بسم الله الرحمن الرحیم

دانشکده علوم ریاضی

رشته آمار ، گرایش آمار ریاضی

پایان‌نامه کارشناسی ارشد

همگرایی کامل برای مجموع موزون متغیرهای تصادفی وابسته زبرجمعی منفی و کاربرد آن در مدل رگرسیون خطی با متغیرهای تبیینی آلوده به خطا

نگارنده: سمانه حبیبی کمنی

استادان راهنما

دکتر نگار اقبال
دکتر حسین باغیشنی

بهمن ۱۳۹۵

تقدیم به

پدر بزرگوارم،

مادر مهربانم و

همسر فداکارم

سپاس‌گزاری

سپاس بی‌منتها سزاوار خداوندی است که به مصلحت از نعمات خویش می‌بخشد و به حکمت از ما می‌ستاند، خدایا آن‌چه داشته‌ام تو داده‌ای و آن‌چه کرده‌ام تو میسر نموده‌ای. همه وجود من زاده توست. من از خود چیزی ندارم و از خود کاری نکرده‌ام پس تو را سپاس می‌گویم که به من منت نهادی تا بتوانم با توکل بر یاری بی‌دریغ و لطف بی‌پایانت این تحقیق را به پایان برسانم. آن‌چه در این مجموعه گرد آمده است حاصل نمی‌شد مگر به یاری عزیزانی که مرا یاری رساندند که به رسم ادب مراتب سپاس و قدردانی خود را تقدیم‌شان می‌نمایم.

اساتید راهنمای محترم، سرکار خانم دکتر نگار اقبال و جناب آقای دکتر حسین باغیشنی که راهنمایی‌های ارزنده ایشان در تمام مراحل پژوهش باعث شد تا این پایان‌نامه را با موفقیت به اتمام برسانم نهایت سپاس را دارم.

همچنین از اساتید محترم، دکتر احمد نزاکتی رضازاده و دکتر محمدرضا ربیعی که داوری این پایان‌نامه را بر عهده گرفته‌اند تشکر می‌نمایم.

در پایان هم از پدر و مادر بزرگوارم که همیشه پشت و پناهم، خواهر و برادرم که حامیانم و تکیه‌گاه زندگیم همسر عزیزم کمال تشکر و سپاس‌گزاری را دارم.

سمانه حبیبی کمنی

بهمن ۱۳۹۵

تعهد نامه

این‌جانب **سمانه حبیبی کمنی** دانشجوی کارشناسی ارشد رشته **آمار** دانشکده **علوم ریاضی** دانشگاه صنعتی شاهرود، نویسنده پایان‌نامه با عنوان **همگرایی کامل برای مجموع موزون متغیرهای تصادفی وابسته زیرجمعی منفی و کاربرد آن در مدل رگرسیون خطی با متغیرهای تبیینی آلوده به خطا** ، تحت راهنمایی **نگار اقبال** و **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 شده است.

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 شده است.

سمانه حبیبی کمنی
بهمن ۱۳۹۵

مالکیت نتایج و حق نشر

چکیده

مفهوم آماری وابستگی سبب شده است تا رشد چشم‌گیری در آمار و احتمال و دیگر شاخه‌های علم پدید آید. مفاهیم متنوع وابستگی در احتمال، محققان را بر آن داشت تا در سه دهه اخیر کاربردهای آن را در زمینه‌های مختلف جست وجو کنند. یکی از این نوع وابستگی‌ها، وابستگی متغیرهای تصادفی زبرجمعی منفی است که می‌توان به برخی ویژگی‌های اساسی آن از جمله نابرابری نمایی کلموگروف و نابرابری ماکسیمال روزنتال اشاره کرد. بنابراین با این ویژگی‌ها و همچنین همگرایی کامل، متغیرهای تصادفی مستقل را در تناظر با متغیرهای تصادفی وابسته مورد بررسی قرار خواهیم داد. از این تناظر، تحت شرایطی، سازگاری کامل برآوردگرهای کم‌ترین توان‌های دوم در مدل رگرسیونی تبیینی آلوده به خطا با خطای وابسته زبرجمعی منفی را بازگو می‌کنیم و با یک مطالعه شبیه‌سازی عملکرد نتایج نظری را ارزیابی خواهیم کرد.

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

پیش‌گفتار

متغیرهای وابسته زبرجمعی منفی، رده بسیار گسترده از دنباله متغیرهای تصادفی است که دنباله‌های مستقل و دنباله‌های پیوندی منفی از متغیرهای تصادفی را به‌عنوان حالت خاص شامل می‌شود. در این پایان‌نامه به کمک همگرایی کامل و همچنین نابرابری‌های احتمالی و گشتاوری، از جمله نابرابری ماکسیمال روزنتال و نابرابری نمایی کلموگروف، به نتایجی برای دنباله‌ای از متغیرهای تصادفی وابسته زبرجمعی منفی اشاره می‌کنیم و کاربرد آن را در رگرسیون خطی تبیینی آلوده به خطا مورد بررسی قرار می‌دهیم. با این مقدمه مطالب هر فصل به‌صورت زیر تنظیم شده‌اند:

- در فصل اول برخی مقدمات و تعاریف لازم را که در فصل‌های بعد مورد استفاده قرار می‌گیرند، بیان می‌کنیم.
- در فصل دوم نابرابری‌های ماکسیمال و نمایی را در حالت استقلال، پیوندی منفی و وابسته زبرجمعی منفی مورد مطالعه قرار می‌دهیم.
- در فصل سوم همگرایی کامل برای مجموع موزون متغیرهای تصادفی وابسته زبرجمعی منفی را با استفاده از ویژگی‌های این دسته از متغیرهای تصادفی بیان و اثبات می‌کنیم.
- فصل چهارم شامل مقدمات و توصیفی از رگرسیون آلوده به خطا است که در حالت خطی و در رده خطای اندازه‌گیری کلاسیک بررسی می‌شود.
- سرانجام، در فصل پنجم کاربردی از همگرایی کامل را ارایه می‌کنیم و سازگاری کامل برآوردگرهای کم‌ترین توان‌های دوم را در مدل رگرسیونی تبیینی آلوده به خطا با خطای وابسته زبرجمعی منفی بررسی می‌کنیم.

فهرست مطالب

فهرست تصاویر

فهرست جداول

فصل ۱

مفاهیم اولیه

در این فصل، برخی تعاریف و قضایای مورد نیاز برای بیان مطالب نظری اصلی پایان‌نامه را بیان می‌کنیم. برای گردآوری عمده مطالب این فصل از هو (۲۰۰۰)، گات (۱۹۹۴) و ماری و کاتز (۲۰۰۱) استفاده کرده‌ایم.

۱.۱ همگرایی

در این بخش چند تعریف از همگرایی‌ها و مطالب مرتبط را ذکر می‌کنیم و از آن‌ها در ادامه به دفعات استفاده خواهیم کرد.

تعریف ۱.۱.۱. دنباله‌ای از متغیرهای تصادفی $\{X_n, n \geq ۱\}$ به‌طور تقریبا حتمی[۱] به متغیر تصادفی X همگراست، هرگاه

$$P(\{\omega : X_n(\omega) \longrightarrow X(\omega); n \longrightarrow \infty\}) = ۱$$

یا به‌طور معادل

$$P(\{\omega : X_n(\omega) \nrightarrow X(\omega); n \longrightarrow \infty\}) = ۰$$

و با نماد $X_n \xrightarrow{a.s.} X$ نشان می‌دهیم.

همگرایی کامل[۲] نقش مهمی را در احتمال، قضایای حدی، آمار ریاضی و به‌ویژه در تعیین نسبت همگرایی ایفا می‌کند. مفهوم همگرایی کامل اولین بار توسط هسو و روبین (۱۹۴۷) بیان شد. در زیر تعریف همگرایی

[۱] Almost surely

[۲] Complete convergence

کامل را ارایه می‌کنیم.

تعریف ۲.۱.۱. دنباله‌ای از متغیرهای تصادفی $\{X_n, n \geq ۱\}$ به متغیر تصادفی X به‌طور کامل همگراست، اگر به ازای هر $\varepsilon > ۰$ داشته باشیم

$$\sum_{n=۱}^{\infty} P(|X_n - X| > \varepsilon) < \infty$$

و آن را با نماد $X_n \xrightarrow{c.} X$ نشان می‌دهیم.

قضیه ۳.۱.۱. (هسو و روبین، ۱۹۴۷) دنباله‌ای از میانگین حسابی متغیرهای تصادفی مستقل و هم‌توزیع به‌طور کامل به مقدار مورد انتظار همگرا خواهد بود و این در صورتی است که مجموع واریانس متناهی باشد.

در ادامه قضایا و لم‌هایی را از گات (۱۹۹۴)، بدون برهان، بیان می‌کنیم.

قضیه ۴.۱.۱. فرض کنید دنباله‌ای از متغیرهای تصادفی $\{X_n, n \geq ۱\}$ به متغیر تصادفی X به‌طور کامل همگرا باشد؛ در این صورت دنباله مورد نظر به متغیر تصادفی X به‌طور تقریبا حتمی نیز همگراست.

قضیه ۵.۱.۱. دنباله‌ای از متغیرهای تصادفی $\{X_n, n \geq ۱\}$ را در نظر بگیرید که به‌طور تقریبا حتمی به متغیر تصادفی X همگراست. این دنباله به‌طور کامل نیز به متغیر تصادفی X همگرا خواهد بود.

