(t)\right)$, $0\le t\le T$, with the differenttial

$$d\widetilde{\xi}_n(t) = \left(\beta_n(t,\widetilde{\xi}_n,\theta) + c_n(\theta)H_c^{\theta,n}(t,\widetilde{\xi}_n)\right)dt + dW_n(t), \quad \widetilde{\xi}_n(0)=0.$$

Hence the main objects are given by the following equalities:

$$N_\theta^n(t) = \int_0^t H_c^{\theta,n}(s,x)(dx_s - \beta_n(s,x,\theta)ds), \quad 0\le t\le T,$$

$$\mu_c^{\theta,n}(dt,dx) := c_n^2(\theta)dtP_\theta^n(dx), \quad x\in C_{[0,T]},$$

$D_n(\psi^n, H^n;\theta) =$

$$= \frac{\left(\int_{C_{[0,T]}}\int_0^T \psi_c^{\theta,n}(t,x)H_c^{\theta,n}(t,x)dtP_\theta^n(dx)\right)^2 + \int_{C_{[0,T]}}\int_0^T \left(\psi_c^{\theta,n}(t,x)\right)^2 dtP_\theta^n(dx)}{\left(\int_{C_{[0,T]}}\int_0^T \psi_c^{\theta,n}(t,x)l_c^{\theta,n}(t,x)dtP_\theta^n(dx)\right)^2},$$

$$H_R^n = \left\{H_c^{\theta,n} : \int_{C_{[0,T]}}\int_0^T |H_c^{\theta,n}(t,x)|dtP_\theta^n(dx)\le R\right\}.$$

Finally, the optimal score martingale is

This case is covered by a general scheme of statistical models [illegible] $)_{n\ge 1}$ in the following manner.

Put $\Omega^n = C_{[0,T]}$, the space of continuous functions (x_t), $0 \le t \le T$, with $x_0 = 0$, $\mathcal{F}^n = \mathcal{B}_T = \sigma(x : x_t,\, t \le T)$, $F^n = (\mathcal{F}_t^n = \sigma(x : x_s,\, s \le t))$, $0 \le t \le T$, P^n is a Wiener measure, P_θ^n is a distribution of the process ξ_n (with a given θ). In other words, the coordinate process $x = \{x_t(\omega),\, \omega \in C_{[0,T]},\, 0 \le t \le T\}$, with $x_t(\omega) = \omega_t$, is a (P^n, F)-semimartingale with a triplet $(0, t, 0)$, $0 \le t \le T$, and a (P_θ^n, F)-semimartingale with a triplet $(\int_0^t \beta_n(s, x, \theta) ds,\, t,\, 0)$, $0 \le t \le T$.

Assume that for each $n \ge 1$,

$$P^n \left(\int_0^T \beta_n^2(t, x, \theta) dt < \infty \right) = P_\theta^n \left(\int_0^T \beta_n^2(t, x, \theta) dt < \infty \right) = 1.$$

Assume that there exists the unique weak solution of equation (50), $P_\theta^n \sim P^n$ and the likelihood ratio process has the form

$$\rho_\theta^n(t) = \exp \left(\int_0^t \beta_n(s, x, \theta) dx_s - \frac{1}{2} \int_0^t \beta_n^2(s, x, \theta) ds \right) = \mathcal{E}_t(M_\theta^n),$$

where $M_\theta^n(t) = \int_0^t \beta_n(s, x, \theta) dx_s$, $\quad 0 \le t \le T$, is a local (P^n, F)-martingale.

Further, let for each $n \ge 1$, $x \in C_{[0,T]}$ and $t \in [0, T]$, the mapping $\theta \rightsquigarrow \beta_n(t, x, \theta)$ be continuously differentiable $\left(\frac{\partial}{\partial \theta} \beta := \dot{\beta} \right)$, and

$$\frac{\partial}{\partial \theta} \int_0^t \beta_n(s, x, \theta) dx_s = \int_0^t \dot{\beta}_n(s, x, \theta) dx_s, \quad \frac{\partial}{\partial \theta} \int_0^t \beta_n(s, x, \theta) ds = \int_0^t \dot{\beta}_n(s, x, \theta) d$$

$$0 < I_T^n(\theta) := E_\theta \int_0^T \left(\dot{\beta}_n(t, x, \theta) \right)^2 dt < \infty.$$

Then regularity conditions are satisfied, and

$Q^{\delta,n}_{\omega,t}(\cdot,\theta)$ satisfies the following conditions: a) $Q^{\delta,n}_{\omega,t}(\bullet,\theta)$ has the unique median; b) $Q^{\delta,n}_{\omega,t}(\{\beta^{\delta,n}(0;\theta)\},\theta) = 0$, where $\beta^{\delta,n}(0;\theta) = \lim_{m\to 0}\beta^{\delta,n}(m;\theta)$.

Theorem 1. *(i) There exists the optimal $\psi^{*,n}$ equal to*

$$\psi^{*,n} = \begin{cases} \left[l_\alpha^{\theta,n}\right]_{-m_n^*(\theta)}^{m_n^*(\theta)}, & if \quad \alpha \neq \delta, \\ \\ \left[l_\delta^{\theta,n} - \beta^{\delta,n}(m_n^*(\theta),\theta)\right]_{-m_n^*(\theta)}^{m_n^*(\theta)}, & if \quad \alpha = \delta, \end{cases} \tag{48}$$

and, thus, optimal ***score martingale*** *has the form*

$$L_\theta^{*,n} = [l_c^{\theta,n}]_{-m_n^*(\theta)}^{m_n^*(\theta)} \cdot (X^c - \beta_\theta^n \cdot C^n) +$$

$$+\Big([l_\pi^{\theta,n}]_{-m_n^*(\theta)}^{m_n^*(\theta)} I_{\{a_\theta^n=0\}} + [l_\delta^{\theta,n} - \beta^{\delta,n}(m_n^*(\theta),\theta)]_{-m_n^*(\theta)}^{m_n^*(\theta)} \cdot I_{\{a_\theta^n>0\}} + [l_\delta^{\theta,n}]_{-m_n^*(\theta)}^{m_n^*(\theta)} I_{\{0<a_\theta^n<1\}}\Big) * (\mu - \nu_\theta^n),$$

where $m_n^*(\theta)$ is the unique solution of the equation

$$\begin{aligned} R^2 m^2 = \sum_{\alpha\neq\delta}\left\{[l_\alpha^{\theta,n}]_{-m}^m l_\alpha^{\theta,n} * \mu_\alpha^{\theta,n} - \left([l_\alpha^{\theta,n}]_{-m}^m\right)^2 * \mu_\alpha^{\theta,n}\right\} + \\ + \left\{[l_\delta^{\theta,n} - \beta^{\delta,n}(m,\theta)]_{-m}^m l_\delta^{\theta,n} * \mu_\delta^{\theta,n} - \left([l_\delta^{\theta,n} - \beta^{\delta,n}(m,\theta)]_{-m}^m\right)^2 * \mu_\delta^{\theta,n}\right\}. \end{aligned} \tag{49}$$

(ii) This $\psi^{*,n}$ is unique (up to the constant factor).

Proof is quite similar to that of Theorem 1 from [1], pp. 82–89, and we omit it here.

7. Special Models

Diffusion Type Processes

We consider this case in detail.

Let for each $n \geq 1$, $\xi_n = (\xi_n(t))$, $0 \leq t \leq T$, be a diffusion type process with differential

$$A_\varepsilon^1 = \left\{ Y > 0,\ Y \geq \operatorname*{ess\,sup}_{\mu} |Y| - \varepsilon \right\}, \quad A_\varepsilon^2 = \left\{ Y < 0,\ -Y \geq \operatorname*{ess\,sup}_{\mu} |Y| - \varepsilon \right\}.$$

From definition of $esssup$ it follows that $\mu\,(A_\varepsilon^1 \cup A_\varepsilon^2) > 0$.

Suppose, e.g., $\mu(A_\varepsilon^1) > 0$. Then we have $\mu(A_\varepsilon^1) = \sum_\alpha \mu_\alpha(A_\varepsilon^1 \cap \Omega_\alpha) > 0$. Hence $\mu_\alpha(A_\varepsilon^1 \cap \Omega_\alpha) > 0$ for some $\alpha = 1, 2, 3, 4$. Let for definiteness, $\mu_1(A_\varepsilon^1 \cap \Omega_1) > 0$.

Consider the function $Z^\varepsilon = (Z_1^\varepsilon, 0, 0, 0)$ with $Z_1^\varepsilon = \frac{RI_{\{A_\varepsilon^1 \cap \Omega_1\}}}{\mu_1(A_\varepsilon^1 \cap \Omega_1)}$.

Now we get $|Z^\varepsilon| * \mu = Z^\varepsilon * \mu = Z_1^\varepsilon I_{\{\Omega_1\}} * \mu = Z_1^\varepsilon * \mu_1 = R$ and

$$|YZ^\varepsilon * \mu| = |YZ^\varepsilon I_{\{\Omega_1\}} * \mu| = |YZ_1^\varepsilon * \mu_1| =$$

$$= |YZ_1^\varepsilon \cdot I_{\{A_\varepsilon^1\}} * \mu_1 + YZ_\varepsilon^1 \cdot I_{\{(A_\varepsilon^1)^c\}} * \mu_1| =$$

$$= YZ_\varepsilon^1 I_{\{A_\varepsilon^1\}} * \mu_1 \geq \left(\operatorname*{ess\,sup}_{\mu} |Y| - \varepsilon \right) R,$$

for each $\varepsilon > 0$, where A^c is a complement of a set A.

Now from (44) we get that the optimal Y has the form (see also [15])

$$Y^* = \text{const.}[X - \beta]_{-m}^{m}, \quad m > 0, \qquad \beta = \beta(\omega, t) I_{\{\Omega_3\}}(\omega, t, x).$$

Our problem is to find an equations for pair (β, m). $\square$

Remark 4. Above mentioned optimization problem is an analitic problem and have not a statistical meaning. More exact specification of sets of score functions and alternatives are needed. $\square$

As in discrete time case, denote

$$Q_{\omega,t}^{\delta,n}(\bullet, \theta) = \int I_{\{x: l_\delta^{\theta,n} \in \bullet\}} q_{\omega,t}^{\theta,n}(dx), \tag{45}$$

where the probability $q_{\omega,t}^{\theta,n}$ is given by (31), and consider the equation with respect to β ($m > 0$ is a number)

$$\int [y - \beta]_{-m}^{m} Q_{\omega,t}^{\delta,n}(dy, \theta) = 0, \tag{46}$$

Remark 3. Consider the following simple construction. Let the Hilbert spaces

$$L^2_\alpha(\Omega_\alpha, \mathcal{F}_\alpha, \mu_\alpha), \ \ \alpha = 1, \dots, 4; \ \ \Omega_\alpha \cap \Omega_j = \emptyset, \quad \alpha \neq j.$$

be given.

We denote $\Omega = \cup \Omega_\alpha$, $\quad \mathcal{F} = \cup \mathcal{F}_\alpha = \{\cup_\alpha A_\alpha : A_\alpha \in \mathcal{F}_\alpha, \ \alpha = 1, \dots, 4\}$,, $\mu(A) = \sum_\alpha \mu_\alpha(A \cap \Omega_\alpha)$, $\forall A \in \mathcal{F}$ and consider a new Hilbert space $L^2(\Omega, \mathcal{F}, \mu)$ with a scalar product $\langle \circ, \circ \rangle$. If $X, Y \in L_2(\Omega, \mathcal{F}, \mu)$, then, obviously $\langle X, Y \rangle = \sum_\alpha \langle X_\alpha, Y_\alpha \rangle_\alpha$, where $X_\alpha = X I_{\{\Omega_\alpha\}}$ $Y_\alpha = Y I_{\{\Omega_\alpha\}}$, $\langle \circ, \circ \rangle_\alpha$ is the scalar product in L^2_α.

Now for each $n \geq 1$ and θ let

$$\Omega_1 = \Omega^n \times [0, T], \qquad \Omega_2 = \{\Omega^n \times [0, T]\} \cap \{(\omega, t) : a^n_\theta = 0\} \times R_1,$$

$$\Omega_3 = \{\Omega^n \times [0, T]\} \cap \{(\omega, t) : a^n_\theta > 0\} \times R_1,$$

$$\Omega_4 = \{\Omega^n \times [0, T]\} \cap \{(\omega, t) : 0 < a^n_\theta < 1\} \times \{\chi\}, \quad \chi \notin R_1,$$

with the corresponding σ-algebras and let $\mu = (\mu_\alpha, \ \alpha = 1, \dots, 4) = (\mu^{\theta,n}_\alpha, \ \alpha = c, \pi, \delta, ; \ \mu^{\theta,n}_b \times \delta_{\{\chi\}})$; $\delta_{\{\circ\}}$ is the Dirac measure.

Let, further, $X = (X_\alpha, \ \alpha = 1, \dots, 4) = (l^{\theta,n}_\alpha, \ \alpha = c, \pi, \delta, b)$, $Y = (Y_\alpha, \ \alpha = 1, \dots, 4) = (\psi^{\theta,n}_\alpha, \ \alpha = c, \pi, \delta, b)$, $Z = (Z_\alpha, \ \alpha = 1, \dots, 4) = (H^{\theta,n}_\alpha, \ \alpha = c, \pi, \delta, b)$.

Then right-hand side of (42) may be written as follows

$$A = \inf_{Y \in \mathcal{Y}} \sup_{|Z| * \mu \leq R} \frac{(YZ * \mu)^2 + Y^2 * \mu}{(YX * \mu)^2}, \tag{43}$$

where $\mathcal{Y} = \{Y : VY * \mu = 0 : \ \text{ for all } \ V = V(\omega, t) I_{\{\Omega_3\}}(\omega, t, x) \ \text{ with } \ |VY| * \mu < \infty\}$. But

$$A = \inf \frac{R^2 (\operatorname{ess\,sup}_\mu |Y|)^2 + Y^2 * \mu}{\cdots} \tag{44}$$

$\overline{t\in B}$

Assume that for each $\alpha = c, \pi, \delta, b$ and $n \geq 1$, $l_\alpha^{\theta,n}$, $\psi_\alpha^{\theta,n}$, $H_\alpha^{\theta,n} \in L_2(\mu_\alpha^{\theta,n})$. Denote

$$
\begin{aligned}
l^n &= \left(l_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b\right), \quad \psi^n = \left(\psi_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b\right), \\
H^n &= \left(H_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b\right), \quad \mu^n = \left(\mu_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b\right).
\end{aligned} \tag{38}
$$

Then the simple calculation results in

$$
D_n(L^n, N^n; \theta) := D_n(\psi^n, H^n; \theta) = \frac{\left(\sum\limits_\alpha \psi_\alpha^{\theta,n} H_\alpha^{\theta,n} * \mu_\alpha^{\theta,n}\right)^2 + \sum\limits_\alpha \left(\psi_\alpha^{\theta,n}\right)^2 * \mu_\alpha^{\theta,n}}{\left(\sum\limits_\alpha \psi_\alpha^{\theta,n} l_\alpha^{\theta,n} * \mu_\alpha^{\theta,n}\right)^2}, \tag{39}
$$

where the sign "$*$" denotes the integral, $\alpha = c, \pi, \delta, b$.

6. Fixed-step Optimization Problem for Statistical Models Associated with Semimartingales

Fix the index $n \geq 1$, the real number $R > 0$ and consider the following sets of functions:

$$
\Psi_n^0 = \{\psi^n = (\psi_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b) : \psi_\alpha^{\theta,n} \in L_2(\mu_\alpha^{\theta,n}), \widehat{\psi}_\delta^{\theta,n} = 0\}. \tag{40}
$$

$$
\begin{aligned}
\mathcal{H}_R^n = \ & \Big\{ H^n = (H_\alpha^{\theta,n},\ \alpha = c, \pi, \delta, b) : H_\alpha^{\theta,n} \in L_2(\mu_\alpha^{\theta,n}),\ \alpha = c, \pi, \delta, b, \\
& \widehat{H}_\delta^{\theta,n} = 0,\ \ H_\delta^{\theta,n}(\omega, t, x) \geq -\lambda^{\theta,n}(\omega, t),\ \ \lambda^{\theta,n}(\omega, t) \geq 0, \\
& \sum |H_\alpha^{\theta,n}| * \mu_\alpha^{\theta,n} + \lambda^{\theta,n} * \mu_\delta^{\theta,n} \leq R \Big\}.
\end{aligned} \tag{41}
$$

(33)

ther, let $\widetilde{P}_\theta \sim P$ and $\frac{d\widetilde{P}_\theta}{dP} = \mathcal{E}(\widetilde{M}_\theta)$ with

$$\widetilde{M}_\theta = \widetilde{\beta}_\theta \cdot X^c + \left(\widetilde{Y}_\theta - 1 + \frac{\widehat{\widetilde{Y}}_\theta - a}{1-a}\right) * (\mu - \nu).$$

Then it is easy to see that $\frac{d\widetilde{P}_\theta}{dP_\theta} = \mathcal{E}\left(L(\widetilde{M}_\theta - M_\theta, M_\theta)\right) = \mathcal{E}(\widetilde{N}_\theta)$ with $\in \mathcal{M}_{loc}(P_\theta)$, and thus

$$\widetilde{N}_\theta = (\widetilde{\beta}_\theta - \beta_\theta)(X^c - \beta_\theta \cdot C) + \left[\left(\frac{\widetilde{F}_\theta}{F_\theta} - 1\right) I_{\{a_\theta = 1\}} + \right.$$
$$\left. + \left(\frac{\widetilde{a}_\theta \widetilde{f}_\theta}{a_\theta f_\theta} - 1\right) I_{\{a_\theta > 0\}} + \frac{\widetilde{a}_\theta - a_\theta}{1 - a_\theta} I_{\{0 < a_\theta < 1\}}\right] * (\mu - \nu_\theta). \tag{34}$$

Starting from equalities (33) and (34) it is natural to represent the core rtingale L_θ and martingale N_θ (for definitions see previous Subsections 2 d 3) as follows:

$$_\theta = \psi_c^\theta(X^c - \beta_\theta \cdot C) + \left(\psi_\pi^\theta I_{\{a_\theta = 0\}} + \psi_\delta^\theta I_{\{a_\theta > 0\}} + \psi_b^\theta I_{\{0 < a_\theta < 1\}}\right) * (\mu - \nu_\theta), \tag{35}$$

ere $\psi_\alpha^\theta = \psi_\alpha^\theta(\omega, t)$, if $\alpha = c, b$; $\psi_\alpha^\theta = \psi_\alpha^\theta(\omega, t, x)$, if $\alpha = \pi, \delta$ and $\widehat{\psi}_\delta^\theta = 0$;

$$I_\theta = H_c^\theta(X^c - \beta_\theta \cdot C) + \left(H_\pi^\theta I_{\{a_\theta = 0\}} + H_\delta^\theta I_{\{a_\theta > 0\}} + H_b^\theta I_{\{0 < a_\theta < 1\}}\right) * (\mu - \nu_\theta), \tag{36}$$

ere $H_\alpha^\theta = H_\alpha^\theta(\omega, t)$, $\alpha = c, b$; $H_\alpha^\theta = H_\alpha^\theta(\omega, t, x)$, $\alpha = \pi, \delta$, and $\widehat{H}_\delta^\theta = 0$.

