بسم الله الرحمن الرحیم

دانشکده علوم ریاضی

رشته آمار، گرایش آمار ریاضی

پایان‌نامه کارشناسی ارشد

همگرایی کامل برآوردگر مدل رگرسیون ناپارامتری با جملات خطای وابسته زبرجمعی منفی

نگارنده: سمانه رحمانی

استادان راهنما

دکتر نگار اقبال
دکتر حسین باغیشنی

بهمن ۱۳۹۵

تقدیم با بوسه بر دستان پدرم

که راه را به من نشان داد و

تقدیم به مادر عزیزتر از جانم

که چگونه رفتن را به من آموخت.

سپاس گزاری

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت‌های او ندانند و کوشندگان، حق او گزاردن نتوانند. خداوندا جان ما را صفای خود ده و دل ما را هوای خود ده و چشم ما را ضیای خود ده، و ما را از فضل و کرم خود آن ده که آن به. بدون شک جایگاه و منزلت معلم طوری است که در مقام قدردانی از زحمات بی‌شائبه او، با زبان قاصر و دست ناتوان، چیزی ننگاریم. اما از آن‌جا که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تأمین می‌کند برحسب وظیفه نهایت قدردانی و تشکر را تقدیم می‌نمایم.

از دو استاد با کمالات و شایسته سرکار خانم دکتر نگار اقبال و جناب آقای دکتر حسین باغیشنی که در کمال صعه‌صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان‌نامه را برعهده گرفتند، تشکر می‌کنم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

از استادان فرزانه و دلسوز جناب آقای دکتر احمد نزاکتی رضازاده و جناب آقای دکتر محمد آرشی که زحمت داوری این پایان‌نامه را متقبل شدند، کمال تشکر و قدردانی را دارم.

در پایان از پدر و مادرم که پس از پروردگار، مایه هستی‌ام بوده‌اند و از خواهران و برادرانم که در سختی‌ها و دشواری‌های زندگی همواره تکیه‌گاه من و وجودشان مایه دلگرمی من بود، سپاسگزاری می‌نمایم.

سمانه رحمانی

بهمن ۱۳۹۵

تعهدنامه

این‌جانب **سمانه رحمانی** دانشجوی کارشناسی ارشد رشته **آمار** دانشکده **علوم ریاضی** دانشگاه **صنعتی شاهرود**، نویسنده پایان‌نامه با عنوان **همگرایی کامل برآوردگر مدل رگرسیون ناپارامتری با جملات خطای وابسته زبرجمعی منفی** تحت راهنمایی دکتر **نگار اقبال** و دکتر **حسین باغیشنی** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط این‌جانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام **"دانشگاه صنعتی شاهرود"** یا **"Shahrood University of Technology"** به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آن‌ها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده‌اند.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده‌اند.

سمانه رحمانی
بهمن ۱۳۹۵

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته‌شده) متعلق به دانشگاه **صنعتی شاهرود** می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نیست.

چکیده

ویژگی‌های استقلال و وابستگی متغیرهای تصادفی، نقش پایه‌ای و مهمی را در آمار و احتمال ایفا می‌کند. از آن‌جا که در بیشتر موارد استقلال بین متغیرهای تصادفی برقرار نیست، لازم است تا نسخه‌های مختلف وابستگی‌ها مورد مطالعه و استفاده قرار گیرند. در این پایان‌نامه، به‌طور خاص وابستگی زبرجمعی منفی (NSD) و کاربردهای آن مورد نظر ما هستند. از نابرابری‌های گشتاوری می‌توان در بسیاری از مسائل احتمالی مانند پیدا کردن کران‌های بالا یا پایین برای گشتاورهایی که مقدار دقیقشان قابل محاسبه نیستند، استفاده کرد. از جمله نابرابری‌هایی که برای متغیرهای NSD می‌توان استفاده کرد، نابرابری‌های نمایی کولموگروف و ماکسیمال روزنتال می‌باشند. با استفاده از این نابرابری‌ها، در این پایان‌نامه، همگرایی کامل آرایه‌های سطری متغیرهای تصادفی NSD و سازگاری کامل برآوردگر تابع رگرسیون ناپارامتری را بر اساس خطاهای NSD بازگو و عملکرد نتایج نظری را در یک مطالعه شبیه‌سازی ارزیابی می‌کنیم.

کلمات کلیدی: متغیرهای تصادفی وابسته زبرجمعی منفی، نامساوی ماکسیمال روزنتال، نابرابری نمایی کولموگروف، همگرایی کامل، سازگاری کامل.

پیش‌گفتار

فرض مستقل بودن متغیرهای تصادفی یک ویژگی قوی است که در اغلب شرایط واقعی برای پدیده‌های مورد علاقه برقرار نیست و مشاهدات یک پدیده به هم وابسته‌اند. بنابراین، در نظر گرفتن ساختار وابستگی بین مشاهدات، مسأله‌ای مهم و قابل توجه است. وابستگی زبرجمعی منفی برای متغیرهای تصادفی، رده نسبتاً بزرگی از ساختارهای وابستگی است که شامل متغیرهای تصادفی مستقل نیز می‌شود. همگرایی کامل برآوردگرهای حاصل از این متغیرها، معمولاً، به کمک نابرابری‌های گشتاوری مانند نابرابری نمایی کولموگروف و نابرابری ماکسیمال روزنتال نمایش داده می‌شود که می‌توان کاربرد آن‌ها را در یک مدل رگرسیون ناپارامتری با جملات خطای وابسته زبرجمعی منفی، بررسی کرد. محتوای مطالب این پایان‌نامه به صورت زیر تدوین شده است:

- در فصل اول به بیان برخی تعاریف که در اثبات نتایج اصلی به آن‌ها نیاز داریم، می‌پردازیم.
- در فصل دوم دو نابرابری ماکسیمال روزنتال و نمایی کولموگروف را برای متغیرهای تصادفی وابسته زبرجمعی منفی ارایه می‌کنیم.
- در فصل سوم ابتدا همگرایی متغیرهای تصادفی مستقل را مطرح می‌کنیم. سپس همگرایی متغیرهای تصادفی وابسته زبرجمعی منفی را، تحت شرایط گوناگون، بیان می‌کنیم.
- در فصل چهارم، پس از یادآوری رگرسیون پارامتری و ناپارامتری، به معرفی روش‌های رگرسیونی ساده ناپارامتری می‌پردازیم. سپس دو روش برآورد تابع رگرسیون به روش هسته نادارایا–واتسون و رگرسیون چندجمله‌ای موضعی را شرح می‌دهیم و همگرایی آن‌ها را با استفاده از قضایای مطرح‌شده نتیجه می‌گیریم.
- در فصل پنجم، با استفاده از یک مطالعه شبیه‌سازی به ارزیابی نتایج نظری در همگرایی (سازگاری) کامل برآوردگر تابع رگرسیون ناپارامتری با استفاده از دو روش هسته و چندجمله‌ای موضعی، در حالت حجم نمونه محدود، می‌پردازیم.
- پیوست پایان‌نامه نیز کدهای نوشته‌شده در نرم‌افزار R برای بازتولید نتایج شبیه‌سازی‌ها را در بر دارد.

فهرست مطالب

فهرست شکل‌ها

فهرست جداول

فصل ۱

تعاریف و مفاهیم مورد نیاز

در این فصل به تعاریف و مفاهیمی که در فصل‌های بعدی به آن‌ها نیاز داریم، می‌پردازیم.

۱.۱ برخی مفاهیم وابستگی

مسائل مربوط به وابستگی زوج متغیرهای تصادفی (X, Y)، بیش‌تر در مورد توزیع نرمال دومتغیره و جداول ۲×۲ مورد مطالعه قرار گرفتهاند. اگرچه در درجه اول این موارد دارای اهمیت هستند، اما در واقع آن‌ها نمایش ساده‌ای از وابستگی می‌باشند. مطالعه مربوط به حالت کلی وابستگی، عمدتاً پیرامون دو مسأله متمرکز شده است:

۱) آزمون استقلال

۲) تعریف و تخمین میزان وابستگی

فرض کنید $(\Omega, \mathcal{F}, \mathcal{P})$ یک فضای احتمال و $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی تعریف‌شده روی آن باشد. در نظر گرفتن رابطه استقلال یا وابستگی برای $\{X_n, n \geq ۱\}$ نقش مهمی را در آمار و احتمال ایفا می‌کند. استقلال متغیرهای تصادفی یک ویژگی قوی است اما در واقعیت بسیاری از پدیده‌ها به یکدیگر وابسته‌اند. بنابراین علاوه بر مطالعه استقلال متغیرهای تصادفی، برخی از رده‌های دیگر وابستگی

مانند مارتینگل [۱] ، m-وابستگی [۲]، ρ-آمیخته [۳]، ϕ-آمیخته [۴]، وابستگی منفی [۵] و وابستگی مثبت [۶] نیز باید در نظر گرفته شوند. برای آشنایی بیشتر با این مفاهیم می‌توان به ترتیب به گات (۲۰۰۵)، جانسون (۱۹۸۴)، پلیگارد (۱۹۸۷)، سن (۱۹۷۱)، دابهاشی (۱۹۹۶) و آلام و سکنا (۱۹۸۱) مراجعه کرد. ما در این بخش به دو نوع از ساختارهای وابستگی از جمله پیوندی منفی [۷] و وابسته زبرجمعی منفی [۸] برای مجموعه متغیرهای تصادفی که اخیراً علاقه بسیاری از نویسندگان را به خود جلب کرده اند، می‌پردازیم. ابتدا چند وابستگی مربوط به این دو ساختار وابستگی است را معرفی می‌کنیم. مطالب این بخش از جاج و پروشان (۱۹۸۳) و هو (۲۰۰۰) استخراج شده اند.

۱.۱.۱ وابسته منفی توسعه‌یافته

تعریف ۱.۱.۱. (لیو، ۲۰۰۹) یک مجموعه متناهی از متغیرهای تصادفی حقیقی $\{X_i, 1 \leq i \leq n\}$ وابسته منفی توسعه‌یافته [۹] (END) است، اگر $M > 0$ وجود داشته باشد به‌طوری که دو نابرابری زیر برای هر x_i حقیقی، $1 \leq i \leq n$ برقرار باشند:

$$P(X_1 > x_1, \dots, X_n > x_n) \leq M \prod_{i=1}^{n} P(X_i > x_i) \tag{1.1}$$

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) \leq M \prod_{i=1}^{n} P(X_i \leq x_i). \tag{2.1}$$

ملاحظه ۲.۱.۱. (جاج و پروشان، ۱۹۸۳). یک دنباله متناهی از متغیرهای تصادفی END است، اگر هر زیردنباله متناهی از آن END باشد.

تعریف ۳.۱.۱. (لهمن، ۱۹۶۶) اگر دو نابرابری (۱.۱) و (۲.۱) به ازای $M = 1$ برقرار باشند، آن‌گاه دنباله $\{X_i, 1 \leq i \leq n\}$ وابسته منفی [۱۰] (ND) است.

---

[۱] Martingale
[۲] m-dependence
[۳] ρ-mixed
[۴] ϕ-mixed
[۵] Negative dependence
[۶] Positive dependence
[۷] Negative association
[۸] Negatively superadditive dependence
[۹] Extended negatively dependent
[۱۰] Negatively dependence

۲.۱.۱ وابسته ربعی منفی

تعریف ۴.۱.۱. (لهمن، ۱۹۶۶) متغیرهای تصادفی X و Y دارای وابستگی ربعی منفی [۱۱] (NQD) هستند، اگر برای هر x و y حقیقی داشته باشیم

$$P(X \leq x, Y \leq y) \leq P(X \leq x)P(Y \leq y) \quad (۳.۱)$$

یا به‌طور معادل

$$F(x,y) \leq F_1(x)F_2(y). \quad (۴.۱)$$

اگر علامت نابرابری روابط (۳.۱) و (۴.۱) را برعکس کنیم به وابستگی ربعی مثبت [۱۲] (PQD) تبدیل می‌شود. شرایط تعریف‌شده برای NQD نشان می‌دهند که $cov(x,y) \leq 0$ است.

۳.۱.۱ $j-$ وابسته

تعریف ۵.۱.۱. (اسچونفلید، ۱۹۷۱). مجموعه $\{\mathbf{X}_t, t \geq 1\}$ از بردارهای تصادفی $\mathbf{X}_t$، $j-$وابسته نامیده می‌شود، اگر یک j صحیح و نامنفی وجود داشته باشد به‌طوری که زیرمجموعه $\{\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \ldots, \mathbf{X}_{t_p}\}$ به طور تصادفی مستقل از هر زیرمجموعه $\{\mathbf{X}_{\tau_1}, \mathbf{X}_{\tau_2}, \ldots, \mathbf{X}_{\tau_p}\}$ باشد که مجموعه‌های اندیس $\{t_m\}_{m=1,2,\ldots,p}$ و $\{\tau_m\}_{m=1,2,\ldots,p}$ تحت شرایط زیر انتخاب شوند.

$$\min_m\{t_m\} - \max_m\{\tau_m\} > j$$

۴.۱.۱ پیوندی منفی

نتایج مختلفی در آمار و احتمال تحت فرض این‌که برخی از متغیرهای تصادفی دارای وابستگی پیوندی منفی (NA) هستند، توسط آلام و سکنا (۱۹۸۱) معرفی شده‌اند و جاج و پروشان (۱۹۸۳) آن‌ها را به دقت مورد مطالعه قرار داده‌اند.

تعریف ۶.۱.۱. یک دنباله متناهی از متغیرهای تصادفی $\{X_i, 1 \leq i \leq n\}$ NA نامیده می‌شود، اگر برای هر جفت زیرمجموعه جدا از هم A_1 و A_2 از $\{1, 2, \ldots, n\}$ بتوان نوشت

$$cov\left(f_1(X_i, i \in A_1), f_2(X_j, j \in A_2)\right) \leq 0 \quad (۵.۱)$$

هرگاه دو تابع f_1 و f_2 صعودی باشند و کواریانس تعریف شده در (۵.۱) وجود داشته باشد.

نکته ۷.۱.۱. (جاج و پروشان، ۱۹۸۳). اگر توابع f_1 و f_2 نزولی باشند، نابرابری (۵.۱) نیز برقرار است.

---

[۱۱] Negative quadrant dependence

[۱۲] Positive quadrant dependence

یکی از مزیت‌های وابستگی NA این است که توابع $f_۱$ و $f_۲$ در رابطه (۵.۱) می‌توانند صورت‌های مختلفی داشته باشند و این نتیجه به بسیاری از نابرابری‌های چندمتغیره مفید منجر می‌شود. بدون از دست دادن کلیت، می‌توان فرض کرد $A_۱ \cup A_۲ = \{۱, ۲, \ldots, n\}$. همچنین یک دنباله نامتناهی، پیوندی منفی است اگر هر زیردنباله متناهی از آن پیوندی منفی باشد.

وابستگی NA یک نسخه کیفی از وابستگی است. برای نسخه‌های دیگر وابستگی منفی از قبیل وابستگی متعامد کنجی از بالا (پایین) منفی[۱۳]، منظم معکوس در زوج مرتب‌ها[۱۴]، کاهشی شرطی در دنباله[۱۵] و وابسته منفی در دنباله به لهمن (۱۹۶۶)، جاج و پاتیل (۱۹۷۵)، کارلین و رینوت (۱۹۸۰)، ابراهیمی و گوش (۱۹۸۱)، بلوک و همکاران (۱۹۸۵) مراجعه کنید. در میان انواع وابستگی‌های منفی فقط نسخه NA از خاصیت مهم بسته بودن نسبت به ایجاد توابع صعودی متغیرهای تصادفی مستقل برخوردار است.

جاج و پروشان (۱۹۸۳) نشان دادند تعدادی از توزیع‌های چندمتغیره شناخته‌شده از قبیل توزیع چندجمله‌ای، پیچش[۱۶] چندجمله‌ای‌های غیرمتشابه، فوق هندسی چندمتغیره، دیریکله[۱۷]، ترکیب دیریکله و چندجمله‌ای، توزیع نرمال وابسته منفی، توزیع جایگشتی[۱۸]، نمونه‌گیری تصادفی بدون جایگذاری و توزیع‌های توأم رتبه‌ها، دارای ویژگی NA هستند.

نکته ۸.۱.۱. توجه کنید که متغیرهای تصادفی NA، ND هستند اما عکس آن برقرار نیست. برای جزئیات بیش‌تر می‌توان به جاج و پروشان (۱۹۸۳) مراجعه کرد.

ویژگی‌های متغیرهای تصادفی پیوندی منفی

چند ویژگی برای متغیرهای تصادفی NA توسط جاج و پروشان (۱۹۸۳) بیان شده‌اند که در زیر آن‌ها را فهرست کرده‌ایم:

(۱) برای هر جفت از متغیرهای تصادفی، NQD معادل NA است.

(۲) فرض کنید $A_۱$، $A_۲$، $\ldots$، A_m زیرمجموعه‌های مجزا از $\{۱, ۲, \ldots, n\}$ و همچنین توابع $f_۱$، $f_۲$، $\ldots$، f_m صعودی مثبت باشند. در این حالت NA بودن $X_۱$، $X_۲$، $\ldots$، X_n دلالت بر این دارد که

$$E\prod_{i=۱}^{m} f_i(X_j, j \in A_i) \leq \prod_{i=۱}^{m} Ef_i(X_j, j \in A_i).$$

(۳) یک نتیجه بدیهی از ویژگی (۲) این است که برای زیرمجموعه‌های جدا از هم $A_۱$ و $A_۲$ از $\{۱, ۲, \ldots, n\}$ و همچنین مقادیر حقیقی $x_۱$، $x_۲$، $\ldots$، x_n می‌توان نوشت

$$\begin{aligned} P(X_۱ > x_۱, \ldots, X_n > x_n) &\leq P(X_i > x_i, i \in A_۱)P(X_j > x_j, j \in A_۲) \\ P(X_۱ \leq x_۱, \ldots, X_n \leq x_n) &\leq P(X_i \leq x_i, i \in A_۱)P(X_j \leq x_j, j \in A_۲). \end{aligned}$$

[۱۳] Negative upper (lower) orthant dependence
[۱۴] Reverse regular of order two in pairs
[۱۵] Conditionally decreasing in sequence
[۱۶] Convolution
[۱۷] Dirichlet
[۱۸] Permutation distribution

(۴) یک زیرمجموعه دوتایی یا بیش‌تر از متغیرهای تصادفی NA، NA است.

(۵) یک مجموعه از متغیرهای تصادفی مستقل، NA است.

(۶) توابع صعودی تعریف‌شده روی زیرمجموعه‌های جدا از هم یک مجموعه از متغیرهای تصادفی NA، NA هستند.

(۷) اجتماع مجموعه‌های مستقل از متغیرهای تصادفی NA، NA است.

ملاحظه ۹.۱.۱. (جاج و پروشان، ۱۹۸۳). ویژگی‌های (۶) و (۷) دامنه کاربردهای متغیرهای تصادفی NA را به‌طور قابل توجهی گسترش داده‌اند. به عنوان مثال برای بررسی NA بودن توزیع‌هایی که از پیچش توزیع‌های نسبتا ساده به وجود آمده‌اند، فقط کافی است بررسی کنیم توزیع‌های ساده NA هستند.

قضیه ۱۰.۱.۱. (جاج و پروشان، ۱۹۸۳). یک توزیع جایگشتی، NA است.

قضیه ۱۱.۱.۱. (شائو، ۲۰۰۰). فرض کنید $\{X_i, 1 \leq i \leq n\}$ n متغیر تصادفی NA و $\{X_i^*, 1 \leq i \leq n\}$ متغیرهای تصادفی مستقل باشند، به‌طوری که X_i و X_i^* برای هر $1 \leq i \leq n$ هم‌توزیع باشند، در این صورت برای هر تابع محدب f روی $\Re$

$$Ef\left(\sum_{i=1}^{n} X_i\right) \leq Ef\left(\sum_{i=1}^{n} X_i^*\right). \tag{۶.۱}$$

نابرابری (۶.۱) معتبر است، هرگاه امیدریاضی سمت راست آن وجود داشته باشد.

اگر f یک تابع محدب نانزولی باشد، آن‌گاه

$$Ef\left(\max_{1\leq k\leq n} \sum_{i=1}^{n} X_i\right) \leq Ef\left(\max_{1\leq k\leq n} \sum_{i=1}^{n} X_i^*\right), \tag{۷.۱}$$

نابرابری (۷.۱) نیز مادامی معتبر است که امیدریاضی سمت راست آن وجود داشته باشد.

قضیه ۱۱.۱.۱ ما را قادر به نشان دادن بسیاری از نابرابری‌های شناخته‌شده برای متغیرهای تصادفی مستقل مانند نابرابری ماکسیمال روزنتال[۱۹] و نابرابری نمایی کولموگروف[۲۰] می‌کند، که برای متغیرهای تصادفی NA نیز برقرار هستند.

برای مطالعه نابرابری‌های گشتاوری برای متغیرهای تصادفی وابسته (مثبت) به بیرکل (۱۹۸۸) و شائو و یو (۱۹۹۶) مراجعه کنید. به نظر می‌رسد دنباله وابسته منفی دارای نابرابری‌های گشتاوری مطلوب‌تر از دنباله وابسته مثبت است. سو و وانگ (۱۹۹۶) برای $p \geq 2$ اثبات کردند که اگر $\{X_i, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی NA با $EX_n = 0$ و $E|X_n|^p < \infty$ برای هر n باشد، آن‌گاه

$$E|\sum_{i=1}^{n} X_i|^p \leq C_p \left\{ \left(\sum_{i=1}^{n} EX_i^2\right)^{p/2} + \sum_{i=1}^{n} E|X_i|^p \right\}$$

[۱۹] Rosenthal-type maximal inequality

[۲۰] Kolmogorov-type exponential inequality

که در آن C_p ثابت مثبت است و فقط به p وابسته است.

شائو و همکاران (۱۹۹۷) نتیجه زیر را برای $p \geq ۲$ زمانی که $\{X_i, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی NA با $EX_n = ۰$ و $E|X_n|^p < \infty$ برای هر n است، ارائه کردند:

$$E \max_{k \leq n} |\sum_{i=۱}^{k} X_i|^p \leq C_p \left\{ n^{p/۲} \max_{i \leq n} \left(EX_i^۲\right)^{p/۲} + n \max_{i \leq n} E|X_i|^p \right\}.$$

۵.۱.۱ مثال‌هایی از متغیرهای تصادفی پیوندی منفی

جاج و پروشان (۱۹۸۳) نشان دادند برای دو مثال زیر ویژگی NA می‌تواند برقرار باشد:

* فرض کنید $\mathbf{Z} = (Z_۱, \ldots, Z_k)$ برداری تصادفی دارای توزیع چندجمله‌ای است که تنها دارای یک مشاهده است. بنابراین فقط یک Z برابر با ۱ می‌باشد و بقیه اعضا صفر هستند. به عبارتی $\sum_{i=۱}^{k} Z_i = ۱$ است. پس با افزایش یک Z_i بقیه یا ثابت می‌مانند یا کاهش پیدا می‌کنند. پس کوواریانس بین آن‌ها منفی است. بنابراین از تعریف ۶.۱.۱ نتیجه می‌شود که ویژگی NA برای $\mathbf{Z}$ بدیهی است. از آن‌جا که چندجمله‌ای عمومی از پیچش جمله‌های مستقل به وجود آمده است، با توجه به ویژگی (۶) به NA بودن $\mathbf{Z}$ پی می‌بریم.

** جعبه‌ای شامل N توپ با رنگ‌های متمایز را در نظر بگیرید. نمونه‌ای شامل n توپ (بدون جایگذاری) از این مجموعه انتخاب کرده و متغیرهای تصادفی X_i، $i = ۱, \ldots, N$، را تابع نشان‌گر وجود یک توپ از رنگ iام در نمونه تعریف کنید. به عبارت دیگر $X_i = ۱$ است اگر توپی با رنگ iام در نمونه موجود باشد و در غیر این‌صورت $X_i = ۰$ است. با توجه به این‌که از هر رنگ فقط یک توپ موجود است، پس بردار تصادفی $(X_۱, X_۲, \ldots, X_n)$ دارای توزیع جایگشتی است. پس NA می‌باشد. به طور کلی، فرض کنید N_i توپ از رنگ iام، $۱ \leq i \leq k$، در جعبه موجود باشد. پس $\sum_{i=۱}^{k} N_i = N$ است. اگر Y_i تعداد توپ‌هایی از رنگ iام در نمونه nتایی باشد، آن‌گاه Y_i را می‌توان به عنوان مجموع N_i متغیر تصادفی نشان‌گری که در قسمت قبل تعریف کردیم، در نظر گرفت. با توجه به ویژگی (۶)، Y_iها هم NA هستند.

نابرابری‌ها

با استفاده از ویژگی (۲) متغیرهای تصادفی NA، می‌توان نابرابری‌های گشتاوری را برای متغیرهای تصادفی NA به دست آورد. برای اطلاعات بیش‌تر به جاج و پروشان (۱۹۸۳) مراجعه کنید.

فرض کنید $X_۱$، $X_۲$، ...، X_k متغیرهای تصادفی مثبت NA باشند، و برای $۱ \leq i \leq m$ که $m \leq k$ داشته باشیم $\alpha_i \geq ۰$. بنابراین

$$\mu_{\alpha_۱, \alpha_۲, \ldots, \alpha_m} \leq \mu_{\alpha_۱} \mu_{\alpha_۲} \cdots \mu_{\alpha_m}$$

که $\mu_{\alpha_1,\alpha_2,\ldots,\alpha_m} = E(X_1^{\alpha_1}, X_2^{\alpha_2}, \ldots, X_m^{\alpha_m})$ و $\mu_{\alpha_i} = E(X_i^{\alpha_i})$ است. به طور خاص،

$$E(X_1, X_2, \ldots, X_m) \leq \prod_{i=1}^{m} E(X_i).$$

در نابرابری‌های بالا، X_1، X_2، ...، X_m می‌تواند جای‌گزین هر زیرمجموعه دیگری از m متغیر تصادفی انتخاب شده از X_1، X_2، ...، X_k باشد.

۶.۱.۱ وابستگی زبرجمعی منفی

مفهوم وابستگی زبرجمعی منفی (NSD) توسط هو (۲۰۰۰) مطرح شد که این نوع وابستگی ضعیف‌تر از پیوندی منفی است. کریستفیدز و واگلاتو (۲۰۰۴) نشان دادند که NA می‌تواند NSD را نتیجه دهد. ساختار وابستگی زبرجمعی منفی تعمیمی از ساختار پیوندی منفی است و گاهی از آن مفیدتر نیز می‌باشد زیرا از آن می‌توان برای به دست آوردن نابرابری‌های مهم احتمالی استفاده کرد. تعریف متغیرهای تصادفی NSD بر پایه توابع زبرجمعی بنا نهاده شده است.

تعریف ۱۲.۱.۱. (کمپرمن، ۱۹۷۷). تابع $\phi : \Re^n \to \Re$ تابع زبرجمعی است، اگر برای هر $\mathbf{x}, \mathbf{y} \in \Re^n$ داشته باشیم

$$\phi(\mathbf{x} \vee \mathbf{y}) + \phi(\mathbf{x} \wedge \mathbf{y}) \geq \phi(\mathbf{x}) + \phi(\mathbf{y})$$

که $\vee$ ماکسیمم مولفه‌ای و $\wedge$ می‌نیمم مولفه‌ای است، یعنی $\mathbf{x} \vee \mathbf{y} = (x_1 \vee y_1, \ldots, x_n \vee y_n)$ و $\mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, \ldots, x_n \wedge y_n)$.

با توجه به این تعریف، بررسی زبرجمعی بودن یک تابع، به آسانی انجام‌پذیر نیست. کمپرمن (۱۹۷۷)، روشی را ارائه کرد که توسط آن زبرجمعی بودن یا نبودن یک تابع را به سادگی می‌توان نشان داد. لم زیر این روش را بیان می‌کند.

لم ۱۳.۱.۱. اگر تابع ϕ دارای مشتق جزیی مرتبه دوم پیوسته باشد، زبرجمعی بودن تابع ϕ معادل است با

$$\frac{\partial^2 \phi(\mathbf{x})}{\partial x_i \partial x_j} \geq 0 \qquad 1 \leq i \neq j \leq n$$

تعریف ۱۴.۱.۱. (هو، ۲۰۰۰). بردار تصادفی $\mathbf{X} = (X_1, \ldots, X_n)$ وابسته زبرجمعی منفی است، اگر برای هر تابع زبرجمعی ϕ داشته باشیم

$$E(\phi(X_1, \ldots, X_n)) \leq E(\phi(X_1^*, \ldots, X_n^*))$$

که در آن X_1^*، X_2^*، ...، X_n^* متغیرهای تصادفی مستقل هستند به‌طوری که X_i و X_i^* برای هر $1 \leq i \leq n$ هم‌توزیع هستند و امید ریاضی تابع ϕ در این رابطه وجود دارد.

حال مفهومی از دنباله متغیرهای تصادفی NSD و آرایه‌های سطری متغیرهای تصادفی NSD را از وانگ و همکاران (۲۰۱۴) معرفی می‌کنیم.

تعریف ۱.۱.۱۵. دنباله $\{X_n, n \geq 1\}$ را دنباله‌ای از متغیرهای تصادفی NSD می‌نامیم اگر برای هر $n \geq 1$، $(X_1, \dots, X_n)$، NSD باشد.

دنباله $\{X_{ni}, i \geq 1, n \geq 1\}$ را نیز آرایه‌ای سطری از متغیرهای تصادفی NSD می‌نامیم اگر برای هر $n \geq 1$، $\{X_{ni}, i \geq 1\}$، NSD باشد.

ویژگی‌های متغیرهای تصادفی وابسته زبرجمعی منفی

نتایج این قسمت از هو (۲۰۰۰) برگرفته شده‌اند. برای متغیرهای تصادفی NSD چند ویژگی زیر را می‌توان فهرست کرد:

(۱) اگر (X, Y) بردار تصادفی NA باشد، NSD نیز هست و برعکس.

(۲) اگر $(X_1, X_2 \dots, X_n)$، NSD باشد، آن‌گاه برای هر جفت از زیرمجموعه‌های جدا از هم A و B از $\{1, 2, \dots, n\}$ و هر x_1، x_2، ...، x_n حقیقی، داریم:

- $P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_m \leq x_m) \leq \prod_{i=1}^{m} P(X_i \leq x_i)$
- $P(X_1 > x_1, X_2 > x_2, \dots, X_m > x_m) \leq \prod_{i=1}^{m} P(X_i > x_i)$
- $$\begin{aligned} P(X_i > x_i, i \in A, X_j \leq x_j, j \in B) - \prod_{i \in A} P(X_i > x_i) \prod_{i \in B} P(X_j \leq x_j) \\ \geq P(X_i > x_i, i \in A) - \prod_{i \in A} P(X_i > x_i) \\ + P(X_j \leq x_j, j \in B) - \prod_{j \in B} P(X_j \leq x_j). \end{aligned}$$

(۳) اگر $(X_1, X_2, \dots, X_n)$ متغیرهای تصادفی NSD و $g_1, \dots, g_n$ توابعی صعودی باشند، آن‌گاه $(g_1(X_1), \dots, g_n(X_n))$ نیز NSD است.

(۴) اگر $(X_1, X_2, \dots, X_n)$ متغیرهای تصادفی NSD باشند، آن‌گاه $(X_{\pi_1}, X_{\pi_2}, \dots, X_{\pi_n})$ نیز برای هر جایگشت $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ از $\{1, 2, \dots, n\}$ نیز NSD است.

(۵) اگر $(X_1, X_2, \dots, X_n)$، متغیرهای تصادفی NSD باشند، آن‌گاه $(X_{i_1}, X_{i_2}, \dots, X_{i_m})$ برای هر $1 \leq i_1 < i_2 < \dots < i_m$، $2 \leq m < n$ نیز NSD است.

(۶) اگر $\mathbf{X} = (X_1, X_2, \dots, X_n)$ و $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ بردارهای تصادفی NSD و مستقل از هم باشند، آن‌گاه $(X_1, X_2, \dots, X_n, Z_1, Z_2, \dots, Z_n)$ نیز NSD است.

(۷) اگر $\mathbf{X} = (X_1, X_2, \dots, X_n)$ و $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ بردارهای تصادفی NSD و مستقل از هم باشند، آن‌گاه $(X_1 + Z_1, X_2 + Z_2, \dots, X_n + Z_n)$ نیز NSD است.