نتیجه ۶.۱.۱. با توجه به تعریف‌های فوق، همگرایی کامل، همگرایی تقریبا حتمی را نتیجه می‌دهد ولی عکس آن همیشه برقرار نیست. از طرفی، وقتی $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشد همگرایی تقریبا حتمی، همگرایی کامل را نتیجه می‌دهد.

لم ۷.۱.۱. **(لم‌های برل-کانتلی)**

۱. (**لم اول برل-کانتلی**[۳]) فرض کنید $\{A_n, n \geq ۱\}$ دنباله‌ای از پیشامدهای دلخواه باشند. آن‌گاه

$$\sum_{n=۱}^{\infty} P(A_n) < \infty \Longrightarrow P(A_n i.o.) = ۰$$

که در آن $i.o.$ به معنی تقریبا بی‌نهایت بار[۴] است.

۲. (**لم دوم برل-کانتلی**[۵]) فرض کنید $\{A_n, n \geq ۱\}$ دنباله‌ای از پیشامدهای مستقل باشند. بنابراین

$$\sum_{n=۱}^{\infty} P(A_n) = \infty \Longrightarrow P(A_n i.o.) = ۱.$$

برای اطلاعات بیش‌تر درباره همگرایی کامل و نتایج آن می‌توانید به ارداس (۱۹۴۹)، اسپیتز (۱۹۵۶)، بوم و کاتز (۱۹۶۵)، وو (۲۰۱۰)، وو (۲۰۱۲)، سانگ (۲۰۱۰ و ۲۰۱۲) وانگ و همکاران (۲۰۱۲) و شن و همکاران (۲۰۱۳) مراجعه کنید.

---

[۳] The first Borel-Cantelli lemma

[۴] Infinitely often

[۵] The second Borel-Cantelli lemma

تعریف ۸.۱.۱. (همگرایی در توزیع) فرض کنید $C(F_X) = \{x :$ تابع F_X در نقطه x پیوسته باشد $\}$. در این صورت $\{X_n\}$ در توزیع به X همگراست، هرگاه برای هر $x \in C(F_X)$، وقتی $n \longrightarrow \infty$، توزیع $F_{X_n}(x)$ به توزیع $F_X(x)$ میل کند که آن را با نماد $X_n \xrightarrow{d} X$ نشان می‌دهیم.

تعریف ۹.۱.۱. (همگرایی در احتمال) دنباله متغیرهای تصادفی $\{X_n; n \geq ۱\}$ در احتمال به متغیر تصادفی X همگراست، هرگاه برای هر $\varepsilon > ۰$، وقتی $n \longrightarrow \infty$، داشته باشیم

$$\lim_{n \longrightarrow \infty} P\left(|X_n - X| < \varepsilon\right) = ۱$$

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

قضیه ۱۰.۱.۱. (قانون قوی اعداد بزرگ[۶]) برای دنباله $\{X_n; n \geq ۱\}$ از متغیرهای تصادفی و متغیر تصادفی X

۱. اگر $E|X| < \infty$، $EX = \mu$ و $S_n = \sum\limits_{i=۱}^{n} X_i$، آن‌گاه
$$\bar{X}_n = \frac{S_n}{n} \xrightarrow{a.s.} \mu.$$

۲. اگر $\frac{S_n}{n} \xrightarrow{a.s.} C$ که در آن C یک مقدار ثابت است، آن‌گاه $EX = C$ و $E|X| < \infty$.

۳. اگر $E|X| = \infty$ آن‌گاه
$$\lim_{n \to \infty} \sup \frac{|S_n|}{n} = \infty.$$

۲.۱ نابرابری‌ها

نابرابری‌ها در نظریه احتمال نقش مهمی را ایفا می‌کنند. زیرا اگر مقدار دقیق احتمال یا گشتاوری قابل محاسبه نباشد، از نابرابری‌های موجود می‌توان استفاده نمود و کران‌هایی برای مقادیر مورد علاقه پیدا کرد. انتخاب یک نابرابری مناسب معمولا کلیدی جهت رفع مشکلات محسوب می‌شود. به عبارتی، هرچه کران انتخابی دقیق‌تر باشد جواب نهایی نیز قابل اعتمادتر خواهد بود. در زیر به برخی از نابرابری‌های قابل استفاده در این پایان‌نامه اشاره می‌کنیم. (گات، ۱۹۹۴)

۱. **نابرابری مارکف**[۷]: فرض کنید برای عدد حقیقی مثبت $r > ۰$ و متغیر تصادفی X، $E|X|^r < \infty$ باشد. آن‌گاه برای هر عدد حقیقی مثبت $x > ۰$
$$P(|X| > x) \leq \frac{E|X|^r}{x^r}.$$

۲. **نابرابری هولدر**[۸]: فرض کنید برای اعداد حقیقی $p > ۱$ و $q > ۱$، $\frac{۱}{p} + \frac{۱}{q} = ۱$ باشد. اگر $E|X|^p < \infty$ و $E|Y|^q < \infty$، آن‌گاه
$$E|XY| \leq (E|X|^p)^{\frac{۱}{p}} (E|Y|^q)^{\frac{۱}{q}}.$$

---

[۶] Strong law of large numbers

[۷] Markov inequality

[۸] Holder inequality

۳. **نابرابری کوشی شوارتز**[۹]: در نابرابری هولدر اگر $p = q = ۱$ باشد، آن‌گاه
$$E|XY|^۲ \leq (E|X|^۲)^{\frac{۱}{۲}}(E|Y|^۲)^{\frac{۱}{۲}}.$$

۴. **نابرابری کلموگروف**[۱۰]: فرض کنید $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل باشند به‌طوری‌که $E[X_j] = ۰$، $\sigma_j^۲ = E[X_j^۲] < \infty$ و $|X_j| \leq Cs_n$ که در آن $j = ۱, \ldots, n$ و $C > ۰$ مقدار ثابت حقیقی است و $s_n^۲ = \sum_{j=۱}^{n} \sigma_j^۲$. فرض کنید $x > ۰$. اگر $xC \leq ۱$، آن‌گاه
$$P(\frac{S_n}{s_n} \geq x) \leq exp\left(-\frac{x^۲}{۲}(۱ - \frac{xC}{۲})\right).$$
اگر $xC \geq ۱$، آن‌گاه
$$P(\frac{S_n}{s_n} \geq x) \leq exp\left(-\frac{x}{۴C}\right).$$

۵. **نابرابری لیاپانوف**[۱۱]: فرض کنید X متغیر تصادفی و s و r اعداد حقیقی باشند که $۰ \leq s \leq r$، آن‌گاه
$$(E|X|^s)^{\frac{۱}{s}} \leq (E|X|^r)^{\frac{۱}{r}}.$$

۶. **نابرابری** C_r[۱۲]: فرض کنید $\{X_n; n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشند. اگر عدد حقیقی مثبت $r > ۰$ وجود داشته باشد که به‌ازای $i = ۱, \ldots, n$، $E|X_i|^r < \infty$ باشد، آن‌گاه
$$E|\sum_{i=۱}^{n} X_i|^r \leq C_r \sum_{i=۱}^{n} E|X_i|^r$$
که C_r مقادیر زیر را اختیار می‌کند
$$C_r = \begin{cases} ۱ & \text{اگر } r \leq ۱ \\ ۲^{r-۱} & \text{اگر } r \geq ۱ \end{cases}$$

۳.۱ برخی از انواع وابستگی‌های متغیرهای تصادفی

در این بخش به معرفی برخی از انواع وابستگی می‌پردازیم که رده بزرگی از دنباله متغیرهای تصادفی وابسته را شامل می‌شوند.

۱.۳.۱ وابستگی پیوندی منفی

وابستگی پیوندی منفی[۱۳] اولین رده از وابستگی‌هاست که توسط جاج و پروشان (۱۹۸۳) ارایه شد.

---

[۹] Cauchy-Shwartz inequality

[۱۰] Kolmogorov inequality

[۱۱] Lyapunov inequality

[۱۲] C_r inequality

[۱۳] Negatively associated

تعریف ۱.۳.۱. به متغیرهای تصادفی $X_1, X_2, \dots, X_n$ پیوندی منفی گویند، اگر برای هر دو زیرمجموعه جدا از هم A_1 و A_2 از مجموعه $\{1, 2, \dots, n\}$ بتوان نتیجه گرفت

$$cov(f(X_i; i \in A_1), g(X_j; j \in A_2)) \leq 0 \tag{1.1}$$

که در آن f و g توابعی صعودی‌اند.

ملاحظه ۲.۳.۱. رابطه (۱.۱) برای توابع f و g نزولی نیز برقرار است، زیرا

$$\begin{aligned} cov(-f(X_i; i \in A_1), -g(X_j; j \in A_2)) &= cov(f(X_i; i \in A_1), g(X_j; j \in A_2)) \\ &\leq 0. \end{aligned}$$

برای اطلاعات بیش‌تر درباره وابستگی پیوندی منفی و ویژگی‌های آن به جاج و پروشان (۱۹۸۳) مراجعه کنید.

۲.۳.۱ وابستگی زبرجمعی منفی

رده دیگری از متغیرهای تصادفی وابسته، متغیرهای تصادفی وابسته زبر جمعی منفی[۱۴] (NSD) است. این نوع از وابستگی ابتدا توسط هو (۲۰۰۰) بیان شد که در پایه آن توابع زبرجمعی می‌باشد. قبل از بیان این نوع وابستگی به مفهوم زبرجمعی که توسط کمپرمن (۱۹۷۷) بیان شد، می‌پردازیم.