Let us endow again with the index n all the objects and introduce the olean measures

$$\mu_c^{\theta,n}(dt, d\omega) = c_n^2(\theta) dC_t^n P_\theta^n(d\omega),$$

$$\mu_\pi^{\theta,n}(dt, dx, d\omega) = c_n^2(\theta) I_{\{a_\theta^n = 0\}} \nu_\theta^n(dt, dx) P_\theta^n(d\omega),$$

(37)

$$= (\psi * \nu_\theta^c)_t + \sum_{s\le t} B_{\omega,s}^\theta(R_1)\left(\int_{R_1}\psi(\omega,s,x)q_{\omega,s}^\theta(dx)\right)\Delta A_s^\theta =$$

$$= (\psi * \nu_\theta^c)_t + \sum_{s\le t} a_\theta(s)\int_{R_1}\psi(\omega,s,x)q_{\omega,s}^\theta(dx), \tag{29}$$

where

$$a_\theta(t) = \Delta A_t^\theta B_{\omega,t}^\theta(R_1) \tag{30}$$

and

$$q_{\omega,t}^\theta(dx)I_{\{a_t(\theta)>0\}} = \frac{B_{\omega,t}^\theta(dx)}{B_{\omega,t}^\theta(R_1)}I_{\{a_t(\theta)>0\}}. \tag{31}$$

Thus $q_{\omega,t}^\theta(dx)$ is a probability measure

$$I_{\{a_\theta(t)>0\}}\int_{R_1} q_{\omega,t}^\theta(dx) = I_{\{a_\theta(t)>0\}}.$$

Denote

$$\frac{d\nu_\theta^c}{d\nu^c} := F_\theta, \quad \frac{q_{\omega,t}^\theta(dx)}{q_{\omega,t}(dx)} = f_\theta(\omega,t,x) \quad (\text{simply } \ f_\theta).$$

Then we have

$$Y_\theta = F_\theta I_{\{a=0\}} + \frac{a_\theta}{a}f_\theta I_{\{a>0\}}, \quad \dot Y_\theta = \dot F_\theta I_{\{a=0\}} + \left(\frac{\dot a_\theta}{a}f_\theta + \frac{a_\theta}{a}\dot f_\theta\right)I_{\{a>0\}}.$$

Therefore (recall that $Y_\theta > 0$)

$$\Phi_\theta = \frac{\dot F_\theta}{F_\theta}I_{\{a_\theta=0\}} + \left(\frac{\dot f_\theta}{f_\theta} + \frac{\dot a_\theta}{a_\theta(1-a_\theta)}\cdot I_{\{a_\theta<1\}}\right)I_{\{a_\theta>0\}}. \tag{32}$$

$$1+\chi-\chi \qquad 1-\chi$$

with

$$\widehat{\psi}(t)=\int_E \psi(t,x)\nu(\{t\},dx)\,,\quad \widehat{\chi}(t)=\int_E \chi(t,x)\nu(\{t\},dx)\,.$$

From (23) we have

$$\dot{M}_\theta=\dot{\beta}_\theta\cdot X^c+\left(\dot{Y}_\theta+\frac{\hat{\dot{Y}}_\theta}{1-a}\right)*(\mu-\nu). \tag{27}$$

Note that $\widehat{Y}_\theta(t)=\int Y_\theta(t,x)\nu(\{t\},dx)\quad(:=a_\theta(t))$. Hence $\hat{\dot{Y}}_\theta(t)=\widehat{\dot{Y}}_\theta(t)=\dot{a}_\theta(t)$.

Now from (23), (27) and (25) we get

$$L(\dot{M}_\theta,M_\theta)=\dot{\beta}_\theta(X^c-\beta_\theta\cdot C)+\Phi_\theta*(\mu-\nu_\theta), \tag{28}$$

where $\Phi_\theta=\dot{Y}_\theta/Y_\theta+\dot{a}_\theta/(1-a_\theta)\cdot I_{\{a_\theta<1\}}$ with $I_{\{a_\theta=1\}}\dot{a}_\theta=0$.

We give more detailed description of the function Φ_θ.

For this purpose recall ([8]) that one can choose a version of characteristics $C(\theta)$ and ν_θ such that

$$C_t(\theta)=(c^\theta\cdot A^{\theta,c})_t,\quad \nu_\theta(\omega,dt,dx)=dA_t^\theta(\omega)B_{\omega,t}^\theta(dx)$$

$$(=\nu_\theta^c(\omega,dt,dx)+\nu_\theta^d(\omega,dt,dx))=\nu_\theta^c(\omega,dt,dx)+dA_t^{\theta,d}(\omega)B_{\omega,t}^\theta(dx),$$

where $\nu_\theta^c=I_{\{a_\theta=0\}}\nu_\theta$ is a continuous part of ν_θ, the process $A^\theta=(A_t^\theta)_{0\le t\le T}\in\mathcal{A}_{loc}^+$, $c^\theta=(c_t^\theta)$, $0\le t\le T$, is a non-negative predictable process and $B_{\omega,t}^\theta(dx)$ is a transition kernel from $(\Omega\times R_+,\mathcal{P})$ into $(R_1,\mathcal{B}(R_1))$ with $B_{\omega,t}^\theta(\{0\})=0$ and $\Delta A_t^\theta B_{\omega,t}^\theta(R_1)\le 1$.

Now for each integrable (with respect to ν_θ) function $\psi=(\psi(\omega,t,x))$ we have

$$= Y_\theta(t) - \int Y_\theta(t,x)\nu(\{t\},dx)$$

If the measure P is such that any (P, F)-local martingale admits the integral representation property with respect to X, then the likelihood-ratio process can be given by the explicit formula

$$\rho_\theta = \frac{dP_\theta}{dP} = \mathcal{E}(M_\theta),$$

where

$$M_\theta = \beta_\theta \cdot X^c + \left(Y_\theta - 1 + \frac{\widehat{Y}_\theta - a}{1-a}\right) * (\mu - \nu) \in \mathcal{M}_{loc}(P) \tag{23}$$

(with the usual convention $\frac{0}{0} = 0$).

Assume that our statistical model is regular. Thus we assume that for almost all (ω, t, x) (with respect to the corresponding Dolean's measure) the functions $\beta_\theta : \theta \rightsquigarrow \beta_\theta(\omega, t)$ and $Y_\theta : \theta \rightsquigarrow Y_\theta(\omega, t, x)$ are continuously differentiable (notations $\dot{\beta}_\theta = \frac{\partial}{\partial\theta}\beta_\theta$ and $\dot{Y}_\theta = \frac{\partial}{\partial\theta}Y_\theta$) and differentiability under the integral sign is possible.

Let us calculate

$$\frac{\partial}{\partial\theta}\ln\rho_\theta = \dot{M}_\theta - \langle \dot{M}^c_\theta, M^c_\theta\rangle - \sum \frac{\Delta\dot{M}_\theta \Delta M_\theta}{1 + \Delta M_\theta} = L(\dot{M}_\theta, M_\theta).$$

Below we use the following

Proposition 3. *Let P-martingales m and M admit integral representations*

$$m = \beta \cdot n + \psi * (\mu - \nu), \qquad M = \gamma \cdot n + \chi * (\mu - \nu), \tag{24}$$

where n is a continuous P-martingale, μ is an integer-valued random measure on $[0,T] \times E$, $E = R_1 \setminus \{0\}$ and ν is its P-compensator.

Let $\widetilde{P}$ be a measure, $\widetilde{P} \sim P$ with

$$\frac{d\widetilde{P}}{dP} = \mathcal{E}(M).$$

$C \geq 0$ is some constant).

Denote

$$D_n(L^n, N^n, \theta) = \frac{(E_\theta^n \langle L_\theta^n, N_\theta^n \rangle_T)^2 + E_\theta^n \langle L_\theta^n \rangle_T}{\left(E_\theta^n \langle L_\theta^n, L(\dot{M}_\theta^n, M_\theta^n) \rangle_T\right)^2}. \tag{20}$$

Obviously, under the condition (19),

$$D(L, N, \theta) = \lim_{n \to \infty} D_n(L^n, N^n, \theta). \tag{21}$$

The classical i.i.d. case is considered e.g. in [12], [13], [14].

5. Calculation of the Explicit Form of Risk Functional $D_n(L^n, N^n, \theta)$

At the begining of this section the index n is fixed and omitted.

Consider the statistical model (1) associated with one-dimensional F-adapted cadlag process $X = (X_t)$, $0 \leq t \leq T$, in the following way: for each $\theta \in \Theta$, P_θ is a unique measure on $(\Omega, \mathcal{F})$ such that the process X is a (P_θ, F)-semimartingale with predictable characteristics $(B(\theta), C(\theta), \nu_\theta)$ (with respect to the truncation function $h(x) = xI_{\{|x| \leq 1\}}$). Assume for convenience that all measures P_θ coincide on $\mathcal{F}_0$. Assume also that under the measure P, X is a semimartingale with triplet $B = B(0)$, $C = C(0)$, $\nu = \nu_0$.

We know ([8], Ch. III) that in this situation there exist a $\widetilde{\mathcal{P}}$-measurable positive function $Y_\theta = \{Y_\theta(\omega, t, x), (\omega, t, x) \in \Omega \times R_+ \times R_1\}$ and a predictable process $(\beta_\theta(t))$, $0 \leq t \leq T$, with

$$|h(Y_\theta - 1)| * \nu \in \mathcal{A}_{loc}^+(P), \qquad \beta_\theta^2 \cdot C \in \mathcal{A}_{loc}^+(P),$$

such that the following is true:

$$1)\ \ B(\theta) = B + \beta_\theta \cdot C + h(Y_\theta - 1) * \nu, \quad 2)\ \ C(\theta) = C, \quad 3)\ \ \nu_\theta = Y_\theta * \nu, \quad Y_\theta > 0. \tag{22}$$

Denote the measure $\widetilde{P}_{\theta}^{n} := P_{\theta}^{n,N}$. For (13) to be valid it is sufficient the existence of the deterministic limit with respect to $\{\widetilde{P}_{\theta}^{n}\}_{n\geq 1}$.

Now (14) implies simply that

$$\lim_{a\to\infty}\lim_{n\to\infty} E_{\theta}^{n,N}\left\{\left(c_n^{-1}(\theta)(T_n^L-\theta)\right)^2\wedge a\right\} = D(L,N,\theta), \tag{15}$$

where $E_{\theta}^{n,N}$ is an expectation with respect to the measure $P_{\theta}^{n,N}$ and

$$D(L,N,\theta) = \frac{\beta_{L,N}^2(\theta)+\Gamma_L(\theta)}{\gamma_L^2(\theta)}. \tag{16}$$

4. Optimization Criteria

Denote $\mathcal{N} = \mathcal{N}\left(\{P_{\theta}^{n}\}_{n\geq 1}\right)$ the class of sequences of processes $\widetilde{N}_{\theta} = \{\widetilde{N}_{\theta}^{n}\}_{n\geq 1}$ satisfying the conditions (9), (12), (13) and, besides, such that the process $\mathcal{E}(\widetilde{N}_{\theta}^{n})$ from (8) is an uniformly integrable martingale with $\inf_t \mathcal{E}_t(\widetilde{N}_{\theta}^{n}) > 0$.

Suppose, further, that for each $n \geq 1$ $\widetilde{N}_{\theta}^{n} \in \mathcal{N}_R^n$, where $\mathcal{N}_R^n$, $R>0$, is some domain in $\mathcal{M}^2(P_{\theta}^{n})$. Consideration of such set $\mathcal{N}_R^n$ is needed if we want to get a nontrivial optimization problem.

Remark 2. Under above mentioned conditions we get that 1) $\widetilde{P}_{\theta}^{n}$ defined by (8) is a probability measure, equivalent to P_{θ}^{n}, 2) $\{\widetilde{P}_{\theta}^{n}\} \lhd \{P_{\theta}^{n}\}$ and thus, $D(L,N,\theta)$, given by (16), has a statistical meaning of a risk functional, see (15). $\square$

The CLAN estimator $T^* = \{T_n^*\}_{n\geq 1}$ is called $(\mathcal{M},\mathcal{N})$-optimal in the minimax sense over the class of CLAN estimators $\left\{T^L,\ L\in\mathcal{M}((\{P_{\theta}^{n}\}_{n\geq 1})\right\}$ if for each $\theta\in\Theta$, $\forall\varepsilon>0$, $\forall L\in\mathcal{M}$, $\exists N^{\varepsilon,L}\in\mathcal{N}$, $\forall N\in\mathcal{N}$

$$\lim_{a\to\infty}\lim_{n\to\infty}\frac{E_{\theta}^{n,N}\left(\left(c_n^{-1}(\theta)(T_n^*-\theta)\right)^2\wedge a\right)}{E_{\theta}^{n,N^{\varepsilon,L}}\left(\left(c_n^{-1}(\theta)(T_n^*-\theta)\right)^2\wedge a\right)} \leq 1+\varepsilon. \tag{17}$$

The score sequence $L_{\theta}^* = \{L_{\theta}^{*,n}\}_{n\geq 1} \in \mathcal{M}\left(\{P_{\theta}^{n}\}_{n\geq 1}\right)$ is called $(\mathcal{M},\mathcal{N})$-optimal in the minimax sense if for each $\theta\in\Theta$

$$\sup D(L^*,N,\theta) = \inf\ \sup D(L,N,\theta). \tag{18}$$

re exists $\widetilde{N}_\theta^n \in \mathcal{M}_{loc}(P_\theta^n)$ such that

$$\frac{d\widetilde{P}_\theta^n}{dP_\theta^n} = \mathcal{E}(\widetilde{N}_\theta^n). \tag{8}$$

If the martingale $\widetilde{N}_\theta^n$ has the form

$$\widetilde{N}_\theta^n = c_n(\theta) N_\theta^n \tag{9}$$

l

$$\{\widetilde{P}_\theta^n\} \lhd \{P_\theta^n\}, \tag{10}$$

n we say that the sequence $\{\widetilde{P}_\theta^n\}_{n\geq 1}$ belongs to the set of shrinking neigh-urhoods of the core sequence $\{P_\theta^n\}_{n\geq 1}$.

For definiteness denote such set by $\mathcal{P}_\theta$ and each element of $\mathcal{P}_\theta$ by $\{\widetilde{P}_\theta^n\}_{n\geq 1}$ by $\{P_\theta^{n,N}\}_{n\geq 1}$.

Proposition 1. *Let $\{P_\theta^{n,N}\}_{n\geq 1} \in \mathcal{P}_\theta$ and $T^L = \{T_n^L\}_{n\geq 1}$ be CLAN esti-tor with asymptotic expansion (7). Then*

$$\mathcal{L}\left(c_n^{-1}(\theta)(T_n^L - \theta) - \frac{\widetilde{B}_T^n}{c_n^2(\theta)\langle L_\theta^n, L(M_\theta^n, M_\theta^n)\rangle_T} | P_\theta^{n,N}\right) \xrightarrow{\omega} \mathcal{N}\left(0, \frac{\Gamma_L(\theta)}{\gamma_L^2(\theta)}\right), \tag{11}$$

ere $\widetilde{B}^n = (\widetilde{B}_t^n, 0 \leq t \leq T)$ is the first characteristic of the process $c_n(\theta)L_\theta^n$ th respect to the measure $P_\theta^{n,N}$.

Proposition 2. *Let the sequence $\{\widetilde{P}_\theta^n\}$ be such that for each $n \geq 1$, $\in \mathcal{P}_\theta^n$, (8)–(10) are satisfied, $N_\theta^n \in \mathcal{M}^2(P_\theta^n)$, and*

$$\lim_{d\to\infty} \overline{\lim_{n\to\infty}} \widetilde{P}_\theta^n \left\{c_n^2(\theta)\langle N_\theta^n\rangle_T > d\right\} = 0. \tag{12}$$

t $T^L = \{T_n^L\}_{n\geq 1}$ be a CLAN estimator with asymptotic expansion (7). Then:

(1) $\{\widetilde{P}_\theta^n\} \lhd \{P_\theta^n\}$,

d thus $\{\widetilde{P}_\theta^n\} \in \mathcal{P}_\theta$;

(2) relation (11) remains true with $c_n^2(\theta)\langle L_\theta^n, N_\theta^n\rangle_T$ instead of $\widetilde{B}_T^n$;

(1) for each $n \geq 1$, $L_\theta^n \in \mathcal{M}^2(P_\theta^n)$;

(2) the sequence $\{L_\theta^n\}_{n\geq 1}$ satisfies the Lindeberg condition

$$\int_0^T \int_{|x|>\varepsilon} x^2 \nu_n(dt, dx) \xrightarrow{P_\theta^n} 0 \quad \text{as} \quad n \to \infty \quad \text{for} \quad \forall \varepsilon \in (0, 1], \tag{3}$$

where ν_n is the compensator (with respect to the measure P_θ^n) of the jump measure of the process $c_n(\theta) L_\theta^n$;

(3)

$$c_n^2(\theta) \langle L_\theta^n \rangle_T \xrightarrow{P_\theta^n} \Gamma_L(\theta) \quad \text{as} \quad n \to \infty, \quad 0 < \Gamma_L(\theta) < \infty; \tag{4}$$

(4)

$$c_n^2(\theta) \langle L_\theta^n, L(\dot{M}_\theta^n, M_\theta^n) \rangle_T \xrightarrow{P_\theta^n} \gamma_L(\theta) \quad \text{as} \quad n \to \infty, \quad 0 < \gamma_L(\theta) < \infty. \tag{5}$$

$\Gamma_L(\theta)$ and $\gamma_L(\theta)$ are deterministic functions.

We assume that $\{L(\dot{M}_\theta^n, M_\theta^n)\}_{n\geq 1} \in \mathcal{M}(\{P_\theta^n\}_{n\geq 1})$. Note that if $L_\theta \in \mathcal{M}(\{P_\theta^n\}_{n\geq 1})$, then

$$\mathcal{L}(c_n(\theta) L_\theta^n(T) | P_\theta^n) \xrightarrow{w} \mathcal{N}(0, \Gamma_L(\theta)) \tag{6}$$

as it simply follows from the central limit theorem (CLT) for the martingales (see e.g., [8]).

The sequence $T^L = \{T_n^L\}_{n\geq 1}$ of $\mathcal{F}_T^n$-measurable random variables with values in Θ is called CLAN estimator if for each $\theta \in \Theta$ there exists $L_\theta \in \mathcal{M}(\{P_\theta^n\}_{n\geq 1})$ such that

$$T_n^L = \theta + \frac{L_\theta^n(T)}{\langle L_\theta^n, L(\dot{M}_\theta^n, M_\theta^n) \rangle_T} + R_n(\theta), \tag{7}$$

where $c_n^{-1}(\theta) R_n(\theta) \xrightarrow{P_\theta^n} 0$ as $n \to \infty$.

It is evident that

$$\rho_\theta := \frac{dP_\theta}{dP} = \mathcal{E}(M_\theta) := \exp\left\{M_\theta - \frac{1}{2}\langle M_\theta^c\rangle\right\}\prod(1+\Delta M_\theta)e^{-\Delta M_\theta},$$

where $M \in \mathcal{M}_{loc}(P)$ is a local P-martingale.

b) Let Q be some other probability measure on $(\Omega,\ \mathcal{F},\ F)$ such that $Q \overset{loc}{\ll} P$ and $\frac{dQ}{dP} = \mathcal{E}(M)$, M is a local P-martingale.

If m is a local P-martingale, then the process (Girsanov transform)

$$L(m, M) := m - \frac{1}{1+\Delta M}\cdot[m, M] = m - \langle M^c, m^c\rangle - \sum\frac{\Delta m \Delta M}{1+\Delta M} \quad (2)$$

is a local Q-martingale.

c) An experiment $\mathcal{E}$ is called *regular* (see also [11]) if:

1. for each $t \geq 0$ (P-a.s.), the function $\theta \to M_\theta(t,\omega)$ is continuously differentiable and the derivative $\dot{M}_\theta := \frac{\partial}{\partial\theta}M_\theta$ is a local P-martingale for $\forall\theta$,

2. for all $t \geq 0$ (P-a.s.), there exists $\frac{\partial}{\partial\theta}\ln\rho_\theta$; $\frac{\partial}{\partial\theta}\ln\rho_\theta = L(\dot{M}_\theta, M_\theta) \in \mathcal{M}^2(P_\theta)$, the class of square integrable P_θ-martingales,
3. the Fisher information $I(\theta) = E_\theta\langle L(\dot{M}_\theta, M_\theta)\rangle$ is finite and positive.

d) Consider the sequence of the regular statistical models

$$\mathcal{E} = \{\mathcal{E}_n\}_{n\geq 1} = \left\{(\Omega^n,\ \mathcal{F}^n,\ F^n = (\mathcal{F}_t^n)_{0<t\leq T},\ \{P_\theta^n, \theta\in\Theta\subset R_1\},\ P^n)\right\}_{n\geq 1},$$

where $T > 0$ is a number.

Denote $c_n(\theta) = (I_T^n(\theta))^{-1/2} = \left(E_\theta^n\langle L(\dot{M}_\theta^n, M_\theta^n)\rangle_T\right)^{-1/2}$ Then, if mining the CLAN estimators and of martingales determining alternatives, is, in a general case, an implicit function of these sequences. Therefore, in general case, it is impossible to obtain any constructive solution of the optimization problem and to construct an optimal sequence of score martingales.

Furthemore, we assumed that the martingales under consideration possess the integral representation property and on the basis of this property we worked out the so-called approximation technique consisting in the following: we first fix index $n \geq 1$ and then consider the optimization problem associated with the risk functional $D_n(L^n, N^n, \theta)$. Stochastic analysis (in the presence of the integral representation property of martingales) allows us to deduce an explicit formula to calculate $D_n(L^n, N^n, \theta)$ and to describe explicitly the classes Ψ_n^0 and $\mathcal{H}_n^R$ of functions for which the optimization minimax problem is formulated and solved. As a result, we obtain a score martingale whose integrands in the integral representation are the Huber functions.