(۸) اگر $(X_1, X_2, \dots, X_n)$ بردار تصادفی از متغیرهای تصادفی NSD باشد و X_i و X_j برای هر $i \neq j$ ناهمبسته باشند، آن‌گاه X_1، X_2، ...، X_n دوبه‌دو مستقل هستند.

۷.۱.۱ مثال‌هایی از متغیرهای تصادفی وابسته زبرجمعی منفی

برای درک مفهوم NSD، ابتدا خانواده پارامتری توزیع‌های بیضوی‌تراز [۲۱] را از هو (۲۰۰۰) و خانواده‌ای از توزیع‌های فارلی-گامبل-مورگنسترن [۲۲] (FGM) را از رامادان و ال داماسس (۲۰۱۵) معرفی می‌کنیم که در آن ویژگی NSD معادل پیوندی منفی است، و نشان می‌دهیم که تعدادی از توزیع‌های چندمتغیره شناخته‌شده دارای ویژگی NSD هستند.

تعریف ۱۶.۱.۱. (توزیع بیضوی‌تراز). فرض کنید $F(\mathbf{x},\Sigma)$ خانواده‌ای از تابع توزیع بیضوی‌تراز براساس تابع حقیقی مقدار g با تابع چگالی احتمال زیر باشد:

$$f(\mathbf{x},\Sigma)=|\Sigma|^{-۱/۲}g(\mathbf{x}^{\tau}\Sigma\mathbf{x}),\qquad x\in\Re^m. \tag{۸.۱}$$

در این‌جا $g(\cdot)$ بسنده است، به‌طوری که $\int_۰^\infty g(t)t^{m/۲-۱}dt<\infty$. بلوک و سمپسون (۱۹۸۸) نشان دادند اگر $\sigma_{ij}\leq ۰$، برای هر $i\neq j$، آن‌گاه $F(\mathbf{x},\Sigma)$، NSD است، که در آن $\Sigma=(\sigma_{ij})$. خانواده نرمال چندمتغیره نمونه‌ای از خانواده بیضوی‌تراز است. بنابراین اگر عناصر خارج از قطر اصلی ماتریس کواریانس نامثبت باشند، آن‌گاه توزیع‌های نرمال چندمتغیره، NSD هستند.

تعریف ۱۷.۱.۱. (توزیع FGM). خانواده توزیع‌های نمایی خطی دو متغیره FGM از توزیع‌های حاشیه‌ای به دست آمده است. توزیع FGM در اصل توسط مورگنسترن(۱۹۵۶) معرفی شد که یک توزیع شامل دو مولفه وابسته را توصیف می‌کند. توزیع FGM دارای تابع توزیع تجمعی توأم زیر است:

$$\begin{aligned}F(x,y) &= F_X(x)F_Y(y)\left[۱+\lambda(۱-F_X(x))(۱-F_y(y))\right] \\ &= (۱-e^{-(\alpha_۱x+\frac{۱}{۲}\beta_۱x^۲)})(۱-e^{-(\alpha_۲x+\frac{۱}{۲}\beta_۲y^۲)})[۱+\lambda e^{-(\alpha_۱x+\alpha_۲y+\frac{۱}{۲}(\beta_۱x^۲+\beta_۲y^۲))}]\end{aligned} \tag{۹.۱}$$

که در آن $|\lambda|\leq ۱$ و $x,y\geq ۰$ است و $F_X(x)$ و $F_Y(y)$ تابع توزیع‌های تجمعی حاشیه‌ای متغیرهای تصادفی X و Y هستند.

تابع چگالی توأم این توزیع با استفاده از رابطه (۹.۱) به صورت زیر تعریف می‌شود:

$$\begin{aligned}f(x,y) &= f_X(x)f_Y(y)[۱+\lambda(۲F_X(x)-۱)(۲F_Y(y)-۱] \\ &= (\alpha_۱+\beta_۱x)(\alpha_۲+\beta_۲y)e^{-(\alpha_۱x+\alpha_۲y+\frac{۱}{۲}(\beta_۱x^۲+\beta_۲y^۲))} \\ &\quad [۱+\lambda(۱-۲e^{-(\alpha_۱x+\frac{۱}{۲}\beta_۱x^۲)})(۱-۲e^{-(\alpha_۲x+\frac{۱}{۲}\beta_۲y^۲)})]\end{aligned}$$

که در آن

$$f_X(x)=(\alpha_۱+\beta_۱x)e^{-(\alpha_۱x+\frac{۱}{۲}\beta_۱x^۲)},\quad F_X(x)=۱-e^{-(\alpha_۱x+\frac{۱}{۲}\beta_۱x^۲)}$$

[۲۱] Elliptically contoured distribution

[۲۲] Farlie-Gumbel-Morgenstern distribution

و

$$f_Y(y) = (\alpha_2 + \beta_2 y)e^{-(\alpha_2 x + \frac{1}{2}\beta_2 y^2)}, \quad F_Y(y) = 1 - e^{-(\alpha_2 x + \frac{1}{2}\beta_2 y^2)}.$$

جانسون و کاتز (۱۹۷۵) توزیع FGM تعمیم‌یافته را با تابع توزیع $F(x)$ زیر تعریف کردند:

$$F(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i)\left[1 + \sum_{i<j}\theta_{ij}(1 - F_i(x_i))(1 - F_j(x_j))\right] \tag{1.10}$$

که در آن $\{\theta_{ij}. i \neq j\} \in \Theta$ است و Θ فضای پارامتر چندمتغیره را طوری مشخص می‌کند که تابع چگالی احتمال این توزیع نامنفی باشد. بلوک و سمپسون (۱۹۸۸) نشان دادند که برای هر $i \neq j$ اگر $\theta_{ij} \leq 0$ باشد، آن‌گاه این توزیع، NSD است.

تعریف ۱.۱.۱۸. (فراوانی پولیای مرتبه ۲[۲۳]). (افرون، ۱۹۶۵). یک تابع چگالی احتمال روی خط حقیقی $r(x)$ یک تابع فراوانی پولیای مرتبه ۲ (PF_2) گفته می‌شود، اگر برای $x_2 \geq x_1$ و $z_2 \geq z_1$ داشته باشیم

$$\begin{vmatrix} r(x_1 - z_1) & r(x_1 - z_2) \\ r(x_2 - z_1) & r(x_2 - z_2) \end{vmatrix} \geq 0.$$

تابع PF_2 دارای ویژگی‌های زیر می‌باشد:

۱) بسیاری از چگالی‌های احتمالی PF_2 هستند. به‌طور خاص تابع چگالی نرمال را می‌توان مثال زد به‌طوری که

$$r(x) = (2\pi\sigma^2)^{-1} exp\{-\frac{1}{2}[(x - m)/\sigma]^2\}$$

که در آن $\sigma > 0$ است.

۲) یک تابع PF_2 کران‌دار است و روی یک فاصله مقدار غیرصفر و خارج از فاصله مقدار صفر می‌گیرد.

۳) یک تابع PF_2 لگ مقعر[۲۴] است و همچنین تقریباً همه‌جا دارای مشتق اول و دوم است.

۴) فرض کنید $r(x)$ یک تابع PF_2 باشد. در این صورت

$$r_{ab}(x) = \begin{cases} \dfrac{r(x)}{\int_a^b r(z)} & a \leq x \leq b \\ 0 & \text{سایر جاها} \end{cases}$$

نیز PF_2 است که در آن a یا b می‌توانند به ترتیب $-\infty$ و $+\infty$ باشند.

۵) فرض کنید $r(x)$ و $q(x)$ توابع PF_2 باشند، آن‌گاه پیچش آن‌ها که به صورت

$$p(x) = \int_{-\infty}^{\infty} r(t)q(x - t)dt$$

تعریف می‌شود، نیز PF_2 است.

---

[۲۳] Poly frequency of order 2

[۲۴] Log-concave

از آن‌جایی که توابع توزیع پواسون، دوجمله‌ای، گاما و دوجمله‌ای منفی دارای خاصیت $PF_۲$ هستند، بلاک و همکاران (۱۹۸۲) توزیع‌های چندمتغیره زیر را در نظر گرفتند که دارای ویژگی NSD می‌باشند:

۱. (چندجمله‌ای). اگر $\mathbf{X}$ دارای توزیع چندجمله‌ای با بردار پارامترهای $(p_۱, \ldots .p_n)$ باشد، آن‌گاه تابع احتمال آن به صورت زیر است:

$$P(X_۱ = x, \ldots, X_n = x_n) = \frac{N!}{x_۱! \ldots x_n!} \prod_{j=۱}^{n} p_j^{x_j}$$

که در آن $p_j \geq ۰$، $۱ \leq j \leq n$، $\sum_{j=۱}^{n} p_j = ۱$ و برای هر $x_j \geq ۰$، $\sum_{j=۱}^{n} x_j = N$ است.

۲. (فوق‌هندسی چندمتغیره). فرض کنید $\mathbf{X}$ دارای توزیع فوق هندسی با پارامترهای صحیح مثبت $(N, M_۱, \ldots, M_n)$ باشد. در این صورت تابع احتمال آن به شکل

$$P(X_۱ = x_۱, \ldots, X_n = x_n) = \prod_{j=۱}^{n} \frac{\binom{M_j}{x_j}}{\binom{M}{N}}$$

است، که در آن $x_j \geq ۰$، $۱ \leq j \leq m$، $\sum_{j=۱}^{n} x_j = N$ و $\sum_{j=۱}^{n} M_j = M$ است.

۳. (دیریکله). اگر X دارای توزیع دیریکله با پارامترهای $(\theta_۱, \ldots, \theta_n)$ باشد، آن‌گاه دارای تابع چگالی احتمالی زیر است

$$f(\mathbf{x}) = \frac{\Gamma(\sum_{j=۱}^{n} \theta_j)}{\prod_{j=۱}^{n} \Gamma(\theta_j)} \prod_{j=۱}^{n} x_j^{\theta_j - ۱}$$

که در آن برای $۱ \leq j \leq n$، $\theta_j \geq ۱$ و برای $x_j \geq ۰$، $\sum_{j=۱}^{n} x_j = ۱$ است.

۴. (ترکیب دیریکله با چندجمله‌ای). فرض کنید $\mathbf{X}$ دارای تابع احتمال زیر باشد:

$$P(X_۱ = x_۱, \ldots, X_n = x_n) = \frac{\Gamma(\sum_{j=۱}^{n} \theta_j)}{\Gamma(N + \sum_{j=۱}^{n} \theta_j)} \prod_{j=۱}^{n} \frac{\Gamma(x_j + \theta_j)}{x_j! \Gamma(\theta_j)}$$

که در آن $۱ \leq j \leq n$، $\theta_j \geq ۱$، $x_j \geq ۰$، $\sum_{j=۱}^{n} x_j = N$ است و N صحیح مثبت می‌باشد. در این صورت $\mathbf{X}$ دارای توزیع ترکیبی دیریکله با چندجمله‌ای است.

مثال ۱۹.۱.۱. (نمونه‌گیری تصادفی بدون جایگذاری). برای $n \leq N$، فرض کنید X_1، X_2، ...، X_n یک نمونه تصادفی بدون جایگذاری از یک جامعه متناهی با مقادیر x_1، x_2، ...، x_N باشد. در این صورت $(X_1, X_2, \ldots, X_n)$ یک زیرمجموعه از X_1، X_2، ...، X_N می‌باشد، که دارای یک توزیع جایگشتی است و همچنین NSD می‌باشد.

مثال ۲۰.۱.۱. (توزیع توأم رتبه‌ها). فرض کنید X_1، X_2، ...، X_n یک نمونه تصادفی از یک جامعه باشد. همچنین فرض کنید R_i رتبه‌ای از X_i، $1 \leq i \leq n$، باشد، در این صورت R_1، R_2، ...، R_n دارای توزیع جایگشت و نیز NSD است. به عنوان مثال اگر (X_1, X_2, X_3) را داشته باشیم و X_1 دارای رتبه R_1، X_2 دارای رتبه R_2 و X_3 دارای رتبه R_3 باشد، زمانی‌که ترتیب به صورت (X_3, X_2, X_1) عوض شود آن‌گاه X_3 دارای رتبه R_1، X_2 دارای رتبه R_2 و X_1 دارای رتبه R_3 می‌شود.

نابرابری‌ها

حال در زیر به برخی از نمونه‌های توابع زبرجمعی، که در به دست آوردن نابرابری‌ها مفید هستند، اشاره می‌کنیم:

لم ۲۱.۱.۱. (هو، ۲۰۰۰). توابع زیر زبرجمعی هستند:

۱) $\phi_1(\mathbf{x}) = f(\max_{k=1}^{m} \sum_{i=1}^{k} x_i)$، که در آن $f : \Re \to \Re$ صعودی و محدب است.

۲) $\phi_2(\mathbf{x}) = \sum_{i=1}^{k} g(x_{(i)})$، که در آن $1 \leq k \leq m$ و $g : \Re \to \Re$ صعودی بود و $x_{(i)}$ آماره ترتیبی iام می‌باشد.

۳) $\phi_3(\mathbf{x}) = -\sum_{i=1}^{k} 1_{(-\infty,t]}(-x_{(n-i+1)})$، که $1 \leq k \leq n$ می‌باشد و $1_{(-\infty,t]}(.)$ یک تابع نشان‌گر از مجموعه $(-\infty, t]$ است. یعنی در این فاصله مقدار ۱ و در خارج آن مقدار ۰ را دارد.

۴) $\phi_4(\mathbf{x}) = g(x_1, x_1 + x_2, \ldots, x_1 + x_2 + \ldots, x_n)$، که $g : \Re^n \to \Re$ در هر شناسه زبرجمعی و محدب است.

با انتخاب صحیح تابع زبرجمعی ϕ بر اساس مطالبی که گفته شد، می‌توانیم نابرابری‌های مفیدی برای بردار تصادفی X که NSD است، به‌دست آوریم. با مثالی که در زیر ارائه می‌کنیم این مسأله را نشان می‌دهیم.

مثال ۲۲.۱.۱. (هو، ۲۰۰۰). فرض کنید $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ یک بردار تصادفی NSD است و همچنین $\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_n^*)$ بردار تصادفی متناظر با مولفه‌های مستقل هستند به‌طوری که $X_i^* \stackrel{d}{=} X_i$، $1 \leq i \leq n$، می‌باشد. نماد $\stackrel{d}{=}$ به معنای هم‌توزیع بودن است. برای هر تابع محدب و صعودی f داریم:

$$E\left[f\left(\max_{k=1}^{n} \sum_{i=1}^{k} X_i\right)\right] \leq E\left[f\left(\max_{k=1}^{n} \sum_{i=1}^{k} X_i^*\right)\right]. \tag{۱۱.۱}$$

به‌عنوان مثال می‌توانیم از $\phi = \phi_1$ استفاده کنیم. شائو (۲۰۰۰) رابطه (۱۱.۱) را برای متغیرهای تصادفی NA نیز اثبات کرد. در این‌جا نشان داده شده است که گشتاورهایی از مجموع جزئی[۲۵] (یا ماکسیمم مجموع جزئی)

[۲۵] Partial sum

متغیرهای تصادفی NSD با استفاده از متغیرهای تصادفی مستقل کران‌دار شده است. رابطه (۱۱.۱) به ما این امکان را می‌دهد تا بسیاری از نابرابری‌های شناخته‌شده برای متغیرهای تصادفی مستقل مانند نابرابری‌های ماکسیمال روزنتال را برای متغیرهای تصادفی NSD تعمیم دهیم.

از رابطه (۱۱.۱) نابرابری کولموگروف، برای متغیرهای تصادفی NSD را به صورت زیر می‌توان بیان کرد.

نتیجه ۲۳.۱.۱. (هو، ۲۰۰۰). فرض کنید X_1، X_2، ...، X_n متغیرهای تصادفی *NSD* با میانگین صفر و گشتاور دوم متناهی باشد. برای هر $k \geq 1$، $S_k = \sum_{i=1}^{k} X_i$ را در نظر می‌گیریم. برای هر $\epsilon > 0$ داریم

$$P\left(\max_{1\leq k\leq n} |S_k| > \epsilon\right) \leq \frac{8}{\epsilon^2}\sum_{i=1}^{n} Var(X_i).$$

برهان. از آن‌جایی که

$$\max_{1\leq k\leq n} |S_k| \leq \max(0, S_1, S_2, \ldots, S_n) + \max(0, -S_1, -S_2, \ldots, -S_n)$$

داریم

$$\begin{aligned}
& P\left(\max_{1\leq k\leq n} |S_k| > \epsilon\right) \\
\leq\ & P(\max(0, S_1, S_2, \ldots, S_n) > \frac{\epsilon}{2}) + P(\max(0, -S_1, -S_2, \ldots, -S_n) > \frac{\epsilon}{2}) \\
\leq\ & \frac{E[\max(0, S_1, S_2, \ldots, S_n)]^2}{\frac{\epsilon^2}{4}} + \frac{E[\max(0, -S_1, -S_2, \ldots, -S_n)]^2}{\frac{\epsilon^2}{4}} \\
=\ & \frac{4}{\epsilon^2}\{E[\max(0, S_1, S_2, \ldots, S_n)]^2 + E[\max(0, -S_1, -S_2, \ldots, -S_n)]^2\} \\
\leq\ & \frac{4}{\epsilon^2}\{E[\max(S_1, S_2, \ldots, S_n)]^2 + E[\max(-S_1, -S_2, \ldots, -S_n)]^2\} \qquad (12.1)
\end{aligned}$$

با استفاده از (۱۱.۱) و $f(x) = x^2$ می‌توان به دست آورد:

$$\begin{aligned}
E\left[\max_{1\leq k\leq n} S_k\right]^2 &\leq E\left[\max_{1\leq k\leq n} \sum_{i=1}^{k} X_i^*\right]^2 \\
&\leq \sum_{i=1}^{n} E(X_i^*)^2 \\
&= \sum_{i=1}^{n} EX_i^2 \\
&= \sum_{i=1}^{n} Var(X_i) \qquad (13.1)
\end{aligned}$$

به آسانی میبینیم که $-X_1$، $-X_2$، ...، $-X_n$ نیز NSD هستند. بنابراین به دست می‌آوریم:

$$E[\max(-S_1, -S_2, \ldots, -S_n)]^2 \leq \sum_{i=1}^{m} EX_i^2 = \sum_{i=1}^{n} Var(X_i). \qquad (14.1)$$

حال با جایگذاری (۱۳.۱) و (۱۴.۱) در رابطه (۱۲.۱) داریم

$$P\left(\max_{1\leq k\leq n}|S_k|>\epsilon\right)\leq\frac{4}{\epsilon^2}\left\{\sum_{i=1}^{n}Var(X_i)+\sum_{i=1}^{n}Var(X_i)\right\}=\frac{8}{\epsilon^2}\sum_{i=1}^{n}Var(X_i)$$

و برهان کامل می‌شود. □

۲.۱ ترتیب‌های تصادفی و غیرتصادفی

ون دروارت (۱۹۹۸) نمادهای ترتیبی تصادفی $O_p(\cdot)$ و $o_p(\cdot)$ و ترتیب‌های غیرتصادفی $O(\cdot)$ و $o(\cdot)$ را با توجه به تعریف همگرایی در احتمال به صورت زیر بیان کرده است:

تعریف ۱.۲.۱. یک دنباله از متغیرهای تصادفی $\{X_n, n\geq 1\}$ و دنباله‌ای از ثابت‌های $\{a_n, n\geq 1\}$ را در نظر بگیرید. گوییم $X_n=o_p(a_n)$ است، اگر وقتی $n\to\infty$، آن‌گاه

$$\frac{X_n}{a_n}\xrightarrow{p}0.$$

که در آن منظور از $\xrightarrow{p}$ همگرایی در احتمال است.

تعریف ۲.۲.۱. یک دنباله از متغیرهای تصادفی $\{X_n, n\geq 1\}$ و دنباله‌ای از ثابت‌های $\{a_n, n\geq 1\}$ را در نظر بگیرید. گوییم $X_n=O_p(a_n)$ است، اگر برای هر $\varepsilon>0$، $K(\varepsilon)>0$ و $n_0(\varepsilon)$ وجود داشته باشد به‌طوری‌که به ازای همه مقادیر $n>n_0(\varepsilon)$ داشته باشیم

$$P\left(\left|\frac{X_n}{a_n}\right|\leq K(\varepsilon)\right)>1-\varepsilon.$$

تعریف ۳.۲.۱. دو دنباله از اعداد ثابت $\{a_n, n\geq 1\}$ و $\{b_n, n\geq 1\}$ را در نظر بگیرید. گوییم $b_n=o(a_n)$، هرگاه

$$\lim_{n\to\infty}|b_n/a_n|=0.$$

تعریف ۴.۲.۱. دو دنباله از اعداد ثابت $\{a_n, n\geq 1\}$ و $\{b_n, n\geq 1\}$ را در نظر بگیرید. گوییم $b_n=O(a_n)$، هرگاه برای هر $\varepsilon>0$، $K(\varepsilon)>0$ و عدد صحیح $N(\varepsilon)$ وجود داشته باشند، به‌طوری‌که برای $n\geq N(\varepsilon)$ داشته باشیم

$$|b_n|<K(\varepsilon)|a_n|.$$

با استفاده از تعاریف بالا، برای هر ثابت حقیقی c، ترتیب‌های $O_p(a_n)$، $o_p(a_n)$، $O(a_n)$ و $o(a_n)$ به ترتیب معادل با $ca_nO_p(1)$، $ca_no_p(1)$، $ca_nO(1)$ و $ca_no(1)$ هستند. همچنین $X_n=O_p(n^c)$، برقراری $X_n=O_p(n^{c+1})$ را نتیجه می‌دهد، اما $X_n=O_p(n^{c+1})$ لزوماً به معنای برقراری $X_n=O_p(n^c)$ نیست.

جدول ۱.۱: برخی ویژگی‌های ترتیب‌های تصادفی و غیرتصادفی

	عملگر ضرب	عملگر جمع	عملگر ترکیب
۱	$o(n^a)o(n^b) = o(n^{a+b})$	$o(n^a) + o(n^b) = o(n^k)$	$O_p(O(n^a)) = O_p(n^a)$
۲	$O(n^a)O(n^b) = O(n^{a+b})$	$O(n^a) + O(n^b) = O(n^k)$	$O_p(O_p(n^a)) = O_p(n^a)$
۳	$O(n^a)o(n^b) = o(n^{a+b})$	$O(n^a) + o(n^b) = O(n^k)$	$O(O(n^a)) = O(n^a)$
۴	$o_p(n^a)o_p(n^b) = o_p(n^{a+b})$	$o_p(n^a) + o_p(n^b) = o_p(n^k)$	$O(O_p(n^a)) = O(n^a)$
۵	$O_p(n^a)O_p(n^b) = O_p(n^{a+b})$	$O_p(n^a) + O_p(n^b) = O_p(n^k)$	$o_p(O_p(n^a)) = o_p(n^a)$
۶	$O_p(n^a)o_p(n^b) = o_p(n^{a+b})$	$O_p(n^a) + o_p(n^b) = O_p(n^k)$	$o(O_p(n^a)) = o_p(n^a)$
۷	$o_p(n^a)o(n^b) = o_p(n^{a+b})$	$O_p(n^a) + O(n^b) = O_p(n^k)$	$O_p(o_p(n^a)) = O_p(n^a)$
۸	$O_p(n^a)o(n^b) = o_p(n^{a+b})$	$O(n^a) + o_p(n^b) = O_p(n^k)$	
۹	$O(n^a)o_p(n^b) = o_p(n^{a+b})$		
۹	$O(n^a)O_p(n^b) = O_p(n^{a+b})$		

در جدول ۱.۱ به برخی از ویژگی‌های ترتیب‌های تصادفی و غیرتصادفی اشاره شده است. در این جدول a و b اعدادی حقیقی و $k = \max(a, b)$ فرض شده است. به‌عنوان نمونه از هر یک از ستون‌های جدول، یکی را اثبات نموده و برای بقیه برهان مشابهی وجود دارند. برای مثال، سومین سطر از ستون عملگر ترکیب را در نظر بگیرید:

$$a_n = O(O(n^a)).$$

اگر فرض کنیم $b_n = O(n^a)$، بنابر تعریف نماد O، داریم

$$\left|\frac{b_n}{a_n}\right| < M. \tag{۱۵.۱}$$

حال با توجه به $a_n = O(b_n)$، می‌توان نوشت

$$\left|\frac{a_n}{b_n}\right| = \left|\frac{\frac{a_n}{n^a}}{\frac{b_n}{n^a}}\right| < M'. \tag{۱۶.۱}$$

از طرفی بنابر (۱۵.۱) و (۱۶.۱) که این همان تعریف ترتیب تصادفی $O(\cdot)$ می‌باشد. در نتیجه

$$a_n = O(n^a).$$

برای سطر هشتم از ستون جمع، فرض می‌کنیم $c_n = O(n^a)$ پس

$$\left|\frac{c_n}{n^a}\right| < M.$$

همچنین اگر فرض کنیم $d_n = o_p(n^b)$، در این صورت

$$\left|\frac{d_n}{n^b}\right| < \varepsilon.$$

بنابراین می‌توان گفت

$$\left|\frac{e_n}{n^k}\right| = \left|\frac{c_n + d_n}{n^k}\right| < \left|\frac{c_n}{n^a} + \frac{d_n}{n^b}\right| < \left|\frac{c_n}{n^a}\right| + \left|\frac{d_n}{n^b}\right| = M + \varepsilon = M'$$

و این بدان معناست که

$$e_n = c_n + d_n = O_p(n^k).$$

برای ستون پنجم از ستون ضرب نیز فرض می‌کنیم

$$f_n = O_p(n^a) \qquad\qquad g_n = O_p(n^b).$$

از طرفی بنا به تعریف ترتیب‌های تصادفی می‌توان گفت:

$$\left|\frac{i_n}{n^{a+b}}\right| = \left|\frac{f_n.g_n}{n^{a+b}}\right| = \left|\frac{f_n}{n^a}\right|.\left|\frac{g_n}{n^b}\right| < MM' = A$$

۳.۱ مارتینگل

در این بخش ابتدا مفهوم میدان و سیگما میدان را از گات (۲۰۰۵) مطرح می‌کنیم و سپس به بیان تعریف مارتینگل و ویژگی‌های آن می‌پردازیم. مارتینگل ابتدا توسط ویل (۱۹۳۹) تعریف شده است.

تعریف ۱.۳.۱. (میدان[۲۶]). فرض کنید $A \in C$ یک مجموعه باشد. گردایه‌ای از زیرمجموعه‌های A یک میدان نامیده می‌شود، اگر

الف) $\oslash \in C$، که $\oslash$ مجموعه تهی است.

ب) اگر $B \in C$، آن‌گاه $B^c \in C$،

ج) اگر $B \in C$ و $D \in C$، آن‌گاه $B \cup D \in C$.

بنابراین، یک میدان از زیرمجموعه‌های A، مجموعه‌ای ناتهی است که تحت عمل‌های مکمل و اجتماع بسته است، همچنین چون $(B \cap D)^c = B^c \cup D^c$، آن‌گاه تحت اشتراک نیز بسته است. بنابراین میدان مجموعه‌ای از زیرمجموعه‌های A است که تحت اجتماع متناهی از زیرمجموعه‌ها، بسته است؛ یعنی اگر $A_1, ..., A_n \in C$، آن‌گاه $\bigcup_{i=1}^{n} A_i \in C$. برای مثال، مجموعه‌ای از تمام زیرمجموعه‌های A یا $\{\oslash, A\}$ در تعریف میدان صدق می‌کنند.

تعریف ۲.۳.۱. (سیگما میدان[۲۷]). سیگما میدان بر روی مجموعه A، به مجموعه‌ای از زیرمجموعه‌های A گفته می‌شود که تحت انجام تعداد شمارایی از جبر مجموعه‌ای (مانند اجتماع، اشتراک یا متمم) بسته بماند. یعنی تعداد شمارایی از انجام این گونه جبرها بر روی اعضای سیگما میدان، باز هم عضوی از آن خواهد بود.

تعریف ۳.۳.۱. (مارتینگل). اگر $\{X_n, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی در فضای احتمال $(\Omega, \mathcal{F}, P)$ و

$\{\mathcal{F}_n, n \geq 1\}$ دنباله‌ای از سیگما میدان‌ها در $\mathcal{F}$ باشند و برای هر $n \geq 0$

۱) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$

۲) X_n در $\mathcal{F}_n$ اندازه‌پذیر باشد

۳) $E(|X_n|) < \infty$

---

[۲۶] Field

[۲۷] σ-field

۴) $E(X_{n+1}|\mathcal{F}_n) = X_n$

آن‌گاه دنباله $\{(X_n, \mathcal{F}_n), n \geq 1\}$ یک مارتینگل است.

اگر $\mathcal{F}_n = \sigma(X_1, X_2, ..., X_n)$، بدیهی است شرط اول برقرار است. بنابراین اگر $A \in \mathcal{F}_n$ باشد، در این‌صورت

$$\int_A X_{n+1} dP = \int_A X_n dP. \tag{17.1}$$

اکنون اگر $h > 1$ در نظر گرفته شود، آن‌گاه

$$\int_A X_n dP = \int_A X_{n+1} dP = ... = \int_A X_{n+h} dP$$

. در نتیجه

$$E(X_{n+h}|\mathcal{F}_n) = X_n$$

. با جایگذاری $A = \Omega$ در رابطه (17.1)، تساوی زیر برقرار می‌شود.

$$E(X_1) = E(X_2) = \dots$$

.

تعریف ۴.۳.۱. (زیرمارتینگل[28]). اگر متغیر تصادفی X_n در شرایط زیر صدق کند

۱) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$

۲) X_n در $\mathcal{F}_n$ اندازه‌پذیر باشد

۳) $E(|X_n|) < \infty$

۴) $E(X_{n+1}|\mathcal{F}_n) \geq X_n$

آن‌گاه X_n زیرمارتینگل متناظر با $\mathcal{F}_n$ است.

در این حالت نیز اگر $\mathcal{F}_n = \sigma(X_1, X_2, ..., X_n)$، برای هر $A \in \mathcal{F}_n$

$$\int_A X_{n+1} dP \geq \int_A X_n dP$$

. اکنون اگر $A = \Omega$ باشد، آن‌گاه برای $h > 1$

$$E(X_{n+h}|\mathcal{F}_n) \geq X_n$$

و

$$E(X_1) \leq E(X_2) \leq ...$$

.

تعریف ۵.۳.۱. (زبرمارتینگل[29]). اگر متغیر تصادفی X_n در شرایط

[28] Submartingale

[29] Supermartingale

۱) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$

۲) X_n در $\mathcal{F}_n$ اندازه‌پذیر باشد

۳) $E(|X_n|) < \infty$

۴) $E(X_{n+1}|\mathcal{F}_n) \leq X_n$

صدق کند، آن‌گاه X_n زبرمارتینگل متناظر با $\mathcal{F}_n$ است.

در این حالت نیز اگر $\mathcal{F}_n = \sigma(X_1, X_2, ..., X_n)$، برای هر $A \in \mathcal{F}_n$

$$\int_A X_{n+1} dP \leq \int_A X_n dP$$

. اگر $A = \Omega$ باشد، برای $h > 1$،

$$E(X_{n+h}|\mathcal{F}_n) \leq X_n$$

و

$$E(X_1) \geq E(X_2) \geq ...$$

.

۴.۱ $U-$ آماره

تعریف ۱.۴.۱. ($U-$ آماره[۳۰]). فرض کنید X_1، X_2، ... و X_n متغیرهای تصادفی مستقل با تابع توزیع F و پارامتر $\theta = \theta(F)$ باشند، به‌طوری که یک برآوردگر نااریب برای θ مانند $h(X_1, \ldots, X_n)$ وجود داشته باشد. به عبارتی دیگر

$$\theta = \theta(F) = E\left(h(X_1, \ldots, X_n)\right) \quad m \leq n.$$

هافدینگ (۱۹۴۸) برای پارامتر $\theta = \theta(F)$ یک برآوردگر نااریب با ویژگی‌های مطلوب به نام $U-$آماره به صورت زیر معرفی کرد:

$$U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < i_2 < \ldots, < i_n \leq n} h\left(X_{i1}, \ldots, X_{im}\right)$$

که در آن مجموع $\sum_{1 \leq i_1 < i_2 < \ldots, i_n \leq m}$ روی $\binom{n}{m}$ ترکیب موجود $(i_1, \ldots, i_m)$ از مجموعه $\{1, 2, \ldots, n\}$ است. به $h\left(X_{i1}, \ldots, X_{im}\right)$ هسته $U-$آماره می‌گویند.