تعریف ۳.۳.۱. تابع $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ را زبرجمعی[۱۵] گویند اگر برای هر $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

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

که در آن $\vee$ ماکسیمم مؤلفه‌ای[۱۶] و $\wedge$ مینیمم مؤلفه‌ای[۱۷] است. به عبارتی دیگر

$$\mathbf{x} \vee \mathbf{y} = (x_1 \vee y_1, x_2 \vee y_2, \dots, x_n \vee y_n)$$

و

$$\mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, x_2 \wedge y_2, \dots, x_n \wedge y_n).$$

با توجه به تعریف ۳.۳.۱، کمپرمن (۱۹۷۷) روشی را بیان کرد که با استفاده از آن زبرجمعی بودن یا نبودن هر تابع را می‌توان مشخص نمود.

لم ۴.۳.۱. فرض کنید تابع ϕ دارای مشتق جزیی مرتبه دوم پیوسته باشد، آن‌گاه ϕ زبرجمعی است اگر

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

[۱۴] Negatively superadditive dependent

[۱۵] Superadditive

[۱۶] Maximum componentwise

[۱۷] Minimum componentwise

تعریف ۵.۳.۱. بردار تصادفی $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ وابسته زبرجمعی منفی است، اگر برای هر تابع زبرجمعی ϕ

$$E\phi(X_1, X_2, \ldots, X_n) \leq E\phi(X_1^*, X_2^*, \ldots, X_n^*)$$

که $X_1^*, X_2^*, \ldots, X_n^*$ متغیرهای تصادفی مستقل‌اند. همچنین برای هر i، X_i و X_i^* هم توزیع هستند.

لم ۶.۳.۱. (هو، ۲۰۰۰) فرض کنید بردار $(X_1, X_2, \ldots, X_n)$، NSD باشد. آن‌گاه

۱. $(-X_1, -X_2, \ldots, -X_n)$ نیز NSD هستند.

۲. برای توابع صعودی دلخواه $g_1, g_2, \ldots, g_n$، $(g_1(X_1), g_2(X_2), \ldots, g_n(X_n))$ نیز NSD هستند.

۳. برای هر $x_1, x_2, \ldots, x_n \in \mathbb{R}$، می‌توان نتیجه گرفت

$$P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \leq \prod_{i=1}^{n} P(X_i \leq x_i)$$

و

$$P(X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n) \leq \prod_{i=1}^{n} P(X_i > x_i).$$

برهان. با توجه به فرض لم، داریم

۱. از آن‌جا که بردار $(X_1, X_2, \ldots, X_n)$ NSD است و با توجه به لم ۴.۳.۱ کافی است نشان دهیم که برای هر تابع زبرجمعی ϕ و به ازای هر $i \neq j$، تابع $\phi(-x_1, -x_2, \ldots, -x_n)$ نیز زبرجمعی است. در نتیجه NSD بودن آن روشن است. زیرا

$$\begin{aligned} \frac{\partial^2 \phi(-x_1, -x_2, \ldots, -x_n)}{\partial x_i \partial x_j} &= \frac{\partial^2 \phi(-x_1, -x_2, \ldots, -x_n)}{\partial(-x_i)\partial(-x_j)} \\ &= \frac{\partial^2 \phi(x_1, x_2, \ldots, x_n)}{\partial x_i \partial x_j} \geq 0. \end{aligned}$$

۲. چون g تابعی صعودی است، با توجه به تعریف ۵.۳.۱ کافی است نشان دهیم برای هر تابع زبرجمعی ϕ و هر $i \neq j$، تابع $\phi\left(g_1(x_1), g_2(x_2), \ldots, g_n(x_n)\right)$ نیز زبرجمعی است. در نتیجه برهان کامل می‌شود، زیرا

$$\begin{aligned} \frac{\partial^2 \phi(g_1(x_1), g_2(x_2), \ldots, g_n(x_n))}{\partial x_i \partial x_j} &= \frac{\partial^2 \phi(g_1(x_1), g_2(x_2), \ldots, g_n(x_n))}{\partial g(x_i)\partial g(x_j)} \\ &\quad . \frac{\partial g(x_i)}{\partial x_i} . \frac{\partial g(x_j)}{\partial x_j} \\ &\geq 0. \end{aligned}$$

۳. با توجه به تعریف رابطه وابستگی زبرجمعی منفی، و همچنین بر اساس قسمت دوم این لم با قرار دادن تابع صعودی نشانگر به‌جای تابع صعودی g، می‌توان نشان داد

$$\begin{aligned} P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) &= E\left(I\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\}\right) \\ &= E\left(\prod_{i=1}^{n} I\{X_i \leq x_i\}\right) \\ &\leq E\left(\prod_{i=1}^{n} I\{X_i^* \leq x_i\}\right) \\ &= \prod_{i=1}^{n} E\left(I\{X_i^* \leq x_i\}\right) \\ &= \prod_{i=1}^{n} P(X_i \leqslant x_i). \end{aligned}$$

مسیر برهان نابرابری دوم بند ۳ در لم، مشابه برهان نابرابری اول است. لذا برهان کامل می‌شود.

□

لم ۷.۳.۱. (هو، ۲۰۰۰) فرض کنید $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ و $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ بردارهای تصادفی مستقل و NSD باشند، آن‌گاه $\mathbf{X} + \mathbf{Z} = (X_1 + Z_1, X_2 + Z_2, \ldots, X_n + Z_n)$ نیز NSD هستند.

نتیجه ۸.۳.۱. فرض کنید $(X_1, X_2, \ldots, X_n)$ بردار تصادفی NSD و توابع $g_1, g_2, \ldots, g_n$ صعودی (یا نزولی) باشند، آن‌گاه با توجه به لم ۶.۳.۱، بردارهای

$$(g_1(X_1), g_2(X_2), \ldots, g_n(X_n))$$

و

$$(g_1(-X_1), g_2(-X_2), \ldots, g_n(-X_n))$$

NSD خواهند بود.

لم ۹.۳.۱. فرض کنید $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ و $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ بردارهای تصادفی مستقل باشند. اگر بردارهای $\mathbf{X}$ و $\mathbf{Z}$، NSD باشند، آن‌گاه بردار $\mathbf{X} + \beta\mathbf{Z}$ برای هر $\beta \in \mathbb{R}$، NSD است.

برهان. برای حالت $\beta = 0$ نتیجه برقرار است. برای $\beta > 0$ با استفاده از لم ۶.۳.۱ بردار $(\beta Z_1, \ldots, \beta Z_n)$ NSD و مستقل از بردار $\mathbf{X}$ است. با لم ۷.۳.۱ ثابت می‌شود که بردار تصادفی $\mathbf{X}+\beta\mathbf{Z}$، NSD خواهد بود. □

در زیر برخی ویژگی‌های متغیرهای تصادفی وابسته زبرجمعی منفی را از هو (۲۰۰۰)، بدون برهان، بیان می‌کنیم.

ملاحظه ۱۰.۳.۱. دنباله $\{X_n, n \geq ۱\}$ را دنباله‌ای از متغیرهای تصادفی NSD گویند، اگر برای هر $n \geq ۱$، بردار تصادفی $(X_۱, \dots, X_n)$، NSD باشد.

ملاحظه ۱۱.۳.۱. آرایه $\{X_{ni}, i \geq ۱, n \geq ۱\}$ را آرایه‌ای از متغیرهای تصادفی وابسته زبرجمعی منفی سطری گویند، اگر برای هر $n \geq ۱$، $\{X_{ni}, i \geq ۱\}$، وابسته زبرجمعی منفی باشد.

ملاحظه ۱۲.۳.۱. وابستگی پیوندی منفی یک زوج متغیر تصادفی با وابستگی زبرجمعی منفی آن معادل است.

لم ۱۳.۳.۱. (کریستوفیدز و واگلاتو، ۲۰۰۴) اگر $X_۱, X_۲, \dots, X_n$ پیوندی منفی باشند، آن‌گاه وابسته زبرجمعی منفی هستند ولی عکس آن همیشه برقرار نیست.

۳.۳.۱ مثال‌ها و کاربردهایی از وابسته زبرجمعی منفی

خانواده‌ای از توزیع‌های چندمتغیره با ویژگی NSD

- خانواده توزیع‌های بیضی‌گون: برای این منظور، در ابتدا به تعریفی از خانواده توزیع‌های بیضی‌گون می‌پردازیم.

تعریف ۱۴.۳.۱. (بولک و سامپسون، ۱۹۸۸) خانواده $F(x, \Sigma)$ خانواده‌ای از توزیع‌های بیضی‌گون[۱۸] پایه‌گذاری شده بر اساس تابع g است، هرگاه تابع چگالی آن به صورت زیر تعریف شود:

$$f(x, \Sigma) = |\Sigma|^{-\frac{۱}{۲}} g(x^T \Sigma x) \qquad x \in \mathcal{R}^m$$

که در آن $\int_۰^\infty g(t) t^{\frac{m}{۲}-۱} dt < \infty$.

بولک و سامپسون (۱۹۸۸) نشان دادند که $F(x, \Sigma)$ دارای وابستگی از نوع NSD خواهد بود اگر برای هر $i \neq j$، $\delta_{ij} \leq ۰$ باشد به‌طوری‌که δ_{ij}ها مؤلفه‌های ماتریس Σ هستند. در این میان، توزیع نرمال چندمتغیره نیز به عنوان حالت خاصی از توزیع‌های بیضی‌گون محسوب می‌شود. بنابراین توزیع نرمال چندمتغیره NSD خواهد بود اگر عناصر خارج از قطر اصلی ماتریس کواریانس آن نامثبت باشند.

- خانواده توابع فراوانی پولیای نوع ۲[۱۹]: به جهت آشنایی با این دست از توزیع‌ها، تعریفی از این را ارایه می‌کنیم.