Then we construct classes Ψ and $\mathcal{H}_\Psi$ of sequences of functions, which determine the score martingales and martingales determine the alternatives, respectively, in such a way that the score martingale, which is optimal for each fixed step n with respect to $D_n(L^n, N^n, \theta)$, would form a sequence of optimal martingales with respect to the risk functional $D(L, N, \theta)$.

2. Specification of the Model. Regularity. Ergodicity

a) Let

$$\mathcal{E} = (\Omega,\ \mathcal{F},\ F = (\mathcal{F}_t)_{t\geq 0},\ \{P_\theta,\ \theta \in \Theta\},\ P) \tag{1}$$

be a general statistical model with filtration (see also [2]–[6]. This means that $(\Omega,\ \mathcal{F},\ F,\ P)$ is a stochastic basis, i.e., a complete probability space with filtration $F = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, P_θ is a probability measure depending on the parameter θ to be estimated and Θ is an open subset of R_1. It is assumed that $P_\theta \overset{loc}{\sim} P$ for $\forall\theta \in \Theta$.

Remark 1. All notations concerning the martingale theory and used ontamination neighbourhoods. If now the sequence of alternative measures $\tilde{P}^n_\theta\}$ is contiguous to that of basic measures $\{P^n_\theta\}$, then we get the shrinking ontamination neighbourhoods with contiguous alternatives.

The above-described scheme can be generalized at least in two principal rections. 1) passage to a infinite dimentional parameter set, i.e., to semi-arametric models and 2) passage to dependent observations (composition of ese two directions is also possible).

In our work, the parametric set Θ is one-dimensional and the observations e dependent.

As it turned out, when passing from the i.i.d. observations to the dependent oservation the contamination of measures and the contamination of trajecto-es or replacement model not coincide but, on the contrary, radically differ.

Contamination of measures for dependent discrete time observations, more ecisely, for a stationary ergodic $AR(p)$ model, has been first investigated by ünsch. Under certain assumptions on the model, he has constructed so-called ›timal B-robust ("B" is the abreviation of "bias") estimate and shown that is estimate is given by the Huber function. Contamination of trajectories s been first considered by Martin and Yohai for stationary ergodic time ·ies. They show that in considering the process of moving averrage, the iber function provides no robust estimates.

Our approach generalizes contamination of measures and, thus, continues vestigations started by Künsch. Moreover, we consider shrinking contamina-n neighbourhoods with contiguous alternatives for statistical models with tration associated with semimartingales.

We consider an array scheme which is formalized by the consideration of equence of statistical models $\mathcal{E} = \{\mathcal{E}_n\}_{n\geq 1}$. Note that for every $n \geq 1$ the odel $\mathcal{E}_n$ is supposed to be regular, whil. the very sequence $\mathcal{E}$ is ergodic. We ve introduced the conception of CL. N estimators and then, on the basis notion of an exponential martingale, we introduced the set of shrinking ntamination neighbourhoods of t core sequence of measures. We also in-stigated the asymptotic properties of CLAN estimators under the sequence alternative measures belonginig to this set and obtained the "biased" esti-ıtes.

Further, we investigate the question under what kind of conditions the

The necessity of extension of a class MLE to a class of M- or CLAN-estimators dictate by the main assumption of the theory: the parametric family of distributions of observations is not exactly specified, observations are assumed to be distributed with density from some neighbourhood of basic (core) density $f(x,\theta)$.

More frequently the so-called Huber's gross error (contamination) model is considered. In this case the "neighbourhood" is given by the formula

$$\Phi_\varepsilon^H = \left\{ \widetilde{f}(x,\theta) : \widetilde{f}(x,\theta) = (1-\varepsilon)f(x,\theta) + \varepsilon h(x,\theta) \right\}, \quad \varepsilon > 0,$$

where $h(x,\theta)$ is a density from some class H. A class H is specified depending on the statistical problem under consideration. Measures generated by densities $\widetilde{f}(x,\theta)$ are called alternative measures or, simply, alternatives, and H is called a class determining alternatives.

This model has a clear statistical meaning. Let X_i, $i \geq 1$, be i.i.d. observations with density $f(x,\theta)$, W_i, $i \geq 1$, be i.i.d. observations with density $h(x,\theta)$. Consider the i.i.d. sequence of 0–1 random variables Z_i, $i \geq 1$, with

$$P(Z_i = 1) = \varepsilon, \quad i \geq 1.$$

If the sequences (X_i), (W_i), (Z_i) are mutually independent, then the random variables

$$Y_i = (1 - Z_i)X_i + Z_i W_i$$

form i.i.d. observations with density $\widetilde{f}(x,\theta)$. Thus the observations (X_i) are "contaminated" by the observations (W_i).

Introduce a criterion of comparison of estimates based on a risk functional. Frequently, as this functional there occurs an asymptotic mean square error and the estimator is called optimal if it is a minimax with respect to the risk functional, where maximum is taken over the class H, whereas minimum is taken over the class $\Psi = \{\psi(x,\theta)\}$ of functions, which determine the estimators (in particular, over the class determines the CLAN estimators).

In such a statement, MLE's are now not optimal. Optimal estimators are prescribed by the Huber functions, included in the class of functions $\psi(x,\theta)$ determining M-estimators.

variables defined on a Markov chain. (in Russian) Liet. Mat. Rink., 1966, T. VI, No. 1, 15–22.

4. *G.L. O'Brien* Limit theorems for sums of chain-dependent processes. J. Appl. Probab., 1974, vol. 11, 582–587.
5. *S. Grigorescu, G. Oprisan* Limit theorems for $J-X$ processes with general state space. Z. Wahrscheinlichkeistheorie verw. Geb., 1976, B. 35, H. 1, 65–73.
6. *I.V. Bokuchava* Limit theorems for conditionally independent sequences. (in Russian) Investigations in Probability Theory and Mathematical Statistics (in Russian), 3–15, Metsniereba, Tbilisi, 1982.
7. *I.V. Bokuchava* Limit theorems for conditionally independent sequences. (in Russian) Teor. Veroyatnost. i Primenen., 1984, T. XXIX, 1, 192–193.
8. *I.V. Bokuchava, R.Ya. Chitashvili, T.L. Shervashidze* On the limit theorem for conditionally independent random variables and on its application to statistics. Limit Theorems and Stochastic Equations. (in Russian), 54–70, Metsniereba, Tbilisi,1984.
9. *I.V. Bokuchava, Z.A. Kvatadze, T.L. Shervashidze* On limit theorems for random vectors controlled by a Markov chain. Probability Theory and Mathematical Statistics. (Proc. 4th Vilnius Conference), v. I, 231–250, VNU Science Press, Utrecht, 1986.
10. *Z.A. Kvatadze, T.L.Shervashidze* On limit theorems for conditionally independent random variables controlled by a finite Markov chain. Probability Theory and Mathematical Statistics (Proc. 5th Japan-USSR Symposium). (Lect. Notes Math., vol. 1299). 250–258, Springer-Verlag, Berlin etc., 1987.
11. *Z.A. Kvatadze* On limit theorems for conditionally m-dependent random variables. Bull. Georgian Acad.Sci.,150 (1994), No. 1, 29–32.
12. *T. Shervashidze* Local limit theorems for conditionally independent random variables controlled by a finite Markov chain (in Russian) (submitted to Teorija Verojatn. i Primenen.)
13. *I.A. Ibragimov, Yu.v. Linnik* Independent and stationary sequences of random variables. Wolters-Noordhoff,Groningen, 1971.
14. *T.W. Anderson* The statistical analysis of time series. John Wiley, N.Y. etc.,1971.
15. *S.V. Nagaev* Some limit theorems for homogeneous Markov chains

ments denoting by $\mu(\alpha)$ and $R(\alpha)$ the suitable vector of means and covariance matrix:

$$\mu(\alpha) = \mathrm{E}(X_1|\xi_1 = \alpha) = (\mu_1(\alpha), \ldots, \mu_k(\alpha)),$$

$$R(\alpha) = \mathrm{cov}(X_1|\xi_1 = \alpha) = \mathrm{E}\left\{X_1 - \mu(\alpha))(X_1 - \mu(\alpha))^\top|\xi_1 = \alpha\right\}.$$

Evidently,

$$\mu = \sum_{\alpha=1}^{s} \pi_\alpha \nu(\alpha), \quad R_0 = \sum_{\alpha=1}^{s} \pi_\alpha R(\alpha).$$

Take in Theorem 3

$$f(\xi_n) = \mu(\xi_n) = (\mu_1(\xi_n), \ldots, \mu_k(\xi_n))^\top \tag{7}$$

and introduce the $k \times s$ matrix

$$M = (\mu_\alpha)_{i=1,\ldots,k,\ \alpha=1,\ldots,s}. \tag{8}$$

Theorem 4. *If in conditions of Theorem 3 f is defined by (7), i.e., as a vector of means of conditional distributions off X_1 given the states of ξ_1, then the limity distribution of U_n is Φ_{T_μ} is obtained by substitution of μ by (7) instead of the vector f into (5) or by substitution of N defined by (8) instead of the matrix f into (6).*

Finally we obtain the following theorem which specifies Theorem 2 for the case of a finite Markov chain as a controlling sequence.

Theorem 5. *If a stationary sequence $X = (X_j)_{j1}$ of m-dependent random vectors controlled by a finite homogenous ergodic Markov chain $\xi = (\xi_j)_{j1}$ and $\mathrm{tr}(R_m) < \infty$ then*

(a) $P_{Sn1}|\xi_{1n} \xrightarrow{w} \Phi_{R_m}$ *a.s.;*

(b) $P_{Sn1} \xrightarrow{w} \Phi_{R_m}$;

(c) $P_{Sn} \xrightarrow{w} \Phi_{R_m+T_\mu}$,

where R_m is determined by (1) and T_μ is specified in Theorem 4.

Note that some useful references concerning limit theorems in a similar setting one can find in [18] and [19].

References

$$\Big(\sum_{j=1} \quad \Big)_c \quad \Big(\sum_{j=1} \quad \Big)_c$$

Let us express V_0 and V_j by means of Π and Z. Simple calculations lead the following matrix equality (Π_{dg} means the diagonal matrix obtained by ishing the nondiagonal elements of Π)

$$Ef(\xi_1) = \sum_{\alpha=1}^{s} \pi_\alpha f_\alpha = \Big(\sum_{\alpha=1}^{s} \pi_\alpha f_1(\alpha), \ldots, \sum_{\alpha=1}^{s} \pi_\alpha f_k(\alpha) \Big)^\top = f \cdot \pi,$$

$$E\left\{ f(\xi_1) f^\top(\xi_1) \right\} = \left(Ef_i(\xi_1) \cdot f_j(\xi_1) \right)_{i,j=1,\ldots,k}$$

$$= \Big(\sum_{\alpha=1}^{s} \pi_\alpha f_i(\alpha) f_j(\alpha) \Big)_{i,j=1,\ldots,k} = f \cdot \Pi_{dg} \cdot f^\top,$$

$$E\left\{ f(\xi_1) f^\top(\xi_{1+j}) \right\} = \left(Ef_i(\xi_1) f_r(\xi_{1+j}) \right)_{i,r=1,\ldots,k}$$

$$= \Big(\sum_{\alpha,\beta=1}^{s} \pi_\alpha f_i(\alpha) P_{\alpha\beta} f_r(\beta) \Big)_{i,r=1,\ldots,k} = f \cdot \Pi_{dg} \cdot P^j \cdot f^\top.$$

Using these equalities we obtain for V_0 and V_j, $j > 0$, that

$$V_0 = \mathrm{E}\left\{ f(\xi_1) f^\top(\xi_1) \right\} - \mathrm{E}\left\{ f(\xi_1) \right\} \mathrm{E}\left\{ f^\top(\xi_1) \right\}$$

$$= f\Pi_{dg} \cdot f^\top - f \cdot \pi (f\pi)^\top = f\Pi_{dg} f^\top - f\pi\pi^\top f^\top = f(\Pi_{dg} - \Pi_{dg}\Pi) f^\top,$$

$$= \mathrm{E}\left\{ f(\xi_1) f^\top(\xi_{1+j}) \right\} - \mathrm{E}\left\{ f(\xi_1) \right\} \mathrm{E}\left\{ f^\top(\xi_{1+j}) \right\} = f\Pi_{dg} P^j f^\top - f\Pi_{dg}\Pi f^\top$$

Using now the representation of the fundamental matrix we obtain

$$\Big(\sum_{j=1}^{\infty} v_j\Big)_c = f\Pi_{dg} \Big(\sum_{j=1}^{\infty} (P^j - \Pi) \Big)_c f^\top = f\Pi_{dg}(Z - I) f^\top = f\left(\Pi_{dg} z - \Pi_{dg}\right) f^\top,$$

$$\Big(\sum_{j=1}^{\infty} v_{-j} \Big)_c = \left[\Big(\sum_{j=1}^{\infty} v_j \Big) \right]^\top = f\left[(\Pi_{dg} z)^\top - \Pi_{dg} \right] f^\top,$$

re the limiting covariation matrix has the form

$$U_n = \frac{1}{\sqrt{n}} \sum_{j=1} [f(\xi_i) - Ef(\xi_j)]$$

has the limit distribution P_{T_f} *where* $T_f = (t_{f:i,j})$, $i, j = 1, \ldots, k$, *is the matrix with the elements*

$$t_{f:i,j} = \sum_{\alpha,\beta=1}^{s} \left(\pi_\alpha z_{\alpha\beta} + \pi_\beta z_{\beta\alpha} - \pi_\alpha \pi_\beta - \pi_\alpha \delta_{\alpha\beta}\right) f_i(\alpha) f_j(\beta), \tag{5}$$

where $\delta_{\alpha\beta}$ *is Kroneker's delta.*

For the case of a scalar function $f(\xi_i)$ the fact of asymptotic normality as well as the representation (5) are well known (see, e.g., [15] and [6,7]). The limiting variance can be written as

$$t = \lim_{n\to\infty} \mathrm{D}(U_n) = \sum_{\alpha,\beta=1}^{s} [\pi_\alpha z_{\alpha\beta} + \pi_\beta z_{\beta\alpha} - \pi_\alpha \pi_\beta - \pi_\alpha \delta_{\alpha\beta}] f(\alpha) f(\beta)$$

as it was calculated, e.g., in [7] and then, using Cramer–Wold technique, one can derive (1.4) for the vector case.

Now we obtain the expression of T_f by means of the characteristics of the chaim. Introduce the $k \times s$ matrix

$$f = (f_{ij})_{i=1,\ldots,k;\ j=1,\ldots,s}, f_{ij} = f_i(j).$$

Denote

$$V_0 = \mathrm{cov}[f(\xi_1)], V_j = \mathrm{E}\left\{[f(\xi_1) - \mathrm{E}f(\xi_1)][f(\xi_{1+j}) - \mathrm{E}f(\xi_{1+j})]^\top\right\}, V_{-j}$$
$$= \mathrm{E}\left\{[f(\xi_{1_j}) - \mathrm{E}f(\xi_{1+j})][f(\xi_1) - \mathrm{E}f(\xi_1)]^\top\right\}, \quad j > 0.$$

Since $\{\xi_j\}$ is a stationary sequence, we'll have the following limit as $n \to \infty$:

$$E(U_n U_n^\top) = \frac{1}{n}[nV_0 + \sum_{j=1}^{n-1}(n-j)(V_j + V_{-j})]$$

$$= V_0 + \frac{1}{n}\sum^{n-1}(n-j)V_j + \frac{1}{n}\sum^{n-1}(n-j)V_{-j} \to$$

3. Markov chain as a controlling sequence.

Consider now a stationary, finite, homogeneous, ergodic Markov chain $= (\xi_j)_{j1}$ with a single class of ergodicity (which can contain the cyclic sub-lasses). Let $\{1, \ldots, s\}$ be the state space of ξ, $P = (p_{\alpha,\beta})$ $\alpha, \beta = 1, \ldots, s$ the natrix of transition probabilities and $\pi = (\pi_1, \ldots, \pi_s)^{\top}$ the vector of station-ry probabilities. For the sake of simplicity we have supposed that the initial istribution coincides with the stationary one: $\mathrm{P}(\xi_1 = \alpha) = \pi_\alpha$, $\alpha = 1, \ldots, s$, e., for any $j1$

$$\mathrm{P}(\xi_j = \alpha) = \pi_\alpha, \quad \alpha = 1, \ldots, s.$$

Let $\mathrm{P}^{(n)} = (p^{(n)}_{\alpha,\beta})_{\alpha,\beta=1,\ldots,s} = P^n$ be the n-step transition matrix and Π be he matrix all the s rows of which equal π.

As it is known [17] if there are no cyclic subclasses, all the π_α are positive nd

$$\max_{\alpha,\beta} \left| p^{(n)}_{\alpha,\beta} - \pi_\beta \right| < cr^n$$

vith positive c and $0 < r < 1$. When there are d cyclic subclasses one can onsider the sum $\sum_{i=0}^{d-1} (p^{(n)}_{\alpha,\beta} - \pi_\beta)$, $n1$, and write the same estimate caused by he Cesaro convergence of P^n to Π. In both cases one can write an estimate

$$\mathrm{E}\left[\nu_n(\alpha)/n - \pi_\alpha\right]^2 = O\left(1/n\right)$$

vhere $\nu_n(\alpha)$ stands for a frequency of the state α in first n moments, i.e., the umber of α's among components of ξ_{1n}, $\alpha = 1, \ldots, s$.

Let us denote $Z = (z_{\alpha,\beta})_{\alpha,\beta} = 1, \ldots, s$, the fundamental matrix of the chain :

$$Z = [I - (P - \Pi)]^{-1} = I + \left(\sum_{j=1}^{\infty} (P^j - \Pi) \right)_c \tag{4}$$

vhere I stands for $s \times s$ unit matrix and $(\quad)_c$ meens Cesaro convergence, as the nentioned above convergence of P^n to Π when there are cyclic subclasses. In he regular case (when there are no cyclic subclasses) both P^n and series in the xpression of Z converge in a common sense of matrix convergence mentioned tends to zero matrix, we obtain that a.s.

$$\mathrm{E}\{\|V_{sn}\|^2|\xi_{1n}\} \to 0.$$

Hence V_{sn} does not influence on the limiting conditional distribution of T_{sn} and since $n/(Ns) \to 1$ we obtain a.s.

$$P_{T_n+V_{sn}|\xi_{1n}} \xrightarrow{w} \Phi_{R_m(s)}.$$

Now let us study a role of of U_{sn} in the sum S_{n1}. Fix s. It is clear that

$$E(\overline{\overline{Y}}_i\overline{\overline{Y}}_i^{\top}|\xi_{1n}) = \sum_{j=is-m+1}^{is} R(\xi_j,\xi_j) + +\sum_{l=1}^{m-1}\sum_{j=is-m+1}^{is-l}\big[R(\xi_j,\xi_{j+l}) + R(\xi_{j+l},\xi_j)\big],$$

$$E(E(\overline{\overline{Y}}_i\overline{\overline{Y}}_i^{\top}|\xi_{1n}) = mR_0^{(0)} + \sum_{l=1}^{m-1}(m-l)\big[R_0^{(l)} + R_0^{(-l)}\big] = \mathrm{R}(m).$$

ergodicity of ξ and conditional independence of the summands of U_{sn}, given ξ_{1n}, lead to a.s. convergence

$$\frac{n}{N}\mathrm{E}(U_{sn}U_{sn}^{\top}|\xi_{1n}) \to \mathrm{R}(m).$$

Note that the condition $\mathrm{tr}(R_m) < \infty$ guarantees the existence of the constant C such that for sn

$$\mathrm{E}(\|U_{sn}\|^2|\xi_{1n}) = \frac{N}{n}\mathrm{tr}\left\{\frac{n}{N}\mathrm{E}(U_{sn}U_{sn}^{\top}|\xi_{1n})\right\}\frac{N}{n}C\frac{C}{s}$$

and since for $\mathrm{E}(\|U_{sn}\|^2|\xi_{in}) = 0$, for $s > n$ we have that $\mathrm{E}(\|U_{sn}\|^2|\xi_{in})$ is uniformly small for large enough s .

The proof of the assertion (a) of Theorem 2 will be finished with multidimensional version of the following lemma (Corrolary 7.7.1 from [14]).

To prove the assertion (b) note that for each Borel set B in R^k

$$\mathrm{P}(S_{n1} \in B) = \mathrm{EP}(S_{n1} \in B|\xi_{1n}) \to^{W} \Phi_{\mathrm{R}_m}.$$

As for (c), we have $\mathrm{E}\exp[(U, S_{n2})] \to \phi(u)$ $u \in \mathrm{R}^k$ where $\phi(u)$ denotes the

$$j=1 \qquad\qquad j=1$$

re conditionally independent, given ξ_{1n}.