---

۳۰. $-U$ -Statistics

فصل ۲

نابرابری‌های گشتاوری و احتمالی

۱.۲ مقدمه

نابرابری‌های گشتاوری و احتمالی نه تنها در کارهای تحقیقاتی کاربرد دارند، بلکه از آن‌ها می‌توان در حل بسیاری از مسائل احتمالی یا احتمال‌های دمی از جمله پیدا کردن کران‌های بالا یا پایین برای گشتاورهایی که مقدار دقیقشان قابل محاسبه نیستند، استفاده کرد. نابرابری مناسب باید طوری انتخاب شود که تا حد ممکن به گشتاور مورد نظر یا احتمال دلخواه نزدیک باشد. در این صورت جواب نهایی قابل اعتمادتر می‌شود. در این بخش ابتدا به چند نابرابری متداول اشاره می‌کنیم. برای موارد دیگر نیز به لین و بای (۲۰۱۱) مراجعه کنید.

قضیه ۱.۱.۲. (نابرابری C_r[۱]). فرض کنید $\{X_i, ۱ \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی مستقل باشند. اگر عدد حقیقی مثبت $r > ۰$ وجود داشته باشد که به ازای $۱ \leq i \leq n$، $E|X_i|^r < \infty$ باشد، آن‌گاه

$$E|\sum_{i=۱}^{n} X_i|^r \leq C_r \sum_{i=۱}^{n} E|X_i|^r$$

که در آن C_r مقادیر زیر را اختیار می‌کند:

$$C_r = \begin{cases} ۱ & r \leq ۱ \\ ۲^{r-۱} & r \geq ۱ \end{cases}$$

قضیه ۲.۱.۲. (نابرابری ینسن[۲]). فرض کنید $g(\cdot)$ یک تابع محدب روی $\Re$ باشد. همچنین فرض کنید امید

[۱] C_r inquality

[۲] Jensen inequality

ریاضی متغیرهای تصادفی X و $g(X)$ وجود داشته باشند. در این صورت

$$g(EX) \leq Eg(X).$$

قضیه ۳.۱.۲. (نابرابری مارکف[۳]). فرض کنید برای $r > ۰$، $E|X| < \infty$ باشد، آن‌گاه برای هر $x > ۰$

$$P\left(|X| \geq x\right) \leq \frac{E|X|^r}{x^r}.$$

۲.۲ نابرابری ماکسیمال روزنتال

یکی از نابرابری‌های مهم گشتاوری که برای متغیرهای تصادفی وابسته نیز مطرح می‌شود، نابرابری ماکسیمال روزنتال می‌باشد که اولین بار توسط روزنتال (۱۹۷۰) در حالت استقلال به صورت زیر بیان شد:

اگر $X_۱$، $X_۲$، ...، X_n متغیرهای تصادفی و $S_n = \sum_{i=۱}^{n} X_i$ مجموع متغیرهای تصادفی باشد، آن‌گاه نابرابری روزنتال یک کران برای گشتاور $E|S_n|^p$ ایجاد می‌کند. فرض کنید $۱ < p < \infty$ و $X_۱$، $X_۲$، ...، X_n متغیرهای تصادفی مستقل باشند، آن‌گاه

$$(E|S_n|^p)^{۱/p} \quad \leq \quad ۲^p \max\left\{(E|x_۱|^p + ... + E|x_n|^p)^{۱/p}, E|x_۱| + ... + E|x_n|\right\}.$$

یک نتیجه شناخته‌شده از مطالعات روزنتال (۱۹۷۰) در مورد این نابرابری به این صورت است که اگر علاوه بر این‌که $۱ \leq p < \infty$ و $X_۱$، $X_۲$، ...، X_n متغیرهای تصادفی مستقل باشند، برای هر $i \geq ۱$، $E(X_i) = ۰$ باشد، آن‌گاه

۱. اگر به ازای هر $۱ \leq i \leq n$، ϵ_iها مقادیر $۱-$ یا $۱+$ باشند، آن‌گاه

$$(E|\epsilon_۱ X_۱ + ... + \epsilon_n X_n|^p)^{۱/p} \leq ۲\left(E|X_۱ + ... + X_n|^p\right)^{۱/p}. \qquad (۱.۲)$$

۲. برای $p < ۲$

$$(E|X_۱ + ... + X_n|^p)^{۱/p} \leq ۲\left(E|X_۱|^p + ... + E|X_n|^p\right)^{۱/p}$$

و برای $p > ۲$

$$(E|X_۱ + ... + X_n|^p)^{۱/p} \geq \frac{۱}{۲}\left(E|X_۱|^p + ... + E|X_n|^p\right)^{۱/p}.$$

در ادامه فرض کنید که $۲ < p < \infty$ باشد، آن‌گاه یک ثابت $B(p)$ که فقط به p وابسته است، وجود دارد، به‌طوری‌که اگر $X_۱$، $X_۲$، ... X_nمتغیرهای تصادفی مستقل باشند و برای هر $i \geq ۱$، $E(X_i) = ۰$، آن‌گاه

$$\left(E|\sum_{i=۱}^{n} x_i|^p\right)^{۱/p} \leq B(p)\max\left\{\left(\sum_{i=۱}^{n} E|x_i|^p\right)^{۱/p}, \left(\sum_{i=۱}^{n} E|x_i|^۲\right)^{۱/۲}\right\} \qquad (۲.۲)$$

[۳] Markov inequality

و

$$\left(E|\sum_{i=1}^{n} x_i|^p\right)^{1/p} \geq \frac{1}{2}\max\left\{\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}, \left(\sum_{i=1}^{n}E|x_i|^2\right)^{1/2}\right\}. \qquad (3.2)$$

با توجه به نتایج به دست آمده آیبراجیموف و شاراخمیتوف (۲۰۰۲) نابرابری‌های (۲.۲) و (۳.۲) را به صورت کلی زیر بیان کرد:

$$E(|S_n|^p) \leq B(p)\max\left\{\sum_{i=1}^{n}E|X_i|^p, \left(\sum_{i=1}^{n}EX_i^2\right)^{p/2}\right\}.$$

مقدار دقیق ثابت $B(p)$ در نابرابری روزنتال برای $p = 2m$ ، $m \in \mathcal{N}$ توسط جانسن و همکاران (۱۹۸۵) به صورت زیر نشان داده شده است:

$$E(|S_n|^p) \leq \frac{14/5p}{logp}\max\left\{\sum_{i=1}^{n}E|X_i|^p, \left(\sum_{i=1}^{n}EX_i^2\right)^{p/2}\right\}.$$

در بسیاری از موارد برای پیدا کردن کران گشتاورها، نمی‌توان از نتایج موجود برای متغیرهای تصادفی مستقل بهره برد چون شرط استقلال برقرار نیست. بنابراین باید از تعمیم نابرابری‌های احتمال و گشتاوری برای متغیرهای تصادفی وابسته استفاده کرد.

سو و همکاران (۱۹۹۶) یک نابرابری ماکسیمال برای متغیرهای تصادفی NA را به صورت زیر بیان کرد:

فرض کنید $\{X_i, 1 \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی NA با $EX_i = 0$، $1 \leq i \leq n$ باشد. اگر برای $p \geq 2$، $\beta_p = sup_n E|X_n|^p < \infty$ باشد، آن‌گاه یک ثابت مثبت $B(p)$ که فقط به p وابسته است، وجود دارد به‌طوری که برای هر n صحیح می‌توان نوشت

$$E\left(\max_{1\leq k\leq n}|S_k|^p\right) \leq B(p)\left\{n\beta_p + (n\beta_2)^{p/2}\right\}.$$

نوع دیگری از نابرابری روزنتال در دسته خاصی از وابستگی به نام $j-$وابسته را لیو و همکاران (۲۰۱۳) این‌گونه بیان کردند که فرض کنید $E(X_1) = 0$ و به ازای $p > 2$، $E(|X_1|^p) < \infty$ و S_n^* ماکسیمم قدر مطلق مجموع‌های جزئی باشد، آن‌گاه

$$\begin{aligned} E\left(\max_{1\leq k\leq n}|S_n^*|^p\right) &\leq n^{1/2}\left[\frac{87p}{logp}\sum_{j=1}^{n}\theta_{j,2} + 3(p-1)^{1/2}\sum_{j=n+1}^{\infty}\theta_{j,p} + \frac{29p}{logp}|X_1|^2\right] \\ &+ 1/p\left[\frac{87p(p-1)^{1/2}}{logp}\sum_{j=1}^{n}j^{1/2-1/p}\theta_{j,p} + \frac{29p}{logp}|X_1|^p\right]. \end{aligned}$$

تعریف ۱.۲.۲. (امید ریاضی زیرخطی[۴]). زانگ (۲۰۱۴) امید ریاضی زیرخطی $\hat{E}$ را معرفی کرد که روی $\mathcal{H}$ یک تابع $\hat{E} : \mathcal{H} \to \bar{\Re} := [-\infty, \infty]$ با خواص مطلوب زیر است که $(\Omega, \mathcal{F})$ یک فضای اندازه‌پذیر و $\mathcal{H}$ یک فضای خطی از توابع حقیقی تعریف‌شده روی $(\Omega, \mathcal{F})$ است.

برای هر $X, Y \in \mathcal{H}$ داریم

[۴] Sub-linear expectations

۱) یکنواختی[۵]: اگر $X \geq Y$ باشد، آن‌گاه $\hat{E}[X] \geq \hat{E}[Y]$ است.

۲) حافظ ثابت[۶]: $\hat{E}[c] = c$.

۳) زیرجمعی بودن: $\hat{E}[X+Y] \leq \hat{E}[X]+\hat{E}[Y]$ است، زمانی که $\hat{E}[X]+\hat{E}[Y]$ به صورت $+\infty-\infty$ یا $-\infty+\infty$ نباشد.

۴) همگنی مثبت[۷]: $\hat{E}[\lambda X] = \lambda \hat{E}[X]$، $\lambda \geq ۰$.

در این تعریف $(\Omega, \mathcal{H}, \hat{E})$ یک فضای امید ریاضی زیرخطی نامیده می‌شود.

همچنین زانگ (۲۰۱۴) نابرابری روزنتال برای متغیرهای تصادفی وابسته منفی و مستقل تحت امید ریاضی زیرخطی را به صورت زیر بیان کرد:

فرض می‌کنیم $\{X_i, ۱ \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی در $(\Omega, \mathcal{H}, \hat{E})$ باشد و تعریف می‌کنیم $S_k = \sum_{i=۱}^{k} X_i$، $S_۰ = ۰$. در این صورت

(۱) فرض کنید برای هر $۱ \leq k \leq n-۱$، X_k وابسته منفی به $(X_{k+۱}, \ldots, X_n)$ و برای هر $۱ \leq k \leq n$، $\hat{E}[X_k] \leq ۰$ باشد. آن‌گاه

$$\hat{E}\left[|\max_{k\leq n} S_k|^p\right] \leq ۲^{۲-p} \sum_{k=۱}^{n} \hat{E}\left[|X_k|^p\right] \qquad ۱ \leq p \leq ۲$$

و

$$\hat{E}\left[|\max_{k\leq n} S_k|^p\right] \leq B(p) n^{p/۲-۱} \sum_{k=۱}^{n} \hat{E}\left[|X_k|^p\right] \qquad p \geq ۲.$$

(۲) فرض کنید برای هر $۱ \leq k \leq n-۱$، X_k مستقل از $(X_{k+۱}, \ldots, X_n)$ و برای هر $۱ \leq k \leq n$، $\hat{E}[X_k] \leq ۰$ باشد. آن‌گاه برای $p \geq ۲$ داریم

$$\hat{E}\left[|\max_{k\leq n} S_k|^p\right] \leq B(p) \left\{ \sum_{k=۱}^{n} \hat{E}\left[|X_k|^p\right] + \left(\sum_{k=۱}^{n} \hat{E}\left[|X_k|^۲\right] \right)^{p/۲} \right\}.$$

(۳) به‌طور کلی، فرض کنید برای هر $۱ \leq k \leq n-۱$، X_k وابسته منفی به $(X_{k+۱}, \ldots, X_n)$ یا $X_{k+۱}$ وابسته منفی به $(X_۱, \ldots, X_k)$ باشد. آن‌گاه

$$\begin{aligned}\hat{E}\left[|\max_{k\leq n} S_k|^p\right] \leq \; & B(p) \left\{ \sum_{k=۱}^{n} \hat{E}\left[|X_k|^p\right] + \left(\sum_{k=۱}^{n} \hat{E}\left[|X_k|^۲\right] \right)^{p/۲} \right. \\ & + \left. \left(\sum_{k=۱}^{n} [(\hat{E}|X_k|)^- + (\hat{E}|X_k|)^+] \right)^p . \right\}\end{aligned}$$

برای اطلاع بیشتر در مورد وابستگی‌ها تحت امید ریاضی زیرخطی می‌توان به زانگ (۲۰۱۴) مراجعه کرد.

---

[۵] Monotonicity

[۶] Constant preserving

[۷] Positive homogeneity

۱.۲.۲ نابرابری ماکسیمال روزنتال برای متغیرهای تصادفی وابسته زبرجمعی منفی

اقبال و همکاران (۲۰۱۰)، نابرابری ماکسیمال برای شکل درجه دوم $T_n = \sum_{1\le i<j\le n} X_iX_j$ و در حالت کلی‌تر $Q_n = \sum_{1\le i<j\le n} a_{ij}X_iX_j$ را ثابت کردند، که در آن‌ها $\{X_i; i\ge 1\}$ یک دنباله از متغیرهای تصادفی وابسته زبرجمعی منفی است.

بدون از دست دادن کلیت، فرض می‌کنیم $\{X_i; i\ge 1\}$ دنباله‌ای از متغیرهای تصادفی نامنفی باشد، زیرا اگر فرض نامنفی بودن متغیرهای تصادفی حذف شود، آن‌گاه با استفاده از $|X| = X^+ + X^-$ و

$$\begin{aligned} |T_n| &\le \sum_{1\le i<j\le n} |X_iX_j| = \sum_{1\le i<j\le n} |X_i||X_j| \\ &= \sum_{1\le i<j\le n} X_i^+X_j^+ + \sum_{1\le i<j\le n} X_i^+X_j^- + \sum_{1\le i<j\le n} X_i^-X_j^- + \sum_{1\le i<j\le n} X_i^-X_j^+ \end{aligned}$$

همه نتایج ما معتبر است، به‌طوری که $X^+ = max\{X, 0\}$ و $X^- = max\{0, -X\}$.

قضیه ۲.۲.۲. فرض کنید $\{X_i; i\ge 1\}$ دنباله‌ای نامنفی از متغیرهای تصادفی وابسته زبرجمعی منفی با $E(X_i^r) < \infty$ برای هر $i\ge 1$ و $r > 1$ باشد، آن‌گاه برای هر $\varepsilon > 0$

$$P\left[\max_{2\le k\le n} T_k > \varepsilon\right] \le C\left(\frac{r}{\varepsilon(r-1)}\right)^r \left(\sum_{i=2}^{n} E(X_j^r)\sum_{i=1}^{j-1} E(X_i^r)\right).$$

شائو (۲۰۰۰)، نابرابری ماکسیمال روزنتال را برای متغیرهای تصادفی NA به صورت زیر بیان کرد

قضیه ۳.۲.۲. فرض کنید $p\ge 1$ و $\{X_i, 1\le i\le n\}$ مجموعه‌ای از متغیرهای تصادفی NA با میانگین صفر و $E|X_i|^p < \infty$ برای هر $1\le i\le n$ و $\{X_i^*, 1\le i\le n\}$ مجموعه‌ای از متغیرهای تصادفی مستقل باشند به‌طوری که X_i^* و X_i برای هر $1\le i\le n$ هم‌توزیع هستند. آن‌گاه برای هر $n\ge 1$ می‌توان گفت

$$\begin{aligned} E|\sum_{i=1}^{n} X_i|^p &\le E|\sum_{i=1}^{n} X_i^*|^p \\ E\left(\max_{1\le k\le n} |\sum_{i=1}^{k} X_i|^p\right) &\le 2E|\sum_{i=1}^{n} X_i^*|^p \\ E\left(\max_{1\le k\le n} |\sum_{i=1}^{k} X_i|^p\right) &\le 2^{3-p}\sum_{i=1}^{n} E|X_i|^p \quad 1\le p\le 2 \end{aligned}$$

و

$$E\left(\max_{1\le k\le n} |\sum_{i=1}^{k} X_i|^p\right) \le 2\left(\frac{15p}{lnp}\right)^p \left[\sum_{i=1}^{n} E|X_i|^p + \left(\sum_{i=1}^{n} EX_i^2\right)^{p/2}\right] \quad p > 2.$$

وانگ و همکاران (۲۰۱۴) نیز نابرابری ماکسیمال روزنتال را برای متغیرهای تصادفی NSD بیان کردند که ما آن را در قالب قضیه زیر بیان و اثبات می‌کنیم.

قضیه ۴.۲.۲. فرض کنید برای $p > ۱$، $\{X_i, ۱ \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی NSD با $E|X_i|^p < \infty$ برای هر $i \geq ۱$ باشد و $\{X_i^*, ۱ \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی مستقل باشد به‌طوری که X_i^* و X_i برای هر $i \geq ۱$ هم‌توزیع هستند. آن‌گاه برای هر $n \geq ۱$ می‌توان نوشت

$$\begin{aligned} E\left(\max_{۱\leq k\leq n}|\sum_{i=۱}^{k}X_i|^p\right) &\leq ۲E|\sum_{i=۱}^{n}X_i^*|^p \\ E\left(\max_{۱\leq k\leq n}|\sum_{i=۱}^{k}X_i|^p\right) &\leq ۲^{۳-p}\sum_{i=۱}^{n}E|X_i|^p \quad ۱\leq p\leq ۲ \end{aligned}$$

و

$$E\left(\max_{۱\leq k\leq n}|\sum_{i=۱}^{k}X_i|^p\right) \leq ۲\left(\frac{۱۵p}{lnp}\right)^p\left[\sum_{i=۱}^{n}E|X_i|^p+\left(\sum_{i=۱}^{n}EX_i^۲\right)^{p/۲}\right] \quad p>۲.$$

برهان. ابتدا نشان می‌دهیم

$$E|\sum_{i=۱}^{k}X_i|^p \leq E|\sum_{i=۱}^{k}X_i^*|^p.$$

در این‌جا فقط کافی است ثابت کنیم برای $p \geq ۰$، $\phi(x_۱, x_۲, \ldots, x_i, x_j, \ldots, x_n) = |\sum_{i=۱}^{n}x_i|^p$ یک تابع زبرجمعی منفی است. ابتدا از تابع ϕ نسبت به x_i مشتق می‌گیریم. داریم

$$\frac{\partial|(x_۱+x_۲+\cdots+x_i+\cdots+x_n)|^p}{\partial x_i} = p\frac{\sum_{i=۱}^{n}x_i}{|\sum_{i=۱}^{n}x_i|}|\sum_{i=۱}^{n}x_i|^{p-۱} = p(\sum_{i=۱}^{n}x_i)(|\sum_{i=۱}^{n}x_i|)^{p-۲}.$$

این بار مشتق دوم را نسبت به x_j می‌گیریم و نتیجه می‌شود

$$\begin{aligned} \frac{\partial p(\sum_{i=۱}^{n}x_i)(|\sum_{i=۱}^{n}x_i|)^{p-۲}}{\partial x_j} &= p\left((|\sum_{i=۱}^{n}x_i|^{p-۲})+(p-۲)\frac{(\sum_{i=۱}^{n}x_i)^۲}{|\sum_{i=۱}^{n}x_i|}|\sum_{i=۱}^{n}x_i|^{p-۳}\right) \\ &= p\left((|\sum_{i=۱}^{n}x_i|^{p-۲})+(p-۲)(\sum_{i=۱}^{n}x_i)^۲|\sum_{i=۱}^{n}x_i|^{p-۴}\right). \end{aligned}$$

در آخر داریم

$$\frac{\partial^۲\phi(\mathbf{x})}{\partial x_i\partial x_j} = p\left((|\sum_{i=۱}^{n}x_i|^{p-۲})+(p-۲)(\sum_{i=۱}^{n}x_i)^۲|\sum_{i=۱}^{n}x_i|^{p-۴}\right) \geq ۰.$$

بنابراین طبق لم ۱۳.۱.۱، تابع ϕ زبرجمعی منفی است. حال با استدلالی مشابه برهان قضیه ۳.۲.۲ در هو (۲۰۰۰) برهان قضیه کامل می‌شود. □

۳.۲ نابرابری نمایی کولموگروف

نرخ نمایی بهترین شکل از نابرابری‌ها در احتمال است. برخی از این نابرابری‌ها مانند نابرابری نمایی کولموگروف به خوبی شناخته شده و اغلب در آمار و احتمال از آن‌ها استفاده می‌شود. کولموگروف (۱۹۶۳) برای متغیرهای تصادفی مستقل و هم‌توزیع $X_۱$، $X_۲$، ... از توزیع برنولی با پارامتر $۰ \leq p \leq ۱$ نابرابری زیر را ارائه کرد:

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq ۲\exp\left\{-۲n\varepsilon^۲(۱-\varepsilon)\right\} \qquad (۴.۲)$$

که در آن $\bar{X}_k = \frac{۱}{k}\sum_{i=۱}^{k} X_i$ و $\varepsilon > ۰$ است.

تکمیل، تعمیم و نتایج مربوط به این نابرابری را می‌توان در هافدینگ (۱۹۶۳) یافت. اثبات این نابرابری توسط بانجویس (۱۹۸۴) ارایه شده است. یانگ و همکاران (۱۹۸۷) این نابرابری را به شکل زیر برای $۱ \leq \varepsilon \leq ۰$ و $۱ \leq p \leq ۰$ ساده کردند:

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq ۲\exp\left\{-۲n\varepsilon^۲\right\}.$$

در حالت خاص که $U-$آماره یک میانگین حسابی از متغیرهای تصادفی برنولی است، تورنر و همکاران (۱۹۹۵) نشان دادند برای هر $\varepsilon > ۰$ اگر $۰ < p+\varepsilon < ۱$، $p \geq \frac{۱}{۲}$ یا $p+\varepsilon < \frac{۱}{۲}$ و $n \geq ۰$، آن‌گاه

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq (۱-۲\varepsilon)^{n\varepsilon+۱/۲}(۱+۲\varepsilon)^{n\varepsilon+۱/۲}.$$

در این قسمت با استفاده از یک قضیه دو نابرابری کولموگروف برای متغیرهای تصادفی برنولی مستقل اما نه لزوماً هم‌توزیع ارائه می‌کنیم.

قضیه ۱.۳.۲. (مائوریکو، ۲۰۰۷). فرض کنید $\{Y_i, i \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل برنولی با $E(Y_i) = p_i$ است و همچنین فرض کنید $\bar{p} = \frac{۱}{k}\sum_{i=۱}^{k} p_i$ و $\bar{Y} = \frac{۱}{k}\sum_{i=۱}^{k} Y_i$ باشد، آن‌گاه

۱) برای هر $\varepsilon < \frac{۱}{۲}$ و $\bar{p_k} + \varepsilon < \frac{۱}{۲}$ یا $\bar{p_k} \geq \frac{۱}{۲} + \frac{۱}{۳}\varepsilon$ می‌توان نوشت

$$P\left[\sup_{k\geq n}(\bar{Y_k} - \bar{p}_k) > \varepsilon\right] \leq \left(۱-۴\varepsilon^۲\right)^{\frac{n}{۲}} \qquad n \geq ۰.$$

۲) برای هر $\varepsilon < ۱$ و $\bar{p_k} + \varepsilon > \frac{۱}{۲}$ یا $\bar{p_k} < \frac{۱}{۲}$ می‌توان گفت

$$P\left[\sup_{k\geq n}(\bar{Y_k} - \bar{p}_k) > \varepsilon\right] \leq \exp\{-n(۲\varepsilon^۲ + \frac{۱}{۳}\varepsilon^۴\exp\{\frac{\varepsilon^۲}{۴}\})\} \qquad n \geq ۰.$$

ملاحظه ۲.۳.۲. قضیه ۱.۳.۲ کران‌های واضح‌تری نسبت به قضیه اصلی تورنر و همکاران (۱۹۹۵) تحت محدودیت‌های $\bar{p}$ و ε ارائه کرده است.

تعمیمی از نابرابری کولموگروف

قبل از شروع اثبات نتیجه اصلی، ابتدا دو لم مقدماتی از تورنر و همکاران (۱۹۹۵) را بیان می‌کنیم.

لم ۳.۳.۲. فرض کنید $\{X_n; n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشند، به طوری که تابع مولد گشتاور $\phi_i(t) = E[e^{tX_i}]$ برای $i \geq ۱$ موجود باشد و همچنین فرض کنید $\varepsilon \geq ۰$ است. اگر $t > ۰$ وجود داشته باشد به‌طوری که برای هر i داشته باشیم $E\left[e^{t(X_i-\varepsilon)}\right] \leq ۱$، آن‌گاه

$$P\left(\sup_{k\geq n} \bar{X}_k \geq \varepsilon\right) \leq e^{-tn\varepsilon} \prod_{i=۱}^{n} \phi_i(t).$$

لم ۴.۳.۲. فرض کنید X یک متغیر تصادفی برنولی دارای میانگین p، $۰ \leq p \leq ۱$ و همچنین فرض کنید $Z = X - p$ است. آن‌گاه برای $t \geq ۰$ داریم

$$E\left[\exp(tZ)\right] \leq \exp(t^۲/۸).$$

حال قضیه اصلی را که در مورد نابرابری کولموگروف برای مجموع جزئی دنباله‌ای از متغیرهای تصادفی برنولی، بدون محدودیت برای پارامترها، ارائه می‌دهیم.

قضیه ۵.۳.۲. (تورنر و همکاران، ۱۹۹۵). فرض کنید $\{X_i, i \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل برنولی باشد به طوری که X_i دارای پارامتر p_i با $۰ \leq p_i \leq ۱$ باشد. هم‌چنین فرض کنید $\bar{X}_k = \dfrac{۱}{k}\sum_{i=۱}^{k} X_i$ و $\bar{p}_k = \dfrac{۱}{k}\sum_{i=۱}^{k} p_i$ است. آن‌گاه برای هر ثابت $\varepsilon \geq ۰$ و هر عدد مثبت $n \geq ۰$، داریم

$$P\left(\sup_{k\geq n}(\bar{X}_k - \bar{p}_k) \geq \varepsilon\right) \leq \exp\{-۲n\varepsilon^۲\}.$$

برهان. اگر $\varepsilon = ۰$ یا $n = ۰$ باشد نتیجه بدیهی است، بنابراین فرض می‌کنیم $\varepsilon > ۰$ و $n \geq ۱$ باشد. همچنین $Z_i = X_i - p_i$ را در نظر می‌گیریم. آن‌گاه از لم‌های ۳.۳.۲ و ۴.۳.۲ برای هر $۰ \leq t \leq ۸\varepsilon$ داریم

$$\begin{aligned} P\left(\sup_{k\geq n}(\bar{X}_k - \bar{p}_k) \geq \varepsilon\right) &\leq \exp\{-nt\varepsilon\} \prod_{i=۱}^{n} E\left[\exp\{tZ_i\}\right] \\ &\leq \exp\{(-nt\varepsilon) + nt^۲/۸\}. \end{aligned}$$

فرض می‌کنیم $۰ \leq t \leq ۸\varepsilon$ و $g(t) = n\{(-t\varepsilon) + t^۲/۸\}$ باشد، آن‌گاه $g(t)$ در $t = ۴\varepsilon$ به کمترین مقدار خود می‌رسد. زیرا

$$g^{'}(t) = n(-\varepsilon + \frac{۲t}{۸}) = n(-\varepsilon + \frac{t}{۴}) = ۰.$$

و چون مشتق دوم $g''(t) = \dfrac{n}{۴}$ مثبت است، بنابراین $t = ۴\varepsilon$، $g(t)$ را کمینه می‌کند. پس

$$P\left(\sup_{k\geq n}(\bar{X}_k - \bar{p}_k) \geq \varepsilon\right) \leq \exp\{-۲n\varepsilon^۲\}$$

و برهان کامل می‌شود. □

۱.۳.۲ نابرابری نمایی کولموگروف برای متغیرهای تصادفی وابسته زبرجمعی منفی

اقبال و همکاران (۲۰۱۱) نابرابری کولموگروف را برای برخی از صورت‌های درجه دوم

$$T_n = \sum_{1 \le i < j \le n} X_i X_j$$

و صورت‌های درجه دوم وزنی

$$Q_n = \sum_{1 \le i < j \le n} a_{ij} X_i X_j$$

به دست آوردند که در آن‌ها $\{X_i, i \ge 1\}$ دنباله‌ای نامنفی از متغیرهای تصادفی NSD است که به‌طور یکنواخت کران‌دار هستند و $\{a_{ij} : 1 \le i < j \le n\}$ یک آرایه از اعداد حقیقی نامنفی است.

قضیه ۶.۳.۲. (اقبال و همکاران، ۲۰۱۱). فرض کنید $\{X_i, i \ge 1\}$ دنباله‌ای از متغیرهای تصادفی نامنفی NSD با $P(a \le X_i \le b) = 1$، برای $1 \le i \le n$، باشد. آن‌گاه برای هر $x > 0$ می‌توان گفت

(i)

$$P(T_n - E(T_n) \ge x) \le exp\left\{-\frac{2(x-A_n)^2}{B_n}\right\}$$

که در آن $A_n = \dfrac{n(n-1)b(b-a)}{2}$ و $B_n = \dfrac{n(n-1)(2n-1)b^2(b-a)^2}{6}$.

(ii)

$$P(T_n - E(T_n) \le -x) \le \exp\left\{-\frac{\binom{n}{2}(b^2 - x + \sum_{i=3}^{n} a_i - a^2)^2}{2\sum_{i=3}^{n}(\sigma_i^2 + a_i^2)}\right\}$$

که در آن $a_i = (i-1)b(b-a)$ و $\sigma_i^2 = (i-1)^2 b^2 (b^2 - a^2)$.

به‌عنوان کاربردی از این قضیه، می‌توانیم یک نرخ همگرایی کامل برای دنباله $\{T_n - E(T_n), n \ge 2\}$ به صورت زیر به دست آوریم:

نتیجه ۷.۳.۲. فرض کنید $\{\gamma_n, n \ge 1\}$ دنباله‌ای نانزولی از اعداد حقیقی مثبت باشد، به‌طوری که $\gamma_n \to \infty$، $n \to \infty$ و $x\gamma_n - A_n = Cn^{\alpha}$ که در آن $(0 < C < \infty)$. برای هر $x > 0$ و هر $\alpha > \frac{3}{2}$ می‌توان نوشت

$$\lim_{n\to\infty} \frac{1}{\gamma_n}(T_n - E(T_n)) \xrightarrow{c.} 0$$

که در آن منظور از $\xrightarrow{c.}$ همگرایی کامل است. به‌طور خاص اگر $\gamma_n = \binom{n}{2}$، آن‌گاه

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}}(T_n - E(T_n)) \xrightarrow{c.} 0.$$

حال ما نتایج مشابه را در قضیه زیر برای صورت‌های درجه دوم وزنی Q_n که $\{a_{ij}, 1 \leq i < j \leq n\}$ یک آرایه از اعداد حقیقی نامنفی است، توضیح می‌دهیم.

قضیه ۸.۳.۲. فرض کنید $\{X_i, i \geq 1\}$ دنباله‌ای از متغیرهای تصادفی نامنفی NSD و $\{a_{ij}, 1 \leq i < j \leq n\}$ یک آرایه از اعداد حقیقی نامنفی باشد. در این صورت برای هر $x > 0$

(i)

$$P(Q_n - E(Q_n) \geq x) \leq \exp\left\{-\frac{2(x-C_n)^2}{D_n}\right\}$$

که در آن $C_n = b(b-a)\sum_{j=2}^{n}\sum_{i=1}^{j-1} a_{ij}$ و $D_n = b^2(b-a)^2\sum_{j=2}^{n}(\sum_{i=1}^{j-1} a_{ij})^2$.

(ii)

$$P(Q_n - E(Q_n) \leq -x) \leq exp\left\{-\frac{(b^2\sum_{1\leq i<j\leq n} a_{ij} - x + \sum_{i=3}^{n} c_i - a_{12}a^2)^2}{2\sum_{i=3}^{n}(\sigma_i^2 + c_i^2)}\right\}$$

که در آن $c_i = b(b-a)\sum_{j=1}^{i-1} a_{ij}$ و $\sigma_i^2 = b^2(\sum_{j=1}^{i-1} a_{ij})^2(b^2-a^2)$.