تعریف ۱۵.۳.۱. (افرون، ۱۹۶۵) تابع چگالی احتمال روی خط حقیقی $r(x)$ را تابع فراوانی پولیای نوع ۲ $(PF_۲)$ نامند، هرگاه به ازای $x_۲ \geq x_۱$ و $z_۲ \geq z_۱$ داشته باشیم

$$\begin{vmatrix} r(x_۱ - z_۱) & r(x_۱ - z_۲) \\ r(x_۲ - z_۱) & r(x_۲ - z_۲) \end{vmatrix} \geq ۰.$$

این خانواده از توزیع‌ها دارای ویژگی‌هایی است که در زیر به‌طور مختصر به آن‌ها اشاره می‌نماییم.

---

[۱۸] Elliptically contoured distribution

[۱۹] Polya frequency function of order 2

آ. تعداد زیادی از توزیع‌های رایج در نظریه احتمال، $PF_۲$ هستند که به‌عنوان حالت خاص می‌توان به توزیع‌های پواسون، دوجمله‌ای، گاما، دوجمله‌ای منفی و همچنین توزیع نرمال با تابع چگالی

$$r(x) = (۲\pi\sigma^۲)^{-\frac{۱}{۲}} \exp\left\{-\frac{۱}{۲}\left[\frac{x-m}{\sigma}\right]^۲\right\}$$

با $\sigma > ۰$، اشاره کرد.

ب. توابع چگالی $PF_۲$ به‌طور لگاریتمی مقعرند[۲۰]. بنابراین تقریبا همه جا دارای مشتق اول و دوم هستند.

پ. توابع چگالی $PF_۲$ کران‌دارند و همچنین بر روی یک فاصله، غیرصفر و در خارج از آن، صفر هستند.

ت. فرض کنید $r(x)$، یک $PF_۲$ باشد، آن‌گاه

$$r(x) = \begin{cases} \dfrac{r(x)}{\int_a^b r(z)dz} & a \leq x \leq b \\ ۰ & \text{سایر نقاط} \end{cases}$$

نیز $PF_۲$ است (a و b می‌توانند نامتناهی باشند).

ث. فرض کنید $r(x)$ و $q(x)$، $PF_۲$ باشند آن‌گاه پیچش آن‌ها با شکل

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

نیز $PF_۲$ است.

بنابراین توزیع‌هایی مانند پواسون، دوجمله‌ای، گاما و دوجمله‌ای منفی دارای ویژگی‌های $PF_۲$ هستند.

بر اساس آن‌چه هو (۲۰۰۰) ارایه کرده است، اگر $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل با تابع چگالی احتمال $PF_۲$ باشند، $S_n = \sum_{i=۱}^{n} X_i$ به‌طور تقریبا حتمی NSD است. لذا با توجه به مطالب فوق، توزیع‌های زیر نیز NSD خواهند بود.

۱. (**توزیع چندجمله‌ای**[۲۱]) در این توزیع بردار $\mathbf{X} = (X_۱, X_۲, \ldots, X_n)$ دارای تابع احتمال

$$P(X_۱ = x_۱, \ldots, X_n = x_n) = \frac{N!}{x_۱! \ldots x_n!}\prod_{i=۱}^{n} p_i^{x_i} \qquad \forall\ x_i \geq ۰,\quad \sum_{i=۱}^{n} x_i = N$$

است، که در آن برای هر $i = ۱, ۲, \ldots, n$، $p_i \geq ۰$ و $\sum_{i=۱}^{n} p_i = ۱$. لازم به ذکر است که توزیع چندجمله‌ای از احتمال شرطی متغیرهای تصادفی مستقل با توزیع پواسون به شرط مجموع این متغیرهای تصادفی حاصل می‌شود.

[۲۰] Log-concave

[۲۱] Multinomial distribution

۲. (**توزیع فوق هندسی چندمتغیره**[۲۲]) در این توزیع بردار $\mathbf{X}$ دارای تابع احتمال زیر با پارامترهایی با مقادیر صحیح مثبت $(N, M_1, \dots, M_n)$ است:

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} \frac{\binom{M_i}{x_i}}{\binom{M}{N}}$$

که در آن $x_i \geq 0$، $\sum_{i=1}^{n} x_i = N$ و همچنین $\sum_{i=1}^{n} M_i = M$. این توزیع نیز از احتمال شرطی متغیرهای تصادفی مستقل با توزیع دوجمله‌ای به شرط مجموع متغیرهای تصادفی مستقل به‌دست می‌آید.

۳. (**توزیع دیریکله**[۲۳]) تابع چگالی احتمال دیریکله برای $\mathbf{X}$ به‌صورت:

$$f(x_1, \dots, x_n) = \frac{\Gamma\left(\sum_{i=1}^{n} \theta_i\right)}{\prod_{i=1}^{n} \Gamma(\theta_i)} \prod_{i=1}^{n} x_i^{\theta_i - 1}$$

است، که در آن برای $i = 1, 2, \dots, n$، $x_i \geq 0$، $\sum_{i=1}^{n} x_i = 1$ و $\theta_i \geq 1$. ذکر این یادآوری ضروری است که احتمال شرطی متغیرهای تصادفی مستقل با توزیع گاما به شرط مجموع‌شان، توزیع دیریکله را نتیجه می‌دهد.

۴. (**توزیع چندجمله‌ای مرکب دیریکله**[۲۴]) در این توزیع بردار $\mathbf{X}$ دارای تابع احتمال

$$P(X_1 = x_1, \dots, X_n = x_n) = \frac{N!\Gamma\left(\sum_{i=1}^{n} \theta_i\right)}{\Gamma\left(N + \sum_{i=1}^{n} \theta_i\right)} \prod_{i=1}^{n} \frac{\Gamma(x_i + \theta_i)}{x_i!\Gamma(\theta_i)}$$

است، که در آن برای $i = 1, 2, \dots, n$، $\theta_i \geq 1$ و $x_i \geq 1$ و برای عدد صحیح و مثبت N، $\sum_{i=1}^{n} x_i = N$ است. توزیع چندجمله‌ای مرکب دیریکله نیز به مانند دیگر توزیع‌ها، از احتمال شرطی متغیرهای تصادفی مستقل با توزیع دوجمله‌ای منفی[۲۵] به شرط مجموع متغیرهای تصادفی نیز ساخته می‌شود.

- توزیع جایگشتی[۲۶]: توزیع $\mathbf{X} = (X_1, X_2, \dots, X_n)$ یک توزیع جایگشتی روی $\mathbf{x} = (x_1, x_2, \dots, x_n)$ است اگر X همه مقادیر $n!$ جایگشت $\mathbf{x}$ را با احتمال $\frac{1}{n!}$ بگیرد که برای هر $n > 1$، $x_1, x_2, \dots, x_n$ اعداد حقیقی‌اند. چنین توزیع‌هایی دارای ویژگی NSD هستند. برای مثال، می‌توان به نمونه تصادفی بدون جای‌گذاری اشاره کرد.

فرض کنید $X_1, X_2, \dots, X_n$ نمونه‌ای تصادفی، بدون جای‌گذاری از یک جامعه متناهی باشد که شامل مقادیر $x_1, x_2, \dots, x_N$ است. در این صورت برای هر $n < N$، $(X_1, X_2, \dots, X_n)$ زیرمجموعه $(X_1, X_2, \dots, X_N)$ بوده که یک توزیع جایگشتی دارد. بنابراین نمونه تصادفی بدون جای‌گذاری، NSD خواهد بود.

---

[۲۲] Multivariate hypergeometric distribution

[۲۳] Dirichlet distribution

[۲۴] Dirichlet compound multinomial distribution

[۲۵] Negative binomial

[۲۶] Permutation distribution

نابرابری‌هایی با ویژگی NSD

از دیگر کاربردهای مفهوم NSD، می‌توان به نابرابری ماکسیمال اشاره نمود. برای پرداختن به این مطلب، لازم است در ابتدا یک تابع زبرجمعی خاص را معرفی می‌کنیم. به همین منظور تعریفی را ارایه می‌کنیم. (هو، ۲۰۰۰)

تعریف ۱.۳.۱۶. بردار $\mathbf{a} = (a_1, a_2, \dots, a_n)$ به وسیله بردار $\mathbf{b} = (b_1, b_2, \dots, b_n)$ بیشنه می‌شود و با نماد $(\mathbf{b} \succ^n \mathbf{a})$ نشان داده می‌شود، اگر $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$ و برای هر $k = 1, \dots, n$ رابطه $\sum_{i=1}^{k} a_{(i)} \geq \sum_{i=1}^{k} b_{(i)}$ برقرار باشد که در آن $a_{(i)}$ آماره ترتیبی iام بردار $\mathbf{a}$ است.

قضیه ۱.۳.۱۷. برای هر تابع صعودی و محدب $f : \mathbb{R} \longrightarrow \mathbb{R}$، تابع $\phi_1(x) = f\left(\max_{1\leq k\leq n} \sum_{i=1}^{k} x_i\right)$ زبرجمعی است.