Introduce for $s \le n$ the following sums

$$T_{sn} = \frac{1}{\sqrt{n}}\sum_{i=1}^{N}\overline{Y}_i, \quad U_{sn} = \frac{1}{\sqrt{n}}\sum_{i=1}^{N}\overline{\overline{Y}}_i,$$

$$V_{sn} = \frac{1}{\sqrt{n}}\left(\overline{X}_{sN+1} + \overline{X}_{sN+2} + \cdots + \overline{X}_{sN+r}\right).$$

For $s > n$ denote $T_{sn} = S_{n1}$, $U_{sn} = V_{sn} = 0$. We have $S_{n1} = T_{sn} + U_{sn} + V_{sn}$.

For $s < n$ the random vector T_{sn} is a sum of N conditionally independent andom vectors each of which has a conditional covariation

$$\mathrm{E}(\overline{Y}_i\overline{Y}_i^\top|\xi_{1n}) = \sum_{j=(i-1)s+1}^{is-m} R(\xi_j,\xi_j) + + \sum_{l=1}^{m}\sum_{j=(i-1)s+1}^{is-m-l}\left[R(\xi_j,\xi_{j+l}) + R(\xi_{j+l},\xi_j)\right].$$

We have

$$\mathrm{E}\{\mathrm{E}(\overline{Y}_i\overline{Y}_i^\top|\xi_{1n})\} == (s-m)R_0^{(0)} + \sum_{l=1}^{m}(s-m-l)[R_0^{(l)} + R_0^{(-l)}] \equiv sR_m(s).$$

t is evident, that $R_m(s) \to R_m$ as $s \to \infty$ (convergence of matrices is undertood as convergence of their elements to the suitable elements of the limiting natrix).

Since ξ is an ergodic sequence and $\mathrm{tr}(R_m) < \infty$, for the sequence $(\overline{Y}_i)_{i\ge 1}$ he conditions of Theorem 1 are fulfilled and we obtain that a.s.

$$P_{\sqrt{n/N}T_{sn}|\xi_{1n}} \xrightarrow{w} \Phi_{sR_m(s)},$$

e., a.s.

$$P_{\sqrt{n/(Ns)}T_{sn}|\xi_{1n}} \xrightarrow{w} \Phi_{R_m(s)}.$$

Consider now V_{sn}. It is clear that

$$\mathrm{E}\{||V_{sn}||^2|\xi_{1n}\} = \mathrm{tr}\mathrm{E}\{V_{sn}V_{sn}^\top|\xi_{1n}\}.$$

$\operatorname{tr}(R_0) < \infty$, *then*

(a) $P_{S_{n1}|\xi_{1n}} \xrightarrow{w} \Phi_{R_0}$ P-a.s.;

(b) $P_{S_{n1}} \xrightarrow{w} \Phi_{R_0}$;

(c) if $P_{S_{n2}} \xrightarrow{w} Q$, then $P_{S_n} \xrightarrow{w} \Phi_{R_0} * Q$.

The following generalization of Theorem 1 is valid.

Theorem 2. *If $X = (X_j)_{j\geq 1}$ is a stationary conditionally m-dependent sequence of random vectors controlled by an ergodic sequence $\xi = (\xi_j)_{j\geq 1}$ and $\operatorname{tr}(R_m) < \infty$, then*

(a) $P_{S_{n1}|\xi_{1n}} \xrightarrow{w} \Phi_{R_m}$ P-a.s.;

(b) $P_{S_{n1}} \xrightarrow{w} \Phi_{R_m}$;

(c) if $P_{S_{n2}} \xrightarrow{w} Q$, then $P_{S_n} \xrightarrow{w} \Phi_{R_m} * Q$.

Proof. It proceeds according to the scheme by which Theorem 7.7.6 is proved in Anderson's book [14] using the ergodicity of the controlling sequence ξ and conditionally m-dependence of X.

It is evident, that $ES_{n1} = E(S_{n1}|\xi_{1n}) = ES_{n2} = 0$ and using the property of conditionally m-dependence we have

$$E(S_{n1}S_{n1}^{\top}|\xi_{1n}) = \frac{1}{n}\sum_{j,l=1}^{n} E\big[(X_j - \mu(\xi_j))(X_j - \mu(\xi_j))^{\top}|\xi_{1n}\big] =$$

$$= \frac{1}{n}\sum_{j=1}^{n} R(\xi_j, \xi_j) + \frac{1}{n}\sum_{l=1}^{m}\sum_{j=1}^{n-l}\big[R(\xi_j, \xi_{j+l}) + R(\xi_{j+l}, \xi_j)\big]. \tag{2}$$

Since $\{\xi_j\}_{j\geq 1}$ is assumed to be ergodic and $\operatorname{tr}(R_m) < \infty$, we have

$$P\Big\{\lim_{n\to\infty}\frac{1}{n-l}\sum_{j=1}^{n-l} R(\xi_j, \xi_{j+l}) = R_0^{(l)}\Big\} = 1,$$

$$P\Big\{\lim_{n\to\infty}\frac{1}{n-l}\sum_{j=1}^{n-l} R(\xi_{j+l}, \xi_j) = R_0^{(-l)}\Big\} = 1, \quad l = 0, 1, \ldots, m, \tag{3}$$

which together with (1) and (2) give that

$$\mathrm{P}\Big\{\lim_{n\to\infty}\mathrm{E}(S_{n1}S_{n1}^{\top}|\xi_{1n}) = R_m\Big\} = 1.$$

.dom vectors, having for a fixed i the same distribution for any j, and $(\xi_i)_{i\geq 1}$ ι Markov chain with the states $\{1,\dots,s\}$ independent from the mentioned .trix of random vectors (cf. [2]).

Denote (here and below T stands for transposition),

$$\mu(\xi_j) = \mathrm{E}(X_j|\xi_{1n}),$$

$$R(\xi_j,\xi_k) = \mathrm{E}\Big\{[X_j - \mu(\xi_j)]\,[X_k - \mu(\xi_k)]^{\mathrm{T}}\,|\xi_{1n}\Big\} = R(\xi_k,\xi_j)^{\mathrm{T}},$$

$$R(\xi_j) = R(\xi_j,\xi_j), \quad \mu = \mathrm{E}X_j = \mathrm{E}\mu(\xi_j), \tag{1}$$

$$, = \mathrm{E}\Big[R(\xi_1)+\sum_{l=1}^{m}[R(\xi_1,\xi_{l+1})+R(\xi_{l+1},\xi_1)]\Big] \equiv R_0^0+\sum_{l=1}^{m}(R_0^{(l)}+R_0^{-l}), R_0^0 \equiv R_0.$$

Consider the sum

$$S_n = n^{-1/2}\sum_{j=1}^{n}(X_j - \mu),$$

ιich can be splitted into two parts

$$S_n = S_{n1} + S_{n2}$$

th

$$S_{n1} = n^{-1/2}\sum_{j=1}^{n}(X_j - \mu(\xi_j)), \quad S_{n2} = n^{-1/2}\sum_{j=1}^{n}(\mu(\xi_j) - \mu).$$

Our approach is based on the technique and results developed for indepen- nt and m-dependent random summands (see, e.g., [13,14]) when treating $_{n1|\xi_{1n}}$ and the conjecture of existence of a limit distribution Q for $P_{S_{n2}}$. ιere exists such a Q when a finite Markov chain is choosed as a controlling quence in which case we can use the results on the limiting behavior of sums random vectors connected in a Markov chain [15,16,6,7]. As our predeces- rs we are treating the case of single ergodic class of a controlling Markov ain considered, but as in [6,7] due to the ergodicity and using Cesaro limits ; write out the parameters of the limiting normal distribution even when ere exist several periodic subclasses. Again due to ergodicity the limit conerty we use below when studying asymptotic normality of sums of stationary m-dependent sequences of random vectors. The preliminary version of Theorem 2 proved below was announced in the survey paper [9], involving a wide spectrum of results concerning limiting behavior of sums and convergence rate estimates, in the case where the controling sequence is a finite regular Markov chain([9] was continued by the papers [10–12]).

The obtained central limit theorem can be used for studying the asymptotic normality of empirical moments of a conditionally independent sequence controlled by a finite Markov chain. The suitable calculations, as well as the results concerning the limiting distribution of empirical process connected with the latter sequence when both the controlled and controlling sequences are observable published briefly in Russian in [8] will be soon the subject of another publication of the authors.

1.Basic Definitions. On a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ consider a stationary two-component sequence $(\xi_j, X_j)_{j \geq 1}$, where $X_j \in \mathrm{R}^k$ and ξ_j takes his values in a given finite set S. Let $\xi = (\xi_j)_{j \geq 1}$ be an *ergodic sequence* in the sence that for some $m \in \mathrm{N}$ and some functions $f : S^m \to \mathrm{R}^1$ (each specific case will show how wide the set of possible m and f should be) we have the P-a.s. convergence

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\xi_{k+1}, \ldots, \xi_{k+m}) \to \mathrm{E} f(\xi_1, \ldots, \xi_m)$$

as $n \to \infty$, provided that the latter expectation exists.

For random vectors Y and η denote by P_Y and $P_{Y|\eta}$ the absolute and conditional for the given η distributions of Y, respectively. Let Φ_B be the normal distribution in R^k with parameters $(0, B)$. Denote $\xi_{1n} = (\xi_1, \ldots, \xi_n)$.

The sequence of second components $X = (X_j)_{j \geq 1}$ of the mentioned stationary two-component sequence is called the *conditional m-dependent random vectors controlled by the sequence* ξ if for any n and given ξ_{1n} the random vectors $X_1, \ldots, X_n$ become m-dependent, i.e., for any r and any $j_1, \ldots, j_r$ such that $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ the random vectors $X_{j_1}, \ldots, X_{j_r}$ are independint of $X_1, \ldots, X_{j_1-m-1}$ and $X_{j_r+m+1}, \ldots, X_n$ and

$$P_{\ldots} \text{[illegible]} = P_{\ldots} \text{[illegible]}.$$

m-DEPENDENT RANDOM VECTORS CONTROLLED BY A FINITE MARKOV CHAIN

T. Shervashidze[1], Z. Kvatadze[2], I. Bokuchava[3]

1-A. Razamdze Mathematical Institute of the Georgian Sciences .

2- Georgian Technical University, 3-Tbilisi State University

Abstract: Central limit theorem is proved for a stationary conditionally m-dependent sequence of random vectors controlled by an ergodic sequence and in particular by a finite Markov chain. One can deduce from it the asymptotic normality of the joint distribution of empirical moments which are estimates of the theoretical ones of a stationary conditionally independent sequence of random variables controlled by a finite Markov chain.

Key words: Conditionally m-dependent random vectors, controlled sequence, ergodic controlling sequence, controlling Markov chain, central limit theorem.

1 Introduction

It is well-known that the lengths of time intervals between the changes of states of a semi-Markov process are conditionally independent, given the sequence of states, and the distributions of those lengths are determined by the states of the imbedded Markov chain, as though the conditionally independent sequence of lenghts is controlled by that imbedded Markov chain.

This example shows that it is important to consider controlled sequences of conditionally independent random variables and for the sake of generality even those of such a random vectors $(X_j)_{j\geq 1}$ the partial sums of which were

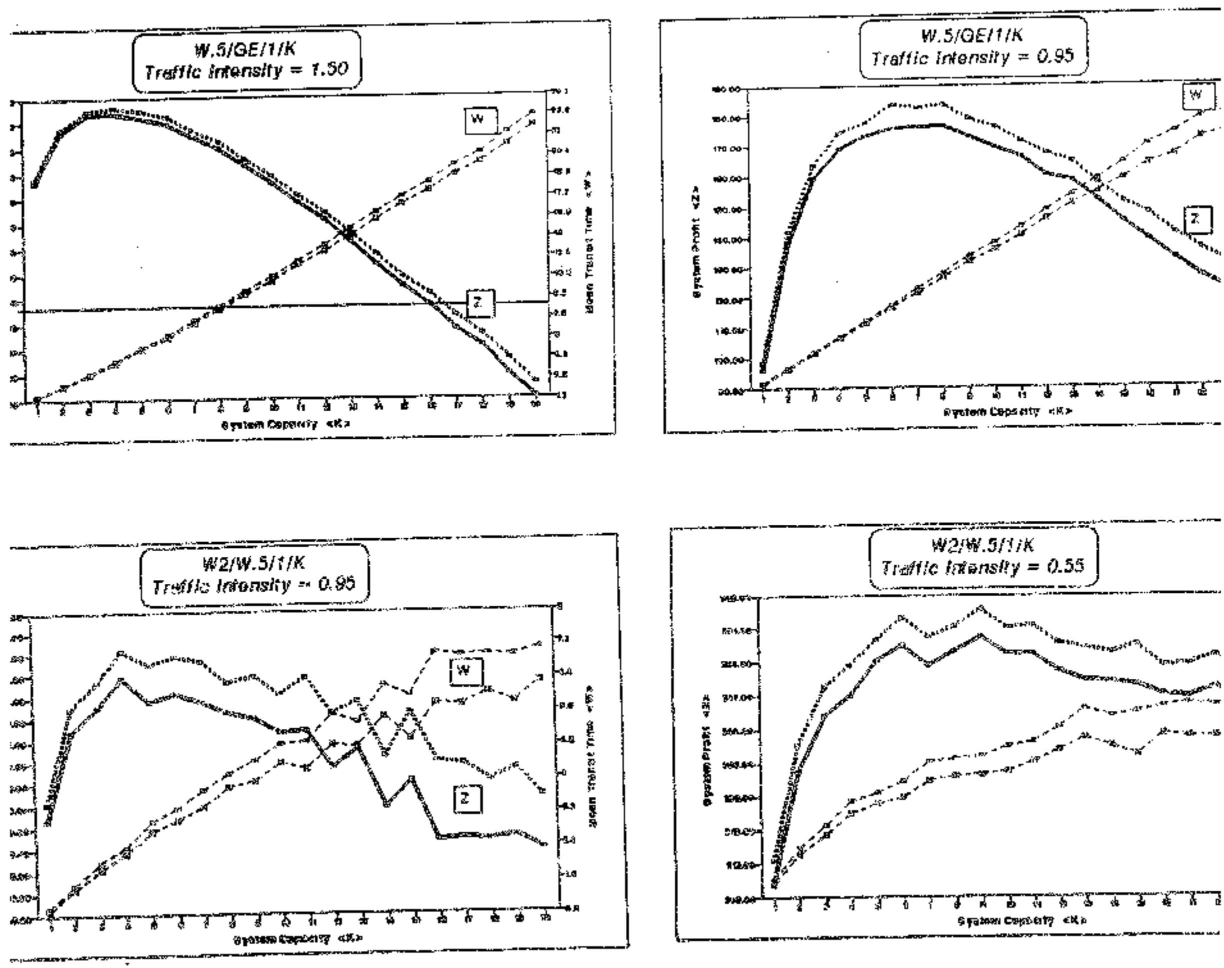
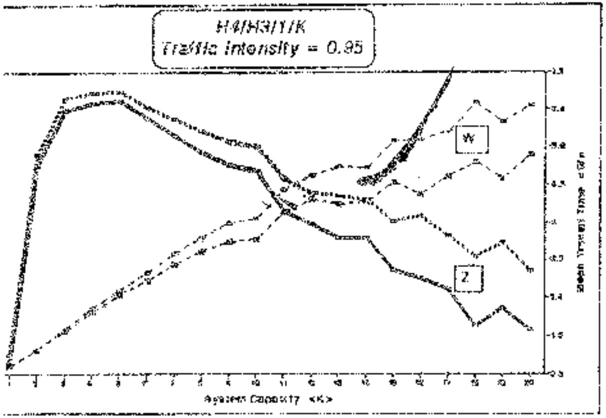
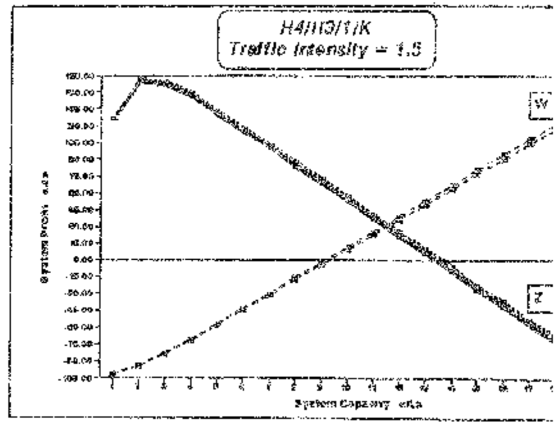
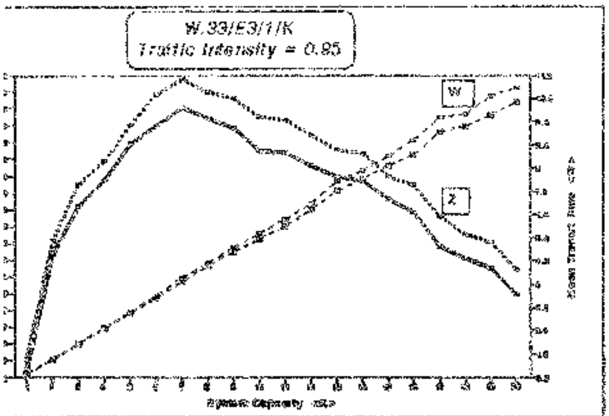
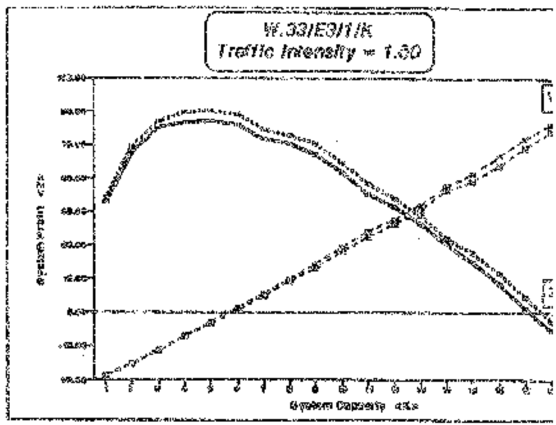

W.5/GE/1/K
Traffic Intensity = 1.50
W
Z
System Capacity <K>
Mean Transit Time <W>
W.5/GE/1/K
Traffic Intensity = 0.95
W
Z
System Profit <Z>
System Capacity <K>
W2/W.5/1/K
Traffic Intensity = 0.95
W
Z
System Capacity <K>
Mean Transit Time <W>
W2/W.5/1/K
Traffic Intensity = 0.55
System Profit <Z>
System Capacity <K>

Fig 5.2


H4/H3/1/K
Traffic Intensity = 0.95
W
Z
System Capacity <K>
Mean Transit Time <W>


H4/H3/1/K
Traffic Intensity = 1.5
W
Z
System Capacity <K>
System Profit <Z>


W.33/E3/1/K
Traffic Intensity = 0.85
W
Z
System Capacity <K>
Mean Transit Time <W>


W.33/E3/1/K
Traffic Intensity = 1.30
System Capacity <K>
System Profit <Z>

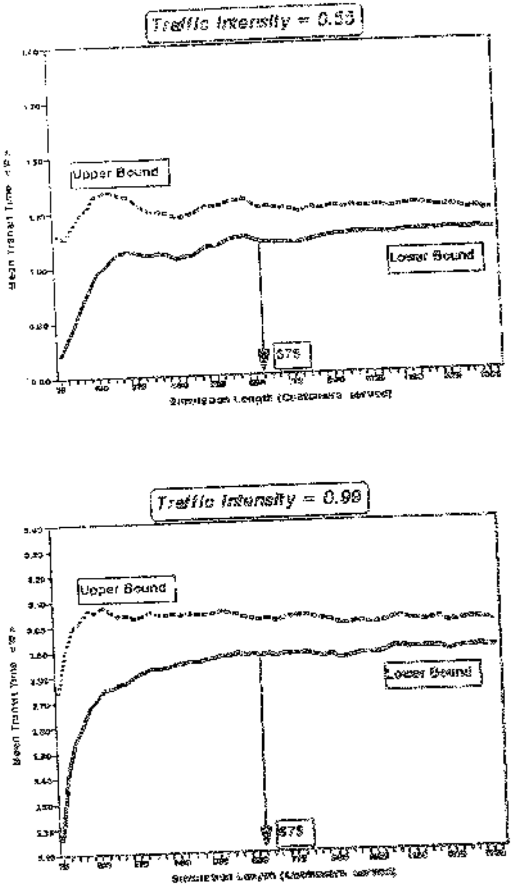

Traffic Intensity =
Upper Bound
Lower Bound
575
Traffic Intensity = 0.99
Upper Bound
Lower Bound
575