شائو (۲۰۰۰) نابرابری نمایی کولموگروف را برای متغیرهای تصادفی NA اثبات کرده است. در ادامه این نابرابری را با استفاده از تعمیم متغیرهای تصادفی NA به متغیرهای تصادفی NSD در قالب قضیه ۱۱.۳.۲ بیان و اثبات می‌کنیم.

ابتدا به بیان دو لم از شائو (۲۰۰۰) برای کمک گرفتن در اثبات قضیه می‌پردازیم.

لم ۹.۳.۲. (شائو، ۲۰۰۰). اگر $\{T_i, 1 \leq i \leq n\}$ زیرمارتینگل نامنفی باشد، آن‌گاه برای هر $0 < \alpha < 1$ می‌توان گفت

$$E\max_{1\leq i\leq n} T_i^\alpha \leq \frac{(ET_1)^\alpha}{(1-\alpha)}.$$

لم ۱۰.۳.۲. (شائو، ۲۰۰۰). برای هر $x \geq 0$ داریم

$$\ln(1+x) \geq \frac{x}{1+x} + \frac{x^2}{2(1+x)^2}\left(1 + \frac{2}{3}\ln(1+x)\right).$$

قضیه ۱۱.۳.۲. فرض کنید $\{X_i, 1 \leq i \leq n\}$ مجموعه‌ای از متغیرهای تصادفی NSD با میانگین صفر و واریانس متناهی باشند. همچنین برای هر $n \geq 1$، تعریف می‌کنیم $S_k = \sum_{i=1}^{k} X_i$ و $B_n = \sum_{i=1}^{n} EX_i^2$. در این صورت برای هر $x > 0, a > 0$، $n \geq 1$ و $0 < \alpha < 1$ می‌توان نوشت

$$\begin{aligned} P\left(\max_{1\leq k\leq n} S_k \geq x\right) &\leq P\left(\max_{1\leq k\leq n} X_k \geq a\right) \\ &+ \frac{1}{1-\alpha}\exp\left\{-\frac{x^2\alpha}{2(ax+B_n)}.\left(1+\frac{2}{3}\ln\left(1+\frac{ax}{B_n}\right)\right)\right\} \end{aligned} \tag{۵.۲}$$

و

$$P\left(\max_{1\le k\le n}|S_k|\ge x\right) \le 2P\left(\max_{1\le k\le n}|X_k|\ge a\right) + \frac{2}{1-\alpha}\exp\left\{-\frac{x^2\alpha}{2(ax+B_n)}.\left(1+\frac{2}{3}\ln\left(1+\frac{ax}{B_n}\right)\right)\right\}. \quad (۶.۲)$$

به ویژه می‌توان نتیجه گرفت

$$P\left(\max_{1\le k\le n}|S_k|\ge x\right) \le 2P\left(\max_{1\le k\le n}|X_k|\ge a\right) + 4\exp\left\{-\frac{x^2}{8B_n}\right\}+4\left(\frac{B_n}{xa+B_n}\right)^{\frac{x}{12a}}. \quad (۷.۲)$$

برهان. نابرابری (۶.۲) مستقیماً از نابرابری (۵.۲) نتیجه گرفته می‌شود. بنابراین برای قسمت اول فقط کافی است نابرابری (۵.۲) را اثبات کنیم.

فرض می‌کنیم

$$t=a^{-1}\ln(1+xa/B_n),\ \ \tilde{S}_i=\sum_{j=1}^{i}min(X_j,a),\ \ U_i=\sum_{j=1}^{i}Y_j,\ \ 1\le i\le n \quad (۸.۲)$$

که در آن Y_1، Y_2، ...، Y_n متغیرهای تصادفی مستقل هستند به‌طوری‌که Y_i و $\min(X_i,a)$ برای هر $1\le i\le n$ هم‌توزیع هستند. از آن‌جا که $\{\min(X_i,a), 1\le i\le n\}$ مجموعه‌ای از متغیرهای تصادفی NSD هستند، بنابراین با استفاده از رابطه (۱۱.۱) که هم برای متغیرهای تصادفی NA و هم برای متغیرهای تصادفی NSD برقرار است، می‌توانیم به دست آوریم

$$\begin{aligned} P(\max_{1\le i\le n}S_i\ge x) &\le P(\max_{1\le i\le n}X_i>a)+P(\max_{1\le i\le n}\tilde{S}_i\ge X)\\ &\le P(\max_{1\le i\le n}X_i>a)+e^{-t\alpha x}Eexp\{t\alpha\max_{1\le i\le n}\tilde{S}_i\}\\ &\le P(\max_{1\le i\le n}X_i>a)+e^{-t\alpha x}E\exp\{t\alpha\max_{1\le i\le n}U_i\}. \end{aligned} \quad (۹.۲)$$

قرار می‌دهیم

$$J_1=e^{-t\alpha x}E\exp\{t\alpha\max_{1\le i\le n}U_i\},\quad T_i=exp\{tU_i-(e^{ta}-1-ta)a^{-2}B_i\}.$$

توجه کنید که $\frac{(e^x-1-x)}{x^2}$ تابعی ناکاهشی از x روی $\Re$ است و $EY_i\le 0$ و $tY_i\le ta$ می‌باشد. بنابراین

$$\begin{aligned} Ee^{tY_i} &= 1+tEY_i+E\left(\frac{(e^{ty_i}-1-tY_i)}{Y_i^2}Y_i^2\right)\\ &\le 1+(e^{ta}-1-ta)a^{-2}EY_i^2=1+(e^{ta}-1-ta)a^{-2}EX_i^2\\ &\le \exp\{(e^{ta}-1-ta)a^{-2}EX_i^2\}. \end{aligned}$$

عبارت آخر در نابرابری فوق از رابطه $۱ + x < e^x$ نتیجه گرفته شده است. همچنین $\{T_i, ۱ \leq i \leq n\}$ یک زبرمارتینگل است. با استفاده از لم ۱۰.۳.۲ می‌توان نوشت

$$\begin{aligned}
& (xa + B_n)a^{-۲} ln(۱ + xa/B_n) - x/a \\
\geq\ & (xa + B_n)a^{-۲}\left(\frac{xa}{xa + B_n} + \frac{(xa)^۲}{۲(xa + B_n)^۲}\left(۱ + \frac{۲}{۳}\ln(۱ + xa/B_n)\right)\right) - x/a \\
=\ & \frac{x^۲}{۲(xa + B_n)}\left(۱ + \frac{۲}{۳}\ln(۱ + xa/B_n)\right)
\end{aligned}$$

بنابراین با استفاده از لم ۹.۳.۲ و رابطه (۸.۲)، داریم

$$\begin{aligned}
J_۱ &= e^{-t\alpha x} E\{\max_{۱\leq i\leq n} T_i e^{(e^{ta}-۱-ta)a^{-۲}B_i}\}^\alpha \\
&\leq \exp\{-t\alpha x + \alpha(e^{ta} - ۱ - ta)a^{-۲}B_n\} E\{\max_{۱\leq i\leq n} T_i\}^\alpha \\
&\leq (۱-\alpha)^{-۱}\exp\{-t\alpha x + \alpha(e^{-ta} - ۱ - ta)a^{-۲}B_n\} \\
&= (۱-\alpha)^{-۱}\exp\{-\alpha(a^{-۱}x\ln(۱ + xa/B_n) \\
&- (xa/B_n - \ln(۱ + xa/B_n))a^{-۲}B_n)\} \\
&= (۱-\alpha)^{-۱} exp\{-\alpha((xa + B_n)a^{-۲}\ln(۱ + xa/B_n) - x/a)\} \\
&\leq (۱-\alpha)^{-۱}\exp\left(-\frac{\alpha x^۲}{۲(xa + B_n)}\left(۱ + \frac{۲}{۳}\ln(۱ + xa/B_n)\right)\right). \qquad (۱۰.۲)
\end{aligned}$$

با قرار دادن (۱۰.۲) در (۹.۲) به نابرابری (۵.۲) می‌رسیم. نابرابری (۷.۲) با گرفتن $\alpha = \frac{۱}{۲}$ و در نظر گرفتن این‌که $B_n \leq xa$ یا $B_n > xa$ است، از (۶.۲) نتیجه می‌شود. داریم

$$\begin{aligned}
& ۴\exp\left\{-\frac{x^۲}{۴(ax + B_n)}\left(۱ + \frac{۲}{۳}\ln(۱ + \frac{ax}{B_n})\right)\right\} \\
\leq\ & ۴\exp\left\{-\frac{x^۲}{۴(ax + B_n)}\left(\frac{۲}{۳}\ln(۱ + \frac{ax}{B_n})\right)\right\} \\
=\ & ۴\exp\left\{\ln(۱ + \frac{ax}{B_n})^{-\frac{۲}{۳}\frac{x^۲}{۴(ax + B_n)}}\right\}
\end{aligned}$$

زمانی که $ax \geq B_n$

$$۴(\frac{B_n + ax}{B_n})^{-\frac{۲}{۳}\frac{x^۲}{۴(۲ax)}} = ۴(\frac{B_n}{ax + B_n})^{\frac{x}{۱۲a}} \qquad (۱۱.۲)$$

زمانی که $ax < B_n$

اگر $ax < B_n$ باشد آنگاه $۲B_n > ax + B_n > ۲ax$

$$
\begin{aligned}
۴\exp\{-\frac{۱}{۴(ax+B_n)}\} &\leq ۴\exp\{-\frac{۱}{۴(۲B_n)}\} \\
&= ۴\exp\{-\frac{۱}{۸B_n}\} \\
&\leq ۴\exp\{-\frac{x^۲}{۸B_n}\} \qquad (۱۲.۲)
\end{aligned}
$$

با جایگذاری روابط (۱۱.۲) و (۱۲.۲) در نابرابری (۶.۲) می‌توان به نابرابری (۷.۲) رسید و برهان کامل می‌شود.
□

فصل ۳

همگرایی کامل آرایه‌های سطری از متغیرهای تصادفی

۱.۳ مقدمه

در ابتدا لازم است دو نوع همگرایی را تعریف و رابطه بین آن‌ها را بیان کنیم. برای مطالعه در مورد همگرایی‌های دیگر و روابط بین انواع همگرایی‌ها به گات (۲۰۰۵) مراجعه کنید.

تعریف ۱.۱.۳. (همگرایی در احتمال[۱]). دنباله متغیرهای تصادفی $\{X_n, n \geq 1\}$ در احتمال به متغیرهای تصادفی X همگراست، هرگاه برای هر $\varepsilon > 0$، وقتی $n \to \infty$ داشته باشیم

$$\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1 \qquad (۱.۳)$$

که آن را با نماد $X_n \xrightarrow{p} X$ نشان می‌دهیم.

تعریف ۲.۱.۳. (همگرایی تقریباً همه‌جا[۲]). دنباله $\{X_n, n \geq 1\}$ را به صورت تقریباً همه‌جا همگرا به متغیر تصادفی X است، وقتی $n \to \infty$ آن‌گاه

$$P\{\omega \in \Omega : X_n(\omega) \to X(\omega), n \to \infty\} = 1$$

[۱] Convergence in probability

[۲] Almost surely

یا

$$P\{\omega \in \Omega : X_n(\omega) \nrightarrow X(\omega), n \to \infty\} = 0.$$

این همگرایی را با نماد $X_n \xrightarrow{a.s.} X$ نشان می‌دهیم.

تعریف ۳.۱.۳. دنباله $\{X_n, n \geq 1\}$ به‌طور کامل همگرا[۳] به متغیر تصادفی X است، هرگاه به ازای هر $\varepsilon > 0$ وقتی $n \to \infty$، داشته باشیم

$$\sum_{n=1}^{\infty} P\left(|X_n - X| > \varepsilon\right) < \infty$$

که با نماد $X_n \xrightarrow{c.} X$ نشان می‌دهیم.

تعریف ۴.۱.۳. فرض کنید $\{A_n, n \geq 1\}$ دنباله‌ای از پیشامدها باشد که زیرمجموعه‌های اندازه‌پذیر از Ω هستند. با استفاده از تعاریف

$$A_* = \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m \quad \text{و} \quad A^* = \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m$$

اگر $\omega \in \Omega$ متعلق به مجموعه‌ی $\liminf_{n\to\infty} A_n$ باشد، آن‌گاه ω به ازای حداقل یک n متعلق به $\bigcap_{m=n}^{\infty} A_m$ است. به عبارت دیگر برای $m \geq n$، $\omega \in A_m$ است. به‌طور خاص، اگر A_n پیشامدی باشد که در زمان n رخ دهد، آن‌گاه $\liminf_{n\to\infty} A_n$ به این معنی است که برخی nها در این خاصیت هرگز اتفاق نمی‌افتند. به‌طور مشابه اگر $\omega \in \Omega$ متعلق به مجموعه $\limsup_{n\to\infty} A_n$ باشد، آن‌گاه ω به ازای همه $n \geq 1$ متعلق به $\bigcup_{m=n}^{\infty} A_m$ است، که مهم نیست ما n را چقدر بزرگ انتخاب می‌کنیم زیرا همیشه $m \geq n$ می‌باشد، به‌طوری که $\omega \in A_m$ است یا به‌طور معادل برای بی‌نهایت از مقادیر m، $\omega \in A_m$ است. یک راه مناسب برای بیان این مسئله به صورت استفاده از نماد $i.o.$[۴] است، به این معنی که اگر حدود بالایی و پایینی هم‌زمان وجود داشته باشند و

$$A = A^* = A_* = \lim_{n\to\infty} A_n$$

آن‌گاه $\omega \in A^*$ اگر و فقط اگر A_n بی‌نهایت بار رخ می‌دهد که آن را با نماد $\omega \in \{A_n; i.o.\}$ نشان می‌دهند.

لم ۵.۱.۳. (لم اول بورل-کانتلی[۵]). اگر $\{A_n, n \geq 1\}$ دنباله‌ای از پیشامدهای Ω باشد و $\sum_{n=1}^{\infty} P(A_n) < \infty$، آن‌گاه

$$P(A_n; i.o.) = 0.$$

[۳] Complete convergence

[۴] Infinitely often

[۵] Borel-kantly

برهان. می‌توان نوشت

$$
\begin{aligned}
P(A_n; i.o.) = P(\limsup_{n\to\infty} A_n) &= P(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_m) \leq P(\bigcup_{n=m}^{\infty} A_n) \\
&\leq \sum_{n=m}^{\infty} P(A_n) \xrightarrow{m\to\infty} \lim_{m\to\infty} P(A_m) = 0.
\end{aligned}
$$

□

قضیه ۶.۱.۳. اگر X_1، X_2، ... متغیرهای تصادفی مستقل باشند و C یک عدد ثابت باشد، آن‌گاه هنگامی‌که $n \to \infty$، $X_n \xrightarrow{c.} C$ اگر و فقط اگر $X_n \xrightarrow{a.s.} C$.

برهان. چون X_n ها مستقل‌اند پس $X_n - C$ نیز از هم مستقل هستند و بنابر لم اول بورل-کانتلی نتیجه می‌گیریم

$$
\sum_{n=1}^{\infty} P\left(|X_n - C| > \varepsilon\right) < \infty.
$$

بنابراین $P\left(|X_n - C| > \varepsilon \;\; i.o\right) = 0$ و برهان کامل می‌شود. □

قضیه همگرایی کامل از آرایه‌های سطری متغیرهای تصادفی مستقل، ابتدا توسط هو و همکاران (۱۹۹۸) مطرح شد. نتیجه همگرایی کامل برای آرایه‌ای از متغیرهای تصادفی مستقل با میانگین صفر توسط کروگلو و همکاران (۲۰۰۶) به‌دست آمده است. این نتیجه به صورت جزئی به متغیرهای تصادفی وابسته منفی و همچنین وابسته منفی با میانگین صفر توسط چن و همکاران (۲۰۰۷) و دهوا و همکاران (۲۰۱۱) تعمیم داده شده است.

سانگ و همکاران (۲۰۰۵) قضیه همگرایی کامل برای آرایه‌های سطری از متغیرهای تصادفی مستقل $\{X_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ را که در آن $\{k_n, n \geq 1\}$ دنباله‌ای از اعداد صحیح مثبت است، به صورت زیر به دست آوردند.

قضیه ۷.۱.۳. فرض کنید $\{X_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ یک آرایه سطری از متغیرهای تصادفی و $\{a_n, n \geq 1\}$ دنباله‌ای از ثابت‌های نامنفی باشند. همچنین فرض کنید شرایط زیر برقرار باشند:

۱) برای هر $\varepsilon > 0$

$$
\sum_{n=1}^{\infty} a_n \sum_{i=1}^{k_n} P(|X_{ni}| > \varepsilon) < \infty.
$$

۲) $J \geq 2$ و $\delta > 0$ وجود داشته باشند، به طوری که

$$
\sum_{n=1}^{\infty} a_n \left(\sum_{i=1}^{k_n} E(X_{ni}^2 I(|X_{ni}| \leq \delta))\right)^J < \infty. \qquad (۲.۳)
$$

۳) اگر وقتی $n \to \infty$ داشته باشیم

$$\sum_{i=1}^{k_n} E(X_{ni} I(|X_{ni}| \leq \delta)) \to 0.$$

آن‌گاه برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} a_n P\left(\left|\sum_{i=1}^{k_n} X_{ni}\right| > \varepsilon\right) < \infty.$$

یکی از فرضیات اساسی ما در این فصل علاوه بر شرایط اولیه‌ای که برای دنباله به آن اشاره می‌کنیم، این است که متغیر تصادفی X بر این دنباله غلبه تصادفی[۶] داشته باشد. بنابراین لازم است مفهوم غلبه تصادفی را در این‌جا بیان کنیم.

تعریف ۸.۱.۳. (وو، ۲۰۰۶). متغیر تصادفی X بر دنباله $\{X_n, n \geq 1\}$ از متغیرهای تصادفی غلبه تصادفی دارد، اگر یک ثابت مثبت C وجود داشته باشد به‌طوری که برای هر $x \geq 0$ و $n \geq 1$ داشته باشیم

$$P(|X_n| > x) \leq CP(|X| > x).$$

به‌طور مشابه متغیر تصادفی X بر $\{X_{ni}, i \geq 1, n \geq 1\}$ که آرایه‌ای از متغیرهای تصادفی است غلبه تصادفی دارد، اگر یک ثابت مثبت C وجود داشته باشد به‌طوری که برای هر $x \geq 0$، $i \geq 1$ و $n \geq 1$ داشته باشیم

$$P(|X_{ni}| > x) \leq CP(|X| > x).$$

مثال ۹.۱.۳. فرض کنید متغیرهای تصادفی X و Y به ترتیب دارای توابع چگالی زیر باشند:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(X-\mu_1)^2}{2\sigma_1^2}}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(Y-\mu_2)^2}{2\sigma_2^2}}.$$

در این صورت، اگر شرایط

i) $\mu_1 > \mu_2$

ii) $\sigma_1^2 \leq \sigma_2^2$

برقرار باشند، آن‌گاه X بر Y غلبه تصادفی دارد.

لم ۱۰.۱.۳. (وو، ۲۰۰۶). فرض کنید متغیر تصادفی X بر $\{X_{ni}, i \geq 1, n \geq 1\}$ که آرایه‌ای از متغیرهای تصادفی است، غلبه تصادفی دارد. برای هر $\alpha > 0$ و $b > 0$ دو جمله زیر برقرار هستند:

$$\begin{aligned} E|X_{ni}|^\alpha I(|X_{ni}| \leq b) &\leq C_1 \left[E|X|^\alpha I(|X_{ni}| \leq b) + b^\alpha P(|X| > b)\right] \\ E|X_{ni}|^\alpha I(|X_{ni}| > b) &\leq C_2 E|X|^\alpha I(|X| \leq b) \end{aligned}$$

که در آن‌ها C_1 و C_2 ثابت‌های مثبت هستند.

---

[۶] Stochastic domination

۲.۳ همگرایی کامل آرایه‌های سطری از متغیرهای تصادفی وابسته زیرجمعی منفی

اخیرا چن و همکاران (۲۰۰۷) قضیه همگرایی کامل را برای آرایه‌های سطری متغیرهای تصادفی NA، $\{X_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ به دست آورده‌اند، که در آن $\{a_n, n \geq 1\}$ دنباله‌ای از اعداد صحیح مثبت است و $\lim\limits_{n\to\infty} k_n = \infty$.

قضیه ۱.۲.۳. فرض کنید $\{X_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ آرایه‌ای سطری از متغیرهای تصادفی NA با میانگین صفر و $\{a_n, n \geq 1\}$ دنباله‌ای از اعداد ثابت مثبت باشند. فرض کنید شرایط زیر برقرار باشند:

۱) برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} a_n \sum_{i=1}^{k_n} P(|X_{ni}| > \varepsilon) < \infty. \tag{۳.۳}$$

۲) $J \geq 1$ وجود داشته باشد، به طوری که

$$\sum_{n=1}^{\infty} a_n \left(\sum_{i=1}^{k_n} EX_{ni}^2 \right)^J < \infty. \tag{۴.۳}$$

۳) اگر وقتی $n \to \infty$ داشته باشیم

$$\sum_{i=1}^{k_n} EX_{ni}^2 \to 0$$

آن‌گاه برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} a_n P\left(\max_{1\leq m\leq k_n} \left| \sum_{i=1}^{m} X_{ni} \right| > \varepsilon \right)^J < \infty. \tag{۵.۳}$$

وانگ و همکاران (۲۰۱۴) قضیه ۱.۲.۳ را به آرایه‌های سطری متغیرهای تصادفی NSD تعمیم و گسترش دادند ولی برای اثبات آن به یک لم از وانگ و همکاران (۲۰۱۴) نیاز داریم که ابتدا آن را بیان و اثبات می‌کنیم.

لم ۲.۲.۳. فرض کنید $\{X_n, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی NSD با میانگین صفر و گشتاور دوم متناهی باشد. برای هر $n \geq 1$ در نظر می‌گیریم $S_n = \sum\limits_{i=1}^{n} X_i$ و $B_n = \sum\limits_{i=1}^{n} EX_i^2$ باشد. آن‌گاه برای هر $x > 0$، $y > 0$ و $n \geq 1$، داریم

$$P\left(\max_{1\leq k\leq n} |S_k| \geq x \right) \leq 2P\left(\max_{1\leq k\leq n} |X_k| \geq y \right) + 8\left(\frac{2B_n}{3xy} \right)^{x/12y}. \tag{۶.۳}$$

برهان. با استفاده از نابرابری کولموگروف در قضیه ۱۱.۳.۲، می‌توان نوشت

$$\begin{aligned} P\left(\max_{1\leq k\leq n} |S_k| \geq x \right) &\leq 2P\left(\max_{1\leq k\leq n} |X_k| \geq y \right) \\ &+ 4\exp\left\{ -\frac{x^2}{8B_n} \right\} + 4\exp\left\{ -\frac{B_n}{(xy+B_n)} \right\}^{x/12y}. \end{aligned}$$

اگر از $\left\{\frac{x^۲}{۸B_n}\right\}$ عبارت $x/۱۲y$ را فاکتور بگیریم، نتیجه می‌شود

$$\exp\left\{-\frac{x^۲}{۸B_n}\right\} = \left[\exp\left\{-\frac{۳xy}{۲B_n}\right\}\right]^{x/۱۲y}.$$

از این عبارت لگاریتم می‌گیریم و از آن‌جا که $\frac{۱}{x} \geq \frac{۱}{(۱+x)} \geq e^{-x}$، برای هر $x > ۰$، $y > ۰$ و $n \geq ۱$ داریم

$$\exp\left\{-\frac{x^۲}{۸B_n}\right\} = \left[\exp\left\{-\frac{۳xy}{۲B_n}\right\}\right]^{x/۱۲y} \leq \left(-\frac{۳xy}{۲B_n}\right)^{x/۱۲y}.$$

عبارت آخر را معکوس می‌کنیم و نتیجه می‌شود

$$\exp\left\{-\frac{x^۲}{۸B_n}\right\} = \left[\exp\left\{-\frac{۳xy}{۲B_n}\right\}\right]^{x/۱۲y} \leq \left(\frac{۲B_n}{۳xy}\right)^{x/۱۲y}. \tag{۷.۳}$$

از طرف دیگر چون $B_n \geq ۰$، پس

$$\left(\frac{B_n}{(xy+B_n)}\right)^{x/۱۲y} \leq \left(\frac{B_n}{۴xy}\right)^{x/۱۲y} \leq \left(\frac{۲B_n}{۳xy}\right)^{x/۱۲y}. \tag{۸.۳}$$

با جایگذاری روابط (۷.۳) و (۸.۳) در نابرابری کولموگروف به معادله (۶.۳) می‌رسیم و برهان کامل است. □

قضیه ۳.۲.۳. فرض کنید $\{X_{ni}, ۱ \leq i \leq k_n, n \geq ۱\}$ آرایه‌ای سطری از متغیرهای تصادفی *NSD* با $EX_{ni} = ۰$ و $EX_{ni}^۲ < \infty$ برای $۱ \leq i \leq k_n$، $n \geq ۱$ باشد. همچنین فرض کنید $\{a_n, n \geq ۱\}$ دنباله‌ای از اعداد ثابت نامنفی باشد. فرض می‌کنیم شرایط ۱ و ۲ از قضیه ۱.۲.۳ برقرار باشند، آن‌گاه معادله (۵.۳) نیز برای هر $\varepsilon > ۰$ برقرار است.

برهان. با به کار بردن لم ۲.۲.۳ و جایگذاری $x = \varepsilon$ و $y = \frac{\varepsilon}{۱۲J}$، داریم

$$\begin{aligned}
& P\left(\max_{۱\leq m\leq k_n}\left|\sum_{i=۱}^{m} X_{ni}\right| > \varepsilon\right) \\
\leq\ & ۲P\left(\max_{۱\leq i\leq k_n}|X_{ni}| > \frac{\varepsilon}{۱۲J}\right) + ۸\left(\frac{۸J}{\varepsilon^۲}\right)^J\left(\sum_{i=۱}^{k_n} EX_{ni}^۲\right)^J \\
\leq\ & ۲\sum_{i=۱}^{k_n} P\left(|X_{ni}| > \frac{\varepsilon}{۱۲J}\right) + ۸^{J+۱}J^J\varepsilon^{-۲J}\left(\sum_{i=۱}^{k_n} EX_{ni}^۲\right)^J.
\end{aligned} \tag{۹.۳}$$

همچنین با استفاده از معادله‌های (۳.۳) و (۴.۳) به همان نتیجه قضیه ۱.۲.۳ یعنی

$$\sum_{n=۱}^{\infty} a_n P\left(\max_{۱\leq m\leq k_n}\left|\sum_{i=۱}^{m} X_{ni}\right| > \varepsilon\right) = \sum_{n=۱}^{\infty} a_n P\left(\max_{۱\leq m\leq k_n}\left|\sum_{i=۱}^{m} X_{ni}\right| > \varepsilon\right)^J < \infty$$

می‌رسیم و برهان کامل است. □

همچنین وانگ و همکاران (۲۰۱۴) مشابه برهان قضیه ۱ از چن و همکاران (۲۰۰۷) و با استفاده از لم ۲.۲.۳ همگرایی کامل آرایه‌های سطری متغیرهای تصادفی NSD را به صورت زیر به دست آوردند:

قضیه ۴.۲.۳. فرض کنید $\{X_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ یک آرایه سطری از متغیرهای تصادفی NSD با میانگین صفر و $\{k_n, n \geq 1\}$ دنباله‌ای از اعداد ثابت مثبت باشند. همچنین فرض کنید شرایط زیر برقرار باشند:

۱) برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} a_n \sum_{i=1}^{k_n} P(|X_{ni}| > \varepsilon) < \infty. \tag{۱۰.۳}$$

۲) وجود دارند $J \geq 1$ و $\delta > 0$ به‌طوری که

$$\sum_{n=1}^{\infty} a_n \left(\sum_{i=1}^{k_n} Var(X_{ni} I(|X_{ni}| \leq \delta)) \right)^J < \infty \tag{۱۱.۳}$$

آن‌گاه برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} a_n P\left(\max_{1 \leq m \leq k_n} \left| \sum_{i=1}^{m} (X_{ni} - EX_{ni} I(|X_{ni}| \leq \delta)) \right| > \varepsilon \right) < \infty. \tag{۱۲.۳}$$

نتیجه ۵.۲.۳. (وانگ و همکاران، ۲۰۱۴). فرض کنید شرایط قضیه ۴.۲.۳ برقرار باشند. همچنین فرض کنید

$$\max_{1 \leq m \leq k_n} \left| \sum_{i=1}^{m} E(X_{ni} I(|X_{ni}| \leq \delta)) \right| \to 0$$

به‌طوری که $n \to \infty$، آن‌گاه رابطه (۵.۳) برای هر $\varepsilon > 0$ برقرار است.

ملاحظه ۶.۲.۳. در قضیه ۳.۲.۳ ما فقط به شرایط ۱ و ۲ از قضیه ۱.۲.۳ نیاز داریم و به شرط ۳ نیازی نداریم.

ملاحظه ۷.۲.۳. باید توجه کنیم که همه نتایج بالا در مورد $k_n = \infty$ برای هر $n \geq 1$ نیز صدق می‌کند. ثابت شده است که سری‌های $\sum_{i=1}^{\infty} X_{ni}$ به‌طور تقریبا همه‌جا همگرا هستند. البته باید در نظر بگیریم که در جمع‌های نامتناهی به جای ماکسیمم از سوپریمم استفاده کنیم.

نتایج ارائه‌شده از همگرایی کامل، برای مجموع موزون آرایه‌های سطری متغیرهای تصادفی سطری NSD نتیجه‌گیری می‌شود.

قضیه ۸.۲.۳. (وانگ و همکاران، ۲۰۱۴). فرض کنید $\beta \geq -1$ باشد. همچنین فرض کنید $\{a_{ni}, i \geq 1, n \geq 1\}$ که یک آرایه از ثابت‌هاست و متغیر تصادفی X بر $\{X_{ni}, i \geq 1, n \geq 1\}$، آرایه‌ای سطری از متغیرهای تصادفی NSD، غلبه تصادفی دارد، به‌طوری که به ازای $r > 0$ داریم

$$\sup_{i \geq 1} |a_{ni}| = O(n^{-r}) \tag{۱۳.۳}$$

و برای یک $0 < \theta < 2$ و α به‌طوری که $\theta + \frac{\alpha}{r} < 2$، داریم

$$\sum_{i=1}^{\infty} |a_{ni}|^{\theta} = O(n^{\alpha}). \tag{۱۴.۳}$$

(۱) اگر $1 + \alpha + \beta < 0$ و $E|X|^{\theta} < \infty$ باشد، آن‌گاه برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} n^{\beta} P\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} a_{ni}X_{ni}\Big| > \varepsilon\right) < \infty \tag{۱۵.۳}$$

و

$$\sum_{n=1}^{\infty} n^{\beta} P\left(\Big|\sum_{i=1}^{\infty} a_{ni}X_{ni}\Big| > \varepsilon\right) < \infty. \tag{۱۶.۳}$$

(۲) اگر $1 + \alpha + \beta > 0$، $\beta > -1$ و $s = \theta + \frac{1+\alpha+\beta}{r}$، آن‌گاه

$$E|X|^{s} < \infty. \tag{۱۷.۳}$$

افزون بر این فرض کنید وقتی $s \geq 1$ است $EX_{ni} = 0$ باشد، آن‌گاه معادله‌های (۱۵.۳) و (۱۶.۳) برقرار هستند.

(۳) اگر $1 + \alpha + \beta = 0$ و

$$E|X|^{\theta} log|X| < \infty \tag{۱۸.۳}$$

و همچنین وقتی $1 \leq \theta < 2$، $EX_{ni} = 0$، آن‌گاه معادله‌های (۱۵.۳) و (۱۶.۳) برقرار هستند.