برهان. با توجه به تعریف توابع زبرجمعی، کافی است نشان دهیم که برای هر $1 \leq p < q \leq n$، تابع $\phi_1(x_p, x_q; x^{(p,q)})$ یک تابع زبرجمعی است که در آن $x^{(p,q)}$ نمادی برای نشان دادن انتخاب دو متغیر x_p و x_q با ثابت بودن متغیرهای تصادفی دیگر معرفی می‌شود.

$$a = \max_{1\leq k\leq n} \sum_{i=1}^{k} x_i \qquad\qquad u(x_q) = \sum_{i=1}^{p-1} x_i + \max_{1\leq k\leq n} \left\{ 0, \sum_{i=p+1}^{k} x_i \right\}.$$

در این‌جا لازم است به این نکته توجه شود که برای هر s, t

$$\max\{s, s+t\} = s + \max\{0, t\}.$$

در نتیجه

$$\begin{aligned}
\phi_1(x_p, x_q; x^{(p,q)}) &= f\left(\max\left\{\max_{1\leq k\leq p-1} \sum_{i=1}^{k} x_i,\ \max_{1\leq k\leq n} \sum_{i=1}^{k} x_i\right\}\right) \\
&= f\left(\max\left\{a, x_p + \sum_{i=1}^{p-1} x_i + \max_{p+1\leq k\leq n} \left\{0, \sum_{i=p+1}^{k} x_i\right\}\right\}\right) \\
&= f\left(\max\left\{a, x_p + u(x_q)\right\}\right).
\end{aligned}$$

چون f و $\max\{a, x\}$ محدب و f صعودی است، بنابراین $f(\max\{a, x\})$ نیز محدب است. برای هر $x_p < x_q^*$ و $x_q < x_q^*$ می‌توان نشان داد

$$\left(x_q + u(x_q), x_p^* + u(x_q^*)\right) \succ^n \left(x_p + u(x_q^*), x_p^* + u(x_q)\right)$$

که نماد $\succ^n$ در زیر تعریف شده است.

حال از آن‌جا که $u(x_q)$ در x_q صعودی است، بنابراین

$$\phi_1(x_p, x_q; x^{(p,q)}) + \phi_1(x_p^*, x_q^*; x^{(p,q)}) \geq \phi_1(x_p, x_q^*; x^{(p,q)}) + \phi_1(x_p^*, x_q; x^{(p,q)}).$$

رابطه بالا نشان می‌دهد که ϕ یک تابع زبرجمعی است. □

با توجه به معرفی تابع زبرجمعی فوق، این امکان به‌وجود می‌آید که برای یک بردار تصادفی وابسته زبرجمعی منفی $\mathbf{X}$ نابرابری‌های مفیدی ارایه شود.

فرض کنید $\mathbf{X} = (X_۱, X_۲, \ldots, X_n)$ یک بردار NSD و $\mathbf{X}^* = (X_۱^*, X_۲^*, \ldots, X_n^*)$ یک بردار تصادفی مستقل باشند که برای هر $i = ۱, ۲, \ldots, n$، X_i و X_i^* هم‌توزیع‌اند. در این صورت برای هر تابع صعودی و محدب f

$$E\left[f\left(\max_{۱\leq k\leq n}\sum_{i=۱}^{k} X_i\right)\right] \leq E\left[f\left(\max_{۱\leq k\leq n}\sum_{i=۱}^{k} X_i^*\right)\right]. \qquad (۲.۱)$$

این نابرابری را می‌توان با استفاده از $\phi = \phi_۱$ به‌دست آورد.

نابرابری فوق نشان می‌دهد که گشتاورهای مجموع جزیی (یا ماکسیمم مجموع جزیی) متغیرهای تصادفی NSD می‌توانند به‌وسیله متغیرهای تصادفی مستقل کران‌دار شوند. از نابرابری (۲.۱) می‌توان نابرابری کلموگروف برای متغیرهای تصادفی NSD را به‌صورت زیر تعمیم داد.

نتیجه ۱۸.۳.۱. فرض کنید $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی NSD با میانگین صفر و گشتاور دوم متناهی باشند. برای هر $k \geq ۱$، قرار می‌دهیم $S_k = \sum\limits_{i=۱}^{k} X_i$. بنابراین به ازای هر $\varepsilon > ۰$

$$P\left(\max_{۱\leq k\leq n}|S_k| > \varepsilon\right) \leq \frac{۸}{\varepsilon^۲}\sum_{i=۱}^{n} Var(X_i).$$

برهان. از آن‌جا که $\max\limits_{۱\leq k\leq n}|S_k| \leq \max(۰, S_۱, S_۲, \ldots, S_n) + \max(۰, -S_۱, -S_۲, \ldots, -S_n)$، خواهیم داشت

$$\begin{aligned}P\left(\max_{۱\leq k\leq n}|S_k| > \varepsilon\right) &\leq P\left(\max(۰, S_۱, S_۲, \ldots, S_n) > \frac{\varepsilon}{۲}\right)\\ &+P\left(\max(۰, -S_۱, -S_۲, \ldots, -S_n) > \frac{\varepsilon}{۲}\right).\end{aligned}$$

بر اساس نابرابری مارکف، نابرابری فوق به‌صورت زیر تغییر می‌یابد:

$$\begin{aligned}P\left(\max_{۱\leq k\leq n}|S_k|>\varepsilon\right) &\leq \frac{E\left[\max(۰, S_۱, S_۲, \ldots, S_n)\right]^۲}{\frac{\varepsilon^۲}{۴}}\\ &+\frac{E\left[\max(۰, -S_۱, -S_۲, \ldots, -S_n)\right]^۲}{\frac{\varepsilon^۲}{۴}}\\ &= ۴\varepsilon^{-۲}\left\{E\left[\max(S_۱, S_۲, \ldots, S_n)\right]^۲\right.\\ &\left.+E\left[\max(-S_۱, -S_۲, \ldots, -S_n)\right]^۲\right\}. \qquad (۳.۱)\end{aligned}$$

حال با استفاده از (۲.۱) و با در نظر گرفتن $f(x)=x^2$ می‌توان رابطه زیر را به‌دست آورد:

$$\begin{aligned}E\left[\max_{1\le k\le n} S_k\right]^2 &\le E\left[\max_{1\le k\le n}\sum_{i=1}^{k} X_i^*\right]^2\\ &\le \sum_{i=1}^{n} E(X_i^*)^2\\ &= \sum_{i=1}^{n} EX_i^2\\ &= \sum_{i=1}^{n} Var(X_i). \qquad (۴.۱)\end{aligned}$$

تساوی آخر در عبارت فوق از رابطه هم‌توزیع بودن X_i و X_i^* به‌دست آمده است.

برای $E[max(-S_1,-S_2,\dots,-S_n)]^2$ نیز می‌توان نتیجه مشابه را ثابت کرد، زیرا با توجه به این‌که $(X_1,X_2,\dots,X_n)$ برداری از متغیرهای تصادفی NSD است، در نتیجه $(-X_1,-X_2,\dots,-X_n)$ نیز NSD خواهد بود. بنابراین

$$\begin{aligned}E[\max(-S_1,-S_2,\dots,-S_n)]^2 &\le \sum_{i=1}^{n} EX_i^2\\ &= \sum_{i=1}^{n} Var(X_i). \qquad (۵.۱)\end{aligned}$$

در پایان با جای‌گذاری (۴.۱) و (۵.۱) در رابطه (۳.۱) داریم

$$\begin{aligned}P\left(\max_{1\le k\le n}|S_k|>\varepsilon\right) &\le 4\varepsilon^2\left\{\sum_{i=1}^{n} Var(X_i)+\sum_{i=1}^{n} Var(X_i)\right\}\\ &= \frac{8}{\varepsilon^2}\sum_{i=1}^{n} Var(X_i)\end{aligned}$$

و برهان کامل می‌شود. □

۴.۱ برخی دیگر از مفاهیم وابستگی

در این بخش برخی دیگر از مفاهیم وابستگی و ارتباط بین آن‌ها را از ماری و کاتز (۲۰۰۱)، به‌طور مختصر، بیان می‌کنیم.

تعریف ۱.۴.۱. دو متغیر تصادفی X و Y را وابسته ربعی منفی[۲۷] (NQD) گویند، هرگاه به ازای اعداد حقیقی دلخواه x و y داشته باشیم

$$P(X\le x, Y\le y)\le P(X\le x)P(Y\le y)$$

[۲۷] Negative quadrant dependent

یا
$$P\left(X>x,Y>y\right)\leq P\left(X>x\right)P\left(Y>y\right).$$
این تعریف معادل است با این‌که، دومتغیر تصادفی وابسته ربعی منفی هستند اگر $Cov(X,Y)\leq ۰$.

تعریف ۲.۴.۱. دنباله $\{X_n; n\geq ۱\}$ از متغیرهای تصادفی را وابسته ربعی منفی خطی توسعه‌یافته[۲۸] (ELNQD) نامند اگر برای هر زیر مجموعه جدا از هم A و B از $\{۱,۲,\ldots,n\}$ و r_jها مثبت حقیقی، دو مجموع $\sum_{i\in A} r_iX_i$ و $\sum_{j\in B} r_jX_j$ متغیرهای تصادفی NQD باشند.

تعریف ۳.۴.۱. دو متغیر تصادفی X و Y را وابسته ربعی منفی توسعه‌یافته[۲۹] (ENQD) می‌نامند، هرگاه ثابتی حقیقی مانند $M>۰$ موجود باشد به‌طوری‌که به ازای همه اعداد حقیقی x و y
$$P\left(X\leq x,Y\leq y\right)\leq MP\left(X\leq x\right)P\left(Y\leq y\right)$$
یا
$$P\left(X>x,Y>y\right)\leq MP\left(X>x\right)P\left(Y>y\right).$$
حال فرض کنید $\mathbf{X}=(X_۱,X_۲,\ldots,X_n)$ برداری از متغیرهای تصادفی باشد، آن‌گاه تعاریف زیر را ارایه می‌کنیم.