Fig 3.2

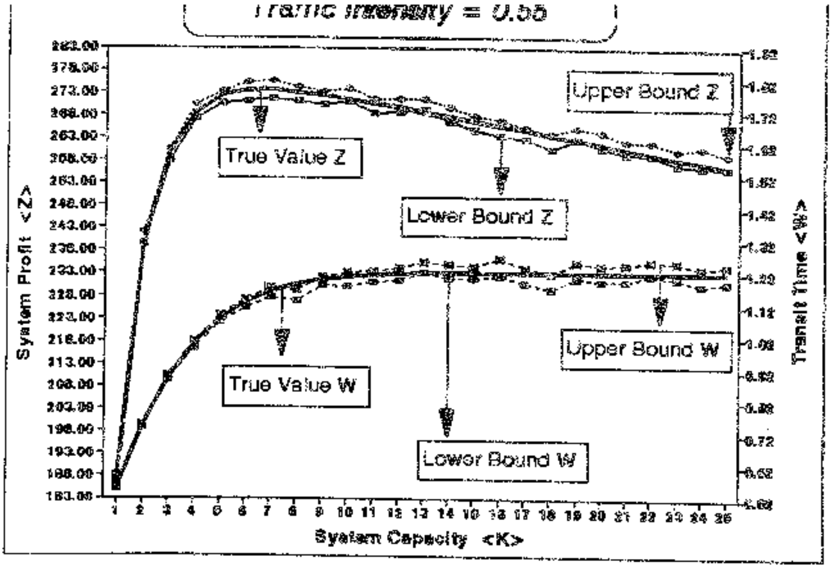
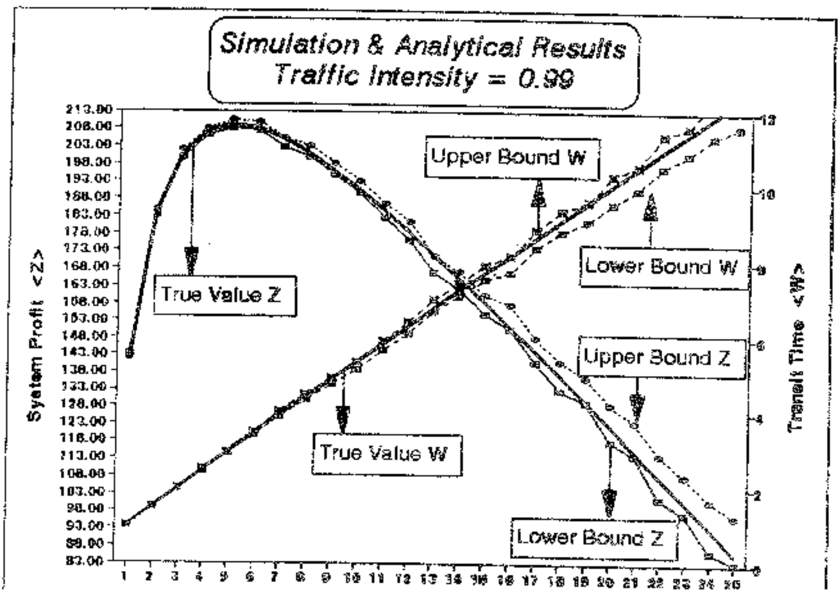

Traffic Intensity = 0.55
Upper Bound Z
True Value Z
Lower Bound Z
Upper Bound W
True Value W
Lower Bound W
System Profit <Z>
Transit Time <W>
System Capacity <K>

Fig 1.2


Simulation & Analytical Results
Traffic Intensity = 0.99
Upper Bound W
Lower Bound W
True Value Z
Upper Bound Z
True Value W
Lower Bound Z
System Profit <Z>
Transit Time <W>

[4] Law, A.M. and Kelton, W.D., (1991) *Simulation Modeling and Analysis.* McGraw Hill.

[5] Pidd, M. (1986) *Computer Simulation in Management Science.* John Wiley, New York.

[6] Rue, R.C., and Rosenshine, M., (1981) *Some Properties of Optimal Control Policies for Entries to an M/M/1 Queue.* Nav.Res. Log. Quart. 28, pp 225 - 232.

[7] Sargent, R.G., (1983) *Validating Simulation Models* Proceedings of the 1983 Winter Simulation Conference, Syracuse Univ. Syracuse, New York.

[8] Shannon, R.E., (1975) *Systems Simulation the Art and Science* Prentice Hall.

[9] Stidham, S.Jr., (1982) *Optimal Control of Arrivals to Queues and Network of Queues.* 21st IEEE Conf. on Decision Control.

$M/E_1/1/K$	36	9	1.22	1.22	242.8	257.3
$M/E_5/1/K$	36	5	0.95	0.88	248.3	281.2
$M/E_\infty/1/K$	36	5	0.89	0.86	251.5	282.7
$GE/E_5/1/K$	36	9	1.22	1.22	242.8	257.3
$E_5/GE/1/K$	36	9	1.22	1.22	242.8	257.3
$GE/GE/1/K$	36	9	1.22	1.22	242.8	257.3

Considering graphs in Figs 4.2 & 4.3, it is observed that as K increases, the confidence intervals of Z and W widens. This means that the variance of transit time is an increasing function of the system capacity.

5. Conclusions

From this study and for the queueing models that were considered the following conclusions are drawn:

1. Irrespective of traffic intensity and queueing models, Z and W are concave in K.
2. The K obtained from social optimization is always equal or less than the K obtained from individual optimization.
3. The social optimization preserves the individual gains and prevents customer from leaving the system.

It is interesting to analyze the sensitivity of the service and interarrival time distributions' moments on the distribution of transit times. The authors are working on this subject at present.

References

[1] Fishman, G.S., (1978) *Principles of Discrete Event Simulation*. John Wiley. New York.

Model	System Capacity		Mean Transit Time		Mean System Profit	
	Indv.	Soci.	Indv.	Soci.	Indv.	Soci.
$M/E_1/1/K$	36	9	1.22	1.22	242.8	257.3
$M/E_5/1/K$	36	5	0.95	0.88	248.3	281.2
$M/E_\infty/1/K$	36	5	0.89	0.86	251.5	282.7
$GE/E_5/1/K$	36	9	1.22	1.22	242.8	257.3
$E_5/GE/1/K$	36	9	1.22	1.22	242.8	257.3
$GE/GE/1/K$	36	9	1.22	1.22	242.8	257.3

$\star$ M : *Exponential interarrival time*

E_k : *Erlang service time distribution with parameter k*

GE : *General Erlang distribution*

As the customers estimate of their gains is based only on $1/\mu$ and not on the arrival and service processes, in some cases they even might incure some losses in entering the system. In light traffic the difference between two mean transit times in individual and social optimization may not be significant in some cases, but in heavy and non-stationary queues the difference is significant (see column 4 of the tables).

Table 4.2 Heavy Traffic ($\rho = 0.95$).

Model	System Capacity		Mean Transit Time		Mean System Profit	
	Indv.	Soci.	Indv.	Soci.	Indv.	Soci.
$M/E_1/1/K$	36	9	1.22	1.22	242.8	257.3
$M/E_5/1/K$	36	5	0.95	0.88	248.3	281.2
$M/E_\infty/1/K$	36	5	0.89	0.86	251.5	282.7
$GE/E_5/1/K$	36	9	1.22	1.22	242.8	257.3
$E_5/GE/1/K$	36	9	1.22	1.22	242.8	257.3
$GE/GE/1/K$	36	9	1.22	1.22	242.8	257.3

; analytical results. In the M/M/1/K model, the following assumptions : made ($l = 1; R = 300; C = 15; D = 1$). Two values of traffic intensity,).55&0.99) that is the mean number of arrivals per mean service time are o considered. As it can be seen from Figs (3.1 & 3.2) the true values of Z d W are always contained within their respective 95% confidence intervals.

Any queueing model that has been considered in this study is fundamenly based on the validated M/M/1/K model. The only difference is that for ch model the interarrival and service times are sampled from the correspond- g distributions. In addition, for each queueing model the fact that for $K = 1$ $V = \frac{1}{\mu}$ has been checked.

4. Discussion of results

The simulation model was run with different arrival and service processes, d also with various values of ρ, R, C and D. The results of some of these ns are presented in Fig 4.1. In analysing the results, it was observed that espective of the aforementioned parameters, W and Z are always concave. ierefore, in comparing the social and individual optimization, it was not felt cessary to change the values of R, C and D.

The results of simulating different queueing models with different traffic ensities for R=300, C=15, and D=1 are given in tables 4.1, 4.2 and 4.3. mparing mean transit times in these tables, one can observe that irrespec- e of ρ and queueing model the social optimization of the system always eserves the customer's benefits more than the individual optimization. The ly exception is M/M/1/K in light traffic where transit times are the same. though in light traffic one does not observe a significant difference between ne of the individual or social transit times (see column 4 of table 4.1), in lividual optimization, the system's profit is significantly reduced (see colm 5 of table 4.1). In heavy traffic and nonstationary queues the differences tween transit times and systems's profits when optimized individually or cially are more significant.

$$GE(x) = 1 - \frac{\quad}{\mu_1 - \mu_2} e \quad + \frac{\quad}{\mu_1 - \mu_2} e \qquad (\mu_1 \neq \mu_2) \qquad (10)$$

$$GE(x) = 1 - e^{-\mu x}(1 + \mu x) \qquad (\mu_1 = \mu_2 = \mu) \qquad (11)$$

we have

$$1 - RN = \frac{\mu_1}{\mu_1 - \mu_2} e^{-\mu_2 x} - \frac{\mu_2}{\mu_1 - \mu_2} e^{-\mu_1 x} \qquad (\mu_1 \neq \mu_2) \qquad (12)$$

where $RN \sim U(0, 1)$. From (11) and by taking $y = e^{-x}$, we get

$$y = exp\left\{ \frac{1}{\mu_2} \ln \left(\frac{\mu_2}{\mu_1} y^{\mu_1} + \frac{(1 - RN)(\mu_1 - \mu_2)}{\mu_1} \right) \right\} \qquad (13)$$

Samples from GE are drawn by solving (12) using numerical methods. And for the Erlang family of random variables, S with probability distribution function:

$$f(x) = \frac{(\mu k)^k}{(k-1)!} x^{k-1} e^{-k\mu_2} \qquad (14)$$

we consider it as the sum of k identically independent exponentially distributed random variables, X . The distribution function of each X is

$$F(x) = 1 - e^{-\lambda x} \qquad , x > 0. \qquad (15)$$

Hence a random sample from the Erlang distribution is generated as:

$$S_k = \frac{-1}{\lambda} \ln \left(\prod_{i=1}^{k} (1 - RN_i) \right). \qquad (16)$$

3. Model Validation

Various authors suggest (For example see [7] and [8]) that for model validation the three steps of 1) Face validation, 2) Testing of model assumptions istributions are considered and simulated. The Erlang distribution is used ecause of its flexibility, and the generalised Erlang distribution (GE) is used ecause it is considered as a good approximation to any distribution. It can e shown that any cumulative distribution function can be approximated to most any degree of accuracy by the convolution of exponential distributions ee e.g. [2]).

2. Simulation

There are three reasons why the system under study is analyzed by simu- tion rather than analytical methods. First, deriving formulae for the system erformance measures (e.g. expected profit rate or mean waiting time) be- omes too complicated when the arrival and service time distributions are as e have considered. Second, the analysis of some of the distribution functions .g. convolution of exponential distributions) requires numerical approaches. hird, we desired to develop an interactive model to enable us to easily test e effect of various input parameters on the system performance.

The system performance measures are collected when the system is in eady state. In order to determine the time period at which the system aches steady state, Welch's method [4] is used. Using Welch's method, 30 dependent replications of the simulation model are made and a plot of mean aiting time is drawn against time. The results of two such simulation runs e shown in Fig 2.1. It is observed that irrespective of the values of r and K eady state is reached at approximately 675 units of time.

Each simulation run consists of arrival and departure of 1500 served cus- mers. The sample size required to establish a given statistical significance a given level of precision is obtained using Central Limit Theorem and quential sampling [8].

The simulation model is constructed based on the event processing ap- oach [5]. To simulate the interarrival and service times, samples are drawn om the corresponding distributions. The method of inverse transformation] is used for this purpose when applicable. For the GE distribution used in is paper, i.e,

f^x

system, and p is the probability of the system being empty.

It can be shown that (see [2] and [3]):

$$p_0 = \frac{\pi_0^*}{\pi_0^* + \rho} \tag{4}$$

where generally π_0^* is the steady state probability of n customers in the system at a departure point.

In order to optimize Z, we need L, the average system size which is obtained from:

$$L = \sum_{n=0}^{K-1} n\pi_n^* \tag{5}$$

where π_n^* can be obtained by solving K equations of:

$$\pi_i^* = \begin{cases} \pi_0^* k_i + \sum_{j=1}^{i+1} \pi_j^* k_{i-j+1} & i = 0, 1, 2, .., K-2, \\ 1 - \sum_{n=0}^{K-2} \pi_n^* & i = K-1. \end{cases} \tag{6}$$

and:

$$k_n = \Pr\{\text{n arrivals during a service time S} = \text{t}\}$$

Although λ' and L, needed to optimize (1) for a M/G/1/K can be obtained theoretically from (3) and (5) respectively, even for the simplest service process, achieving numerical results for M/G/1/K, is complicated [see pp 263 of 2].

In the M/G/1/K system the interarrival time is exponential which has the Markovian property. In some cases the memoryless property of the exponential distribution may not be appropriate. For instance, if the interarrival time distribution is exponential, then due to the memoryless property of this distribution, the time expired since the last arrival has no effect on the remaining interarrival time. This is not appropriate for many situations such as tandom queues.

In an GI/G/1/K queueing system, we consider the following objective func-
on:

$$\text{Max } Z = \text{Max}(R\lambda' - Cl - DK) \tag{1}$$

which:

= Expected system profit per unit of time.

= A fixed fee charged per customer.

= The effective arrival rate,i.e., the mean rate of customers actually entering
e system.

= The cost incurred per customer per unit time.

= The expected system size.

= The cost of providing one unit of space, and

= The system capacity.

We seek the value of K that maximizes Z. This process is referred to as
ocial Optimization, that is, it maximizes benefits to the entire system, rather
an the personal gain of an individual customer.

Now consider the case where customers may renege. If an arriving customer
nds N customers ahead of himself in the system, and if he estimates the
ean service time as $1/\mu$, then his estimates of his waiting time is $(N +$
$/\mu$. Therefore from an individual customer's standpoint we get the following
jective function:

$$\text{Max}Z = R - \frac{C(\mu + 1)}{\mu} \tag{2}$$

The customer will serve his own interest (positive gain) by joining the queue
long as $C(N + 1)/\mu \leq R$, or equivalently, $N \leq R\mu/C - 1$. Hence, if every
stomer uses this strategy, an GI/G/1/K queue results, where K = $[R\mu/C]$.
his process is referred to as Individual Optimization.

Rue and Rosenshine have shown that for the M/M/1/K queueing system,
is concave in K. They have also shown that the K obtained from social
timization of the system is always less than the K obtained when the criterion
the individual gains.

In the rest of this paper, we will try to examine the possibility of generaliz-
g Rue & Roshenshine's remarks to the general system of GI/G/1/K. Results
tained from this study can be valuable in the design and control of queueing

DIFFERENT OPERATIONAL POLICIES USING SIMULATION

G.H. Shahkar, H.R. Tareghian

Ferdosi University of Mashhad

Abstract: In this paper we determine and compare the optimal capacity (K) of a GI/G/1/K queueing system under social and individual optimization. It is shown by simulation that irrespective of the traffic intensity, ρ, and arrival and service time distributions , the K obtained from social optimization of the system is always equal or less than the K obtained from its individual optimization.

keywords:*Simulation, Queueing Theory, Optimization.*

1. Introduction

Usually human beings arriving at a queueing system estimate their waiting time in the system. Based on this estimate, they decide whether to stay or leave the system. However, due to the short period of time during which customers are observing the behaviour of the queue, they can not correctly guess the waiting time distribution. Hence their estimate of their waiting time may be incorrect. If we consider that all the customers follow the same strategy, then the system reaches a self determined capacity (K).

If we determine the optimum capacity (say K') with respect to systems profit without considering the customers estimate of their waiting time, do we still satisfy the individual customer's benefit? Does not the customers estimate of their waiting time in the *optimized system* discourage them from staying in and getting served? In other words, is $K' \leq K$? And finally, if the system is designed with capacity equal to K, does the server's capabilities optimally utilized?

Rue and Rosenshine [6] have shown that if a M/M/1/K system is designed

knowledgement

I would like to thank my supervisor Professor John Biggins for his unfailing
istance throughout the duration of this work.

References

[1] BIGGINS, J.D. (1976). The first-and last-birth problems for a multitype
-dependent branching process. *Adv. Appl. Probab.* **8**, 446–459.

[2] BIGGINS, J.D. (1977). Martingale convergence in the branching random
k. *J. Appl. Probab.*.**14**, 25–37.

[3] BRAMSON, M., NEY, P. AND TAO, J. (1992). The population compo-
on of a multitype branching random walk. *Ann. Appl. Probab.* **2**, 575-596.

[4] COX, D.R. AND [11], H.D. (1965). *The Theory of Stochastic Process.*
ndon: Chapman-Hall.

[5] DONEY, R.A. (1972). A limit theorem for a class of supercritical branch-
processes. *J. Appl. Probab.* **9**, 707-724.

[6] DONEY, R.A. (1973). On a functional equation for general branching
cesses. *J. Appl. Probab.* **10**, 198-205.

[7] FREEDMAN, D. (1971). *Markov Chains.* Holden-Day, San Francisco.

[8] JENSEN, J.L. (1995). *Saddlepoint Approximations.* Oxford University
ss.

[9] KESTEN, H. AND STIGUM, B.P. (1966). A limit theorem for multi-
nensional Galton-Watson process *Ann. Math. Stat.* **37**, 1211-1223.

[10] KESTEN, H. AND STIGUM, B.P. (1967). Limit theorems for decom-
sable multi-dimensional Galton-Watson processes. *J. Math. Ann. Appl.* **17**,
-338.

[11] MILLER, H.D. (1961). A convexity property in the theory of random
iables defined on a finite Markov chain. *Ann. Math. Statist.*, **32**, 1260-1270.

[12] MODE, C.J. (1971). *Multitype Branching Processes.* New York, Amer-
n Elsevier.

[13] NEY, P. AND NUMELLIN, E. (1987, I, II). Markov Additive Processes,
Ann. Probab. **15**, 561-592, 593-609.

[14] QUANSHENG LIU (1998). Fixed point of genralized smoothing trans-
mation and application s to the branching random walk. *J. Appl. Probab.*

and

$$-\theta\frac{\rho'(\theta)}{\rho(\theta)}+\log\rho(\theta)>0. \tag{29}$$

Moreover if either of these fail then

$$E[W_i(\theta)]=0 \quad \forall i=1,\ldots,p, \tag{30}$$

hence all $W_i(\theta)$ are degenerate at zero.

Proof. In Proposition 5 we proved that, when the equation $\phi=H\phi$ has no $\mathcal{L}_1^p$ solution then for all $i=1$, $E[W_i(\theta)]=0$. Also, $\chi<0$ is equivalent to the supercriticality of the process. The infinite differentiability of the Laplace transforms of the measures $\mu_{ij}(.,\theta)$ implies the finiteness of the expectations of the r. variables $|X_{ij}|$ and $(X_{ij}^+)^2$ (similar to the proof in Lemma 4 in [2]).

Then the proof of the theorem is completed by showing that the existence of a solution for the functional equation $\phi=H\phi$ is equivalent to the existence of a non-degenerate limit for the martingale $\{W_i^n(\theta)\}$.

Let $\theta\in\Omega$ and $\theta>0$ be fixed. The equations

$$E\left[N_{ij}(\infty)|\log N_{ij}(\infty)|\right]<\infty \quad \forall i,j=1,\ldots,p$$

are equivalent to

$$E[W_i^1(\theta)|\log W_i^1(\theta)|]<\infty \quad \forall i=1,\ldots,p.$$

Therefore, the conditions of the theorem are equivalent to the conditions of Theorem 1; and hence are equivalent to the existence of an $\mathcal{L}_1^p$ solution for the functional equation $\phi=H\phi$.