برهان. برهان معادله (۱۶.۳) شبیه به برهان معادله (۱۵.۳) است. بنابراین فقط به اثبات معادله (۱۵.۳) بسنده می‌کنیم. بدون از دست دادن کلیت، فرض می‌کنیم برای هر $i \geq 1$ و $n \geq 1$، $a_{ni} > 0$ باشد (به عبارت دیگر به جای a_{ni} به ترتیب از a^{+}_{ni} و a^{-}_{ni} استفاده می‌کنیم و به خاطر می‌سپاریم که $a_{ni} = a^{+}_{ni} - a^{-}_{ni}$). با استفاده از شرایط (۱۳.۳) و (۱۴.۳) فرض می‌کنیم که

$$\sup_{i\geq 1} a_{ni} = n^{-r}, \quad \sum_{i=1}^{\infty} a^{\theta}_{ni} = n^{\alpha}, \qquad n \geq 1 \tag{۱۹.۳}$$

(۱) اگر $1 + \alpha + \beta < 0$ باشد، آن‌گاه

$$\begin{aligned}
\sum_{n=1}^{\infty} n^{\beta} P\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} a_{ni}X_{ni}\Big| > \varepsilon\right) &\leq C\sum_{n=1}^{\infty} n^{\beta} E\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} a_{ni}X_{ni}\Big|^{\theta}\right)\\
&\leq C\sum_{n=1}^{\infty} n^{\beta} \sum_{i=1}^{n} E|a_{ni}X_{ni}|^{\theta}\\
&= C\sum_{n=1}^{\infty} n^{\beta} \underbrace{\sum_{i=1}^{n} a^{\theta}_{ni}}_{n^{\alpha}} E|X_{ni}|^{\theta}\\
&\leq C\sum_{n=1}^{\infty} n^{\alpha+\beta} E|X|^{\theta} < \infty.
\end{aligned}$$

نابرابری اول با استفاده از نابرابری مارکف و نابرابری دوم با استفاده از نابرابری ماکسیمال روزنتال حاصل شده‌اند و چون $\alpha+\beta<-1$، پس عبارت $\sum_{n=1}^{\infty} n^{\alpha+\beta}$ متناهی است. همچنین چون $E|X|^{\theta}$ متناهی است، پس نتیجه حاصل می‌شود.

در ادامه، نتیجه را وقتی $1+\alpha+\beta\geq 0$ است، اثبات می‌کنیم. قرار می‌دهیم

$$X'_{ni}=\begin{cases} -1 & a_{ni}X_{ni}<-1\\ a_{ni}X_{ni} & -1\leq a_{ni}X_{ni}<1\\ 1 & a_{ni}X_{ni}>1\end{cases}$$

به عبارتی برای $n\geq 1$ و $i\geq 1$ می‌توان نوشت

$$X'_{ni}=-I(a_{ni}X_{ni}<-1)+a_{ni}X_{ni}I(|a_{ni}X_{ni}|\leq 1)+I(a_{ni}X_{ni}>-1).$$

بنابراین با استفاده از ویژگی (۳)، وابستگی زیرجمعی منفی برای $n\geq 1$ ثابت شده و $\{X'_{ni}, i\geq 1\}$ دنباله‌ای از متغیرهای تصادفی NSD است، که به آسانی برای هر $\varepsilon>0$ صدق می‌کند:

$$\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} a_{ni}X_{ni}\Big|>\varepsilon\right)\subset\bigcup_{i=1}^{n}\left(\Big|a_{ni}X_{ni}\Big|>1\right)\bigcup\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} X'_{ni}\Big|>\varepsilon\right)$$

نتیجه می‌دهد

$$\begin{aligned}\sum_{n=1}^{\infty} n^{\beta}P\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} a_{ni}X_{ni}\Big|>\varepsilon\right) &\leq \sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n}P(|a_{ni}X_{ni}|>1)\\ &+\sum_{n=1}^{\infty} n^{\beta}P\left(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} X'_{ni}\Big|>\varepsilon\right)=I+J.\end{aligned} \tag{۲۰.۳}$$

از این رو برای اثبات معادله (۱۵.۳) کافی است اثبات کنیم $I<\infty$ و $J<\infty$ هستند.

(۲) اگر $1+\alpha+\beta>0$ باشد، در ابتدا با استفاده از ویژگی غلبه تصادفی می‌توان گفت

$$\begin{aligned} I=\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n}P(|a_{ni}X_{ni}|>1) &\leq C\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n}P(|a_{ni}X|>1)\\ &= C\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n}P\{\omega\in\Omega, |a_{ni}X(\omega)|>1\}\\ &= C\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n}P\{\omega\in\Omega, |a_{ni}I(|X|>\frac{1}{a_{ni}})X(\omega)|>1\}.\end{aligned}$$

همچنین با به‌کار بردن نابرابری مارکف در این‌جا داریم

$$\begin{aligned} I &\leq C\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n} a_{ni}^{\theta}E|X|^{\theta}I^{\theta}\left(|X|>\frac{1}{a_{ni}}\right)\\ &= C\sum_{n=1}^{\infty} n^{\beta}\underbrace{\sum_{i=1}^{n} a_{ni}^{\theta}}_{n^{\alpha}}E|X|^{\theta}I\left(|X|>\frac{1}{a_{ni}}\right)\\ &\leq C\sum_{n=1}^{\infty} n^{\alpha+\beta}E|X|^{\theta}I(|X|>\underbrace{n^{r}}_{\sup_{i\geq 1} 1/a_{ni}=n^{r}}). \end{aligned}$$

می‌توان $I(|X|>n^{r})$ را به صورت $I(n^{r}<|X|<(n+1)^{r}$ یا $I((n+1)^{r}<|X|<(n+2)^{r}$ یا $\ldots$ و در نهایت به صورت $\sum_{k=n}^{\infty} I(k^{r}\leq|X|<(k+1)^{r})$ نمایش داد. پس

$$\begin{aligned} I &\leq C\sum_{n=1}^{\infty} n^{\alpha+\beta}\sum_{k=n}^{\infty} E|X|^{\theta}I\left(k^{r}\leq|X|<(k+1)^{r}\right)\\ &\leq C\sum_{k=1}^{\infty}\sum_{n=1}^{k} n^{\alpha+\beta}E|X|^{\theta}I\left(k^{r}\leq|X|<(k+1)^{r}\right). \end{aligned}$$

در این‌جا $\sum_{n=1}^{k} n^{\alpha+\beta}$ معادل است با

$$\sum_{n=1}^{k} n^{\alpha+\beta}=1^{\alpha+\beta}+2^{\alpha+\beta}+\cdots+k^{\alpha+\beta}\leq\underbrace{k^{\alpha+\beta}+k^{\alpha+\beta}+\cdots+k^{\alpha+\beta}}_{k\text{بار}}=k^{1+\alpha+\beta}.$$

بنابراین

$$I \leq C\sum_{k=1}^{\infty} k^{1+\alpha+\beta}E|X|^{\theta}I\left(k^{r}\leq|X|<(k+1)^{r}\right).$$

با توجه به بازه تعریف‌شده در تابع نشانگر می‌توان نوشت

$$I \leq C\sum_{k=1}^{\infty} E|X|^{\theta+(1+\alpha+\beta)/r}I\left(k^{r}\leq|X|<(k+1)^{r}\right)$$

و در نهایت با توجه به تبدیلات رابطه (۱۷.۳)، داریم

$$\sum_{n=1}^{\infty} n^{\beta}\sum_{i=1}^{n} P(|a_{ni}X_{ni}|>1) \leq CE|X|^{\theta+(1+\alpha+\beta)/r}<\infty. \tag{۲۱.۳}$$

در نتیجه از (۲۱.۳) می‌توان نتیجه گرفت که $I<\infty$.

اکنون به ترتیب برای $s\geq 1$ و $s<1$ ثابت می‌کنیم $J<\infty$ است.

مورد اول. $s \geq ۱$: ابتدا نشان می‌دهیم وقتی $n \to \infty$، داریم

$$\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} EX'_{ni}\right| \to ۰. \tag{۲۲.۳}$$

در واقع، با استفاده از شرط $EX_{ni} = ۰$ و لم ۱۰.۱.۳، داریم

$$\begin{aligned}
\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} EX'_{ni}\right| &\leq \max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} Ea_{ni}X_{ni}I(\left|a_{ni}X_{ni}\right| \leq ۱)\right| + \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱)\\
&= \max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} Ea_{ni}X_{ni}[۱ - I(|a_{ni}X_{ni}| \geq ۱)]\right| + \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱)\\
&= \max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} Ea_{ni}X_{ni} - \sum_{i=۱}^{j} Ea_{ni}X_{ni}I(|a_{ni}X_{ni}| > ۱)\right|\\
&+ \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱)\\
&= \max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}\underbrace{EX_{ni}}_{۰}\right| + \left|\sum_{i=۱}^{j} Ea_{ni}X_{ni}I(|a_{ni}X_{ni}| > ۱)\right| \qquad (۲۳.۳)\\
&+ \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱)\\
&= \max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} Ea_{ni}X_{ni}I(|a_{ni}X_{ni}| > ۱)\right| + \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱)\\
&\leq C\sum_{i=۱}^{n} E|a_{ni}X_{ni}|^{s}I(|a_{ni}X_{ni}| > ۱)\\
&\leq C\sum_{i=۱}^{n} a_{ni}^{s}E|X|^{s}I\left(|X| > \frac{۱}{a_{ni}}\right)\\
&\leq C\left(\sup_{i\geq ۱} a_{ni}\right)^{s-\theta}\sum_{i=۱}^{n} a_{ni}^{\theta}E|X|^{s}I\left(|X| > n^{r}\right)\\
&\leq C(n^{-r})^{s-\theta}n^{\alpha}E|X|^{s}I\left(|X| > n^{r}\right)\\
&\leq Cn^{-(۱+\beta)}E|X|^{s}I\left(|X| > n^{r}\right) \to ۰. \qquad (۲۴.۳)
\end{aligned}$$

دو جمله آخر به ترتیب از معادله (۱۹.۳) و رابطه (۱۷.۳) به دست آمده‌اند و چون $n \to \infty$ پس به صفر همگرا است. همچنین برای اثبات $J < \infty$، فقط لازم است نشان دهیم برای هر $\varepsilon > ۰$ داریم

$$J^{*} = \sum_{n=۱}^{\infty} n^{\beta}P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j}(X'_{ni} - EX'_{ni})\right| > \frac{\varepsilon}{۲}\right) < \infty. \tag{۲۵.۳}$$

با استفاده از نابرابری مارکف، داریم

$$J^{*} \leq C\sum_{n=۱}^{\infty} n^{\beta}E\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j}(X'_{ni} - EX'_{ni})\right|^{p}\right)$$

و همچنین با استفاده از نابرابری ینسن و C_r، داریم

$$
\begin{aligned}
J^* &\leq C\sum_{n=1}^{\infty} n^{\beta}\left[\left(\sum_{i=1}^{n} E|X'_{ni}|^2\right)^{p/2}+\sum_{i=1}^{n} E|X'_{ni}|^p\right]\\
&= J_1+J_2. \qquad (۲۶.۳)
\end{aligned}
$$

فرض می‌کنیم

$$
p > max\left(2, \frac{2(1+\beta)}{r(2-\theta)-\alpha}, \theta+\frac{1+\alpha+\beta}{r}\right)
$$

که مبتنی بر $\beta-[r(2-\theta)-\alpha]p/2 < -1$ و $\alpha+\beta-r(p-\theta) < -1$ پیشنهاد می‌شود. با استفاده از نابرابری C_r و لم ۱۰.۱.۳، می‌توان گفت

$$
\begin{aligned}
J_1 &= C\sum_{n=1}^{\infty} n^{\beta}\left(\sum_{i=1}^{n} E|X'_{ni}|^2\right)^{p/2}\\
&\leq C\sum_{n=1}^{\infty} n^{\beta}\left[\sum_{i=1}^{n} P(|a_{ni}X|>1)+\underbrace{\sum_{i=1}^{n} E|a_{ni}X|^2 I(|a_{ni}X|\leq 1)}_{\text{یک جمله متناهی است}}\right]^{P/2} \qquad (۲۷.۳).
\end{aligned}
$$

اگر $2 > s \geq 1$باشد، آن‌گاه با استفاده از نابرابری مارکف، $E|X|^s < \infty$ و معادله (۱۹.۳) داریم

$$
\begin{aligned}
J_1 &\leq C\sum_{n=1}^{\infty} n^{\beta}\left(\sum_{i=1}^{n} a_{ni}^s E|X|^s\right)^{p/2}\\
&\leq C\sum_{n=1}^{\infty} n^{\beta}\left[\left(\sup_{i=1} a_{ni}\right)^{s-\theta}\sum_{i=1}^{n} a_{ni}^{\theta}\right]^{P/2}\\
&\leq C\sum_{n=1}^{\infty} n^{\beta}\left[n^{-r(s-\theta)}.n^{\alpha}\right]^{p/2}\\
&= C\sum_{n=1}^{\infty} n^{\beta-(1+\beta)p/2} < \infty. \qquad (۲۸.۳)
\end{aligned}
$$

جمله آخر با استفاده از تبدیلات به دست آمده و از آن‌جا که $1+\beta > 0$، نتیجه شده است.

اگر $s \geq ۲$ باشد، آن‌گاه با استفاده از نابرابری مارکف، $E|X|^s < \infty$ و معادله (۱۹.۳) مجدداً داریم

$$\begin{aligned} J_۱ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\left(\sum_{i=۱}^{n} a_{ni}^{۲} E|X|^{۲}\right)^{p/۲} \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\left[\left(\sup_{i=۱} a_{ni}\right)^{۲-\theta}\sum_{i=۱}^{n} a_{ni}^{\theta}\right]^{P/۲} \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\left[n^{-r(۲-\theta)}.n^{\alpha}\right]^{p/۲} \\ &= C\sum_{n=۱}^{\infty} n^{\beta-[r(۲-\theta)-\alpha]p/۲} < \infty. \end{aligned} \tag{۲۹.۳}$$

از معادله‌های (۲۷.۳) تا (۲۹.۳)، اثبات کردیم که $J_۱ < \infty$ است.

با استفاده از تعریف غلبه تصادفی، می‌توان دید

$$J_۲ = C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} E|X'_{ni}|^{p}$$

و مجدداً با استفاده از لم ۱۰.۱.۳، داریم

$$\begin{aligned} J_۲ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n}\left[E|a_{ni}X_{ni}|^{p} I(|a_{ni}X_{ni}| \leq ۱) + P(|a_{ni}X_{ni}| > ۱)\right] \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} P(|a_{ni}X| > ۱) + C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} E|a_{ni}X|^{p} I(|a_{ni}X| \leq ۱) \\ &= J_۳ + J_۴. \end{aligned} \tag{۳۰.۳}$$

با استفاده از معادله (۲۱.۳) اثبات کردیم $J_۳ < \infty$ است.

در ادامه نشان می‌دهیم $J_۴ < \infty$ است. قرار می‌دهیم

$$I_{nj} = \left\{i : (nj)^r \leq \frac{۱}{a_{ni}} < [n(j+۱)]^r\right\}, \quad n \geq ۱, j \geq ۱. \tag{۳۱.۳}$$

به آسانی می‌بینیم که برای $k \neq j$، $I_{nk} \cap I_{nj} = \varnothing$ و برای هر $n \geq ۱$، $\bigcup_{j=۱}^{\infty} I_{nj} = \mathcal{N}$ است. همچنین

$$\begin{aligned} J_۴ &= C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} E|a_{ni}X|^{p} I(|a_{ni}X| \leq ۱) \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}\sum_{i\in I_{nj}} E|a_{ni}X|^{p} I(|a_{ni}X| \leq ۱). \end{aligned}$$

با توجه به بازه (۳۱.۳)، داریم

$$\begin{aligned}
J_۴ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}(\sharp I_{nj})(nj)^{-rp}E|X|^{p}I(|X|\leq[n(j+۱)]^{r}) \\
&\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}(\sharp I_{nj})(nj)^{-rp}\sum_{k=۰}^{n(j+۱)} E|X|^{p}I(k\leq|X|^{۱/r}<k+۱) \\
&= C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}(\sharp I_{nj})(nj)^{-rp}\sum_{k=۰}^{۲n} E|X|^{p}I(k\leq|X|^{۱/r}<k+۱) \\
&+ C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}(\sharp I_{nj})(nj)^{-rp}\sum_{k=۲n+۱}^{n(j+۱)} E|X|^{p}I(k\leq|X|^{۱/r}<k+۱) \\
&= J_۵+J_۶ \qquad (۳۲.۳)
\end{aligned}$$

که در آن منظور از نماد $\sharp$ تعداد است.

با استفاده از بازه (۳۱.۳) داریم

$$n^{\alpha} = \sum_{i=۱}^{\infty}a_{ni}^{\theta}=\sum_{j=۱}^{\infty}\sum_{i\in\sharp}a_{ni}^{\theta}\geq\sum_{j=۱}^{\infty}(\sharp I_{nj})[n(j+۱)]^{-r\theta}.$$

به آسانی برای هر $m\geq ۱$ نتیجه می‌شود

$$\begin{aligned}
n^{\alpha} &\geq \sum_{j=m}^{\infty}(\sharp I_{nj})[n(j+۱)]^{-r\theta}\geq\sum_{j=m}^{\infty}(\sharp I_{nj})[n(j+۱)]^{-r\theta}\left[\frac{n(m+۱)}{n(j+۱)}\right]^{r(p-\theta)} \\
&= \sum_{j=m}^{\infty}(\sharp I_{nj})[n(j+۱)]^{-r\theta}\left[n(j+۱)\right]^{-r(p-\theta)}\left[n(m+۱)\right]^{r(p-\theta)} \\
&= \sum_{j=m}^{\infty}(\sharp I_{nj})[n(j+۱)]^{-rp}\left[n(m+۱)\right]^{r(p-\theta)}.
\end{aligned}$$

برای هر $m\geq ۱$

$$\sum_{j=m}^{\infty}(\sharp I_{nj})(nj)^{-rp}\leq Cn^{\alpha}.n^{-r(p-\theta)}.m^{-r(p-\theta)}=Cn^{\alpha-r(p-\theta)}.m^{-r(p-\theta)}. \qquad (۳۳.۳)$$

بنابراین

$$J_۵ = C\sum_{n=۱}^{\infty} n^{\beta}\sum_{j=۱}^{\infty}(\sharp I_{nj})(nj)^{-rp}\sum_{k=۰}^{۲n} E|X|^{p}I(k\leq|X|^{۱/r}<k+۱).$$

با توجه به رابطه (۳۳.۳) و چون $m^{-r(p-\theta)}$ وابسته به n نیست، داریم

$$\begin{aligned}
J_5 &\leq C\sum_{n=1}^{\infty} n^{\beta}.n^{\alpha-r(p-\theta)}\sum_{k=0}^{2n} E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&\leq C\sum_{k=0}^{2}\sum_{n=1}^{\infty} n^{\alpha+\beta-r(p-\theta)}E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&+ C\sum_{k=2}^{\infty}\sum_{n=\lceil k/2\rceil}^{\infty} n^{\alpha+\beta-r(p-\theta)}E|X|^{p}I(k\leq |X|^{1/r}<k+1).
\end{aligned}$$

با توجه به تبدیلات ذکرشده داشتیم $\alpha+\beta-r(p-\theta)<-1$. بنابراین

$$J_5 \leq C+C\sum_{k=2}^{\infty} k^{1+\alpha+\beta-r(p-\theta)}E|X|^{p}I(k\leq |X|^{1/r}<k+1)$$

و دوباره با استفاده از تبدیل $\alpha+\beta-r(p-\theta)<-1$ با اضافه کردن $(1+\alpha+\beta)/r$ و کم کردن $(p-\theta)$ (که با هم معادل هستند) به توان $E|X|$، داریم

$$\begin{aligned}
J_5 &\leq C+C\sum_{k=2}^{\infty} E|X|^{p+((1+\alpha+\beta)/r)-(p-\theta)}I(k\leq |X|^{1/r}<k+1)\\
&\leq C+CE|X|^{\theta+((1+\alpha+\beta)/r)}=C+CE|X|^{s}<\infty \qquad (۳۴.۳)
\end{aligned}$$

و

$$\begin{aligned}
J_6 &= C\sum_{n=1}^{\infty} n^{\beta}\sum_{j=1}^{\infty}(\sharp I_{nj})(nj)^{-rp}\sum_{k=2n+1}^{n(j+1)} E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&\leq C\sum_{n=1}^{\infty} n^{\beta}\sum_{k=2n+1}^{\infty}\sum_{j\geq (k/n)-1}(\sharp I_{nj})(nj)^{-rp}E|X|^{p}I(k\leq |X|^{1/r}<k+1).
\end{aligned}$$

با استفاده از رابطه (۳۳.۳) داریم

$$\begin{aligned}
J_6 &\leq C\sum_{n=1}^{\infty} n^{\beta}\sum_{k=2n+1}^{\infty} n^{\alpha-r(p-\theta)}\left(\frac{k}{n}\right)^{-r(p-\theta)} E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&\leq C\sum_{k=2}^{\infty}\sum_{n=1}^{\lceil k/2\rceil} n^{\alpha+\beta}.k^{-r(p-\theta)}E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&\leq C\sum_{k=2}^{\infty} k^{1+\alpha+\beta-r(p-\theta)}E|X|^{p}I(k\leq |X|^{1/r}<k+1)\\
&\leq C\sum_{k=2}^{\infty} E|X|^{p+((1+\alpha+\beta)/r)-(p-\theta)}I(k\leq |X|^{1/r}<k+1)\\
&\leq CE|X|^{\theta+((1+\alpha+\beta)/r)}=CE|X|^{s}<\infty. \qquad (۳۵.۳)
\end{aligned}$$

بنابراین نابرابری (۲۵.۳) از معادله‌های (۲۶.۳) تا (۳۰.۳)، (۳۲.۳)، (۳۴.۳) و (۳۵.۳) نتیجه می‌شود. نتیجه (۱۵.۳) نیز بلافاصله از معادله‌های (۲۰.۳)، (۲۱.۳) و (۲۵.۳) به دست می‌آید.

مورد دوم. $s < ۱$: فرض می‌کنیم $p > ۰$ به‌طوری که $\theta + (۱+\alpha+\beta)/r = s < p < ۱$ باشد. پس $\alpha + \beta - r(p-\theta) < -۱$ است. با استفاده از نابرابری مارکف و نابرابری C_r داریم

$$
\begin{aligned}
J &= \sum_{n=۱}^{\infty} n^{\beta} P\left(\max_{۱\le j\le n}\Big|\sum_{i=۱}^{j} X'_{ni}\Big| > \varepsilon\right) \\
&\le \sum_{n=۱}^{\infty} n^{\beta} E\left(\max_{۱\le j\le n}\Big|\sum_{i=۱}^{j} X'_{ni}\Big|\right)^{P} \\
&= \sum_{n=۱}^{\infty} n^{\beta} \sum_{i=۱}^{n} E|X'_{ni}|^{P}. \qquad (۳۶.۳)
\end{aligned}
$$

مابقی برهان شبیه به فرآیند برهان $J_۲ < \infty$ است. بنابراین از بیان جزییات پرهیز می‌کنیم.

(۳) اگر $۱+\alpha+\beta = ۰$ باشد، آن‌گاه با استفاده از نابرابری مارکف و شبیه به فرآیند معادله (۲۱.۳) می‌توان نوشت

$$
\begin{aligned}
I &= \sum_{n=۱}^{\infty} n^{\beta} \sum_{i=۱}^{n} P(|a_{ni}X_{ni}| > ۱) \\
&\le C\sum_{k=۱}^{\infty}\sum_{n=۱}^{k} n^{\alpha+\beta} E|X|^{\theta} I(k^r \le |X| < (k+۱)^r) \\
&\le C\sum_{k=۱}^{\infty}\sum_{n=۱}^{k} n^{-۱} E|X|^{\theta} I(k^r \le |X| < (k+۱)^r) \\
&\le C\sum_{k=۱}^{\infty} log k E|X|^{\theta} I(k^r \le |X| < (k+۱)^r) \\
&\le C\sum_{k=۱}^{\infty} E|X|^{\theta} log|X| I(k^r \le |X| < (k+۱)^r).
\end{aligned}
$$

با استفاده از رابطه (۱۸.۳) نتیجه می‌شود

$$
I \le CE|X|^{\theta} log|X| < \infty. \qquad (۳۷.۳)
$$

همچنین برای اثبات معادله (۱۵.۳)، فقط کافی است نشان دهیم $J < \infty$ است. و هم‌چنان موارد $s \ge ۱$ و $s < ۱$ را در نظر می‌گیریم. در اینجا $s = \theta$.

مورد اول. $s \ge ۱$: از آن‌جایی‌که $۱+\alpha+\beta = ۰$، $۱+\beta = -\alpha \ge ۰$ و $s = \theta$ است، در نتیجه معادله (۲۳.۳) برقرار است. بنابراین، کافی است نشان دهیم $J^* < \infty$ است. با استفاده از نابرابری مارکف،

نابرابری C_r، نابرابری ماکسیمال روزنتال و معادله (۳۷.۳)، داریم

$$\begin{aligned} J^* &= \sum_{n=۱}^{\infty} n^{\beta} P\left(\max_{۱\leq j\leq n}\Big|\sum_{i=۱}^{j}(X'_{ni}-EX'_{ni})\Big| > \frac{\varepsilon}{۲}\right) \\ &\leq \sum_{n=۱}^{\infty} n^{\beta} P\left(\max_{۱\leq j\leq n}\Big|\sum_{i=۱}^{j}(X'_{ni}-EX'_{ni})\Big|^{۲}\right) \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} E|X'_{ni}|^{۲}\left(I(|a_{ni}X|>۱)+I(|a_{ni}X|\leq ۱)\right) \\ &\leq C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} P(|a_{ni}X|>۱)+C\sum_{n=۱}^{\infty} n^{\beta}\sum_{i=۱}^{n} E|a_{ni}X|^{۲} I(|a_{ni}X|\leq ۱). \end{aligned}$$

جمله اول با توجه به رابطه (۳۷.۳) متناهی است و برای جمله دوم داریم

$$J^* \leq C+J_{۵}^*+J_{۶}^*. \tag{۳۸.۳}$$

در این‌جا، $J_{۵}^*$ و $J_{۶}^*$ همان $J_{۵}$ و $J_{۶}$ در برهان قسمت ۲ به ازای $p=۲$ است. لازم است توجه کنیم که $\alpha+\beta=-۱$ و $\alpha+\beta-r(۲-\theta)<-۱$ می‌باشد. همانند اثبات $J_{۵}<\infty$ داریم

$$J_{۵}^*\leq C+CE|X|^{\theta}<\infty \tag{۳۹.۳}$$

و شبیه به اثبات $J_{۶}<\infty$ داریم

$$\begin{aligned} J_{۶}^* &\leq C\sum_{k=۲}^{\infty}\sum_{n=۱}^{[k/۲]} n^{-۱}.k^{-r(۲-\theta)}E|X|^{۲}I(k\leq |X|^{۱/r}<k+۱) \\ &\leq C\sum_{k=۲}^{\infty}\log k.k^{-۱}.k^{-r(۲-\theta)}E|X|^{۲}I(k\leq |X|^{۱/r}<k+۱) \\ &\leq CE|X|^{\theta}log|X|<\infty. \end{aligned} \tag{۴۰.۳}$$

بنابراین، $J^*<\infty$ سریعاً از معادله‌های (۳۸.۳) تا (۴۰.۳) نتیجه می‌شود.

مورد دوم. $s<۱$: روند برهان همانند مورد ۲ در برهان قسمت ۲ می‌باشد. لذا ما فقط باید نشان دهیم که $J_{۵}<\infty$ و $J_{۶}<\infty$ هستند. در واقع مانند اثبات معادله (۳۴.۳) داریم

$$J_{۵}\leq C+CE|X|^{\theta}<\infty \tag{۴۱.۳}$$

و مانند اثبات معادله (۳۵.۳) داریم

$$J_{۶} \leq CE|X|^{\theta}log|X|<\infty \tag{۴۲.۳}$$

برهان قضیه کامل می‌شود. □

با استفاده از این قضیه می‌توان نتایج بوم و کاتز (۱۹۶۵) برای متغیرهای تصادفی مستقل و هم‌توزیع را به آرایه‌های سطری متغیرهای تصادفی NSD تعمیم داد.

نتیجه ۹.۲.۳. فرض کنید متغیر تصادفی X با $EX_{ni}=0$ برای هر $i\geq 1$، $n\geq 1$ بر روی $\{X_{ni}, i\geq 1, n\geq 1\}$ که یک آرایه سطری از متغیرهای تصادفی NSD است، غلبه تصادفی دارد.

(i) فرض کنید $\gamma>1$ و $1\leq t<2$ است. اگر $E|X|^{\gamma t}<\infty$ باشد، آن‌گاه برای هر $\varepsilon>0$ داریم

$$\sum_{n=1}^{\infty} n^{\gamma-2} P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} X_{ni}\right|>\varepsilon n^{1/t}\right)<\infty. \tag{۴۳.۳}$$

(ii) اگر $E|X|log|X|<\infty$ باشد، آن‌گاه برای هر $\varepsilon>0$ داریم:

$$\sum_{n=1}^{\infty} \frac{1}{n} P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} X_{ni}\right|>\varepsilon n\right)<\infty. \tag{۴۴.۳}$$

برهان.

(i) فرض کنید اگر $n<i$، $a_{ni}=0$ باشد و اگر $i\leq n$، $a_{ni}=n^{-1/t}$ است. همچنین شرایط (۱۳.۳) و (۱۴.۳) برای $\theta=1$، $r=1/t$ برقرار است و $\alpha=1-1/t<r$، $\beta=\gamma-2>-1$ باشد. به آسانی می‌بینیم که

$$1+\alpha+\beta=1+1-\frac{1}{t}+\gamma-2=\gamma-\frac{1}{t}>0$$

و به دلیل این‌که $\gamma>1$ است و $\frac{1}{2}<\frac{1}{t}\leq 1$ است، بنابراین عبارت بالا مثبت است. همچنین

$$1+\frac{1+\alpha+\beta}{r}=1+\frac{\gamma-\frac{1}{t}}{\frac{1}{t}}=1+\gamma t-1=\gamma t=s.$$

در این قسمت نیز چون $\gamma>1$ و $t\geq 1$ است بنابراین $s\geq 1$ می‌باشد و نتیجه مطلوب (۴۳.۳) بلافاصله از قسمت ۲ قضیه ۸.۲.۳ به دست می‌آید.

(ii) فرض کنید اگر $n<i$، $a_{ni}=0$ باشد و اگر $i\leq n$، $a_{ni}=n^{-1}$ باشد. همچنین شرایط (۱۳.۳) و (۱۴.۳) برای $r=-1$، $\theta=1$، $\alpha=0$، $\beta=-1$ برقرار است، داریم

$$1+\alpha+\beta=1+(-1)=0, \qquad 1+\frac{1+\alpha+\beta}{r}=1+\frac{1-1}{-1}=1=s\geq 1.$$

بنابراین نتیجه مطلوب (۴۴.۳) بلافاصله از قسمت ۳ قضیه ۸.۲.۳ به‌دست می‌آید و برهان نتیجه کامل می‌شود.

□

فصل ۴

رگرسیون ناپارامتری با جمله خطای وابسته زبرجمعی منفی

۱.۴ مقدمه

یکی از مهمترین روش‌های مدل‌بندی آماری، تحلیل رگرسیونی است که هدف اصلی آن بررسی نحوه تأثیرگذاری یک یا چند متغیر تبیینی (توضیحی یا پیش‌بین) بر یک متغیر پاسخ است. مدل‌های رگرسیونی را از لحاظ نوع تابعی که رابطه بین متغیرهای تبیینی و پاسخ را تعیین می‌کند، می‌توان به دو دسته رگرسیون پارامتری و ناپارامتری تقسیم کرد.

در رگرسیون پارامتری، رابطه بین متغیرهای تبیینی و پاسخ از طریق تابع پارامتری f به صورت زیر مشخص می‌شود:

$$y_i = f(\boldsymbol{\beta}', \mathbf{x}_i') + \varepsilon_i, \qquad ۱ \leq i \leq n \tag{۱.۴}$$

که در آن $\mathbf{x}_i' = (x_{i۱}, \ldots, x_{ip})$، $\boldsymbol{\beta}' = (\beta_۱, \ldots, \beta_p)$ برداری از پارامترهای مجهول است و ε_i جمله خطا است که یک متغیر تصادفی با میانگین صفر در نظر گرفته می‌شود. در صورت برآورد $\boldsymbol{\beta}$، تابع پارامتری f و بنابراین مدل به صورت $\hat{y} = f(\hat{\boldsymbol{\beta}}', \mathbf{x}')$ برآورد می‌شود.