تعریف ۴.۴.۱. بردار $\mathbf{X}$ را وابسته متعامد منفی قوی[۳۰] (SNOD) نامند، اگر برای هر مجموعه اندیس‌گذار A و برای هر $\mathbf{x}\in\mathbb{R}^n$
$$P\left(X_i>x_i, i=۱,۲,\ldots,n\right)\leq P\left(X_i>x_i, i\in A\right)P\left(X_j>x_j, j\in A^c\right)$$
یا
$$P\left(X_i\leq x_i, i=۱,۲,\ldots,n\right)\leq P\left(X_i\leq x_i, i\in A\right)P\left(X_j\leq x_j, j\in A^c\right).$$

تعریف ۵.۴.۱. بردار $\mathbf{X}$ را وابسته متعامد منفی فوقانی[۳۱] (NUOD) نامند، اگر برای هر $\mathbf{x}\in\mathbb{R}^n$
$$P\left(X_i>x_i, i=۱,۲,\ldots,n\right)\leq\prod_{i=۱}^{n}P(X_i>x_i).$$

تعریف ۶.۴.۱. بردار $\mathbf{X}$ را وابسته متعامد منفی تحتانی[۳۲] (NLOD) نامند، اگر برای هر $\mathbf{x}\in\mathbb{R}^n$
$$P\left(X_i\leq x_i, i=۱,۲,\ldots,n\right)\leq\prod_{i=۱}^{n}P(X_i\leq x_i).$$

نتیجه ۷.۴.۱. بردار $\mathbf{X}$ را وابسته متعامد منفی گویند، اگر هم NUOD و هم NLOD باشد.

---
[۲۸] Extended Linear negative quadrant dependent
[۲۹] Extended negative quadrant dependent
[۳۰] Strongly negative orthant dependent
[۳۱] Negatively upper orthant dependent
[۳۲] Negatively lower orthant dependent

ذکر این نکته ضروری است که تعاریف بالا برای حالت وابستگی منفی ارایه شده‌اند و به‌طور مشابه برای حالت وابستگی مثبت نیز قابل تعریف هستند.

حال می‌توانیم روابط میان وابستگی‌ها را به‌طور خلاصه به صورت زیر بیان کنیم:

$$NA \begin{array}{c}\Longrightarrow \quad SNOD \quad \Longrightarrow \\ \Longrightarrow \quad NSD \quad \Longrightarrow\end{array} NOD \begin{array}{c}\Longrightarrow \quad NUOD \\ \Longrightarrow \quad NLOD\end{array}$$

ملاحظه ۸.۴.۱. برای یک زوج متغیر تصادفی NQD با NA معادل است، لذا روابطی که میان NA و NSD برقرار هستند، میان NSD و NQD نیز برقرار خواهند بود.

۵.۱ ترتیب‌های تصادفی و غیرتصادفی

نمادهای ترتیبی تصادفی O_p و o_p و ترتیب‌های غیرتصادفی O و o، رایج‌ترین نمادها برای توصیف مرتبه مجانبی کمیت‌های تصادفی و غیرتصادفی هستند. ون در وارت (۱۹۹۸) تعاریفی را از ترتیب‌های تصادفی و غیرتصادفی با توجه به تعریف همگرایی در احتمال به همراه نمادهای مناسب آن‌ها ارایه کرده است که در ادامه به بیان آن‌ها می‌پردازیم.

تعریف ۱.۵.۱. یک دنباله از متغیرهای تصادفی $\{X_n; n \geq ۱\}$ و دنباله‌ای از ثابت‌های $\{a_n; n \geq ۱\}$ را در نظر بگیرید. گوییم $X_n = o_p(a_n)$، اگر وقتی $n \longrightarrow \infty$، آن‌گاه

$$\frac{X_n}{a_n} \xrightarrow{p} ۰.$$

تعریف ۲.۵.۱. یک دنباله از متغیرهای تصادفی $\{X_n; n \geq ۱\}$ و دنباله‌ای از ثابت‌های $\{a_n; n \geq ۱\}$ را در نظر بگیرید. گوییم $X_n = O_p(a_n)$، اگر برای هر $\varepsilon > ۰$، $K(\varepsilon) > ۰$ و $n_۰(\varepsilon)$ وجود داشته باشند به‌طوری‌که به ازای همه مقادیر $n > n_۰(\varepsilon)$،

$$P\left(\left|\frac{X_n}{a_n}\right| \leq K(\varepsilon)\right) > ۱ - \varepsilon.$$

تعریف ۳.۵.۱. دو دنباله از اعداد ثابت $\{a_n; n \geq ۱\}$ و $\{b_n; n \geq ۱\}$ را در نظر بگیرید. گوییم $b_n = O(a_n)$، هرگاه برای هر $\varepsilon > ۰$، $K(\varepsilon) > ۰$ و عدد صحیح $N(\varepsilon)$ وجود داشته باشند به‌طوری‌که برای هر $n \geq N(\varepsilon)$،

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

تعریف ۴.۵.۱. دو دنباله از اعداد ثابت $\{a_n; n \geq ۱\}$ و $\{b_n; n \geq ۱\}$ را در نظر بگیرید. گوییم $b_n = o(a_n)$، هرگاه

$$\lim_{n \longrightarrow \infty} \left|\frac{b_n}{a_n}\right| = ۰.$$

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 $Ca_nO_p(۱)$، $Ca_no_p(۱)$، $Ca_nO(۱)$ و $Ca_no(۱)$ هستند. همچنین $X_n = O_p(n^C)$،

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	$O(n^a)o_p(n^b) = o_p(n^{a+b})$		
۱۰	$O(n^a)O_p(n^b) = O_p(n^{a+b})$		

برقراری $X_n = O_p(n^{C+۱})$ را نتیجه می‌دهد، اما $X_n = O_p(n^{C+۱})$ لزوما به این معنی نیست که رابطه $X_n = O_p(n^C)$ برقرار است.

در جدول ۱.۱ به برخی از ویژگی‌های ترتیب‌های تصادفی و غیرتصادفی اشاره شده که از پایان‌نامه اسحاقی (۱۳۹۲) استخراج 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 (مشابه کار اسحاقی، ۱۳۹۲)، 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

$$\left|\frac{b_n}{a_n}\right| < M. \qquad (۶.۱)$$

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

$$\left|\frac{a_n}{b_n}\right| = \left|\frac{\frac{a_n}{n^a}}{\frac{b_n}{n^a}}\right| < M'. \qquad (۷.۱)$$

از طرفی بنابر (۶.۱) و (۷.۱) که این همان تعریف ترتیب غیرتصادفی $O(\cdot)$ 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:

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

۶.۱ غلبه تصادفی

مفهوم غلبه تصادفی[۳۳] در آمار نقش مهمی را ایفا می‌کند و به توصیف مجموعه ارتباطاتی که بین دو متغیر یا دو توزیع وجود دارند، می‌پردازد. در بسیاری از موارد، محققان اطلاعات محدودی درباره متغیر و یا توزیع مورد نظر در اختیار دارند. در این حالت غلبه تصادفی می‌تواند نقش خود را نشان دهد و اطلاعات تکمیلی دیگری را برای انتخاب مدل مناسب ارایه نماید. به دلیل فهم هرچه بهتر غلبه تصادفی، در ابتدا تعاریفی را بیان می‌کنیم.

تعریف ۱.۶.۱. فرض کنید X و Y متغیرهای تصادفی با توابع توزیع تجمعی به‌ترتیب F و G باشند. گوییم توزیع X روی توزیع Y غلبه تصادفی نوع اول را داراست، اگر و فقط اگر به ازای هر $z \in \mathbb{R}$

$$P\left(X \geq z\right) \geq P\left(Y \geq z\right)$$

یا به‌طور معادل

$$\bar{F}_X(z) \geq \bar{G}_Y(z)$$

که در آن $\bar{F}_X(z) = ۱ - F_X(z)$ و $\bar{G}_Y(z) = ۱ - G_Y(z)$. همچنین می‌توان نوشت

$$F_X(z) \leq G_Y(z).$$

لم ۲.۶.۱. (ولفستتر، ۱۹۹۶) اگر متغیر تصادفی X بر متغیر تصادفی Y غلبه تصادفی داشته باشد، آن‌گاه $EX \geq EY$.

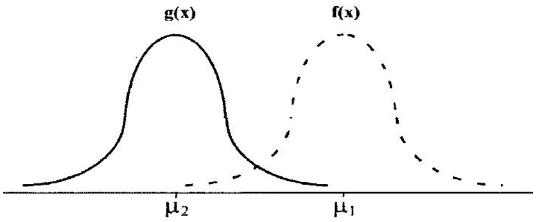
برای مثال، دو متغیر تصادفی X و Y را در نظر بگیرید که دارای توزیع نرمال

$$X \sim N(\mu_۱, \sigma_۱) \qquad\qquad Y \sim N(\mu_۲, \sigma_۲)$$

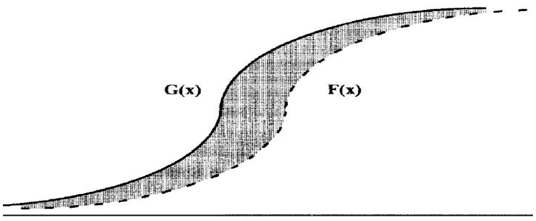
با توابع چگالی احتمال به‌ترتیب $f(\cdot)$ و $g(\cdot)$ هستند (شکل ۱.۱). متغیر تصادفی X روی متغیر تصادفی Y غلبه دارد، اگر و تنها اگر

۱. $\mu_۱ > \mu_۲$

۲. $\sigma_۱ = \sigma_۲$



شکل ۱.۱: تابع چگالی دو متغیر تصادفی نرمال وقتی $\mu_۱ > \mu_۲$



شکل ۲.۱: تابع توزیع تجمعی دو متغیر تصادفی نرمال وقتی $\mu_۱ > \mu_۲$

این موضوع را می‌توان در شکل ۲.۱ نیز مشاهده نمود. لازم به ذکر است که غلبه تصادفی انواع متفاوتی دارد که در این‌جا به مختصری از آن اشاره کردیم. برای اطلاعات بیش‌تر می‌توانید به لوای (۲۰۰۶) مراجعه کنید.