Let the conditions of Theorem hold, then the conditions of the Theorem 1 hold. Define $\hat\phi_i^0(u)=E[e^{-uW_i^0(\theta)}]=e^{-u}$ and $\hat\phi_i^n(u)=E[e^{-uW_i^n(\theta)}]$ for all $n=1,\ 2,\ \ldots$. Then $\{e^{-uW_i^n(\theta)}\}$ is a submartingale so $\hat\phi_i^{n+1}(u)\geq\hat\phi_i^n(u)$. Thus, by applying the results of the lemma 2, as $n\to\infty$, $\hat\phi_i^n(u)$ converges to the Laplace transform $\hat\phi_i(u)=E[e^{-uW_i(\theta)}]$, which is an $\mathcal{L}_1^p$ solution of the equation $\phi=H\phi$. So, for all i, $E[W_i(\theta)]=1$; and by considering $E[W_i^1(\theta)]=1$, we get the convergence of $\{W_i^n(\theta)\}$ in mean.

Recall that, for any fixed $\theta \in \Omega$, the random collection of positive real numbers $\{y_{ij}(s); s\}$ defined in terms of the random discrete measure $\{Z_{ij}(dx, \theta)\}$ (see (6)). We apply Theorem 1 to the random discrete measures $Z_{ij}(., \theta)$, $i, j = 1, \ldots, p$, to obtain the main result, Theorem 2. Let $\theta \in \Omega$ be fixed and let $c_{i,j} = \log\left(v_j(\theta)/(\rho(\theta)v_i(\theta))\right)$. The martingale $\{W_i^n(\theta)\}$ converges to the limit $W_i(\theta)$, as $n \to \infty$, with $E[W_i^n(\theta)] = 1$; and by Fatou's lemma, $E[W_i(\theta)] \in [0, 1]$. Let $\phi_i(u) = E[e^{-uW_i(\theta)}]$ be the Laplace transform of $W_i(\theta)$, then they satisfy the equation

$$\phi_i(u) = E \prod_j \prod_s \phi_j(uy_{ij}(s)), \quad \forall i = 1, \ldots, p. \tag{26}$$

Recall the r. variables X_{ij}, the Markov chain $\{k_n\}$ with transition matrix $P = \{p_{ij}\}$ and stationary distribution $\pi = (\pi_1, \ldots, \pi_p)$; and the additive component $\{S^n\}$ with mean drift χ; all defined in terms of $y_{ij}(s)$. This mean drift satisfies

Proposition 4 *For each $\theta \in \Omega$ we have $\chi = (\theta\rho'(\theta))/\rho(\theta) - \log\rho(\theta)$. Moreover, when $\chi < 0$ the process is supercritical.*

Proof. We use the abbreviations m_{ij}, v_i, u_i, and ρ, the mean drift of the random walk $\{S^n\}$, is $\chi = \sum_i \sum_j (u_i v_i p_{ij}) E[X_{ij}]$, where, $p_{ij}E[X_{ij}] = E[\sum_s y_{ij}(s) \log y_{ij}(s)]$. Writing in terms of $Z_{ij}(s)$ and using $\rho' = uM'v$ gives the result. The supercriticality follows from $\chi < 0$ by considering the function $h(\gamma) = \log\rho(\gamma\theta) - \gamma\log\rho(\theta)$.□

Proposition 5 *If the equation $\phi = H\phi$ has no $\mathcal{L}_1^p$ solution then $E[W_i(\theta)] = 0$ for all i.*

Proof. Suppose the functional equation $\phi = H\phi$ has no solution in $\mathcal{L}_1^p$ so, it has no $\mathcal{L}^p$ solution. But the Laplace transforms ϕ_i, $i = 1, \ldots, p$, satisfy this equation, hence $\phi \notin \mathcal{L}^p$ but $\phi \in \bar{\mathcal{L}}^p$. Consequently, for at least one $i, E[W_i(\theta)] = 0$. Assume for some $k, E[W_k(\theta)] > 0$, then $\phi_k(u) < 1$ for $u > 0$. From positive regularity, for some i and some k, $P(Z_{ik} > 0) > 0$. So, for this i, $1 = \phi_i(u) = E\left[\prod_{j,s} \phi_j(uy_{ij}(s))\right] \leq E\left[\prod_{j,s} 1\right]$. The product must involve ϕ_k so, an argument similar to that used in Proposition 3 gives a contradiction.□

$$E_i\left[h\left(\frac{u}{u_o}e^{S^n}\right)\right] \le \frac{1}{(\log(u/u_o)+n(\chi+\epsilon))^2} + P_i\left(S^n \ge n(\chi+\epsilon)\right). \quad (22)$$

Recall that $\chi+\epsilon<0$, sum over n, then the first sum on the right hand side converges. We apply Proposition 2 to the second sum on the right hand side to get

$$\sum_{n=0}^{\infty} E_i\left[h\left(\frac{u}{u_0}e^{S^n}\right)\right] < \infty. \quad (23)$$

A similar argument gives

$$\sum_{n=0}^{\infty} E_i\left[h\left(\frac{u}{u_0}e^{S^{n+1}}\right)\right] < \infty. \quad (24)$$

Now we show that the first series on the right hand side of (21) diverges to infinity. The matrix P is positive regular, so for each j, $p_{ij}^n \to \pi_j > 0$, as $n\to\infty$. Let N_0 be such that for $n \ge N_0$, $p_{ij}^n > \delta > 0$, for some δ and all i, j. Then we get

$$\sum_{n=0}^{\infty}\sum_k p_{i,k}^n E_i\left[\psi_k\left(\beta u e^{S^n}\right)\middle|\, k_n=k\right] \;\; \ge \;\; \delta \sum_{n=N_0}^{\infty}\sum_k E_i\left[\psi_k\left(\beta u e^{S^n}\right)\middle|\, k_n = k\right].$$

By taking only the terms with $k_n = i$

$$\ldots \ge \delta \sum_{n=N_0}^{\infty} E_i\left[\psi_i\left(\beta u e^{U_i^n}\right)\right]. \quad (25)$$

The random walk $\{U_i^n\}$ has identical increments with mean drift χ/π_i. So, by the week law of Large Numbers, $U_i^n/n \to (\chi/\pi_i)$ in probability. Then for some $\epsilon > 0$, and integer N_0,

$$\sum_{n=0}^{\infty}\sum_k p_{i,k}^n E_i\left[\psi_k\left(\beta u e^{S^n}\right)\middle|\, k_n=k\right] \;\; \ge \;\; \frac{\delta}{2}\sum_{n=N_0}^{\infty} \psi_i\left(\beta u e^{n(\chi/\pi_i-\epsilon)}\right).$$

$$E\left[N_{ij}(\infty)\left|\log N_{ij}(\infty)\right|\right] = \infty, \tag{18}$$

then the equation $\phi = H\phi$ has no solution in $\mathcal{L}_1^p$.

Proof. The proof is similar to Lemma 3 in [2]. Let $\phi = H\phi$. Then conditional starting state i,

$$\phi_i^*(u) + \sum_{l=0}^{n-1}\sum_j \int_R A_j(ue^y)\nu_{i,j}^{l*}(dy) \quad \leq \quad E_i\left[\phi^*(ue^{S^n})\right],$$

and as $n \to \infty$,

$$\phi_i^*(u) + \sum_{l=0}^{\infty}\sum_j \int_R A_j(ue^y)\nu_{i,j}^{l*}(dy) \leq 1. \tag{19}$$

We show that the left hand side is infinity, and this contradiction completes the proof. For this purpose, we find a lower bound for the functions $A_i(.)$. For each i and $u > 0$ define $L_i(u) = u[\beta\psi_i(\beta u) - A_i(u)]$. By following the proof of Lemma 3 in [2], one can show that,

$$L_i(u) \leq Muh(\frac{u}{u_0}) + MuE_i[h(u\beta e^{S^1})], \tag{20}$$

where, M, β, u_0, are positive constants and h is the function defined by $h(u) = (\log u)^{-2}$ for $0 < u < 1/e$, and $h(u) = 1$ otherwise. Using the definition of $L_i(u) = u(\beta\psi_i(\beta u) - A_i(u))$, gives

$$A_i(u) \quad \geq \quad \beta\psi_i(\beta u) - Mh\left(\frac{u}{u_0}\right) - M\sum_j \int_R h\left(\beta u e^x\right)\nu_{ij}(dx).$$

Write the above equation with replacing i by k and u by ue^y, then integrating with respect to $\nu_{i,k}^{n*}(dy)$,

$$\sum_{n=0}^{\infty}\sum_k p_{ik}^n E_i\left[A_k\left(ue^{S^n}\right)\middle| k_n = k\right]$$

∞

$$A_i(u) = \frac{1}{u} E\left[\sum_{j,s}(1-\phi_j(uy_{ij}(s))) - \left(1 - \prod_{j,s}\phi_j\left(uy_{ij}(s)\right)\right)\right], \tag{15}$$

and let $\underline{A}(u) = \min_i A_i(u)$. Then $A_i(u)$ and $\underline{A}(u)$ are positive, bounded and continuous.

Let i be fixed then, we get

$$\phi_i^*(u) + A_i(u) = \sum_j \int_R \phi_j^*(ue^x)\nu_{ij}(dx).$$

Iterating by $\nu_{ij}(.)$ gives:

$$\phi_i^*(u) + \sum_{l=0}^{n-1} E_i\left[A(ue^{S^l})\right] \leq E_i\left[\phi^*(ue^{S^n})\right]. \tag{16}$$

Lemma 3 *If $\chi \geq 0$ then the equation $\phi = H\phi$ has no $\mathcal{L}_1^p$ solution.*

Proof. If $\chi \geq 0$ and $\phi = H\phi$ in $\mathcal{L}_1^p$, then we obtain a contradiction. For any compact set $I \subset (0,\infty)$ there is a constant $C > 0$ such that $u \in I \rightarrow \underline{A}(u) \geq C > 0$. Then we have

$$E_i\left[\underline{A}(ue^{S^n}))\right] \;\; \geq \;\; C\sum_j \int_I \nu_{ij}^{n*}(dx) = CP_i\left(S^n \in I\right),$$

where, P_i means that the probability is conditional on $k_0 = i$. Applying the technique of Lemma 3 in [2], one can complete the proof.□

In the next lemma we will prove the necessity of the second condition in the theorem. The proof requires some lower bounds for the functions $A_i(u)$. Let $\phi = H\phi$. Then $\phi_i(u) > e^{-u}$, $(u > 0)$ for all i. In this case $\min_i\{\phi_i(1)\} > e^{-1}$. Let $\beta > 0$ be such that $e^{-1} < e^{-\beta} < \min_i \phi_i(1)$. Then $0 < \beta < 1$ and for each i, we can define the function

$$\hat{\phi}_i(u) = e^{\beta u}\phi_i(u), \quad (u > 0), \tag{17}$$

$$g_i^n(u) = \frac{1}{u}(\phi_i^n(u) - \phi_i^{n-1}(u)),$$

ınd let $g^n(u) = \max_i \{g_i^n(u)\}$. Then $g_i^n(u)$ and $g^n(u)$ are non-negative and :ontinuous functions on $(0, \infty)$. Let i and $u > 0$ be fixed,

$$g_i^{n+1}(u) \leq \sum_{j=1} \int_R g_j^n(ue^x)\nu_{ij}dx. \tag{14}$$

:he applying proposition 3 to the functions g_i^n implies $g_i^{n+1}(u) \leq E_i g^1(ue^{S^n})$. 'rom $g^1(u) \leq \psi(u)$ we get $g_i^{n+1}(u) \leq E_i[\psi(ue^{S^n})]$, hence

$$\sum_{n=2}^{\infty} g_i^n(u) \;\; \leq \;\; \sum_{n=2}^{\infty} E_i[\psi(ue^{S^n})] \leq \sum_{n=1}^{\infty} P_i(S^n \geq -nc) + \sum_{n=0}^{\infty} \psi(ue^{-nc}).$$

Iow we show each of the last two series converge. Since $-c > \chi$ the conditions f the Lemma 1 hold and $\sum_{n=1}^{\infty} P_i(S^n \geq -nc) < \infty$. If we set $Y_i = \sum_j N_{ij}(\infty)$ hen $E[N_{ij}(\infty)|\log N_{ij}(\infty)|] < \infty$ which implies $E[Y_i|\log Y_i|] < \infty$, for each i. Iow by applying the Lemma 3.4 in [5] to each ψ_i we get $\sum_{n=1}^{\infty} \psi_i(ue^{-nc}) < \infty$. 'hen $\psi(e^{-nc}) \leq \sum_{i=1}^{p} \psi_i(ue^{-nc})$ implies

$$\sum_{n=1}^{\infty} \psi(ue^{-nc}) \leq \sum_{i=1}^{p} \sum_{n=1}^{\infty} \psi_i(ue^{-nc}) < \infty.$$

ow an argument similar to proof of Lemma 2 in [2] shows that $\phi(u)$, the limit f $\{\phi^n(u)\}$, is a solution of $\phi = H(\phi)$ in $\mathcal{L}_1^p$.

To prove the uniqueness, let $\phi = (\phi_1, \ldots, \phi_p)$ and $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_p)$ be vo solutions of the equation $\phi = H(\phi)$ in $\mathcal{L}_1^p$. Define, for $u > 0$, $g_i(u) = |\phi_i(u) - \tilde{\phi}_i(u)|$ and $g(u) = Max_i g_i(u), (i = 1, \ldots, p)$. Then $g_i(u)$ is continuous onnegative and bounded on $[0, \infty)$. With a similar argument to that for g_i^n,

$$g_i(u) \;\; \leq \;\; \sum_j \int_R g_j(ue^x)\nu_{ij}(dx).$$

point in $\mathcal{L}^p$.

(ii). *Let $\phi = (\phi_1, \ldots, \phi_p)$ be a solution of $\phi = H\phi$ in $\mathcal{L}_1^p$. Then for all $i, \phi_i(u) = e^{-u}$ or for all $i, \phi_i(u) > e^{-u}$; for $u > 0$.*

To complete the proof note that if $\phi = H\phi$ in $\mathcal{L}_1^p$ with $\phi_i(u) = E[e^{-u\hat{W}_i}]$ then by Jensen's inequality we have

$$1 = \phi_i(u) = E[e^{-u\hat{W}_i}] \geq e^{-uE[\hat{W}_i]} = e^{-u}.$$

So, by the definition of H and Jensen's inequality

$$e^{-u} = \phi_i(u) = E\prod_j\prod_s \phi_j(uy_{ij}(s)) \geq E\prod_j\prod_s e^{-uy_{ij}(s)} \geq e^{-u}.$$

Since $\{y_{ij}(s); s\}$ is non-degenerate, then for any $u > 0$ with a positive probability $\prod_j\prod_s \phi_j(uy_{ij}(s)) > \prod_j\prod_s e^{-uy_{ij}(s)}$, which implies part (ii).

Define the sequence $\{\phi^n\}$ as follow: For $u \in [0,\infty)$ let $\phi^0(u) = (\phi_1^0(u), \ldots, \phi_p^0(u)) = (e^{-u}, \ldots, e^{-u})$; and define the sequence $\{\phi^n(u)\}$ by induction: $\phi^{n+1}(u) = (H\phi^n)(u)$. Hence each $\phi_i^n(u)$ is convex, decreasing in $u \in [0,\infty)$ and differentiable of each order for $u \in (0,\infty)$; moreover,

$$\phi_i^n(0^+) = 1, \quad (\phi_i^n)'(0^+) = -1, \quad 0 \leq \phi_i^n(u) \leq 1.$$

Also, by induction and using Jensen's inequality, $\{\phi_i^n(u)\}$ is increasing **in** n hence, converges to a function, say $\phi_i(u)$.

Now we are able to apply the proofs in [2] to the multitype case in **proving**:

Theorem 1 *Let for all i, j, $E|X_{ij}| < \infty$ and $E\left[\left(X_{ij}^+\right)^2\right] < \infty$. Then the conditions*

$$-\infty < \chi < 0, \; E\left[N_{ij}(\infty)\left|\log N_{ij}(\infty)\right|\right] < \infty,$$

for all i, j; are necessary and sufficient for the functional equation $\phi = H\phi$ to have a solution in $\mathcal{L}_1^p$.

Proof. The proof of this theorem is divided into three lemmas.□

First we introduce some functions. Similar functions are studied in [5]. For each i and $u > 0$ define $\psi_i(u) = \frac{\phi_i^1(u)+1-u}{u}$ and let $\psi(u) = \max_i \psi_i(u)$. Then by letting $\psi)i(0) = 0$ the function ψ_i and $\psi$are continuous on $[0,\infty)$ an

$$g_i(u) \le \sum_h \int_{R,dz} \left[\sum_l \int_{R,dt} g_l (ue) \nu_{hl}(dt) \right] \nu_{ih}(dz).$$

By changing the variable of integral from t to $t-z$, changing the order of integrals, and applying the definition of convolutions of $\nu_{ij}^{n*}(.)$,

$$g_i^{n+1}(u) \le \sum_l \int_{R,dt} g_l^{n-k-1}(ue^t)\nu_{il}^{(k+2)*}(dt).$$

Therefore, by induction in k, the equation (11), holds for all $1 \le k < n$. So for $k = n-1$, with $g(u) = \max_i g_i^1(u)$,

$$g_i^{n+1}(u) \le \sum_j \int_R g_j^1(ue^x)\nu_{ij}^{n*}(dx) \le \sum_j \int_R g(ue^x)\nu_{ij}^{n*}(dx). \tag{12}$$

Thus $g_i^{n+1}(u) \le E_i\left[g(\hat{u}e^{S^n})\right]$. $\square$

Now we define the p-dimensional space of Laplace transforms of p-vector of distribution functions of non-negative r. variables.

Definition 1 *Let $\mathcal{L}_\mu$, $\mathcal{L}$ and $\bar{\mathcal{L}}$, be the set of Laplace transforms of all real valued r. variables on $[0,\infty)$ with finite mean $\mu(\in(0,\infty))$, finite strictly positive mean, and finite mean, respectively. Define $\mathcal{L}_\mu^p = \mathcal{L}_\mu \times \mathcal{L}_\mu \times \ldots \times \mathcal{L}_\mu$ and similarly define $\mathcal{L}^p$ and $\bar{\mathcal{L}}^p$.*

Definition 2 *The operator $(H\phi)(u) = \tilde{\phi}(u)$ is defined for $\bar{\mathcal{L}}^p$ for $\phi = (\phi_1, \ldots, \phi_p)$ by*

$$\tilde{\phi}_i(u) = E\prod_{j=1}^{p}\prod_s \phi_j(uy_{ij}(s)).$$

It is easy to verify that H is closed in $\mathcal{L}^p$ and $\mathcal{L}_\mu^p$. Moreover, let $\tilde{\phi} = H\phi$ with $\phi = (\phi_1, \ldots, \phi_p) \in \mathcal{L}_{\mu_1} \times \ldots \times \mathcal{L}_{\mu_p}$. Then $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_p) \in \mathcal{L}_{\tilde{\mu}_1} \times \ldots \times \mathcal{L}_{\tilde{\mu}_p}$ where,

$$(\tilde{\mu}_1, \ldots, \tilde{\mu}_p)' = P(\mu_1, \ldots, \mu_p)'.$$

So, by positive regularity of P, $\mu' = P\mu'$ implies $\mu_1 = \mu_2 = \ldots = \mu_p = \mu$, for

$$E\left[\sum_s y_{ij}(s) f(\log y_{ij}(s))\right] = \int_R f(x)\nu_{ij}(dx). \tag{7}$$

Let $\nu(.) = \{\nu_{ij}(.)\} = \{p_{ij}F_{ij}(.)\}$ then it is similar to the matrix of tilted intensity measures $\{\mu_{ij}(.,\theta)\}$. Its n-fold convolution $\nu^{n*}(.) = \{\nu_{ij}^{n*}(.)\}$ is defined by induction

$$\nu_{ij}^{n*}(dx) = \sum_k \int_R \nu_{ik}^{(n-1)*}(dy)\nu_{kj}(dx-y). \tag{8}$$

Let $\varphi_{ij}(u)$ be the Laplace transform of $F_{ij}(.)$, and for $P(u) = \{p_{ij}\varphi_{ij}(u)\}$ we write $P^n(u) = \{p_{ij}^n \varphi_{ij}^{(n)}(u)\}$, where, $\{p_{ij}\}^n = \{p_{ij}^n\}$. Then $p_{ij}^n \varphi_{ij}^{(n)}(u)$ is the Laplace transform corresponding to $\nu_{ij}^{n*}(.)$. From Lemma 9.1.2 in [8] $\phi_{ij}^n(u)$ is the Laplace transform of S^n conditional on $k_0 = i$ and $k_n = j$, where S^n is a random walk with Markov dependent increments related to the random variables X_{ij}. Moreover, conditional on $k_0 = i$,

$$E_i\left[g\left(S^n\right)\right] = \sum_j \int_R g\left(x\right)\nu_{ij}^{n*}(dx). \tag{9}$$

A result on iterating of $\nu(.)$ is:

Proposition 2 *Let $g^n(u) = (g_1^n(u), \ \ldots, \ g_p^n(u))$, $n = 1, \ 2, \ \ldots$, be a sequence of vector valued functions defined on $(0,\infty)$, with each $g_i^n(u)$ being bounded non-negative and continuous in $u > 0$. Assume further that, for each $n \geq 1$, $u > 0$, and i, the sequence $\{g_i^n(u)\}$ satisfies the functional relation*

$$g_i^{n+1}(u) \leq \sum_j \int_R g_j^n(ue^x)\nu_{ij}(dx). \tag{10}$$

Let $g(u) = \max_i g_i^1(u)$ then for all n and $u > 0$,

$$g_i^{n+1}(u) \ \ \leq \ \ E_i\left[g(ue^{S^n})\right],$$

where, E_i means that the expectation is conditional on $k_0 = i$.