اما در رگرسیون ناپارامتری، رابطه بین متغیرهای تبیینی و پاسخ از طریق تابع ناپارامتری $g(\cdot)$ به صورت زیر تعیین می‌شود:

$$y_i = g(\mathbf{x}_i') + \varepsilon_i, \qquad ۱ \leq i \leq n$$

که در آن تابع نامعلوم $g(\cdot)$ در اکثر روش‌های ناپارامتری، تابعی پیوسته و مشتق‌پذیر در نظر گرفته می‌شود.

در ابتدای این فصل ما به معرفی رگرسیون ناپارامتری می‌پردازیم. برای فهم بهتر مطالب مربوط به رگرسیون ناپارامتری، به یادآوری مختصری از آنچه در مورد رگرسیون پارامتری آموخته‌ایم، پرداخته می‌شود. برای مطالعه بیش‌تر می‌توانید به ترابی و بقایی‌پور (۱۳۸۹) مراجعه کنید.

۲.۴ رگرسیون پارامتری

رگرسیون پارامتری را می‌توان به دو دسته رگرسیون پارامتری تک‌متغیره و چندگانه[۱] تقسیم‌بندی کرد.

۱) تک‌متغیره: در این مدل، رابطه تنها یک متغیر تبیینی با متغیر پاسخ سنجیده می‌شود. یعنی

$$y_i = f(\boldsymbol{\beta}', x_i) + \varepsilon_i, \qquad 1 \leq i \leq n.$$

به عنوان مثال می‌توان مدل رگرسیون چندجمله‌ای را مثال زد که در آن

$$f(x_i) = \beta_0 + \beta_1 x_i + \cdots + \beta_q x_i^q, \qquad 1 \leq i \leq n.$$

۲) چندگانه: در این مدل، رابطه بیش از یک متغیر تبیینی، مثلاً pتا، با متغیر پاسخ بررسی می‌شود. بنابراین می‌توان نوشت

$$y_i = f(\boldsymbol{\beta}', x_{i1}, \ldots, x_{ip}) + \varepsilon_i, \qquad 1 \leq i \leq n.$$

به عنوان مثال رگرسیون خطی چندگانه را می‌توان نام برد که در آن

$$f(\boldsymbol{\beta}', x_{i1}, \ldots, x_{ip}) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}, \qquad 1 \leq i \leq n.$$

برای سادگی کار در این قسمت تنها مدل رگرسیونی تک‌متغیره چندجمله‌ای، یعنی مدلی با یک متغیر تبیینی x که با یک متغیر پاسخ y در ارتباط است را مورد بررسی قرار می‌دهیم. داده‌ها به صورت (x_i, y_i)، $1 \leq i \leq n$، هستند و مدل به صورت

$$y_i = \sum_{j=0}^{q} \beta_j x_i^j + \varepsilon_i \qquad 1 \leq i \leq n$$

نوشته می‌شود. یک راه برای برآورد پارامترهای نامعلوم $(\beta_0, \ldots, \beta_q)$ استفاده از روش کمترین توان‌های دوم خطا از طریق می‌نیمم کردن رابطه زیر است:

$$Q = \sum_{i=1}^{n} \left(y_i - \sum_{j=0}^{q} \beta_j x_i^j \right)^2.$$

با به‌کارگیری نمادهای ماتریسی که در ادامه معرفی می‌شوند، مدل رگرسیونی چندجمله‌ای[۲] به صورت زیر بازنویسی می‌شود:

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

[۱] Multiple regression

[۲] Polynomial regression

که در آن

$$X=\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^q \\ 1 & x_2 & x_2^2 & \dots & x_2^q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^q \end{pmatrix}$$

$$\mathbf{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \boldsymbol{\beta}=\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_q \end{pmatrix} \qquad \boldsymbol{\varepsilon}=\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

با مشتق‌گیری از Q برحسب β_jها و به‌کارگیری نمادهای ماتریسی، معادلات نرمال به صورت زیر به دست می‌آیند:

$$\frac{\partial Q}{\partial \beta}=\begin{pmatrix} -\sum_{i=1}^{n}(y_j-\beta_0+\dots+\beta_q x_i^q) \\ \vdots \\ -\sum_{i=1}^{n} x_i^j(y_i-\beta_0+\dots+\beta_q x_i^q) \\ \vdots \\ -\sum_{i=1}^{n} x_i^q(y_i-\beta_0+\dots+\beta_q x_i^q) \end{pmatrix}=X'(\mathbf{y}-X\beta)=0$$

و داریم

$$X'X\boldsymbol{\beta}=X'\mathbf{y}.$$

بنابراین برآورد پارامترهای نامعلوم مدل به صورت یکتا، از رابطه زیر نتیجه می‌شود:

$$\hat{\boldsymbol{\beta}}=(X'X)^{-1}X'\mathbf{y}$$

و در آخر برآورد پاسخ به صورت زیر است:

$$\begin{aligned} \hat{\mathbf{y}} &= X\hat{\boldsymbol{\beta}} \\ &= X(X'X)^{-1}X'\mathbf{y} \\ &= H\mathbf{y}. \end{aligned}$$

به‌طوری که $\hat{\mathbf{y}}=(\hat{y}_1,\dots,\hat{y}_n)$. ماتریس $n\times n$، $H=X(X'X)^{-1}X'$، که در واقع ماتریس تصویر[۳] روی فضای تولیدشده توسط ستون‌های ماتریس X است، ماتریس کلاه[۴] نام دارد که ماتریس برازش نیز نامیده می‌شود.

[۳] Mapping matrix

[۴] Hat matrix

۳.۴ رگرسیون ناپارامتری

هدف از رگرسیون ناپارامتری، برآورد تابع رگرسیون به صورت مستقیم است. رگرسیون ناپارامتری اثر یک یا چند متغیر تبیینی را روی متغیر پاسخ بررسی می‌کند، بدون این‌که از قبل تابع ویژه‌ای را برای ارتباط بین آن‌ها در نظر گرفته باشد. رگرسیون ناپارامتری را می‌توان از لحاظ تعداد متغیرهای تبیینی به دو گروه زیر تقسیم‌بندی کرد:

۱) با یک متغیر تبیینی (رگرسیون ناپارامتری ساده): در این حالت، مدل رگرسیونی به صورت

$$y_i = g(x_i) + \varepsilon_i$$

در نظر گرفته می‌شود.

۲) با k متغیر تبیینی (رگرسیون ناپارامتری جمعی[۵]): در این حالت می‌توان از یک مدل جمعی به صورت

$$y_i = g_۰ + g_۱(x_{i۱}) + g_۲(x_{i۲}) + \cdots + g_k(x_{ik}) + \varepsilon_i$$

استفاده کرد، به طوری که به $g_j(\cdot)$ها، $۱ \leq j \leq k$، توابع رگرسیونی جزئی[۶] گفته می‌شود.

رگرسیون ناپارامتری ساده یکی از ساده‌ترین مدل‌های رگرسیون ناپارامتری است که در آن متغیر تبیینی به صورت تک‌متغیره در نظر گرفته می‌شود. در این مدل، داده‌ها به صورت (x_i, y_i)، $۱ \leq i \leq n$، هستند که غالباً x_iها به صورت صعودی مرتب شده‌اند و

$$g(x_i) = E(Y_i|X_i = x_i).$$

با توجه به این که در این روش، نمودار پراکنش y_i برحسب x_i کاربرد زیادی دارد، در برخی از مراجع نام هموارسازی پراکنش[۷] نیز برای این روش در نظر گرفته شده است (کلولند، ۱۹۷۹). توجه کنید که در انواع روش‌های برازش یک نمودار به داده‌ها، منظور از هموارسازی، برازش یک نمودار مشتق‌پذیر به داده‌ها است. در معادله رگرسیون ناپارامتری ساده $g(x)$ تابعی است که رفتار ذاتی داده‌ها را ارائه می‌دهد و به تصادف یا شانس وابسته نیست. مقدار برآوردشده $g(x)$ را با $\hat{g}(x)$ نمایش می‌دهیم که تابعی هموار از x است.

همان‌طور که در رگرسیون پارامتری گفتیم، همیشه برازش چندجمله‌ای به روش کمترین توان‌های دوم مناسب نیست زیرا در این روش تنها یک چندجمله‌ای به ناحیه‌ای که داده‌ها در درون آن قرار گرفته‌اند، برازش داده می‌شود. برازش تنها یک معادله به تعداد زیادی از داده‌ها با رفتار پیچیده گاهی به نتایج نامطلوبی منجر می‌شود (هاردل و همکاران، ۱۹۴۷).

این مشکلات به مفهوم افراز کردن داده‌ها به چند ناحیه برای برازش منحنی به داده‌ها در هر ناحیه و عطف آن‌ها به طوری که برآوردگرهای مناسب و مفیدی باشند، منتهی می‌شود. در واقع ساده‌ترین روش، برازش

[۵] Additive nonparametric regression

[۶] Partial regression

[۷] Scatterplot smoothing

معادله یک منحنی به صورت پاره‌ای[۸] (تکه‌ای) است به این شرط که برآوردها در انتهای هر زیرناحیه مقادیر یکسانی را اختیار کنند. انواع روش‌هایی که در رگرسیون ناپارامتری ساده برای برآورد $g(\cdot)$ وجود دارند، عبارتند از

۱) نادارایا-واتسون[۹] (برآوردگر هسته[۱۰])

۲) چندجمله‌ای موضعی[۱۱] (LOESS)

۳) هموارساز اسپلاین[۱۲]

۴) هموارسازی پراکنش موضعی[۱۳]

در ادامه تنها با دو روش اول آشنا می‌شویم. برای آشنایی بیش‌تر در مورد رگرسیون ناپارامتری می‌توان به فاکس (۲۰۰۰) ، سیمنوف (۱۹۹۶) و تاکزاوا (۲۰۰۶) مراجعه کرد.

۱.۳.۴ برآوردگر هسته نادارایا-واتسون

مدل رگرسیونی مبتنی بر برآوردگر هسته از میانگین وزنی داده‌ها به‌طور مستقیم برای برآورد $g(\cdot)$ بهره می‌گیرد. فرض کنید متغیرهای تبیینی و پاسخ، هر دو پیوسته باشند. می‌دانیم که هرگاه $f_X(x)$ تابع چگالی احتمال X باشد. آن‌گاه برای هر Δx کوچک

$$\begin{aligned} P(x < X < x + \Delta x) &= \int_x^{x+\Delta x} f_X(t)dt = \frac{F_X(x+\Delta x) - F_X(x)}{\Delta x}\Delta x \\ &\approx f_X(x)\Delta x \end{aligned} \qquad (۲.۴)$$

که علامت $\approx$ به معنای حدوداً برابر می‌باشد.

به همین ترتیب هرگاه $f(x,y)$ تابع چگالی توأم X و Y باشد، آن‌گاه برای هر Δx و Δy کوچک با استدلال مشابه می‌توان نوشت

$$P(x < X < x + \Delta x, y < Y < y + \Delta y) \approx f(x,y)\Delta x \Delta y. \qquad (۳.۴)$$

همچنین هرگاه $f(y|x)$ تابع چگالی احتمال شرطی Y به شرط $X = x$ باشد، آن‌گاه

$$P(y < Y < y + \Delta y | x < X < x + \Delta x) \approx f(y|x)\Delta y.$$

[۸] Piecewise
[۹] Nadaraya-Watson regression
[۱۰] Kernel estimator
[۱۱] Local polynomial regression
[۱۲] Spline smoothers
[۱۳] Locally Scatterplot Smoothing

با توجه به این نکته که مقدار تابع $g(x)$ برای مقدار x هم‌ارز با امید ریاضی متغیر Y به شرط $X = x$ است، نتیجه می‌شود:

$$g(x) = E(Y|X = x) = \int_{-\infty}^{\infty} yf(y|x)dy = \frac{۱}{f_X(x)} \int_{-\infty}^{\infty} yf(x, y)dy. \qquad (۴.۴)$$

رابطه (۴.۴) نشان می‌دهد که برای برآورد تابع $g(x)$ کافی است تابع $f(y|x)$ یا در واقع توابع $f_X(x)$ و $f(x, y)$ را برآورد کنیم. یکی از ساده‌ترین روش‌ها برای برآورد این توابع به کارگیری چگالی هسته[۱۴] است. با استفاده از ایده روش گشتاورها و روابط (۲.۴) و (۳.۴)، برآوردگرهای توابع $f_X(x)$ و $f(x, y)$ به صورت زیر به دست می‌آیند:

$$\hat{f}(x, y) = \frac{۱}{nh_xh_y} \sum_{i=۱}^{n} k_x\left(\frac{x - x_i}{h_x}\right) k_y\left(\frac{y - y_i}{h_y}\right)$$

$$\hat{f}_X(x) = \frac{۱}{nh_x} \sum_{i=۱}^{n} k_x\left(\frac{x - x_i}{h_x}\right)$$

به طوری که $k_x(\cdot)$ و $k_y(\cdot)$ تابع‌هایی مشتق‌پذیر و پیوسته هستند و اگر در شرایط زیر صدق کنند آن را تابع هسته[۱۵] می‌نامند:

$$\int_{-\infty}^{\infty} k_x(u)du = \int_{-\infty}^{\infty} k_y(u)du = ۱ \qquad (۵.۴)$$

$$\int_{-\infty}^{\infty} uk_x(u)du = \int_{-\infty}^{\infty} uk_y(u)du = ۰ \qquad (۶.۴)$$

$$\int_{-\infty}^{\infty} u^۲k_x(u)du < \infty \qquad \int_{-\infty}^{\infty} u^۲k_y(u)du < \infty.$$

مقادیر ثابت و مثبت h_x و h_y به نام پهنای پنجره[۱۶]، پارامتر هموارسازی[۱۷] یا پهنای باند[۱۸] معروف هستند. به ازای کوچک کردن پهنای باند منحنی حاصل از برآورد هسته ناهموارتر شده و جزئیات جعلی بیشتری را از چگالی واقعی به نمایش می‌گذارد و به ازای بزرگ کردن آن منحنی هموار و باعث محو شدن جزئیات واقعی تابع چگالی می‌گردد. بنابراین انتخاب مناسب پهنای باند برای برآورد تابع مورد نظر یک جنبه کلیدی در این روش محسوب می‌شود. چند روش برای انتخاب پارامترهای هموارسازی وجود دارند:

۱. انتخاب ذهنی

[۱۴] Kernel density
[۱۵] Kernel functions
[۱۶] Window width
[۱۷] Smoothing parameter
[۱۸] Bandwidth

۲. معیار اعتبارسنجی متقابل[۱۹] (CV) (کوهاوی، ۱۹۹۵).

۳. معیار اعتبارسنجی متقابل تعمیم‌یافته[۲۰] (GCV) (گلوب و همکاران، ۱۹۷۹).

از معادلات بیان شده در روابط (۵.۴) و (۶.۴) می‌توان نتیجه گرفت

$$\frac{1}{h_x}\int_{-\infty}^{\infty} k_x\left(\frac{x-x_i}{h_x}\right)dx = \frac{1}{h_y}\int_{-\infty}^{\infty} k_y\left(\frac{y-y_i}{h_y}\right)dx = 1$$

$$\frac{1}{h_x}\int_{-\infty}^{\infty} k_x\left(\frac{x-x_i}{h_x}\right)dx = x_i \qquad \frac{1}{h_y}\int_{-\infty}^{\infty} k_y\left(\frac{y-y_i}{h_y}\right)dx = y_i. \tag{۷.۴}$$

به این ترتیب برای به دست آوردن $\hat{g}(x)$ کافی است برآوردهای $f_X(x)$ و $f(x,y)$ را در رابطه (۴.۴) جایگزین کنیم. بنابراین

$$\begin{aligned}\hat{g}(x) &= \int_{-\infty}^{\infty} y\frac{\hat{f}(x,y)}{\hat{f}_X(x)}dy = \int_{-\infty}^{\infty} y\frac{\frac{1}{nh_xh_y}\sum_{i=1}^{n} k_x\left(\frac{x-x_i}{h_x}\right)k_y\left(\frac{y-y_i}{h_y}\right)}{\frac{1}{nh_x}\sum_{i=1}^{n} k_x\left(\frac{x-x_i}{h_x}\right)}dy \\ &= \frac{\sum_{i=1}^{n} k_x\left(\frac{x-x_i}{h_x}\right)\int_{-\infty}^{\infty} yk_y\left(\frac{y-y_i}{h_y}\right)dy}{h_y\sum_{i=1}^{n} k_x\left(\frac{x-x_i}{h_x}\right)}\end{aligned} \tag{۸.۴}$$

با جایگزین کردن رابطه (۷.۴) در (۸.۴) و جایگزین کردن $k_x(\cdot)$ و h_x با $k(\cdot)$ و h به دست می‌آوریم

$$\hat{g}(x) = \sum_{i=1}^{n}\omega_i(x,h)y_i \tag{۹.۴}$$

به طوری که

$$\omega_i(x,h) = \frac{k(\frac{x-x_i}{h})}{\sum_{k=1}^{n} k(\frac{x-x_k}{h})}.$$

برآوردگری که برای برآورد $g(\cdot)$، در رابطه (۸.۴) معرفی شد، برآوردگر نادارایا-واتسون نام دارد. بیش‌تر روش‌های رگرسیونی ناپارامتری به صورت تعمیمی از برآوردگر ارائه‌شده توسط نادارایا-واتسون به دست می‌آیند (کای، ۲۰۰۱).

انواع توابع هسته

در این‌جا انواع توابع هسته متداول در برآورد تابع چگالی را معرفی می‌کنیم:

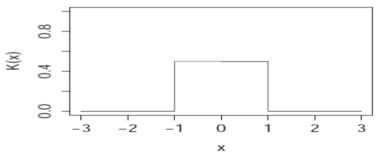
[۱۹] Cross validation

[۲۰] Generalized cross validation

۱) تابع هسته مستطیلی[۲۱] (یکنواخت): ضابطه این تابع هسته به صورت

$$k(x) = \frac{۱}{۲}I(|x| < ۱)$$

است. شکل ۱.۴ تابع هسته مستطیلی را نمایش می‌دهد.

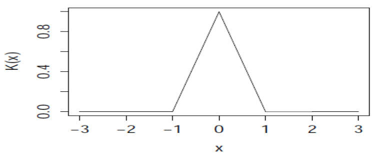


شکل ۱.۴: منحنی تابع هسته مستطیلی

۲) تابع هسته مثلثی[۲۲]: ضابطه این تابع هسته به صورت

$$k(x) = (۱ - |x|)I(|x| < ۱)$$

است و منحنی آن را در شکل ۲.۴ می‌توان مشاهده کرد.

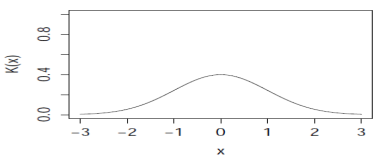


شکل ۲.۴: منحنی تابع هسته مثلثی

۳) تابع هسته گوسی[۲۳]: ضابطه این تابع هسته به صورت

$$k(x) = \frac{۱}{\sqrt{۲\pi}} exp - \frac{۱}{۲}x^{۲}$$

است و منحنی آن در شکل ۳.۴ قابل مشاهده است.



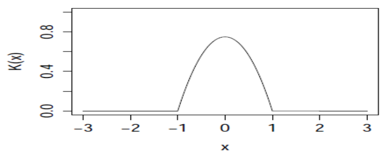
شکل ۳.۴: منحنی تابع هسته گوسی

[۲۱] Rectangular
[۲۲] Triangular
[۲۳] Gaussian

۴) تابع هسته اپانچنیکوف[۲۴]: یکی از متداول‌ترین توابع هسته، تابع اپانچنیکوف است که ضابطه ریاضی آن به صورت

$$k(x) = \frac{3}{4}(1 - x^2)I(|x| < 1)$$

است. منحنی این تابع هسته در شکل ۴.۴ مشهود است.

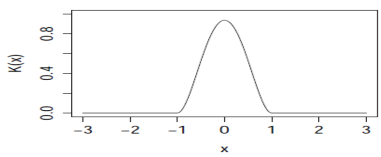


شکل ۴.۴: منحنی تابع هسته اپانچنیکوف

۵) تابع هسته توان چهارم[۲۵] (دو وزنی): ضابطه این تابع هسته به صورت

$$k(x) = \frac{15}{16}(1 - x^2)^2 I(|x| < 1)$$

است و منحنی آن در شکل ۵.۴ قابل مشاهده است.

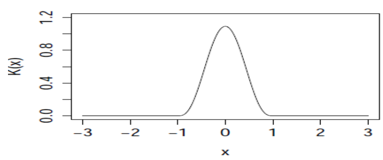


شکل ۵.۴: منحنی تابع هسته توان چهارم (دو وزنی)

۶) تابع هسته سه‌وزنی[۲۶]: ضابطه این تابع هسته به صورت

$$k(x) = \frac{35}{32}(1 - x^2)^3 I(|x| < 1)$$

است و منحنی این تابع را در شکل ۶.۴ می‌توان مشاهده کرد.



شکل ۶.۴: منحنی تابع هسته سه وزنی

[۲۴] Epanechnikov
[۲۵] Biguadratic
[۲۶] Tuesday Weight

۲.۳.۴ رگرسیون چندجمله‌ای موضعی

رگرسیون چندجمله‌ای موضعی یکی از انواع روش‌های رگرسیونی ناپارامتری است. در این روش نیز مانند اکثر روش‌های رگرسیون ناپارامتری، ناحیه‌ای را که داده‌ها در آن قرار گرفته‌اند به زیرناحیه‌هایی تقسیم می‌کنیم و سپس معادله خطی را به هر زیرناحیه برازش می‌دهیم به‌طوری که نمودارها در انتهای هر زیرناحیه پیوسته باشند. این روش تعمیمی از روش نادارایا–واتسون می‌باشد. معادله چندجمله‌ای موضعی وقتی مقادیر x به x^* نزدیک است به صورت زیر تعریف می‌شود:

$$g(x,x^*) = a_۰(x^*) + \sum_{j=۱}^{p} a_j(x^*)(x-x^*)^j \qquad (۱۰.۴)$$

که در آن x^* مرکز مرز و ناحیه را مشخص می‌کند.

ساده‌ترین و عمومی‌ترین حالت در این روش، حالتی است که $p = ۱$ در نظر گرفته شود؛ یعنی از روش رگرسیون خطی موضعی بهره بگیریم.

مقادیر $\{a_۰(x^*), a_۱(x^*), \ldots, a_p(x^*)\}$ از می‌نیمم کردن رابطه زیر به دست می‌آیند:

$$E_{local}(x^*) = \sum_{i=۱}^{n} \left(\omega(\frac{x_i - x^*}{h})(g(x_i,x^*) - y_i)^۲ \right) \qquad (۱۱.۴)$$

که در آن $g(x_i,x^*) = a_۰(x^*) + \sum_{j=۱}^{p} a_j(x^*)(x_i - x^*)^j$. در رابطه (۱۱.۴)، $\omega(\frac{x_i - x^*}{h})$ همان تابع هسته است که درجه اهمیت نقطه (x_i, y_i)، وقتی معادله رگرسیونی مربوط به x^* مد نظر است را معین می‌کند. این تابع وزن نامنفی، بیش‌ترین مقدار خود را در حالت $x_i = x^*$ می‌گیرد. انواع مختلف توابع وزن که در این روش می‌توان از آن‌ها بهره گرفت؛ عبارتند از

$$\omega(\frac{x_i - x^*}{h}) = \exp\left\{ -\frac{۱}{۲}(\frac{x_i - x^*}{h})^۲ \right\}$$

$$\omega(\frac{x_i - x^*}{h}) = \begin{cases} \left(۱ - (\frac{x_i - x^*}{h})^۲\right)^۲ & (\frac{x_i - x^*}{h})^۲ \leq ۱ \\ ۰ & (\frac{x_i - x^*}{h})^۲ > ۱ \end{cases}$$

$$\omega(\frac{x_i - x^*}{h}) = \begin{cases} \left(۱ - (\frac{x_i - x^*}{h})^۲\right)^۳ & (\frac{x_i - x^*}{h})^۲ \leq ۱ \\ ۰ & (\frac{x_i - x^*}{h})^۲ > ۱ \end{cases}$$

$$\omega(\frac{x_i - x^*}{h}) = \begin{cases} \left(۱ - (\frac{|x_i - x^*|}{h})^۳\right)^۳ & (\frac{|x_i - x^*|}{h})^۳ \leq ۱ \\ ۰ & (\frac{|x_i - x^*|}{h})^۳ > ۱ \end{cases}$$

که این توابع وزن به ترتیب گوسی[۲۷]، دو مربعی[۲۸]، سه وزنی[۲۹]، و سه مکعبی[۳۰] نامیده می‌شوند (لودر، ۱۹۹۹).
با بهره‌گیری از نمادهای ماتریسی، $E_{local}(x^*)$ را به صورت ساده‌تر زیر می‌توان بازنویسی کرد:

$$E_{local}(x^*) = (Xa - \mathbf{y})' W (Xa - \mathbf{y}) \tag{۱۲.۴}$$

که در آن ماتریس طرح X، ماتریس وزن W و بردارهای y و a به صورت زیر تعریف می‌شوند:

$$X = \begin{pmatrix} ۱ & (x_۱ - x^*) & \dots & (x_۱ - x^*)^p \\ ۱ & (x_۲ - x^*) & \dots & (x_۲ - x^*)^p \\ \vdots & \vdots & \ddots & \vdots \\ ۱ & (x_n - x^*) & \dots & (x_n - x^*)^p \end{pmatrix} \qquad \mathbf{a} = \begin{pmatrix} a_۰(x^*) \\ \vdots \\ a_p(x^*) \end{pmatrix}$$

$$W = \begin{pmatrix} \omega(\frac{x_۱ - x^*}{h}) & ۰ & \dots & ۰ \\ ۰ & \omega(\frac{x_۲ - x^*}{h}) & \dots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \dots & \omega(\frac{x_n - x^*}{h}) \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} y_۱ \\ \vdots \\ y_n \end{pmatrix}$$

به کمک رابطه (۱۲.۴) برآوردگر a از رابطه $\hat{a} = (X'WX)^{-۱} X'W\mathbf{y}$ به دست می‌آید.

باید توجه داشته باشیم که ماتریس X که در این روش به کار می‌بریم، تابعی از x^* است. در واقع در رگرسیون چندجمله‌ای، معادله رگرسیونی به کل داده‌هایی که در مدل وجود دارند برازش داده می‌شود، در حالی که در رگرسیون چندجمله‌ای موضعی، معادله رگرسیونی را به داده‌هایی که در همسایگی x^* وجود دارند، برازش می‌دهیم. با توجه به رابطه (۱۰.۴) مقدار $\hat{g}(x^*, x^*)$ برابر با $\hat{a}_۰(x^*)$ است که با به کارگیری صورت ماتریسی، می‌توانیم به صورت $\hat{g}(x^*, x^*) = e'_۱ \left((X'WX)^{-۱} X'W\mathbf{y}\right)$ بنویسیم، که در آن $e_۱$ بردار ستونی $(۱, ۰, ۰, \dots, ۰)'$ با $p + ۱$ عضو است. صورت دیگر نمایش $\hat{g}(x^*, x^*)$ به صورت زیر است:

$$\hat{g}(x^*, x^*) = q(x^*)'\mathbf{y} = \sum_{j=۱}^{n} q_j(x^*)\mathbf{y}_j \tag{۱۳.۴}$$

که در آن $q(x^*)$ از رابطه زیر پیروی می‌کند:

$$q'(x^*) = e'_۱ (X'WX)^{-۱} X'W. \tag{۱۴.۴}$$

حال دو طرف رابطه (۱۴.۴) را از سمت راست بردار در X ضرب می‌کنیم و نتیجه می‌شود

$$q'(x^*)X = e'_۱ (X'WX)^{-۱} X'WX = e'_۱.$$

با توجه به این‌که $x_{i۱}$ ها، $۱ \leq i \leq n$، همگی برابر با یک هستند، رابطه زیر برقرار است:

$$\sum_{i=۱}^{n} q_i(x^*) = ۱. \tag{۱۵.۴}$$

---

[۲۷] Gaussian
[۲۸] Bisquare
[۲۹] Triweight
[۳۰] Tricube

با توجه به روابط (۱۳.۴) و (۱۵.۴) نتیجه می‌گیریم که برآورد حاصل از روش رگرسیون چندجمله‌ای موضعی، همان میانگین وزنی y_jها است. اگر $x^* = x_k$ باشد، بنابراین

$$\hat{y}_k = \hat{g}(x_k, x_k) = \sum_{j=۱}^{n} q_j(X_k) y_j \qquad (۱۶.۴)$$

که در آن $q_j(x_k)$، k_jامین عضو ماتریس کلاه است. در این حالت kامین ستون ماتریس X' مانند بردار $e'_۱$ است و می‌توانیم kامین عضو روی قطر ماتریس کلاه را به صورت زیر تعریف کنیم:

$$h_{kk} = q_k(X_k) = \left[(X'WX)^{-۱}\right]_{۱۱} w(\circ).$$

با توجه به اعضای ماتریس کلاه، در روش رگرسیون چندجمله‌ای موضعی، مقادیر CV و GCV را برای تعیین پهنای باند بهینه به کار می‌بریم.

معیارهای ارزیابی برازش

از آن‌جایی که در هر روش برازش مدل معیاری برای ارزیابی درستی برازش لازم است، در این بخش دو معیار میانگین توان دوم خطا[۳۱] (MSE) و میانگین توان دوم خطای تجمعی[۳۲] (MISE) را معرفی می‌کنیم:

میانگین توان دوم خطا: زمانی که برآورد در یک تک نقطه مورد نظر باشد، معیار طبیعی اختلاف میانگین توان دوم خطا است که به صورت زیر تعریف می‌شود:

$$MSE_x(\hat{f}(x)) = E(\hat{f}(x) - f(x))^۲.$$

میانگین توان دوم خطای تجمعی: برای ارزیابی اختلاف $\hat{f}$ و f در کل مجموعه مقادیر x، معیار مناسب MISE می‌باشد که عبارت است از

$$MISE_x(\hat{f}(x)) = \int_{-\infty}^{\infty} E(\hat{f}(x) - f(x))^۲ dx.$$

۴.۴ کاربرد نتایج نظری در رگرسیون ناپارامتری

گئورجیف (۱۹۸۵) مدل $y_i = g(x_i) + \varepsilon_i$ را در نظر گرفت که در آن $g(\cdot)$ یک تابع کران‌دار روی بازه $[\circ, ۱]$ است و x_iها غیرتصادفی و به صورت $\circ = x_\circ < x_۱ < \cdots < x_n \leq ۱$ می‌باشند. همچنین ε_i ها متغیرهای تصادفی مستقل دارای میانگین صفر و واریانس کوچک‌تر از $\sigma^۲$ فرض شده است. او برآورد $g(x)$ را به صورت زیر معرفی کرد:

$$g_n(x) = \sum_{i=۱}^{n} W_{ni}(x) y_i.$$

این برآورد توسط گئورجیف و گربلیکی (۱۹۸۶) برای بعد $p = ۱$ مورد بحث قرار گرفته است و همچنین مدل رگرسیونی برای بعد $p \geq ۱$ توسط گئورجیف (۱۹۸۸) به صورت زیر معرفی شد:

[۳۱] Mean squared error

[۳۲] Mean integral squared error

فرض کنید p یک عدد طبیعی و A یک مجموعه فشرده در $\Re^p$ باشد. گئورجیف (۱۹۸۸) مشاهدات پاسخ را از مدل زیر در نظر گرفت:

$$y_i^{(n)} = g(x_i^{(n)}) + \varepsilon_i^{(n)} \qquad 1 \leq i \leq n$$

که در آن $x_1^{(n)}, \ldots, x_n^{(n)} \in A$ نقاط طرح رگرسیون هستند، $g(\cdot)$ یک تابع حقیقی‌مقدار داده‌شده کران‌دار روی A و خطاهای تصادفی $\varepsilon_1^{(n)}, \ldots, \varepsilon_n^{(n)}$ مستقل ولی نه لزوماً هم‌توزیع با $E(\varepsilon_i^{(n)}) = 0$، $1 \leq i \leq n$، یا متقارن حول صفر فرض شده‌اند. برای جزئیات بیش‌تر در مورد ویژگی‌های برآوردگر بالا می‌توان به میلر (۱۹۸۷)، روساس (۱۹۸۹)، فن (۱۹۹۰)، روساس و همکاران (۱۹۹۲)، ترن و همکاران (۱۹۹۶)، هو و همکاران (۲۰۰۲)، لیانگ و جینگ (۲۰۰۵) و یانگ و همکاران (۲۰۱۲) مراجعه کرد.