تعریف ۳.۶.۱. (وو، ۲۰۰۶) دنباله $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی است که متغیر تصادفی X برآن غلبه تصادفی دارد، اگر ثابت مثبت C موجود باشد به‌طوری‌که برای هر $x \geq ۰$ و $n \geq ۱$،

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

لم ۴.۶.۱. (وو، ۲۰۰۶) فرض کنید $\{X_n, n \geq ۱\}$ آرایه‌ای از متغیرهای تصادفی باشد که متغیر تصادفی X بر آن غلبه تصادفی دارد. آن‌گاه برای هر $\alpha > ۰$ و $b > ۰$ رابطه زیر برقرار است:

$$\begin{aligned} E|X_n|^\alpha I(|X_n| \leq b) &\leq C_۱ [E|X|^\alpha I(|X| \leq b) + b^\alpha P(|X| > b)] \\ E|X_n|^\alpha I(|X_n| > b) &\leq C_۲ E|X|^\alpha I(|X| > b) \end{aligned}$$

که در آن $C_۱$ و $C_۲$ ثابت مثبت هستند. در نتیجه، $E|X_n|^\alpha \leq CE|X|^\alpha$، که C ثابت مثبت است.

در این پایان‌نامه اگر در جایی از ثابتی مانند C سخن به میان آمد، منظور ثابتی است که به n وابسته نمی‌باشد و البته ممکن است در بسیاری از قسمت‌ها متفاوت باشد. همچنین فرض کنید $a_n = O(b_n)$ برقرار باشد، این بدان معناست که برای هر $n \geq ۱$ رابطه $a_n \leq Cb_n$ صدق می‌کند. همچنین تعریف می‌کنیم $\log x = \ln \max(x, e)$، $x^+ = xI(x \geq ۰)$ و $x_- = -xI(x < ۰)$ که در آن e عدد نپر است.

[۳۳] Stochastic dominance

فصل ۲

نابرابری‌های نمایی و ماکسیمال در حالت خاصی از وابستگی

۱.۲ نابرابری ماکسیمال روزنتال

در نابرابری‌های ماکسیمال، یک کران بالا برای احتمال دم ماکسیمم دنباله‌ای از متغیرهای تصادفی به‌دست می‌آید. از این رو، برای بررسی همگرایی دنباله‌ای از متغیرهای تصادفی، استفاده از این نابرابری یک روش عمومی محسوب می‌شود (گات، ۱۹۹۴). از جمله نابرابری‌های ماکسیمال مهم، می‌توان به نابرابری ماکسیمال روزنتال[۱] اشاره نمود. برای اولین بار روزنتال (۱۹۷۰) نابرابری را در حالت استقلال ارایه کرد که به نابرابری روزنتال معروف شد.

قضیه ۱.۱.۲. (روزنتال، ۱۹۷۰) فرض کنید $p \geq ۱$ و همچنین $\{X_i; ۱ \leq i \leq n\}$ متغیرهای تصادفی مستقلی هستند که به ازای هر i، $E|X_i|^p < \infty$ است. آن‌گاه

$$E|S_n^p| \leq \max\left\{ ۲^p \sum_{k=۱}^{n} E|X_k|^p, ۲^{p^۲}\left(\sum_{k=۱}^{n} E|X_k|\right)^p \right\}.$$

برهان. این نابرابری به‌خاطر آن‌که $S_n \leq \sum\limits_{k=۱}^{n} |X_k|$ و همچنین چون محدودیتی را ایجاد نمی‌کند اگر فرض کنیم همه جمع‌وندها نامنفی‌اند، نتیجه می‌شود. حال قرار می‌دهیم $S_n^{(j)} = \sum\limits_{\substack{k=۱\\k\neq j}}^{n} X_k$. با استفاده از نابرابری

[۱] Rosenthal-type maximal inequality

برای متغیرهای تصادفی مستقل نامنفی و نابرابری لیاپانوف، می‌توان نشان داد C_r

$$
\begin{aligned}
E(S_n)^p &= \sum_{j=۱}^{n} E(S_n)^{p-۱} X_j \\
&\leq ۲^{p-۱} \sum_{j=۱}^{n} E\left(\left(X_j^{p-۱} + (S_n^{(j)})^{p-۱}\right) X_j\right) \\
&= ۲^{p-۱} \sum_{j=۱}^{n} \left(EX_j^p + E(S_n^{(j)})^{p-۱} EX_j\right) \\
&\leq ۲^{p-۱} \sum_{j=۱}^{n} \left(EX_j^p + E(S_n)^{p-۱} EX_j\right) \\
&= ۲^{p-۱} \left(\sum_{j=۱}^{n} EX_j^p + E(S_n)^{p-۱} \sum_{j=۱}^{n} EX_j\right) \\
&\leq ۲^{p-۱} \left(\sum_{j=۱}^{n} EX_j^p + (E(S_n)^p)^{\frac{p-۱}{p}} \sum_{j=۱}^{n} EX_j\right) \\
&\leq ۲^p Max \left\{\sum_{j=۱}^{n} EX_j^p, (E(S_n)^p)^{\frac{p-۱}{p}} \sum_{j=۱}^{n} EX_j\right\}.
\end{aligned}
$$

بنابراین

$$E(S_n)^p \leq ۲^p \sum_{j=۱}^{n} EX_j^p$$

و

$$E(S_n)^p \leq ۲^p (E(S_n)^p)^{\frac{p-۱}{p}} \sum_{j=۱}^{n} EX_j$$

حال اگر فرض شود $p > ۲$، در این صورت

$$
E|S_n|^p \left\{
\begin{array}{ll}
\leq & D_p \max\left\{\sum_{k=۱}^{n} E|X_k|^p, (\sum_{k=۱}^{n} EX_k^۲)^{\frac{p}{۲}}\right\} \\
\geq & ۲^p \max\left\{\sum_{k=۱}^{n} E|X_k|^p, (\sum_{k=۱}^{n} EX_k^۲)^{\frac{p}{۲}}\right\}
\end{array}
\right.
$$

یا به‌طور معادل

$$\max\left\{\sum_{k=۱}^{n} E|X_k|^p, (Q_n(x))^p\right\} \leq ۲^p E|S_n|^p \leq D_p^* \max\left\{\sum_{k=۱}^{n} E|X_k|^p, (Q_n(x))^p\right\}.$$

که در آن $Q_n(x) = (\sum_{k=۱}^{n} EX_k^۲)^{\frac{p}{۲}}$ و D_p ثابتی است که فقط به p وابسته است. همچنین $D_p^* = ۲^p D_p$. □

برای متغیرهای مستقل، محققان زیادی به بیان نابرابری روزنتال پرداخته‌اند. از جمله می‌توان به پاپنلس و یوتف (۱۹۸۴) اشاره کرد که نابرابری روزنتال را در حالی که $\{X_i; ۱ \leq i \leq n\}$ متغیرهای تصادفی مستقل و دارای گشتاور مرتبه p متناهی باشند، به‌صورت زیر ارایه کرد:

$$E(|S_n|^p) \leq \frac{p(p-۱)}{۲} \max\left\{۱, ۲^{p-۳}\right\} \max\left\{(\sum_{k=۱}^{n} EX_k^۲)^{\frac{p}{۲}}, \sum_{k=۱}^{n} E|X_k|^p\right\}.$$

جانسن و همکاران (۱۹۸۵) نشان دادند اگر $\{X_i; ۱ \leq i \leq n\}$ متغیرهای تصادفی مستقل و متقارن باشند که به ازای $p > ۲$ دارای گشتاور pام متناهی هستند، آن‌گاه

$$E(|S_n|^p) \leq \frac{۱۴/۵p}{\log p} \max\left\{ (\sum_{k=۱}^{n} EX_k^۲)^{\frac{p}{۲}}, \sum_{k=۱}^{n} E|X_k|^p \right\}.$$

محققان بسیاری برای متغیرهایی که به دسته‌ای خاص از وابستگی تعلق دارند، کران‌هایی را ارایه کرده‌اند که برخی از آن‌ها را در ادامه مطرح می‌کنیم.

شائو (۱۹۹۵) نابرابری ماکسیمال را برای مجموع جزیی دنباله ρ-آمیخته[۲] بیان کرد. برای این منظور در ابتدا به تعریف دنباله ρ-آمیخته می‌پردازیم.

تعریف ۲.۱.۲. فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی تعریف‌شده در $(\Omega, \mathcal{F}, P)$ که در آن Ω یک میدان، همچنین $\mathcal{F}_n^m$ نیز σ-جبر تولیدشده با متغیرهای تصادفی $X_n, X_{n+۱}, \ldots, X_m$ باشند و برای هر $S \subset \mathbb{N}$، $\mathcal{F}_S = \sigma(X_i; i \in S)$ تعریف شود. حال دو σ-جبر $\mathcal{A}$ و $\mathcal{B}$ را در $\mathcal{F}$ در نظر بگیرید و قرار دهید

$$\rho^*(\mathcal{A}, \mathcal{B}) = \sup\{Corr(X,Y); X \in L_۲(\mathcal{A}), Y \in L_۲(\mathcal{B})\}$$

که در آن $Corr(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Vra(Y)}}$ و $L_۲(\cdot)$ یک فضای اندازه‌پذیر است. در این صورت ضرایب ρ-آمیخته به صورت زیر به‌دست می‌آیند:

$$\rho(n) = \sup\{\rho^*(\mathcal{F}_S, \mathcal{F}_T); S, T \subset \mathbb{N}; dist(S,T) \geq n\}$$

که در آن $dist(S,T)$ یک فاصله میان S و T را تعریف می‌کند.