$0 < r < 1$,

$$E\left[\left((\bar{X}_j^k)^+\right)^2\right] \leq \sum_{n=1}^{\infty} Ar^n \left[nM_1 + 2\frac{n(n-1)}{2}M_1\right] \leq \sum_{n=1}^{\infty} Ac' n^2 r^n < \infty.$$

ain by Proposition (79) in [7]) we can bound B.□

3 Functional Equation

In this section we define an operator then we find a necessary and sufisient dition for existing a unique fixed point of this operator.

For any $\theta \in \Omega$, define the tilted measure $\mu_{ij}(.,\theta)$ by

$$\mu_{ij}(dx,\theta) = \frac{v_j(\theta)}{v_i(\theta)} \frac{e^{-\theta x}}{\rho(\theta)} \mu_{ij}(dx). \tag{4}$$

e measure $\mu(.,\theta)$ is the kernel of an operator related to $\mu(.)$ (see [8]). From nma 3.1 in [11], $P = \{p_{ij}\} = \{\mu_{ij}(\mathcal{R},\theta)\}$ is a positive regular stochastic trix with invariant measures

$$\pi(\theta) = (\pi_1(\theta), \ldots, \pi_p(\theta)), \quad \pi_k(\theta) = u_k(\theta)v_k(\theta), \quad (k = 1, \ldots, p).$$

analogy, define the tilted measure $Z_{ij}(.,\theta)$, related to $Z_{ij}(.)$, with

$$Z_{ij}(dx,\theta) = \frac{v_j(\theta)}{v_i(\theta)} \frac{e^{-\theta x}}{\rho(\theta)} Z_{ij}(dx). \tag{5}$$

r any fixed $\theta \in \Omega$, and any i, j, to the tilted measure $Z_{ij}(.,\theta)$ there corre- onds a random collection of positive reals $\{y_{ij}(s); s\}$ where, s is an enumer- on variable, given by the equation:

$$y_{ij}(s) = \frac{v_j(\theta)}{v_i(\theta)} \frac{e^{-\theta Z_{ij}(s)}}{\rho(\theta)}, \quad (s = 1,\ 2,\ \ldots,\ Z_{ij}). \tag{6}$$

e fix $\theta \in \Omega$, then we talk about the *θ-type objects,* so, in all that follows ien we are dealing with $\{y_{ij}(s) :\ s\}$ the objects of our study depend on this each of these groups we get, as $n \to \infty$,

$$\frac{S_j^{N_j}}{N_j}\frac{N_j}{n} \to \chi_j \pi_j, \quad \text{(a.s.)}.$$

If we let $n \to \infty$ in both sides of the equation

$$\frac{S^n}{n} = \sum_{j=1}^{p} \left(\frac{S_j^{N_j}}{N_j}\right)\left(\frac{N_j}{n}\right), \tag{2}$$

Proposition 1 *(i) If for each j, $E|X_j| < \infty$, then, as $n \to \infty$,*

$$\frac{S^n}{n} \longrightarrow \chi = \sum_{j=1}^{p} \pi_j \chi_j \quad \textit{(a.s.)}.$$

(ii) If $\chi = 0$ then S^n is persistent and if $\chi \neq 0$ then S^n is transient. Moreover if $\chi > 0$ then $S^n \to \infty$ (a.s.) and if $\chi < 0$ then $S^n \to -\infty$ (a.s.).

To prove (ii) define $U_j^k = S^{N_j^k}$ then $\{U_j^n\}$ is random walk with identical increments. For any set $F \subset \mathcal{R}$, $\sum_{n=1}^{\infty} P(S^n \in F) = \sum_{j=1}^{p}\sum_{n=1}^{\infty} P(U_j^n \in F)$, then by $\chi = \sum_j \pi_j \chi_j$ and all $\pi_j > 0$, it is easy to complete the proof.

Lemma 1 *Let $c > \chi$. If for each j, $E|X_j| < \infty$, and $E[(X_j^+)^2] < \infty$, then $\sum_{n=1}^{\infty} P(S^n \geq nc) < \infty$.*

Proof. Let $c > \chi$. The sequences $\{N_1^k\}, \ldots, \{N_p^k\}$ make a partition of all positive integers, hence

$$\sum_{n=1}^{\infty} P(S^n \geq nc) = \sum_{j=1}^{p}\sum_{k=1}^{\infty} P\left(U_j^k \geq k\left(\frac{N_j^k}{k}c\right)\right).$$

For each j let $(\chi/\pi_j) < c_j < (c/\pi_j)$. Then we can continue

$$\cdots \leq \sum_{j=1}^{p}\sum_{k=1}^{\infty} P\left(U_j^k \geq kc_j\right) + \sum_{j=1}^{p}\sum_{k=1}^{\infty} P\left(N_j^k c < kc_j\right) = A + B. \tag{3}$$

$\mathbf{H}_p$: *$M(\theta)$ is positive regular for all $\theta \in \mathcal{R}$.*
erefore, by the Perron-Frobenius Theorem (Theorem 1 [16]), for each $\theta \in \Omega$, re exists a simple positive maximum eigenvalue $\rho(\theta)$ and corresponding ctly positive right and left eigenvectors $v(\theta) = (v_1(\theta), \ldots, v_p(\theta))$ and $u(\theta) =$ $(\theta), \ldots, u_p(\theta))$, which are normalized so that $\sum_{j=1}^{p} u_j(\theta) = 1$ and $u(\theta)v(\theta) =$ The right eigenvector v is always assumed to be a *column vector*. Define the l functions

$$W_i^n(\theta) = \sum_{j=1}^{p} \sum_{r=1}^{Z_{ij}^n} \frac{v_j(\theta)}{v_i(\theta)} \frac{e^{-\theta Z_{ij}^n(r)}}{\rho(\theta)^n}. \tag{1}$$

en $\{W_i^n(\theta)\}$ is a non-negative martingale with respect to $\{\mathcal{F}_n\}$, where $\mathcal{F}_n$ the σ-algebra that contains all information up to generation n, (see [1]).

The individuals in the n-th generation are determined by the point pros $\{Z_{ij}^n(r); r\}$ and has the corresponding intensity measures $\mu_{ij}^{n*}(.)$, defined luctively: if $\mu^{n*} = \{\mu_{ij}^{n*}\}_{p \times p}$ then $\mu_{ij}^{(n+1)*} = \sum_{k=1}^{p} \mu_{ik} \otimes \mu_{kj}^{n*}$ where " $\otimes$ " is linary convolution of measures.

2.1 A Tail Result for Markov Dependent R.V.s

In this subsection we give a tail result for the sums of r. variables dending on a Markov chain (Lemma 1), the the extension of Lemma 1 in . After giving the definition of Markov dependent r. variables, we apply e above lemma to the certain subsequences with identical increments. Let $= \{p_{ij}\}$ be a positive regualar stochastic matrix with equilibrium probaities $\pi = (\pi_1, \ldots, \pi_p)$. Let $\{k_n\}$ be a Markov chain with finite state space $, 2, \ldots, p\}$, and transition matrix $P = \{p_{ij}\}$ (positive regular), and starts state $k_0 = i$. Let $G_1, G_2, \ldots, G_p$ be distribution functions of r. variables $, \ldots, X_p$, with finite means, $\chi_1, \ldots, \chi_p$, respectively. $\{X^n\}$ is a sequence r. variables independent conditional on the Markov chain such that, at X^n has distribution function G_j, if $k_n = j$. Thus $\{X^n\}$ are called *Markov pendent* r. variables. Define $S^n = \sum_{k=1}^{n} X^k$ with $S^0 = 0$, then S^n is a ranm walk in random environment with asymptotic mean drift $\chi = \sum_{j=1}^{p} \pi_j \chi_j$. t for each fixed j, $N_j^1 < N_j^2 < \ldots$ be the times that the Markov chain is $T_j^k = N_j^k - N_j^{k-1}$ $(k = 2, \ldots)$ with $T_j^1 = N_j^1$. Then $\{T_j^k\}_{k=1}^{\infty}$

Galton-Watson processes. Biggins ([2]) extend the same result for a single type branching random walk on the real line by employing Laplace transforms related to non-negative r. variables. We follow that paper Biggins ([2]). The proof is rather long and is divided into three sections. In the first step, section 2. we give the notations and some preliminary background results. In the second step, section 3, we give a necessary and sufficient condition for a certain operator to have a fixed point (Theorem 1). More precisely, we define an operator on the p-dimensional space of Laplace transforms of probability distribution functions defined on $[0, \infty)$. Then we show that the unique fixed point of this operator is the solution of the functional equation. This is the technique that is used by many authors (e.g. [2], [5], [6]). All their proofs involve iterating a measuure on the real line to constract a random walk with identical increments. In the multitype case we have a matrix of measures, so we need to apply the Markov Additive results to obtain an additive component, which is a ranom walk with Markov dependent increments (see [11], [4], [13], [3], [8]). And we have a taile result on Markov dependent random valk Lemma 1 in section 2. We apply results from [4], and [8]. Finally in section 4, we prove the main result, Theorem 2, which is a multitype version of Theorem A in [2] .

2 Notations and Some Preliminary Results

In this sectin we introduce some notations. We consider a multitype Branching random walk on $\mathcal{R}$ with a single i-type ancestor located at the origin, and with embedded multitype Galton-Watson process $\{Z_i^n\} = \{(Z_{i1}^n, \ldots, Z_{ip}^n)\}$. The process $\{Z_i^n\}$ is positive regular in the sence that the matrix $M = \{m_{ij}\} = \{E[Z_{ij}]\}$ is positive regular ([12]) and has a simple positive Perron-Frobenus eigenvalue ρ. The process is assumed to be supercritical, $\rho > 1$. The intensity measures $\mu = \{\mu_{ij}(.)\}_{p\times p}$, formed by the counting measures $\{Z_{ij}(.)\}$ of the first generation individuals, defined for all Borel measurable sets $A \subset \mathbf{R}$, $\mu_{ij}(A) = E[Z_{ij}(A)]$.

We keep the small letters i, j, and k to denote the types and assumed to be in $\{1, \ldots, p\}$ in all that follows.

Define the real valued Laplace transforms $m_{ij}(\theta)$ by

BRANCHING RANDOM WALK

A. Rahimzadeh sani

University of Teacher Education

Abstract: A result like Kesten-Stigum Theorem is obtained, i.e. a suitable $X \log X$' conditions are necessary and sufficient for the martingales that arise the processes to converge in mean.

Key Word: Multitype Branching Random Walk, Martingales, Functional quation, Multitype Branching Process, Random Walk in Varying Environ- ent.

1 Introduction

A discrete time multitype Branching Random Walk (BRW) on the real ne is a mathematical description of the growth of a population, in which the articles cosistes of p different types and occupy a position in $\mathcal{R}$. Let the process art with a single 'i'-type ancestor located at the origin. This particle splits to a random number of new particles of different types, with probability w depending on i. Then each of these particles behaves in the same way, dependently of the other particles and the history of the process with their production laws depending on their type. Let $\{Z_{ij}(r);r\}$ be the point process rmed by the positions of the j-type individuals in the first generation. So we ave a matrix of point processes $\{Z_{ij}(r);r\}$; $i,j=1,\ \ldots,\ p$, to determine the osition of first generation individuals of each type for each type of parent.

Let $\{Z^n_{ij}(r);r\}$ be the point process (on R) giving the positions of j-type dividuals of the n-th generation resulting from this construction. Let $Z^n_{ij} = Z^n_{ij}(\mathcal{R})$ then $\{Z^n_i\} = \{(Z^n_{i1},\ldots,Z^n_{ip})\}$ is the *imbedded* multitype Galton-Watson rocess (see [12]). By normalizing the Laplace transforms of $Z^n_{ij}(.)$ we define a

Variables. *Invited Papers, First Iranian Statistical Conference.* 195-201.

[8] Soltani, A. R. (1994). On Spectral Representation of Multivariate Stable Processes. *SIAM Theory of Probability and Its Applications.* **39(3)**, 465-475.

[9] Soltani, A. R. and Moeanaddin R. (1994). On Dispersion of Stable Random Vectors and Its Application in the Prediction of Multivariate Stable Processes. *J. Appl. Prob.* **31**, 691-699.

Corollary 2.1 Let $\mathbf{X}$ be an n-dimensional α-stable random vector, $0 <$ < 2 with independent components, and let $\mathbf{A}_{d\times n} = (a_{ij})$, $d \leq n$ be a matrix real numbers for which $\mathbf{AX}$ is nonsingular. If $\mathbf{A}$ has a column with at least o nonzero elements, then $\mathbf{AX}$ has dependent components.

By using Theorem 2.1 we provide a transparent proof for the following sult of Modarres and Nolan (1994), which is the main ingredient of their mulation procedure.

Corollary 2.2 Let $0 < \alpha < 2$, $\gamma_j > 0$, $j = 1, \cdots, n$, $\mathbf{t}_j \in S_{d-1}$, $j =$ $\cdots, n$, and V_j, $j = 1, \cdots, n$, be independent one dimensional normalized α-able random variables that are totally skewed to the right. If $\mathbf{Y}$ is a α-stable ndom vector with discrete spectral measure $\Gamma_{\mathbf{Y}}(d\mathbf{t}) = \sum_{j=1}^{n} \gamma_j \ I_{\{\mathbf{t}_j\}}(\mathbf{t})$ (I note indicator function) and shift vector μ then

$$\mathbf{Y} \stackrel{d}{=} \sum_{j=1}^{n} \gamma_j^{1/\alpha}(V_j + 1_\alpha \frac{2}{\pi} \ln(\gamma_j))\ \mathbf{t}_j + \mu.$$

emark 2.3 Theorem 2.1 makes the softwares that generate one dimensional able distributions, such as S-PLUS, accessible for the simulation of an stable ndom vector.

Refrences

1] Kuelbs, J. (1973). A Representation Theorem for Symmetric Stable Processes and Stable Measures on *H. Z. Wahr. verw. Geb.* **26** 259-271.

2] Modarres, R. and Nolan, J. P. (1994). A Method for Simulating Stable Random Vectors. *Computational Statistics*, **9**, 11-19.

3] Nolan, J. P. (1997). Multidimensional Stable Distributions: Approximation, Estimation, Simulation and Identification (to appear).

4] Nolan, J. P. and Rajput, B. S. (1995). Calculation of Multidimensional Stable Densities. *Communication in Statistics*, **24(3)**, 551-566.

symmetric on S_{d-1} and ...

$$\phi(\mathbf{t}) = \exp\left\{-\int_{S_{d-1}} |<\mathbf{t},\mathbf{s}>|^{\alpha}\,\Gamma(d\mathbf{s})\right\}, \quad 0<\alpha<2.$$

The measure Γ is unique, and is called the spectral measure of the stable random vector $\mathbf{X}$, Kuelbs (1973) and Paulauskas (1976).

Stable random variables and random vectors have been considered by different authors some of the works are included in the Modarres and Nolan (1994), Nolan (1997), Nolan and Rajput (1995), Samorodnisky and Taqqu (1994), Soltani (1992), (1994).

2 A Linear Transformation

The following theorem concerns the law of a linear transformation of stable random vectors. This theorem for $\mathbf{X}$ 2-dimensional and symmetric was proved in Soltani (1992), and was applied in Soltani and Moeanaddin (1994) for extrapolation of multivariate stable ARMA processes.

Theorem 2.1 Let $\mathbf{X}$ be an n-dimensional α-stable random vector, $0<\alpha<2$, with spectral measure $\Gamma_{\mathbf{X}}$, $\mathbf{A}$ a $d\times n$ matrix of real numbers with $d\leq n$ and $\mathbf{b}$ a $d\times 1$ vector of real numbers. Then the random vector $\mathbf{Y}=\mathbf{AX}+\mathbf{b}$ is α-stable and its spectral measure together with its location parameter are given by

$$\Gamma_{\mathbf{Y}}(E) = \int_{T^{-1}(E)} \|\mathbf{As}\|^{\alpha}\ \Gamma_{\mathbf{X}}(d\mathbf{s}), \quad E\in\mathcal{B}(S_{d-1}), \tag{2.1}$$

and

$$\mu_{\mathbf{Y}} = \mathbf{A}\mu_{\mathbf{X}} + \mathbf{b} - 1_{\alpha}\frac{2}{\pi}\int_{S_{n-1}} \mathbf{As}\ \ln(\|\mathbf{As}\|)\ \Gamma_{\mathbf{X}}(d\mathbf{s}), \tag{2.2}$$

where

$$1_{\alpha} = \begin{cases} 1 & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha\neq 1, \end{cases}$$

$T: S_{n-1}\longrightarrow S_{d-1}$ is a mapping defined by $T(\mathbf{s}) = \frac{\mathbf{As}}{\|\mathbf{As}\|}$ and $\mathcal{B}(S_{d-1})$ stands

VECTORS AND THEIR APPLICATIONS

A. Mohammadpour, A.R. Soltani

Shiraz University

Abstract The transformations of the form **AX+b** are considered and their rule in simulation procedures are realized. The Modarres and Nolan's simulating method on stable random vectors is obtained by using above transformation.

Key words: Linear transformation, stable random vectors, random number generation.

1 Introduction

Let $\mathbf{X} = (X_1, \cdots, X_d)'$ be a random vector. We say that $\mathbf{X}$ is α-stable if its characteristic function $\phi(\mathbf{t})$ satisfies that, for each $c_1 > 0$, $c_2 > 0$, there exist $c > 0$ and $\mathbf{a} \in \mathbf{R}^d$ such that

$$\phi(c_1\mathbf{t})\phi(c_2\mathbf{t}) = \phi(c\mathbf{t})\exp\{i < \mathbf{a}, \mathbf{t} >\}, \quad \mathbf{t} \in \mathbf{R}^d,$$

where c is uniquely determined by $c_1^\alpha + c_2^\alpha = c^\alpha$, with $\alpha \in (0, 2]$ characteristic exponent of $\mathbf{X}$, and $< \mathbf{a}, \mathbf{t} >$ denotes the inner product of $\mathbf{a}$ and $\mathbf{t}$ in $\mathbf{R}^d$. When $\alpha = 2$, the multivariate stable random vectors are the multivariate normal random vectors

Stable random vectors were characterized by Kuelbs (1973), that $\mathbf{X}_{d\times 1}$ is α-stable if and only if its characteristic function is given by

$$\phi(\mathbf{t}) = \begin{cases} \exp\{-\int_{S_{d-1}} \psi(< \mathbf{t}, \mathbf{s} >)\Gamma(d\mathbf{s}) + i < \mathbf{t}, \mu >\}, & 0 < \alpha < 2, \\ \exp\{- < \mathbf{A}\mathbf{t}, \mathbf{t} > + i < \mathbf{t}, \mu >\}, & \alpha = 2, \end{cases}$$

where

branching random walk. Ann. Probab. 25 337–360.

[7] Chauvin, B. (1991) Multiplicative martingales and stopping lines for branching Brownian motion. Ann. Probab. 30 1195–1205.

[8] Chauvin, B. and Roualt, A. (1996) Boltzman-Gibbs weights in the branching random walk. In Classical and Modern Branching Processes (K.B. Athreya and P. Jagers, eds) 84 41–50. Springer, New York.

[9] Durrett, R. and Liggett, M. (1983) Fixed points of the smoothing transform. Z. Wahrsch. verw. Gebeite. 64, 275-301.

[10] Kahane, J.P. and Peyrière, J. (1976) Sur certaines martingales de Benoit Mandelbrot. Adv. Math. 22, 131-145.