همگرایی کامل برآوردگر ناپارامتری تابع رگرسیونی

در این قسمت می‌خواهیم در مورد سازگاری کامل برآوردگر تابع رگرسیون ناپارامتری در حضور مولفه خطای NSD با استفاده از همگرایی کامل، تحقیق کنیم. مدل رگرسیون ناپارامتری زیر را در نظر بگیرید:

$$y_{ni} = g(x_{ni}) + \varepsilon_{ni}, \qquad 1 \leq i \leq n \tag{۱۷.۴}$$

که در آن x_{ni} نقاط طرح شناخته‌شده از A هستند، $A \subset \Re^p$ یک مجموعه فشرده و پیوسته از $p \geq 1$ می‌دهد، $g(\cdot)$ یک تابع رگرسیون ناشناخته است که روی A تعریف شده است و ε_{ni} خطاهای تصادفی NSD هستند. همچنین یک برآوردگر برای $g(\cdot)$، برآوردگر رگرسیون وزنی است که به صورت زیر در نظر می‌گیریم:

$$g_n(x) = \sum_{i=1}^{n} W_{ni}(x) y_{ni}, \qquad x \in A \subset \Re^p \tag{۱۸.۴}$$

که در آن $W_{ni}(x) = W_{ni}(x; x_{n1}, x_{n2}, ..., x_{nn})$ برای $1 \leq i \leq n$، تابع وزن است. تابع وزن توسط پریستلی و چائو (۱۹۷۲) به صورت زیر برآورد شده است:

$$\hat{W}_{ni}(x) = \begin{cases} \dfrac{x_{i+1} - x_i}{h_n} k\left(\dfrac{x - x_i}{h_n}\right) & 1 \leq i \leq n-1 \\ 0 & \text{سایر جاها} \end{cases}$$

که در آن $k(\cdot)$ یک تابع هسته است. یک برآورد دیگر توسط گاسر و میلر (۱۹۷۹) و چنگ و لین (۱۹۸۱) به صورت زیر پیشنهاد شده است:

$$\tilde{W}_{ni}(x) = \frac{k\left(\dfrac{x - x_i}{h_n}\right)}{\displaystyle\sum_{i=1}^{n} k\left(\dfrac{x - x_i}{h_n}\right)}.$$

تحقیقات ما بیشتر در مورد همگرایی کامل برآوردگر رگرسیون $g_n(x)$ بر اساس خطاهای NSD می‌باشد.

برای هر تابع $g(x)$ از نماد $c(g)$ به منظور مشخص کردن نقاط پیوستگی تابع g روی A استفاده می‌کنیم. منظور از $\|\cdot\|$ نیز نرم اقلیدسی است. برای هر نقطه $x \in A$ تابع وزن برآوردگر $g(\cdot)$ یعنی $W_{ni}(x)$ باید در شرایط زیر صدق کند:

(A) زمانی که $n \to \infty$، $\sum_{i=1}^{n} W_{ni}(x) \to 1$.

(B) برای هر n، $\sum_{i=1}^{n} |W_{ni}(x)| \le C < \infty$.

(C) برای هر $a > 0$، زمانی که $n \to \infty$، $\sum_{i=1}^{n} |W_{ni}(x)|.|g(x_{ni} - g(x)|I(\|x_{ni} - x\| > a) \to 0$.

بر اساس فرضیات بالا، سازگاری کامل برآوردگر رگرسیونی $g_n(x)$ را در قضیه ۱.۴.۴ مطرح می‌کنیم.

قضیه ۱.۴.۴. فرض کنید $\{\varepsilon_{ni}, i \ge 1, n \ge 1\}$ یک آرایه سطری از متغیرهای تصادفی *NSD* با میانگین صفر باشد به طوری که توسط یک متغیر تصادفی X غلبه تصادفی شده است و همچنین برای هر $p \ge 1$، $E|X|^{2p} < \infty$ باشد. فرض می‌کنیم شرایط $(A) - (C)$ برقرار هستند. اگر

$$\sup_{i \ge 1} |W_{ni}(x)| = O(n^{-1/p})$$

آن‌گاه برای هر $x \in c(g)$ هرگاه $n \to \infty$ داریم

$$g_n(x) \xrightarrow{c.} g(x). \tag{۱۹.۴}$$

برهان. با استفاده از تعریف همگرایی کامل به ازای هر $\varepsilon > 0$ و $x \in c(g)$، باید نشان دهیم

$$\sum_{n=1}^{\infty} P\left(|g_n(x) - g(x)| > \varepsilon\right) < \infty.$$

با اضافه و کم کردن $E(g_n(x))$ می‌توان نوشت

$$\sum_{n=1}^{\infty} P\left(|g_n(x) - g(x)| > \varepsilon\right) = \sum_{n=1}^{\infty} P\left(|g_n(x) - E(g_n(x)) + E(g_n(x)) - g(x)| > \varepsilon\right).$$

در این‌جا چون دو مقدار نامنفی داریم که می‌خواهیم بزرگ‌تر از ε باشند، باید حداقل یکی از آن‌ها بزرگ‌تر از $\frac{\varepsilon}{2}$ باشد. بنابراین

$$\begin{aligned}\sum_{n=1}^{\infty} P\left(|g_n(x) - g(x)| > \varepsilon\right) \le & \underbrace{\sum_{n=1}^{\infty} P\left(|g_n(x) - E(g_n(x))| > \frac{\varepsilon}{2}\right)}_{(1)} \\ & + \underbrace{\sum_{n=1}^{\infty} P\left(|E(g_n(x)) - g(x)| > \frac{\varepsilon}{2}\right)}_{(2)}.\end{aligned}$$

برای قسمت (۲) طبق قضیه ۶.۱.۳ چون همگرایی کامل با همگرایی تقریباً همه‌جا معادل است، پس کافی است نشان دهیم $g(x) \xrightarrow{a.s.} g_{ni}(x)$. با استفاده از معادلات (۱۷.۴) و (۱۸.۴) برای $x \in c(g)$ و $a > 0$

چون

$$E(g_n(x)) = \sum_{i=1}^{n} W_{ni}(x)E(y_{ni}) = \sum_{i=1}^{n} W_{ni}(x)Eg(x_{ni}) + \sum_{i=1}^{n} W_{ni}(x)E(\varepsilon_{ni})$$

پس داریم

$$E(g_n(x)) = \sum_{i=1}^{n} W_{ni}(x)g(x_{ni}).$$

در این‌جا $E(g(x_{ni})) = g(x_{ni})$ است زیرا فقط تابع $g(\cdot)$ متغیر است و x ثابت می‌باشد. همچنین امید ریاضی ε برابر صفر است. بنابراین

$$E(g_n(x)) - g(x) = \sum_{i=1}^{n} W_{ni}(x)g(x_{ni}) - g(x)$$

که با اضافه و کم کردن $\sum_{i=1}^{n} W_{ni}(x)g(x)$ داریم

$$\begin{aligned}
E(g_n(x)) - g(x) &= \sum_{i=1}^{n} W_{ni}(x)g(x_{ni}) - \sum_{i=1}^{n} W_{ni}(x)g(x) + \sum_{i=1}^{n} W_{ni}(x)g(x) - g(x) \\
&= \sum_{i=1}^{n} W_{ni}(x)(g(x_{ni}) - g(x)) + g(x)(\sum_{i=1}^{n} W_{ni}(x) - 1).
\end{aligned}$$

در این صورت

$$\begin{aligned}
|Eg_n(x) - g(x)| &\leq \sum_{i=1}^{n} |W_{ni}(x)||(g(x_{ni}) - g(x))| \, (I(\|x_{ni} - x\| \leq a) \\
&+ I(\|x_{ni} - x\| > a)) + |g(x)||\sum_{i=1}^{n} W_{ni}(x) - 1| \\
&= \sum_{i=1}^{n} |W_{ni}(x)||g(x_{ni}) - g(x)|I(\|x_{ni} - x\| \leq a) \\
&+ \sum_{i=1}^{n} |W_{ni}(x)||g(x_{ni}) - g(x)|I(\|x_{ni} - x\| > a) \\
&+ |g(x)||\sum_{i=1}^{n} W_{ni}(x) - 1|. \qquad (۲۰.۴)
\end{aligned}$$

عبارت آخر با توجه به شرط (A) برابر صفر می‌شود. از آن‌جایی‌که $x \in c(g)$، برای هر $\varepsilon > 0$، یک $\delta > 0$ وجود دارد به طوری که هرگاه $\|x^* - x\| < \delta$، $|g(x_{ni}) - g(x)| < \varepsilon$ است. اگر مجموعه $a \in (0, \delta)$ در معادله (۲۰.۴) را داشته باشیم، آن‌گاه می‌توانیم به دست آوریم

$$\begin{aligned}
|Eg_n(x) - g(x)| &\leq \varepsilon \sum_{i=1}^{n} |W_{ni}(x)| + |g(x)||\sum_{i=1}^{n} W_{ni}(x) - 1| \\
&+ \sum_{i=1}^{n} |W_{ni}(x)||g(x_{ni}) - g(x)|I(\|x_{ni} - x| > a). \qquad (۲۱.۴)
\end{aligned}$$

بنابراین، با استفاده از معادله (۲۱.۴) و شرایط $(A) - (C)$ داریم

$$\lim_{n\to\infty} Eg_n(x) = g(x) \qquad x \in c(g). \tag{۲۲.۴}$$

برای $x \in c(g)$، بدون از دست دادن کلیت، فرض می‌کنیم که $W_{ni}(x) > ۰$. (در غیر این صورت به جای $W_{ni}(x)$ به ترتیب از $W_{ni}^{+}(x)$ و $W_{ni}^{-}(x)$ استفاده می‌کنیم و توجه می‌کنیم که

$$W_{ni}(x) = W_{ni}^{+}(x) - W_{ni}^{-}(x)$$

است). در قسمت (۱) با امید ریاضی گرفتن از $g_n(x)$ در معادله (۱۸.۴) داریم

$$\sum_{n=۱}^{\infty} P\left(|g_n(x) - E(g_n(x))| > \frac{\varepsilon}{۲}\right) = \sum_{n=۱}^{\infty} P\left(\sum_{i=۱}^{n} W_{ni}(x)\varepsilon_{ni} > \frac{\varepsilon}{۲}\right)$$

که بر اساس قضیه ۸.۲.۳ وقتی $n \to \infty$، $\sum_{i=۱}^{n} W_{ni}\varepsilon_{ni}$ همگرای کامل به صفر است. پس به ازای هر $\varepsilon > ۰$

$$\sum_{n=۱}^{\infty} P\left(|\sum_{i=۱}^{n} W_{ni}(x)\varepsilon_{ni}| > \varepsilon\right) < \infty. \tag{۲۳.۴}$$

حال قضیه ۸.۲.۳ را با $\alpha = ۰$، $\theta = ۱$، $\beta = ۱ - ۱/p \geq ۰$ و $r = ۱/p$ به کار می‌بریم و همچنین تعریف می‌کنیم $۱ + \alpha + \beta = ۲ - ۱/p > ۰$ و $s = \theta + (۱ + \alpha + \beta)/r = ۲p$. برای $x \in c(g)$، در قضیه ۸.۲.۳، $a_{ni} = W_{ni}(x)$ را مشخص می‌کنیم. به آسانی می‌بینیم که شرایط (۱۳.۳)، (۱۴.۳) و (۱۷.۳) در قضیه ۸.۲.۳ کافی هستند. همچنین با استفاده از قسمت (۲) قضیه ۸.۲.۳ داریم

$$\sum_{n=۱}^{\infty} n^{\beta} P\left(\max_{۱\leq j\leq n} |\sum_{i=۱}^{n} W_{ni}(x)\varepsilon_{ni}| > \varepsilon\right) < \infty; \qquad \varepsilon > ۰$$

که به معادله (۲۳.۴) دلالت دارد و برهان قضیه کامل است. □

نتیجه ۲.۴.۴. برای اثبات سازگاری برآوردگر رگرسیونی $g_n(x)$ در دو روش ذکرشده نادارایا–واتسون و رگرسیون چندجمله‌ای موضعی، کافی است ثابت کنیم سه شرط (A) تا (C) برای آن‌ها برقرار هستند.

بررسی شرایط همگرایی کامل برای برآوردگر هسته

از آن‌جایی که

$$\sum_{i=۱}^{n} \frac{k(\dfrac{x - x_i}{h})}{\sum_{k=۱}^{n} k(\dfrac{x - x_k}{h})} = ۱$$

شرط (A) به طور بدیهی برقرار است.

در شرط (B) نیز به دلیل مثبت بودن W_{ni}ها می‌توان قدر مطلق را در نظر نگرفت که در این صورت با حالت قبلی برابر می‌شود و متناهی بودن این سری را نتیجه می‌دهد. زیرا

$$\sum_{i=1}^{n}\left|\frac{k(\dfrac{x-x_i}{h})}{\sum_{k=1}^{n}k(\dfrac{x-x_k}{h})}\right|=1.$$

همچنین زمانی شرط (C) برقرار است که مرتبه $\sup_{1\leq i\leq n}|g_{ni}(x)-g(x)|$ از مرتبه $\sup_{1\leq i\leq n}|W_{ni}|$ کمتر باشد. به عبارت دیگر اگر $\sup_{1\leq i\leq n}|g_{ni}(x)-g(x)|=O(n^{-\alpha})$ و $\sup_{1\leq i\leq n}|W_{ni}|=O(n^{-1/p})$، آن‌گاه

$$\begin{aligned}\sum_{n=1}^{\infty}|W_{ni}(x)||g(x_{ni}-g(x)|I(\|x_{ni}-x\|>a) &\leq \sum_{n=1}^{\infty}O(n^{-1/p})O(n^{-\alpha})\\ &\leq \sum_{n=1}^{\infty}\frac{C_1}{n^{-1/p}}\frac{C_2}{n^{-\alpha}}=\sum_{n=1}^{\infty}\frac{C}{n^{-\alpha-1/p}}\end{aligned}$$

و زمانی این سری به سمت صفر میل می‌کند که همه جملاتش صفر باشد، یعنی $-\alpha-\dfrac{1}{p}>0$ شود که در این‌صورت باید $\alpha<-\dfrac{1}{p}$ شود، آن‌گاه شرط (C) برقرار است.

بررسی شرایط همگرایی کامل برای برآوردگر ناپارامتری چندجمله‌ای موضعی

برای بررسی شرایط همگرایی کامل در روش رگرسیون چندجمله‌ای موضعی، تابع وزن گوسی را در نظر می‌گیریم. برای اثبات شرط (A) واضح است که

$$\lim_{n\to\infty}\sum_{i=1}^{n}\exp\left(-\frac{1}{2}(\frac{x_i-x^*}{h})^2\right)=1$$

توابع هسته همیشه h را آنقدر کوچک در نظر می‌گیرند که x_i به سمت ماکسیمم خود یعنی x^* میل کند. که در این حالت توان صفر می‌شود و $\exp$ به سمت ۱ میل می‌کند.

برای اثبات شرط دوم می‌توانیم بگوییم که چون همه عبارات توابع وزن رگرسیون چندجمله‌ای موضعی نامنفی و به ۱ میل می‌کنند، پس متناهی هستند. بنابراین

$$\sum_{i=1}^{n}\left|\exp\left(-\frac{1}{2}(\frac{x_i-x^*}{h})^2\right)\right|<C$$

این شرط برای حالات مختلف توابع وزن رگرسیون چندجمله‌ای موضعی نیز صدق می‌کند. شرط (C) نیز مشابه روش نادارایا-واتسون به سادگی نتیجه می‌شود.

نتیجه ۳.۴.۴. شرایط (A) تا (C) برای سایر توابع وزن معرفی‌شده، مانند دو مربعی و سه وزنی، نیز برقرار هستند. دلیل آن هم به سادگی با توجه به بسط تابع نمایی و برقراری شرایط برای تابع وزن گوسی، روشن است.

نتیجه ۴.۴.۴. با توجه به برقراری سه شرط ذکرشده برای توابع وزن نادارایا-واتسون و چندجمله ای موضعی، نتیجه می‌گیریم که این برآوردگرها در مدلی که ε آن متغیر تصادفی NSD است، دارای همگرایی کامل است.

فصل ۵

ارزیابی نتایج نظری

به منظور ارزیابی نتایج نظری برآوردگر ناپارامتری ارایه‌شده در فصل چهارم، از مطالعه شبیه‌سازی استفاده کردیم که نتایج آن را در این فصل ارایه می‌دهیم. در مدل رگرسیون ناپارامتری

$$y = g(x) + \varepsilon \tag{۵.۱}$$

فرض می‌کنیم جمله خطا، یعنی ϵ، یک متغیر تصادفی NSD است.

در این مطالعه برای جمله خطای ϵ از دو توزیع نرمال چندمتغیره و تی چندمتغیره استفاده کردیم. برای NSD بودن این دو توزیع باید درایه‌های غیر قطر اصلی ماتریس کواریانس توزیع نرمال و ماتریس مقیاس تی، منفی باشند. در این دو توزیع چندمتغیره برای اطمینان از معتبر بودن ماتریس کواریانس (مقیاس)، در تولید نمونه از جمله خطا محدودیت داریم.

در واقع، اگر مساله تولید درایه‌های غیر قطر اصلی ماتریس کواریانس $\Sigma = (\sigma_{ij})$ با بعد n از توزیع یکنواخت در فاصله $[-a, ۰]$ باشد، سوالی که مطرح می‌شود این است که مقدار ماکسیمم a، به‌عنوان تابعی از n، وقتی که درایه‌های قطر اصلی همگی ۱ هستند، چند می‌تواند باشد؟ پاسخ این سوال مقدار $\frac{۱}{n-۱}$ است. با این واقعیت، ماتریس کواریانس (مقیاس) در توزیع‌های نرمال و تی چندمتغیره را در تمام شبیه‌سازی‌های این فصل به شکل زیر در نظر گرفتیم:

$$\sigma_{ii} = ۱, \quad \sigma_{ij} = -\frac{۱}{n - ۱}.$$

تمام شبیه‌سازی در محیط نرم‌افزار R اجرا شده‌اند و کدهای شبیه‌سازی توسط نویسنده پایان‌نامه نوشته شده‌اند.

۱.۵ برآوردگر هسته تابع رگرسیون ناپارامتری

در فصل قبل به این نتیجه رسیدیم که برآوردگر هسته ناداریا-واتسون، شرایط همگرایی کامل را تحت جمله خطای NSD دارا هست. برای ارزیابی کوچک نمونه این نتیجه، مدل (۱.۵) را با تابع ناپارامتری واقعی

$$g(x) = \sin ۲\pi x$$

شبیه‌سازی کردیم. در شبیه‌سازی چهار حجم نمونه $n = ۲۰, ۴۰, ۸۰, ۱۲۰$ را در نظر گرفتیم و از تابع هسته نرمال استفاده کردیم. متغیر تبیینی x را نیز از توزیع نرمال با میانگین صفر و واریانس ۴ تولید کردیم. همچنین برای جمله خطا دو توزیع نرمال و تی چندمتغیره را در نظر گرفتیم.

به منظور اجرای روش هسته برای برآورد تابع رگرسیونی $g(x)$ از بسته MASS در R کمک گرفتیم.

۱.۱.۵ خطای نرمال چندمتغیره

در شبیه‌سازی اول، جمله خطا را از توزیع نرمال n متغیره با میانگین صفر و ماتریس کواریانس Σ، که نحوه تشکیل آن در بالا تشریح شد، شبیه‌سازی کردیم. منحنی برآوردشده تابع رگرسیون به همراه منحنی واقعی و پراکنش داده‌های شبیه‌سازی‌شده، برای حجم‌های نمونه مختلف، در شکل ۱.۵ گزارش شده‌اند.

با توجه به منحنی‌های برازش‌شده، واضح است که برای حجم نمونه‌های کوچک ۲۰ و ۴۰، اریبی برآورد بزرگ است، اما با افزایش حجم نمونه، منحنی‌های برازش‌شده به مقدار واقعی نزدیک‌تر می‌شوند و برآوردها قابل قبول‌تر از حجم‌های کوچک‌تر هستند. بنابراین در این مثال می‌توان به برقراری نتایج نظری در کوچک نمونه اعتماد کرد.

۲.۱.۵ خطای تی چندمتغیره

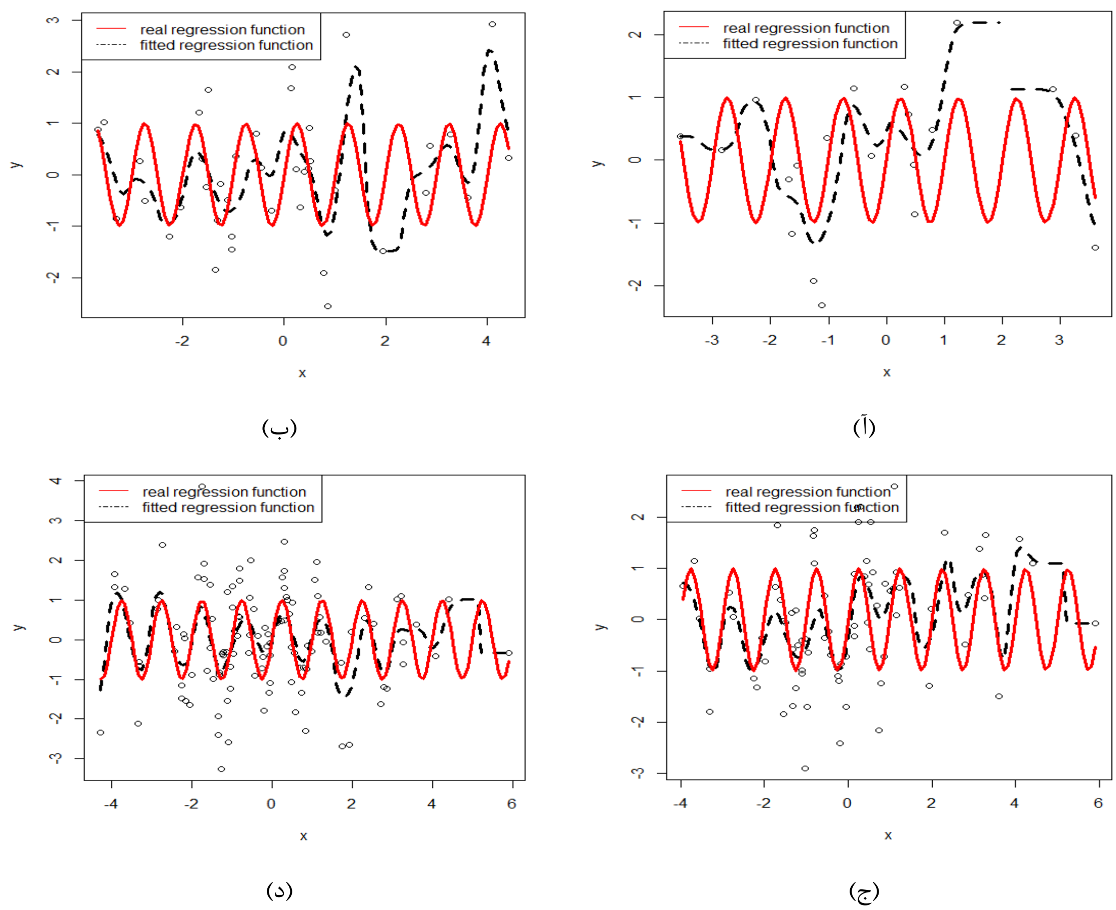
در مثال دوم، جمله خطای رگرسیون را از توزیع تی چندمتغیره با ماتریس مقیاس Σ تولید کردیم. نتایج در شکل ۲.۵ نمایش داده شده‌اند.

در نمودارهای شکل ۲.۵ نیز ملاحظه می‌کنیم که منحنی برازش‌شده زمانی که حجم نمونه کوچک است از منحنی واقعی دورتر و هرچه حجم نمونه بزرگ‌تر می‌شود به منحنی واقعی نزدیک‌تر می‌شود.

برای هر دو توزیع خطا ملاحظه می‌کنید که هرچه حجم نمونه بزرگ‌تر می‌شود، سازگاری برآوردگر هسته تابع رگرسیون ناپارامتری نمود بیشتری دارد و نمودار برازش‌شده، به جز در کران‌های مشاهدات، اریبی کمتر و همواری بیشتری دارد. این هموار‌ی در مدل رگرسیونی با توزیع خطای نرمال چندمتغیره مشهودتر است.

یک عیب همیشگی در برآورد تابع رگرسیونی به روش هسته در کران‌های مشاهدات متغیر تبیینی وجود دارد؛ به‌طور روشن‌تر، برآورد تابع رگرسیونی در نقاطی از متغیر تبیینی که در کران‌های مشاهدات قرار می‌گیرند، به شدت تحت تاثیر مشاهداتی از متغیر تبیینی قرار می‌گیرد که در آن سمت محدود نیستند (یعنی اگر کران بالا باشد تحت تاثیر مشاهدات سمت چپ نقطه مورد نظر قرار می‌گیرد و برعکس). این مساله باعث افزایش اریبی برآورد در آن نقطه خواهد شد. این مشکل در روش هسته به اریبی کران [۱] معروف است.

[۱] Boundary bias



شکل ۱.۵: منحنی واقعی تابع رگرسیون $g(x) = \sin 2\pi x$ با جمله خطای نرمال به همراه منحنی برآوردشده به روش هسته برای آ) $n = 20$، ب) $n = 40$، ج) $n = 80$ و د) $n = 120$

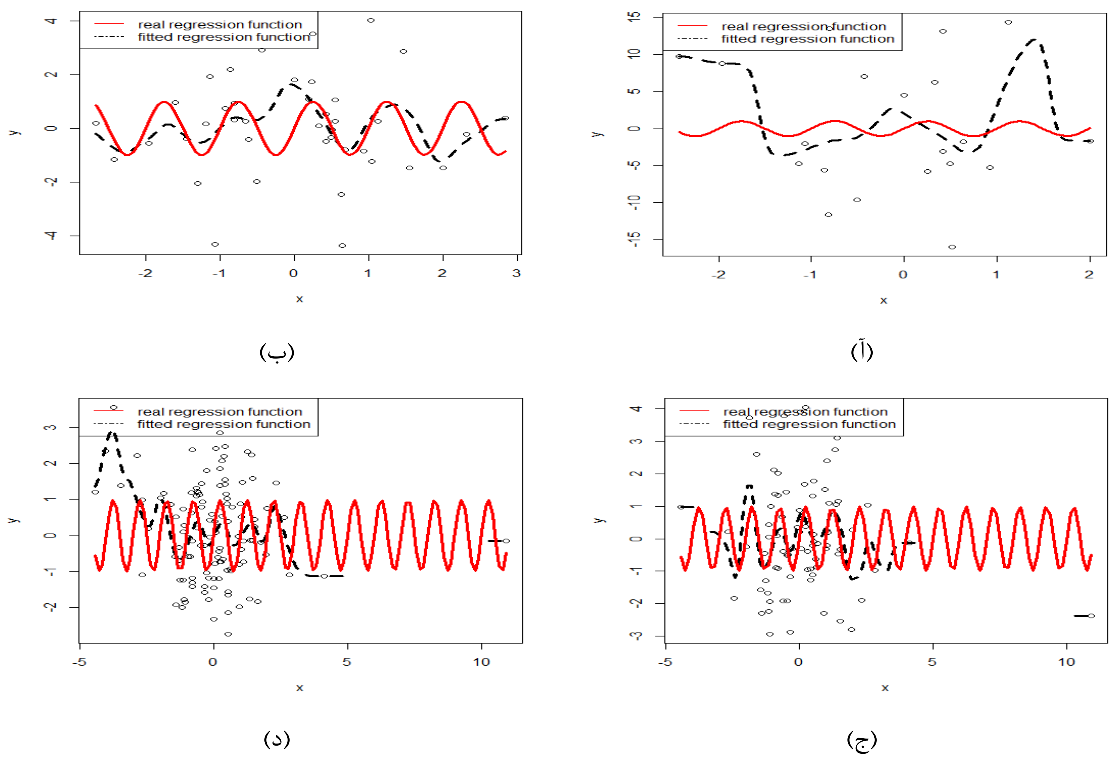
روش رگرسیون چندجمله‌ای موضعی یکی از انواع روش‌های رگرسیون ناپارامتری است که این ایراد را کاهش می‌دهد و می‌توان این روش را تعمیمی از روش رگرسیون نادارایا- واتسون در نظر گرفت.

۲.۵ برآوردگر رگرسیون چندجمله‌ای موضعی

برای برآوردگر تابع رگرسیون ناپارامتری با روش رگرسیون چندجمله‌ای موضعی نیز دیدیم که شرایط همگرایی کامل فصل چهارم برقرار هستند. در این حالت، مدل (۱.۵) را با دو تابع مختلف برای $g(\cdot)$ شبیه‌سازی کردیم:

۱. در مثال اول، تابع رگرسیونی به شکل
$$g_1(x) = \sqrt{x(1-x)}\frac{\sin(2/1\pi)}{x+0/2}$$
در نظر گرفته شد، که در آن متغیر تبیینی x از توزیع یکنواخت استاندارد شبیه‌سازی شد.

۲. در مثال دوم، تابع رگرسیونی به صورت
$$g_2(x) = x + 5\exp(-4x^2)$$
در نظر گرفته شد، که در آن متغیر تبیینی از توزیع یکنواخت در فاصله $[-1, 1]$ تولید شد.



شکل ۲.۵: منحنی واقعی تابع رگرسیون $g(x) = \sin ۲\pi x$ با جمله خطای تی به همراه منحنی برآوردشده به روش هسته برای آ) $n = ۲۰$، ب) $n = ۴۰$، ج) $n = ۸۰$ و د) $n = ۱۲۰$

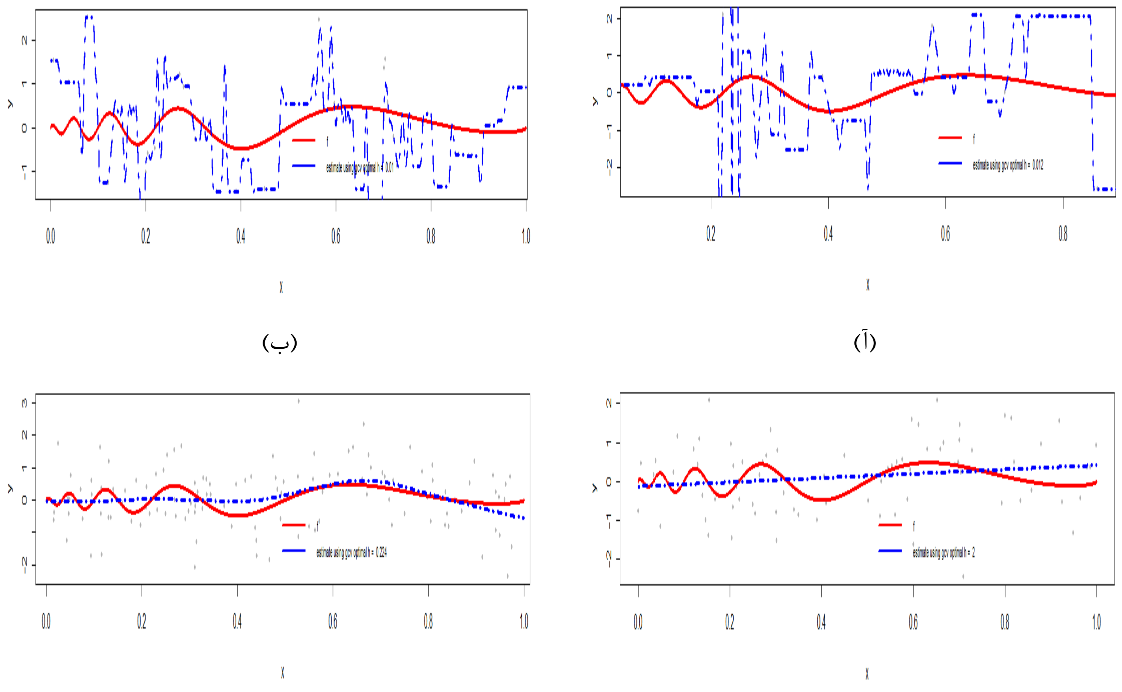
در این روش نیز چهار حجم نمونه $n = ۲۰, ۴۰, ۸۰, ۱۲۰$ را در نظر گرفتیم و برای جمله خطا دو توزیع نرمال و تی چندمتغیره را به کار بردیم.