[11] Kingman, J.F.C. (1975) The first birth problem for an age-dependent branching processes. Ann. Probab. 3, 790-801.

[12] Koukiou, F. (1997) Directed polymers in random media and spin glass models on trees. In Classical and Modern Branching Processes (K.B. Athreya and P. Jagers, eds) 84 ??-??. Springer, New York.

[13] Kyprianou, A.E. (1998a) Slow variation and uniqueness of solutions to the functional equation in the branching random walk. To appear in the Journal of Applied Probability.

[14] Kyprianou, A.E. (1998b) A note on branching Lévy processes. Submitted to Stochastic Processes and their Applications.

[15] McDiarmid, C. (1995) Minimal positions in a branching random walk. Ann. Appl. Probab. 5, 128-139.

[16] Nerman, O. (1981) On the convergence of supercritical general (C-M-J) branching processes. Z.Wahrsch. verw. Gebeite. 57, 365-395.

[17] Neveu, J. (1988) Multiplicative martingales for spatial branching processes. In Seminar on Stochastic Processes 1987, eds E. Çinlar. K.L. Chung, R.K. Getoor. Progress in Probability and Statistics, 15, 223-241.

he limit $W(\theta)$ (of the martingale sequence $W_n(\theta)$).

Note that this theorem gives sufficient conditions also for the supermartingales to be a martingale. We know that they are not necessary conditions ecause of the additive martingale that appears in Nerman (1981) and the existence of the additive martingale in the embedded general branching process lentified within the branching random walk in Biggins and Kyprianou (1997).

From the decomposition (2) we have that for any dissecting stopping line ',

$$\begin{aligned}
\lim_{n\uparrow\infty} E_{F_L}(W_n) &= \lim_{n\uparrow\infty}\left[\sum_{u\in A_L(n)} y_u + \sum_{|u|\le n u\in L} y_u E_{F_L}\sum_{|v|=n-|u|} y_u\circ T_u\right] \\
&= \lim_{n\uparrow\infty}\left[\sum_{u\in A_L(n)} y_u + \sum_{|u|\le n u\in L} y_u\right] \\
&= W_L.
\end{aligned}$$

Also because of Biggins' Martingale Convergence Theorem we have

$$\lim_{n\uparrow\infty} E\left|E_{F_L}(W_n) - E_{F_L}(W)\right| \le \lim_{n\uparrow\infty} E\left|W_n - W\right| = 0$$

ıd hence $W_L = E_{F_L}(W)$. The martingale property thus follows for sequences $L_t\}_{t\ge 0}$ and similarly to Theorem 6.2 of Biggins and Kyprianou (1997) we have ıat $\lim_{t\uparrow\infty} W_{L_t} = W$.

Refrences

1] Biggins, J.D. (1977a) Martingale convergence in the branching random walk. J. Appl. Probab. 14, 25-37.

2] Biggins, J.D. (1977b) Chernoff's Theorem in the branching random walk. J. Appl. Probab. 14, 630-636.

3] Biggins, J.D. (1979) Growth rates in the branching random walk. Z.

$$E_{F_n}\left(W_R^{(n+1)}\right) = \sum_{u\in R|u|\leq n-1} y_u + E_{F_n}\left(\sum_{u\in R|u|=n} y_u + \sum_{u\in A_R(n)} y_u \sum_{|v|=1} y_v \circ T_u\right)$$

$$= \sum_{u\in R|u|\leq n-1} y_u + E_{F_n}\left(\sum_{u\notin D_R|u|=n}\left\{I(u\in R)y_u + I(u\notin R)y_u\sum_{|v|=1} y_v\circ T_u\right\}\right)$$

$$= \sum_{u\in R|u|\leq n-1} y_u + \sum_{u\notin D_R|u|=n}\left\{I(u\in R)y_u + I(u\notin R)y_u E_{F_n}\left(\sum_{|v|=1} y_v\circ T_u\right)\right\}$$

$$= \sum_{u\in R|u|\leq n-1} y_u + \sum_{u\in A_R(n-1)} y_u \sum_{|v|=1} y_v\circ T_u = W_R^{(n)}.$$

Note in particular that for the third equality we have taken advantage of Theorem and in the fourth equality we have used $E\sum_{|u|=1} y_u = EW_1 = 1$; the rest of this computation is manipulating the definitions of $A_R(n)$ and D_R and observing the measurability of the summands. We see now that $\lim_{n\uparrow\infty} W_R^{(n)} = W_R$ almost surely and in expectation because R is dissecting. Fatou's Lemma tells us thus that $EW_R \leq EW_R^{(1)} = EW_1 = 1$.

As an immediate corollary we have the existence of our class of additive supermartingales.

Corollary Let $\{L_t\}_{t\geq 0}$ be an increasing sequence of dissecting stopping lines tending to infinity, then W_{L_t} is a almost surely convergent supermartingale with respect to F_{L_t}.

This result follows simply from the fact that for $s < t$, $L_s = A_{L_t}(L_s) \cup (L_s \cap L_t)$ and $A_{L_s}(L_t) = \emptyset$ almost surely.

What is missing here to get the result up to 'martingale' standard is an understanding of whether $\lim_{n\uparrow\infty} E\sum_{u\in A_R(n)} y_u = 0$. Even though we have the dissecting condition implying that every ancestral 'ray' from $\mho$ hits R, it may be possible that these rays carry a non-zero mass in the previously mentioned limit and hence corrupt the expectation we are after; namely $EW_R = 1$. If

[illegible] of Biggins' Martingale conver-into mutually exclusive subsets

$$L^{(v)} = \{u \in L : u \geq v \in Q\}$$

such that conditional on F_Q, $L^{(v)}$ is a (dissecting) stopping line on the tree $\Gamma \circ T_v$.

Note that this Theorem is not necessarily true for stopping lines with Jagers' definition. The earlier example, $N^{(1)}$, shows this to be the case. It suffices to consider the two stopping lines $N^{(1)}$ and $N^{(2)}$ (the generation in which there is a family of size 27 appearing for the second time). See also the discussion in Kyprianou (1998b).

3 Additive Martingales

For convenience, set $y_u(\theta) = e^{-\theta\zeta_u}/m^{|u|}(\theta)$. Where it is not necessary we will not include the dependence on θ of these terms and functions of these term. For any stopping line L, define $W_L = \sum_{u\in L} y_u$. Clearly as the nth generation is a stopping line, W_n is an additive structure of this form.

Lemma Suppose now that L and Q are two dissecting stopping lines, then

$$E\left(W_Q\,|F_L\right) \leq \sum_{u\in A_L(Q)} y_u + \sum_{u\in Q\cap L} y_u + \sum_{u\in A_Q(L)} y_u \tag{1}$$

It is not difficult to show (a rough sketch may help) that Q can be decomposed according to that part which is 'below' L, the part that intersects with L and the descendents of the part of L that sits 'below' Q. (Note some of these three may be empty sets). Taking advantage of the fact that we can write y_{uv} as $y_u \times (y_v \circ T_u)$ (this decomposition is discussed in more detail in Biggins and Kyprianou (1997)) we have

$$W_Q = \sum_{u\in A_L(Q)} y_u + \sum_{u\in Q\cap L} y_u + \sum_{u\in A_Q(L)} y_u \sum_{v\in Q(u)} y_v \circ T_u. \tag{2}$$

The result thus follows from Theorems and because, as we will now show, for any dissecting stopping line R, $EW_R \leq 1$. Consider then the following mation about all individuals who are neither members of ℓ nor descendent of members of ℓ. Let us call $G_u = \sigma(\Omega_v : v \leq u)$ the sigma algebra describing the life histories of u and all its ancestry. Suppose that $\tau : [\Omega_u : u \in I] \to \{0,1\}^I$ is an ensemble of maps so that τ_u takes the value 0 or 1.

Definition The set of individuals $L(\tau) := \{u \in T : \tau_u = 1\}$ is a stopping line if it is line and τ_u is a optional stopping time for G_u. Consistently with lines, it has associated sigma algebra $F_L = \sigma(\Omega_u : u \notin D_L)$.

This definition is consistent with that of Chauvin (1991) for branching Brownian motion. Jagers (1989) however gives a definition of stopping lines which is less stringent. He defines a stopping line $L \subset T$ to be a random line which is optional in the sense that $\{L \leq \ell\} \in F_\ell$ for all lines ℓ. The following example communicated to me by John Biggins shows nicely a set of nodes that is a stopping line by Jagers' definition, but not by Definition . Let $N^{(1)}$be the first generation in which there is a family of 27 children. For any individual u it cannot be established whether it is a member of the $N^{(1)}$th generation simply by looking back at the history of its ancestry. This is because it may be another group of siblings distinct from u's but in the same generation as u which is first of size 27 and therefore responsible for making $\tau_u = 1$. Hence we can see that Jagers' definition can accommodate for greater correlation between the individual mappings of τ.

Let us now introduce some more notation. For any two stopping lines L and Q, we say that L dominates Q if for each $u \in L$ there exists a $v \in Q$ such that $v \leq u$. A sequence of stopping lines $\{L_t\}_{t\geq 0}$ is said to be increasing to infinity if L_t dominates L_s for all $t \geq s$ and $\lim_{t\uparrow\infty} \inf\{|u| : u \in L\} = \infty$ almost surely. Let

$$A_L(Q) = \{u \in Q : v \leq u, v \in L\}$$

be those individuals in Q who have no ancestor, including themselves, on the stopping line L. We can think of $A_L(Q)$ as being the part of Q that is 'below' L. We will reserve the special notation $A_L(n)$ for the case that Q is the nth generation. A stopping line L is called dissecting if $\sup\{n : A_L(n) = \emptyset\} < \infty$ almost surely. This refers to the fact that under the given condition, every line of descent from $\mho$ will meet L if it does not first terminate due to an individual enhanced form given by Lyons, reads as follows.

Theorem1 Suppose that $m(\theta) < \infty$ and $m'(\theta) := -E\sum_{|u|=1} \zeta_u e^{-\theta\zeta_u}$ exists and is finite, then $W_n(\theta)$ converges in expectation to $W(\theta)$ if and only if $E\left[W_1(\theta)\log^+ W_1(\theta)\right] < \infty$ and $\log m(\theta) - \theta m'(\theta)/m(\theta) < 0$.

Since its first appearance, this result has been identified as being of relevance both within the field of branching processes (see for example Biggins (1977b, 1978, 1979, 1992), Biggins and Kyprianou (1997), Chauvin and Rouault (1988), Kyprianou (1998a), Liu (1977a, 1997b), McDiarmid (1995), Neveu (1988) and Uchiyama (1978, 1982)) and other related topics (see for example Chauvin and Rouault (1997), Durrett and Liggett (1983), Kahane and Peyrière (1976) and Koukiou (1997)) such as statistical physics, turbulence models and smoothing transforms.

With a brief calculation one notes of the martingale W_n that it is essentially a Laplace-Stieltjes transformation of the point process $\{u : |u| = n\}$ weighted by its mean. In this presentation we will introduce the idea of a stopping line as other 'subpopulations' like $\{u : |u| = n\}$ from which supermartingales may be built via the Laplace-Stieltjes transform. A rigorous definition of what exactly a stopping line is follows in the next section. Needless to say however, a stopping line must somehow embrace the idea of the strong Markov property in order for it to form the main ingredient of a martingale. Our exposition of stopping lines in the context of constructing supermartingales will go so far as to generalize one direction of the two way implication in Biggins' Martingale Convergence Theorem. This result is not dissimilar to Corollary 6.6 of Jagers (1989). It differs however in the sense that Jagers worked explicitly with age dependent branching processes (effectively the branching random walk where Z is concentrated on the positive half line) and that the results here fall out with considerably less computation.

2 Stopping Lines

The branching random walk may also be considered as a tree T realized from the space of possible nodes

$$I = \mathbb{N} \cup \bigcup^{\infty} N^n$$

SPATIAL BRANCHING PROCESSES

A.E. Kyprianou

Edinburgh University

Abstract We discuss briefly the construction of stopping lines in the branching random walk and their relevance to the existence of certain additive (super)martingales.
Keywords: Branching Random Walk, Additive Martingales, Martingale Convergence.

1 Introduction

The branching random walk begins with an initial ancestor that we shall label $\mho$. After one unit of time this individual gives birth to a random number of offspring, the first generation, scattered at random points on the real line according to the point process Z. Attached to each child is an independent copy of Z and hence the process evolves following this branching rule. Let us assume in what follows that the average number of children born to each individual is greater than one ($EZ(\Re) > 1$) thus guaranteeing the survival of the process with positive probability. Any individual that appears in the process may be identified by its ancestry using the Ulam-Harris notation. So for example an individual $u = (i_n, i_{n-1}, ..., i_1)$ is the i_nth child of the i_{n-1}th child of ....of the i_1th child of $\mho$. In this way, we understand $|u|$ to mean the generation in which u resides, $u < v$ to indicate that u is a strict ancestor of v and uv to refer to the individual who from u's perspective, has line of descent expressible as v.

Kingman (1975) showed the existence of almost surely convergent additive martingales for a special class of branching random walks. Biggins (1977a) later provided conditions equivalent to the convergence of these martingales

Refrences

[1] Cinlar, E. (1969). Markov renewal Theory. *Adv. Appl. Prob.* 1, 123-187.

[2] Cinlar, E. (1975). Markov Renewal Theory: a Survey. *Management Sci.* 21, 727-752.

[3] Keilson, J. (1969). On the Matrix Renewal Function for Markov Renewal Processes. *Ann. Math. Statist.* 40, 1901-1907.

[4] Keilson, J. (1971). A Process with Chain Dependent Growth Rate. Part II: The Ruin and Ergodic Problems. *Adv. Appl. Prob.* 3, 315-338.

[5] Khorshidian, K. (1997). Asymptotic Behavior of Reward Processes for Semi-Markov Processes. *Ph.D. Thesis, Shiraz University.*

[6] Masuda, Y. and Sumita, U. (1991). A Multivariate Reward Processes Defined on a Semi-Markov Process and its First Passage Time Distributions. *J. Appl. Prob.* 28, 360-373.

[7] Mclean, R. A. and Neuts, M. F. (1967). The Integral of a Step Function Defined on a Semi Markov Process. *SIAM J. Appl. Math.* 15, 726-737.

[8] Parham, G. A. and Soltani, A. R. First Passage Time For Reward Processes With Nonlinear Reward Functions: Asymptotic Behavior. *Pakistan Journal Of Statistics* To be appear.

[9] Royden, H. L. (1968). Real Analysis. *Collier Macmillan Limited, London.*

0] Soltani, A. R. (1996). Reward Processes with nonlinear Reward Functions. *J. Appl. Prob.* 33, 1011-1017

1] Soltani, A. R. and Khorshidian, K. Reward Processes for Semi-Markov Processes: Asymptotic Behavior. *J. Appl. Prob.* (December Issue.)

2] Sumita, U. and Masuda, Y. (1987). An Alternative Approach to the Analysis of Finite Semi-Markov and Related Processes. *COMMUN. STATIST. STOCHASTIC MODELS* 3(1), 67-87.

$$(I-\alpha(s))^{-1}(-\Theta^0_{D:2}+s\,\Theta^1_{D:2}+\circ(s)), \quad (4.15)$$

and by (2.1) and (4.13),

$$\sigma(\underline{w},s)\frac{\partial^2 E_D(\underline{w},s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}} = (I-\alpha(s))^{-1}(\Theta^0_{D:12}-s\,\Theta^1_{D:12}+\circ(s)). \quad (4.16)$$

Using the Keilson's approximation of $(I-\alpha(s))^{-1}$, and substituing (4.14)-(4.16) in (4.3) imply that as $s \longrightarrow 0$,

$$\frac{\partial^2 \phi(0,\underline{w},s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}}\ \underline{e} = \{\frac{1}{s^3}H_1(B_1^0 H_1 B_2^0 + B_2^0 H_1 B_1^0)$$

$$+\frac{1}{s^2}H_0(B_1^0 H_1 B_2^0 + B_2^0 H_1 B_1^0) + \frac{1}{s^2}H_1(B_1^0 H_0 B_2^0 + B_2^0 H_0 B_1^0 - B_1^0 H_1 B_2^1 - B_2^0 H_1 B_1^1$$

$$-B_1^1 H_1 B_2^0 - B_2^1 H_1 B_1^0 + B_1^0 H_1 \Theta^0_{D:2} + B_2^0 H_1 \Theta^0_{D:1} + B_{12}^0)\ \}\ \underline{e} + \circ(\frac{1}{s^2}).$$

giving the result.

5 A Numerical Example

In this section we show that the analyses performed in preceding sections can be implemented and are useful to exploit the partial observations of stochastic systems using a simple example.

Consider a trivariate reward process $\underline{\rho}(k,x)$, defined on a Markov-renewal process with the state space $\mathcal{N}=\{0,1,2\}$ and the semi-Markov matrix

$$a(x)=\frac{d}{dx}A(x)=\begin{bmatrix} 0 & 0.8.2e^{-2x} & 0.2.2e^{-2x} \\ 0.9.e^{-x} & 0 & 0.1.e^{-x} \\ 0.8.e^{-x} & 0.2.e^{-x} & 0 \end{bmatrix}.$$

Let

$$\underline{\rho}(0,x)=\begin{pmatrix} x \\ x^2 \\ e^{\frac{x}{3}}-1 \end{pmatrix},\ \underline{\rho}(1,x)=\begin{pmatrix} 2x \\ 2x^2 \\ 2(e^{\frac{x}{3}}-1) \end{pmatrix},\ \underline{\rho}(2,x)=\begin{pmatrix} 3x \\ 3x^2 \\ 3(e^{\frac{x}{3}}-1) \end{pmatrix}.$$

and the initial probabilities be $p'(0)=(0.3\ ,\ 0.2\ ,\ 0.5)$. It follows from Theorems 3.2. and 4.2. that as $t \longrightarrow \infty$, nd therefore $C_{kj}(\underline{0}, s) = \int_0^\infty e^{-sx}\, dA_{kj}(x)$, or $C(\underline{0}, s) = \alpha(s)$. Also

$$\frac{\partial C(\underline{w}, s)}{\partial w_1}\Big|_{\underline{w}=\underline{0}} = \left[-\int_0^\infty \rho_1(k,x)\, e^{-sx}\, dA_{kj}(x)\right], \tag{4.6}$$

nd

$$\frac{\partial^2 C(\underline{w}, s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}} = \left[\int_0^\infty \rho_1(k,x)\rho_2(k,x)\, e^{-sx}\, dA_{kj}(x)\right]. \tag{4.7}$$

Ioreover

$$E_j(\underline{w}, s) = \int_0^\infty e^{-\sum_{i=1}^{p} w_i \rho_i(k,x) - sx}\, \overline{A}_j(x) dx,$$

giving that $E_j(\underline{0}, s) = \dfrac{1 - \alpha_j(s)}{s}$ or $E_D(\underline{0}, s) = \dfrac{I - \alpha_D(s)}{s}$.

lso

$$\frac{\partial E_j(\underline{w}, s)}{\partial w_1}\Big|_{\underline{w}=\underline{0}} = -\int_0^\infty \rho_1(k,x)\, e^{-sx}\, \overline{A}_j(x) dx, \tag{4.8}$$

nd

$$\frac{\partial^2 E_j(\underline{w}, s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}} = \int_0^\infty \rho_1(k,x)\, \rho_2(k,x)\, e^{-sx}\, \overline{A}_j(x) dx. \tag{4.9}$$

imilar to one dimensional case, it follows from (4.6)-(4.9) that

$$\begin{aligned}
\frac{\partial C(\underline{w}, s)}{\partial w_1}\Big|_{\underline{w}=\underline{0}} &= -B_1^0 + s\, B_1^1 + o(s) && (4.10)\\
\frac{\partial^2 C(\underline{w}, s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}} &= B_{12}^0 - s\, B_{12}^1 + o(s) && (4.11)\\
\frac{\partial E_D(\underline{w}, s)}{\partial w_1}\Big|_{\underline{w}=\underline{0}} &= -\Theta_{D:1}^0 + s\, \Theta_{D:1}^1 + o(s) && (4.12)\\
\frac{\partial^2 E_D(\underline{w}, s)}{\partial w_1 \partial w_2}\Big|_{\underline{w}=\underline{0}} &= \Theta_{D:12}^0 - s\, \Theta_{D:12}^1 + o(s). && (4.13)
\end{aligned}$$

herefore (4.4) together with (4.10) and (4.11) implies that

$$\frac{\partial^2 \sigma(\underline{w}, s)}{\partial w_1 \partial w_2} E_D(\underline{w}, s)\Big|_{\underline{w}=\underline{0}} = (I - \alpha(s))^{-1}\{$$