برای اجرای این روش در R و اجرای شبیه‌سازی‌های این زیربخش از سه بسته locfit، MASS و mvtnorm کمک گرفتیم. بسته locfit یک بسته نرم‌افزاری برای انجام رگرسیون موضعی، و روش‌های هموارسازی با رهیافت مبتنی بر درست‌نمایی است. رگرسیون موضعی در زمینه‌های گوناگون در اواخر قرن ۱۹ و اوایل قرن ۲۰ مورد استقبال قرار گرفت. محبوبیت فعلی این روش عمدتا به دلیل وجود روش‌های LOESS و LOWESS است. از آن‌جایی که پارامتر هموارسازی در این روش، تاثیر زیادی بر منحنی (هموار) برآوردشده دارد، به‌طور پیش‌فرض بسته locfit، بر اساس حجم نمونه، یک پهنای باند بهینه را با استفاده از روش نزدیک‌ترین همسایگی[۲] انتخاب و هموار مناسب را کنترل می‌کند. تابع وزن پیش‌فرض در بسته $locfit$، تابع وزن گوسی است که شرایط قضیه ۱.۴.۴ برای همین تابع وزن بررسی و اعتبار آن نتیجه گرفته شد.

۱.۲.۵ خطای نرمال چندمتغیره

دو شکل ۳.۵ و ۴.۵ نتایج برآورد منحنی‌های به ترتیب $g_۱(x)$ و $g_۲(x)$ را برای حالتی که جمله خطا نرمال چندمتغیره است، بر حسب حجم‌های نمونه مختلف، نشان می‌دهند.

[۲] Nearest neighbours



(ب) (آ)

(د) (ج)

شکل ۳.۵: منحنی واقعی تابع رگرسیون $g_۱(x)$ با جمله خطای نرمال به همراه منحنی برآوردشده به روش LOESS برای آ) $n = ۲۰$، ب) $n = ۴۰$، ج) $n = ۸۰$ و د) $n = ۱۲۰$

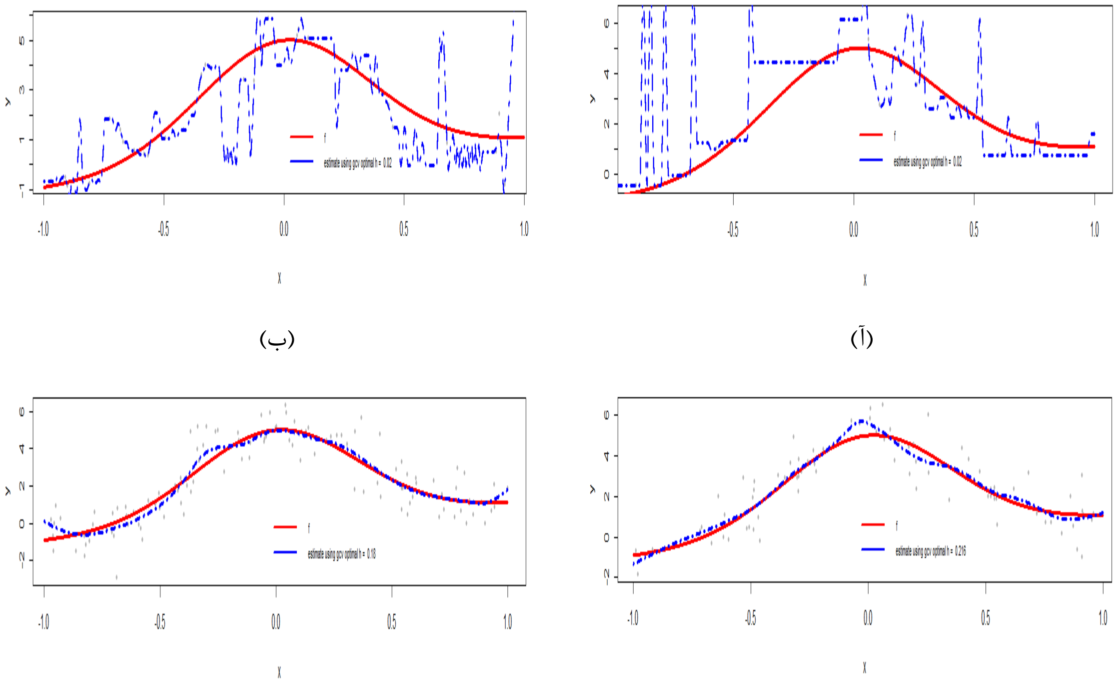
در شکل ۳.۵ ملاحظه می‌شود که برای حجم‌های نمونه کوچک، منحنی برازش‌شده به‌طور پیاپی و به صورت تکه‌های به هم وصل‌شده حاصل شده است. واضح است که این منحنی به شدت تحت تاثیر تغییرات مکانی و اغتشاشات تصادفی داده‌ها قرار دارد. مجموع توان‌های دوم باقیمانده‌ها در نقاط مشاهده‌شده صفر است و منحنی برآوردشده حاصل ناهموار و کاملا شکسته است. با افزایش حجم نمونه منحنی برازش‌شده هموارتر و به مقدار واقعی نزدیک‌تر می‌شود.

شکل ۴.۵ استنباط یکسانی را نتیجه می‌دهد. برای هر دو تابع در نظر گرفته‌شده، برآورد تابع رگرسیونی با افزایش حجم نمونه به منحنی واقعی نزدیک شده و سازگاری برآوردگر روش LOESS مشهود است.

۲.۲.۵ خطای تی چندمتغیره

برای حالتی که جمله خطا از توزیع NSD تی چندمتغیره شبیه‌سازی شده است، نتایج برازش تابع رگرسیونی برای دو تابع واقعی $g_۱(x)$ و $g_۲(x)$ به ترتیب در دو شکل ۵.۵ و ۶.۵ نمایش داده شده‌اند.

برای مدل رگرسیون ناپارامتری با جمله خطای NSD تی چندمتغیره، نتایج مشابه حالت نرمال به دست آمده‌اند؛ با این تفاوت که در مدل رگرسیون با جمله خطای تی چندمتغیره، همواری مدل برازش داده‌شد سریع‌تر از مدل با خطای نرمال است.



(ب) (آ)

(د) (ج)

شکل ۴.۵: منحنی واقعی تابع رگرسیون $g_2(x)$ با جمله خطای نرمال به همراه منحنی برآوردشده به روش LOESS برای آ) $n = 20$، ب) $n = 40$، ج) $n = 80$ و د) $n = 120$

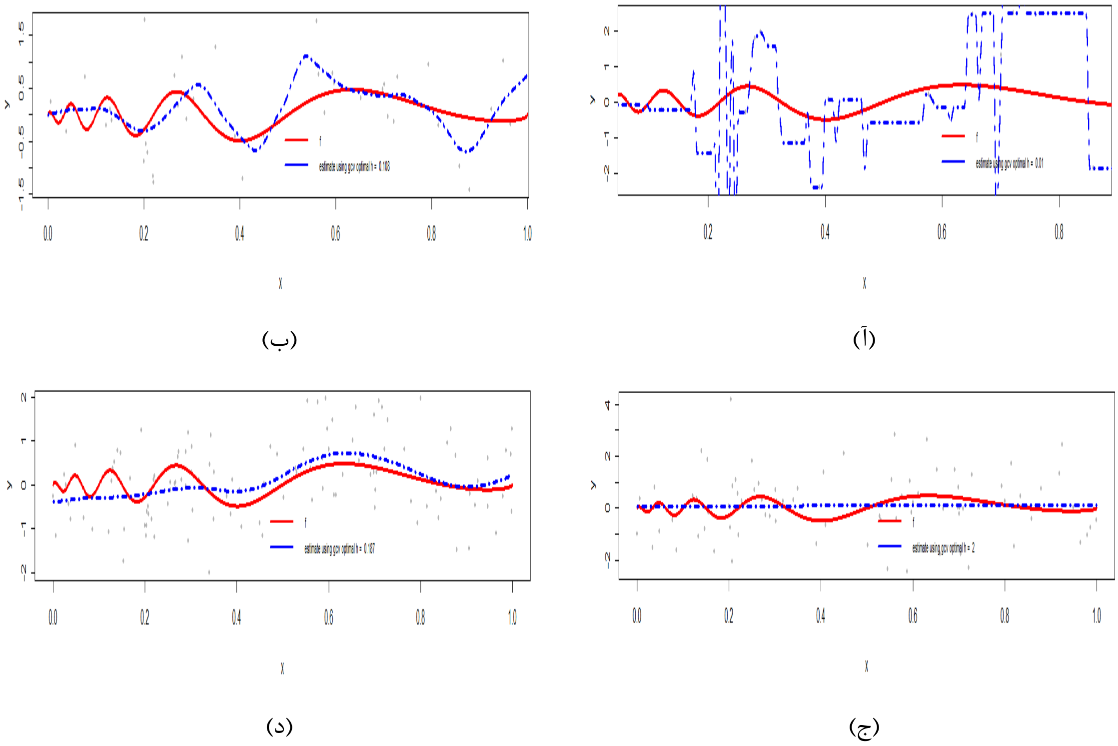
۳.۲.۵ میانگین توان دوم خطای مدل رگرسیونی

برای ارزیابی کمی نمودارهای برآوردشده برای تابع رگرسیونی، از معیار MISE استفاده کردیم. جدول ۱.۵ مقادیر محاسبه‌شده معیار MISE را برای برآوردهای دو تابع $g_1(x)$ و $g_2(x)$، بر حسب حجم‌های نمونه $n = 20, 40, 80, 120$ نشان می‌دهد.

همان‌طور که دیده می‌شود با افزایش حجم نمونه مقادیر MISE برای هر دو تابع $g_1(x)$ و $g_2(x)$ کاهش پیدا کرده‌اند. بنابراین این معیار نیز حاکی از سازگار بودن برآوردهای تابع رگرسیونی با روش LOESS است.

۳.۵ نتیجه‌گیری و آینده تحقیق

در این پایان‌نامه، در یک مدل رگرسیون ناپارامتری فرض کردیم جمله خطا یک متغیر تصادفی وابسته با ساختار وابستگی NSD است. سازگاری کامل یک برآوردگر تابع رگرسیون ناپارامتری را نیز، تحت شرایطی برای تابع وزن موجود در برآوردگر پیشنهادی، بازگو کردیم. همچنین نشان دادیم شرایط مطرح‌شده برای دو برآوردگر ناپارامتری هسته و رگرسیون چندجمله‌ای موضعی برقرار هستند. این نتیجه تضمینی است برای همگرایی کامل برآوردگر ناپارامتری تابع رگرسیونی در یک مدل رگرسیون با جمله خطای NSD وقتی که از روش‌های هسته یا چندجمله‌ای موضعی استفاده می‌کنیم. با استفاده از مطالعه شبیه‌سازی نیز برقراری نتایج نظری را، در حالت حجم نمونه محدود، نمایش دادیم.

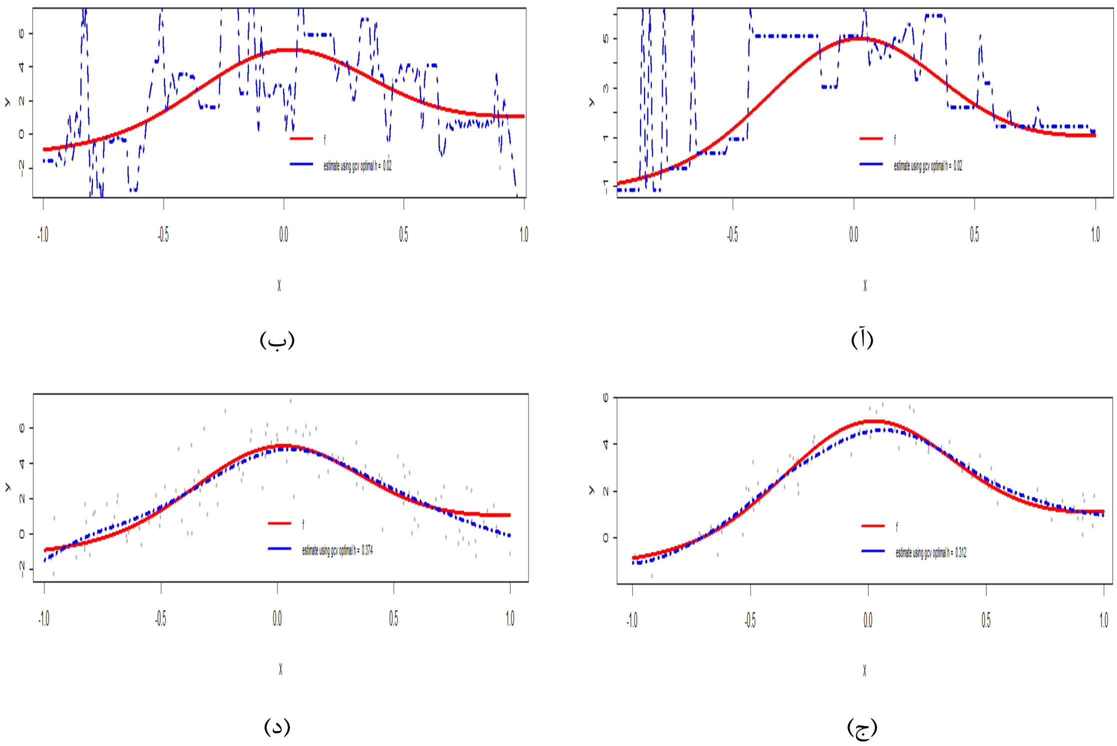


شکل ۵.۵: منحنی واقعی تابع رگرسیون $g_1(x)$ با جمله خطای تی به همراه منحنی برآوردشده به روش LOESS برای آ) $n = ۲۰$، ب) $n = ۴۰$، ج) $n = ۸۰$ و د) $n = ۱۲۰$

با توجه به مطالب نظری و شبیه‌سازی بیان‌شده، می‌توان موارد زیر را به‌عنوان آینده تحقیق خروجی این پایان‌نامه فهرست کرد:

- رده وابستگی NSD یک رده نسبتا بزرگ از وابستگی‌هاست که متغیرهای تصادفی NA و مستقل را شامل می‌شود. بنابراین، نتایج ارایه‌شده در این پایان‌نامه برای این دو رده از متغیرهای تصادفی نیز برقرار هستند. بنابراین می‌توان جمله خطا در مدل رگرسیون ناپارامتری را با این دو نوع متغیر تصادفی در نظر و نتایج همگرایی کامل مشابه را نتیجه گرفت. بررسی همگرایی کامل برآوردگر تابع رگرسیون ناپارامتری برای سایر رده‌های وابستگی یکی از موضوع‌های مورد علاقه به‌عنوان آینده این تحقیق، قابل طرح است.

- بررسی شرایط نظری تابع وزن در برآوردگر ناپارامتری معرفی‌شده در فصل چهارم، برای سایر روش‌های مطرح‌شده برای وزن‌دهی، مانند روش نزدیکترین همسایگی، نیز می‌تواند جالب توجه باشد.



شکل ۶.۵: منحنی واقعی تابع رگرسیون $g_۲(x)$ با جمله خطای تی به همراه منحنی برآوردشده به روش LOESS برای آ) $n = ۲۰$، ب) $n = ۴۰$، ج) $n = ۸۰$ و د) $n = ۱۲۰$

جدول ۱.۵: میانگین توان دوم خطای تجمیع‌شده برای برآوردگرهای دو تابع $g_۱(x)$ و $g_۲(x)$ در روش LOESS

خطای تی چندمتغیره		خطای نرمال چندمتغیره		
$g_۲$	$g_۱$	$g_۲$	$g_۱$	حجم نمونه
۸/۵۲۶	۰/۶۴۰	۸/۳۳۶	۰/۵۰۳	۲۰
۸/۰۳۱	۰/۶۰۵	۷/۸۰۹	۰/۴۳۸	۴۰
۷/۷۹۰	۰/۵۹۴	۷/۵۵۳	۰/۴۰۳	۸۰
۷/۰۷۱	۰/۵۹۱	۷/۴۶۸	۰/۳۹۱	۱۲۰

مراجع

[۱] ترابی، ح. بقایی‌پور، م. (۲۰۱۱). رگرسیون ناپارامتری، مجله اندیشه آماری، (۱)**۱۵**، ۵۵-۴۱.

[2] Alam, K, and Lai Saxena, K. M. (1981). Positive dependence in multivariate distributions. *Comm Stat Theo Meth*, **10**(12), 1183-1196.

[3] Banjevic, D. (1984). On a Kolmogorov inequality. *Theo Prob Appl*, **29**(2), 391-394.

[4] Baum, L. E., and Katz, M. (1965). Convergence rates in the law of large numbers. *Tran American Math Soc*, **120**(1), 108-123.

[5] Birkel, T. (1988). Moment bounds for associated sequences. *Ann Prob*, **16**(3), 1184-1193.

[6] Block, H. W., Savits, T. H., and Shaked, M. (1982). Some concepts of negative dependence. *Ann Prob*, **10**(3), 765-772.

[7] Block, H. W., Savits, T. H., and Shaked, M. (1985). A concept of negative dependence using stochastic ordering. *Stat Prob Lett*, **3**(2), 81-86.

[8] Block, H. W., and Sampson, A. R. (1988). Conditionally ordered distributions. *J Mult Ann*, **27**(1), 91-104.

[9] Cai, Z. (2001). Weighted nadaraya–watson regression estimation. *Stat Prob Lett*, **51**(3), 307-318.

[10] Chang, C. S. (1992). A new ordering for stochastic majorization: Theory and applications. *Adva appl Prob*, **24**(3), 604-634.

[11] Chen, P. Y., Hu, T. C., Liu, X. D., and Volodin, A. (2007). On complete convergence for arrays of row-wise negatively associated random variables. *Theo Prob Appl*, **52**(2), 323-328.

[12] Cheng, K. F., and Lin, P. E. (1981). Nonparametric estimation of a regression function. *Zeitschrift für Wahrscheinlich keits Theoie und verwandte Gebiete*, **57**(2), 223-233.

[13] Christofides, T. C., and Vaggelatou, E. (2004).A connection between supermodular ordering and positive/negative association. *J Mult Ann*, **88**(1), 138-151.

[14] Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *J American stat asso*, **74**(368), 829-836.

[15] Dehua, Q., Chang, K. C., Antonini, R. G., and Volodin, A. (2011). On the strong rates of convergence for arrays of rowwise negatively dependent random variables. *Stoc Ann Appl*, **29**(3), 375-385.

[16] Dubhashi, D. P., and Ranjan, D. (1996). Balls and bins: A study in negative dependence. *BRICS Rep Seri*, **3**(25), 1-27.

[17] Efron, B. (1965). Increasing properties of Polya frequency function. *Ann Math Stat*, **36**(1), 272-279.

[18] Eghbal, N., Amini, M., and Bozorgnia, A. (2011). On the Kolmogorov inequalities for quadratic forms of dependent uniformly bounded random variables. *Stat Prob Lett*, **81**(8), 1112-1120.

[19] Eghbal, N., Amini, M., and Bozorgnia, A. (2010). Some maximal inequalities for quadratic forms of negative superadditive dependence random variables. *Stat Prob Lett*, **80**(7), 587-591.

[20] Fan, Y. (1990). Consistent nonparametric multiple regression for dependent heterogeneous processes: the fixed design case. *J Mult Ann*, **33**(1), 72-88.

[21] Fox, J.(2000). *Noparametric simple regression: smoothing scatterplots*, sage, thousand oaks CA.

[22] Gasser, T., and Müller, H. G. (1979). *Kernel Estimation Of Regression Functions*. In Smoothing techniques for curve estimation (pp. 23-68). Springer Berlin Heidelberg.

[23] Georgiev, A. A. (1985). Local properties of function fitting estimates with applications to system identification. *Math Stat Appl*, **2**, 141-151.

[24] Georgiev, A. A., and Greblicki, W. (1986). Nonparametric function recovering from noisy observations. *J Stat Plan Infe*, **13**, 1-14.

[25] Georgiev, A. A. (1988). Consistent nonparametric multiple regression: the fixed design case. *J Mult Ann*, **25**(1), 100-110.

[26] Ghosh, M. (1981). Multivariate negative dependence. *Comm Stat Theo Meth*, **10**(4), 307-337.

[27] Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Tech*, **21**(2), 215-223.

[28] Gut, A.(2005).*Probility:A Gradvate Course*, Spring T Stat, Sweden.

[29] Hardle, W., and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann Stat*, **21**(4), 1926-1947.

[30] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann Math Stat*, **19**(3), 293-325.

[31] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J American stat Asso*, **58**(301), 13-30.

[32] Hu, T. C., Szynal, D., and Volodin, A. I. (1998). A note on complete convergence for arrays. *Stat Prob Lett*, **38**(1), 27-31.

[33] Hu, T. Z. (2000). Negatively superadditive dependence of random variables with applications. *Chinese J Appl Prob Stat*, **16**(2), 133-144.

[34] Hu, S. H., Zhu, C. H., Chen, Y. B.,Wang, L. C. (2002). Fixed-design regression for linear time series. *Acta Math Sci*, **22**B(1), 9–18.

[35] Ibragimov, R., and Sharakhmetov, S. (2002). The exact constant in the Rosenthal inequality for random variables with mean zero. *Theo Prob Appl*, **46**(1), 127-132.

[36] Joag-Dev, K., and Patil, G. P. (1975). Probability inequalities for certain multivariate discrete distribution. *Sankhya: Indian J Stat, (Seri B)*, **37**(2), 158-164.

[37] Joag-Dev, K., and Proschan, F. (1983). Negative association of random variables with applications. *Ann Stat*, **11**(1), 286-295.

[38] Johnson, N. L., and Kott, S. (1975). On some generalized Farlie-Gumbel-Morgenstern distributions. *Comm Stat-Theo Meth*, **4**(5), 415-427.

[39] Janson, S. (1984). Runs in m-dependent sequences. *Ann Prob*, **12**(3), 805-818.

[40] Karlin, S., and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J Mult Ann*, **10**(4), 467-498.

[41] Kemperman, J. H. B. (1977, January). On the FKG-inequality for measures on a partially ordered space. *Inda Math* (Proceedings) **80**(4), 313-331. North-Holland.

[42] Kruglov, V. M., Volodin, A. I., and Hu, T. C. (2006). On complete convergence for arrays. *Stat Prob Lett*, **76**(15), 1631-1640.

[43] Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai* **14**(2), 1137-1145.

[44] Kolmogorov, A. N. (1963). On tables of random numbers. *Sankhyā: Indian J Stat, (Seri A)*, **25**(4), 369-376.

[45] Lehmann, E. L. (1966). Some concepts of dependence. *Ann Math Stat*, **37**(5), 1137-1153.

[46] Liang, H. Y., and Jing, B. Y. (2005). Asymptotic properties for estimates of non-parametric regression models based on negatively associated sequences. *J Mult Ann*, **95**(2), 227-245.

[47] Lin, Z., and Bai, Z. (2011). *Probability Inequalities*. Springer Scinence and Business Media.

[48] Liu, L. (2009). Precise large deviations for dependent random variables with heavy tails. *Stat Prob Lett*, **79**(9), 1290-1298.

[49] Liu, W., Xiao, H., and Wu, W. B. (2013). Probability and moment inequalities under dependence. *Stat Sini*, **23**(3), 1257-1272.

[50] Mavrikiou, P. M. (2007). Kolmogorov inequalities for the partial sum of independent Bernoulli random variables. *Stat Prob Lett*, **77**(11), 1117-1122.

[51] Müller, H. G. (1987). Weak and universal consistency of moving weighted averages. *Peri Math Hung*, **18**(3), 241-250.

[52] Peligrad, M. (1987). On the central limit Theoem for ρ-mixing sequences of random variables. *Ann Prob*, **15**(4), 1387-1394.

[53] Priestley, M. B., and Chao, M. T. (1972). Non-parametric function fitting. *J Roy Stat Soc. (Series B)* (Methodological), **34**(3), 385-392.

[54] Roussas, G. G. (1989). Consistent regression estimation with fixed design points under dependence conditions. *Stat Prob Lett*, **8**(1), 41-50.

[55] Roussas, G. G., Tran, L. T., and Ioannides, D. A. (1992). Fixed design regression for time series: asymptotic normality. *J Mult Ann*, **40**(2), 262-291.

[56] Rosenthal, H. P. (1970). On the subspaces ofL p (p> 2) spanned by sequences of independent random variables. *Israel J Math*, **8**(3), 273-303.

[57] Sazonov, V. V. E. (1974). On the estimation of moments of sums of independent random variables. *Theo Vero Prim*, **19**(2), 383-386.

[58] Schönfeld, P. (1971). A useful central limit theorem for m-dependent variables. *Metr*, **17**(1), 116-128.

[59] Sen, P. K. (1971). A note on weak convergence of empirical processes for sequences of φ-mixing random variables. *Ann Math Stat*, **42**(6), 2131-2133.

[60] Simonoff, J. S. (1996). *Smoothing Methods In Statistics*. Springer Scinence and Business Media.

[61] Shao, Q. M., and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann Prob*, **24**(4), 2098-2127.

[62] Shao, Q. M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J Theo Prob*, **13**(2), 343-356.

[63] Su, C., Zhao, L., and Wang, Y. (1996). Moment inequalities and weak convergence for negatively associated sequences. *Sci China (Seri A): Math*, **40**(2), 172-182.

[64] Sung, S. H., Volodin, A. I., and Hu, T. C. (2005). More on complete convergence for arrays. *Stat Prob Lett*, **71**(4), 303-311.

[65] Takezawa, K. (2006). *Introduction to nonparametric regression* (Vol. 606). John Wiley and Sons.

[66] Tran, L., Roussas, G., Yakowitz, S., and Van, B. T. (1996). Fixed-design regression for linear time series. *Ann Stat*, **24**(3), 975-991.

[67] Turner, D. W., Young, D. M., and Seaman, J. W. (1995). A Kolmogorov inequality for the sum of independent Bernoulli random variables with unequal means. *Stat Prob Lett*, **23**(3), 243-245.

[68] Van der Vaart, A. W. (2000). *Asymptotic Statistics* (Vol. 3). Cambridge university press.

[69] Ville, J. (1939). *Etude critique de la notion de collectif*. Gauthier-Villars, Paris.

[70] Wang, X., Deng, X., Zheng, L., and Hu, S. (2014). Complete convergence for arrays of rowwise negatively superadditive-dependent random variables and its applications. *Stat*, **48**(4), 834-850.

[71] Wu, Q. Y. (2006). Probability limit theory for mixing sequences.*Sci Press China*, Beijing.

[72] Yang, W., Wang, X., Wang, X., and Hu, S. (2012). The consistency for estimator of nonparametric regression model based on NOD errors. *J Ineq Appl*, **2012**(1), 1.

[73] Young, D. M., Seaman, J. W., and Marco, V. R. (1987). A note on a Kolmogorov inequality. *Stat Prob Lett*, **5**(3), 217-218.

[74] Zhang, L. (2014). Rosenthal's inequalities for independent and negatively dependent random variables under sub-linear expectations with applications. *Sci China Math*, **59**(4), 751-768.

پیوست آ

کدهای نرم‌افزاری

آ.۱ کدهای مربوط به روش هسته

```
rm(list=ls())
library(MASS)
set.seed(2314)
#####
n = 80
a = -1/(n-1)
M = a*matrix(rep(1,n*n), ncol=n)
diag(M) = rep(1,n)
eigen(M)$values
###
#x = rnorm(n,0,2)
x=rt(n,3)
g = sin(2*pi*x)
gen.data <- function(n){
#epsilon = mvrnorm(1, mu=rep(0,n), Sigma=M)
```

```
       epsilon = as.vector(rmvt(1, sigma=M, df=3))
y = g + epsilon
return(y)
}
y = gen.data(n=n)
plot(y~x)

B = 2
out <- list()
 esm <- function(B){
for (i in 1:B){
y <- gen.data(n=n)
nonp = ksmooth(x, y, kernel="normal")
xx <- nonp$x
out[[i]] <- nonp$y
}
return(list('xx' = xx, 'out' = out))
}
natije <- esm(B=2)
#####
nonp1 = ksmooth(x, y, kernel="normal")
nonp2 = ksmooth(x, y, kernel="normal")
nonp3 = ksmooth(x, y, kernel="normal")
plot()
lines(nonp1, lty=2)
lines(nonp2, lty=2)
lines(nonp3, lty=2)

lines(natije$xx, (sin(2*pi*(natije$xx))),col="red")
legend("topleft",c("true variable","estimate variable"),lwd=1,lty=c(1,4),col=c("re
```

۲.آ کدهای مربوط به روش چندجمله‌ای موضعی

```
library(MASS)
library(locfit)
```

```
library(mvtnorm)
set.seed(2314)
######
n = 120
a = -1/(n-1)
M = a*matrix(rep(1,n*n), ncol=n)
diag(M) = rep(1,n)
eigen(M)$values
###################
locfit_simdata = function(f,Sig,n,h,deg,xlo,xhi){
x = runif(n,xlo,xhi)
y = f(x) + mvrnorm(1, mu=rep(0,n), Sigma=Sig)
#y = f(x) + as.vector(rmvt(1, sigma=Sig, df=3)) # replace it with mvrnorm
locfitfit = locfit(y~x, alpha=c(0,h), deg=deg, maxk=1000)
alphamat = matrix(0,ncol=2,nrow=30)
alphamat[,2] = (1.2^(seq(30)-30))*2*(xhi-xlo)
gcvs = gcvplot(y~x, alpha=alphamat,deg=deg,maxk=1000)
optband = max(gcvs$alpha[gcvs$values == min(gcvs$values),2])
locfitopt = locfit(y~x,alpha=c(0,optband),deg=deg,maxk=1000)
xg = seq(xlo,xhi,length=1000)
plot(x,y,xlab="x",ylab="y",pch=16,cex=0.5,col=rgb(0.7,0.7,0.7),
main="Local fitting of nonparametric function f")
lines(xg,f(xg),col=2,lwd=3)
# lines(xg,predict(locfitfit,newdata=xg),col=3,lwd=3)
lines(xg,predict(locfitopt,newdata=xg),col=4,lwd=3)
legend(xhi,min(y),legend=c(expression(f),
paste("estimate using gcv optimal h = ",round(optband,3))),
col=c(2,4),lwd=2,bty="n",cex=.7,yjust=0,xjust=1)
# mtext(paste("degree=",deg,", n=",n,",sigma=",sigma,sep=""),line=1,
# adj=0,cex=1.3)
list(y=y)
}
par(mfrow=c(2,1))
f1 = function(x){sqrt(x*(1-x))*sin((2.1*pi)/(x+0.2))}
```

```
y1=locfit_simdata(f1,Sig=M,n=n,h=0.2,deg=1,xlo=0,xhi=1)
f2 = function(x){x + 5*exp(-4*x^2)}
y2=locfit_simdata(f2,Sig=M,n=n,h=0.2,deg=1,xlo=-1,xhi=1)
########################
#### computing mise of estimators
fit.x <- seq(0,1,length.out=120)
fit.x
deg=1
h=0.2
locfitfit = locfit(y1$y~x, alpha=c(0,h), deg=deg, maxk=1000)
predict(locfitfit,fit.x)
g.hat <- predict(locfitfit,fit.x)
g.real <- f1(fit.x)
mean((g.hat-g.real)^2)

fit.x <- seq(0,1,length.out=120)
fit.x
locfitfit = locfit(y2$y~x, alpha=c(0,h), deg=deg, maxk=1000)
predict(locfitfit,fit.x)
g.hat <- predict(locfitfit,fit.x)
g.real <- f2(fit.x)
mean((g.hat-g.real)^2)
```

Abstract

Independence and dependence properties of random variables play a basic role in statistics and probability. In most situations, the random variables are not independent. Hence, we should study and use the different measures of dependence between random variables. In this thesis, particularly, we focus on negatively superadditive-dependence (NSD) structure and its applications. Moment inequalities could be used in many applications of probability, e.g. finding lower or upper bounds for moments which their exact values are not known. Two useful moment inequalities for NSD random variables are Rosenthal-type maximal and Kolmogorov-type exponential inequalities. In this thesis, by using the mentioned inequalities, we re-introduce the complete convergence of arrays of row-wise NSD random variables and the complete consistency for the estimator of nonparametric regression model based on NSD errors. Then, we examine the performance of theoretical results by a simulation study.

Keywords: Negatively superadditive-dependent random variables, Rosenthal-type maximal inequality, Kolmogrov-type exponential inequality, complete convergence, complete consistency.

Shahrood University of Technology

Faculty of Mathematical Sciences

MSc Thesis in: Mathematical Statistics

Complete Convergence for the Estimator of Nonparametric Regression Model Based on Negatively Superadditive-dependent Errors

By: Samaneh Rahmani

Supervisors

Dr. Negar Eghbal
Dr. Hossein Baghishani

January 2017