بدیهی است $۰ \leq \rho(n+۱) \leq \rho(n) \leq \rho(۰) = ۱$. با این مقدمه، دنباله $\{X_n, n \geq ۱\}$ را ρ-آمیخته می‌نامند، اگر $k \in \mathbb{N}$ موجود باشد به‌طوری که $\rho(k) < ۱$. توجه کنید که اگر $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل باشد، به ازای هر $n \geq ۱$، $\rho(n) = ۰$ است. حال با توجه به تعریف ۲.۱.۲، فرض کنید X_i متغیری تصادفی با میانگین صفر باشد که به ازای $p \geq ۲$، $E|X_i|^p < \infty$. در این صورت، ثابت مثبتی مانند $B = B(p, \rho(\cdot))$ موجود است که فقط به p و $\rho(\cdot)$ وابسته است.

همچنین اگر برای هر $k \geq ۰$ و $n \geq ۱$، $S_n(k) = \sum_{i=k+۱}^{k+n} X_i$ باشد، در این صورت

$$\begin{aligned} E\left(\max_{۱\leq i\leq n}|S_k(i)|^p\right) \quad &\leq \quad B\left[n^{\frac{p}{۲}} \exp\left(B\sum_{k=۰}^{[\log n]}\rho(۲^i)\right)\max_{k\leq i\leq k+n}(E|X_i|)^{\frac{p}{۲}}\right. \\ &\left. +n\exp\left(B\sum_{k=۰}^{[\log n]}\rho^{\frac{۲}{p}}(۲^i)\right)\max_{k\leq i\leq k+n}(E|X_i|)^{\frac{p}{۲}}\right]. \qquad (۱.۲) \end{aligned}$$

این حالت را می‌توان به متغیرهای تصادفی بریده‌شده نیز تعمیم داد. در این صورت، اگر به ازای هر $۰ \leq A \leq \infty$ و $x > ۰$، داشته باشیم

$$۲n \max_{k\leq i\leq k+n} E|X_i|I\{|X_i| \geq A\} \leq x$$

[۲] ρ-mixing

آن‌گاه برای ماکسیمم مجموع جزیی خواهیم داشت

$$\begin{aligned} P\left(\max_{i\leq n}|S_k(i)|\geq x\right) &\leq \sum_{i=k+1}^{k+n} P(|X_i|\geq A) \\ &+Bx^{-p}\left[n^{\frac{p}{2}}\exp\left(B\sum_{k=0}^{[\log n]}\rho(2^i)\right)\max_{k\leq i\leq k+n}\left(E[X_iI\{|X_i|\leq A\}]^2\right)^{\frac{p}{2}}\right. \\ &\left.+n\exp\left(B\sum_{i=0}^{[\log n]}\rho^{\frac{2}{p}}(2^i)\right)\max_{k\leq i\leq k+n}E|X_i|^pI\{|X_i|\leq A\}\right]. \qquad (2.2)\end{aligned}$$

اگر در معادلات (۱.۲) و (۲.۲) به ازای $p\geq 2$، $E(X_i)=0$، $(EX_i^p)^{\frac{p}{2}}<\infty$ و $\sum_{n=1}^{\infty}\rho^{\frac{2}{p}}(2^n)<\infty$ باشند، در این صورت ثابتی مثبت مانند $B=B(p,\rho(\cdot))$ موجود است که فقط به p و $\rho(\cdot)$ بستگی دارد. آن‌گاه به ازای هر $k>0$ و $n\geq 1$ داریم

$$E\left(\max_{i\leq n}|S_k(i)|^p\right)\leq\left(\left(n\max_{k\leq i\leq k+n}(EX_i^2)\right)^{\frac{p}{2}}+n\max_{k\leq i\leq k+n}(E|X_i|^p)\right)$$

و

$$\begin{aligned} P\left(\max_{i\leq n}|S_k(i)|\geq x\right) &\leq \sum_{i=k+1}^{k+n} P(|X_i|\geq A) \\ &+Bx^{-p}\left[n^{\frac{p}{2}}\max_{k\leq i\leq k+n}\left(E[X_iI\{|X_i|\leq A\}]^2\right)^{\frac{p}{2}}\right. \\ &\left.+n\max_{k\leq i\leq k+n}E|X_i|^pI\{|X_i|\leq A\}\right].\end{aligned}$$

شائو (۲۰۰۰) نشان داد که گشتاورهای مجموع جزیی (ماکسیمم مجموع جزیی) متغیرهای تصادفی پیوندی منفی می‌توانند با متغیرهای تصادفی مستقل، محدود شوند.

فرض کنید $\{X_i; 1\leq i\leq n\}$ متغیرهای تصادفی پیوندی منفی باشند. همچنین فرض کنید $\{X_n^*; 1\leq i\leq n\}$ متغیرهای تصادفی مستقل باشند به‌طوری‌که X_i و X_i^* به ازای هر $i=1,2,\ldots,n$ هم‌توزیع‌اند. به ازای هر تابع محدب f در $\mathbb{R}$ می‌توان نوشت

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

مشروط بر آن‌که امید ریاضی‌های نابرابری بالا موجود باشند. اگر f تابعی محدب و ناکاهشی باشد، با شرط وجود داشتن $Ef\left(\max_{1\leq k\leq n}\sum_{i=1}^{k}X_i^*\right)$ می‌توان نوشت

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

در ادامه، فرض کنید $p\geq 1$ و $\{X_i; 1\leq i\leq n\}$ متغیرهای تصادفی پیوندی منفی با میانگین صفر باشند که به ازای هر $1\leq i\leq n$، $E|X_i|^p<\infty$. همچنین متغیرهای تصادفی $\{X_i^*; 1\leq i\leq n\}$ را در

نظر بگیرید که مستقل هستند و X_i و X_i^* به ازای هر $i=1,2,\ldots,n$ هم‌توزیع هستند. بنابراین

$$E|\sum_{i=1}^{n}X_i|^p \leq E|\sum_{i=1}^{n}X_i^*|^p$$

$$E|\max_{1\leq k\leq n}\sum_{i=1}^{k}X_i|^p \leq 2E|\sum_{i=1}^{n}X_i^*|^p$$

و برای $1<p\leq 2$

$$E\max_{1\leq k\leq n}|\sum_{i=1}^{k}X_i|^p \leq 2^{3-p}E\sum_{i=1}^{n}|X_i|^p.$$

همچنین به ازای $p>2$

$$E\max_{1\leq k\leq n}|\sum_{i=1}^{k}X_i|^p \leq 2\left(\frac{15p}{\ln\ p}\right)^p\left\{\left(\sum_{i=1}^{k}X_i^2\right)^{\frac{p}{2}}+\sum_{i=1}^{n}E|X_i|^p\right\}.$$

هو (۲۰۰۰) نشان داد که گشتاورهای مجموع جزیی (یا مجموع جزیی ماکسیمم) متغیرهای تصادفی وابسته زیرجمعی منفی را نیز می‌توان با متغیرهای تصادفی مستقل محدود کرد.

فرض کنید $\mathbf{X}=(X_1,X_2,\ldots,X_n)$ بردار تصادفی NSD و $\mathbf{X}^*=(X_1^*,X_2^*,\ldots,X_n^*)$ بردار تصادفی متناظر با مؤلفه‌هایی مستقل باشند به‌طوری‌که به ازای $i=1,2,\ldots,n$، X_i^* و X_i هم‌توزیع هستند. برای هر تابع صعودی و محدب f می‌توان نوشت

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

پلیگرات و همکاران (۲۰۰۷) نشان دادند که به ازای $p\geq 2$، اگر $E|X_1|^p<\infty$ و قرار دهیم $S_n=\sum_{k=1}^{j}X_k$، آن‌گاه

$$E\left[(\max_{i\leq n}|S_i|)^p\right]\leq C_p^{\frac{1}{p}}n^{\frac{1}{2}}\left(E|X_1|^p+80\sum_{k=1}^{j}j^{-\frac{3}{2}}\left[E|E(S_j|\mathcal{F}_0)|^p\right]^{\frac{1}{p}}\right).$$

حال در زیر نابرابری ماکسیمال روزنتال برای متغیرهای تصادفی NSD را از هو (۲۰۰۰) و ونگ و همکاران (۲۰۱۴) بیان می‌کنیم و سپس به اثبات آن می‌پردازیم.

لم ۳.۱.۲. (نابرابری ماکسیمال روزنتال) فرض کنید $\{X_n, n\geq 1\}$ دنباله‌ای از متغیرهای تصادفی NSD باشد که $EX_n=0$ و برای $p\geq 2$، $E|X_n|^p<\infty$. آن‌گاه ثابت مثبت C_p (وابسته به p) وجود دارد به‌طوری‌که برای هر $n\geq 1$،

$$E\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^{k}X_i\right|^p\right)\leq C_p\left\{\sum_{i=1}^{n}E|X_i|^p+\left(\sum_{i=1}^{n}EX_i^2\right)^{\frac{p}{2}}\right\}. \tag{۳.۲}$$

برهان. برای اثبات کافی است در نتیجه آ.۱.۴ در پیوست آ، $t=2$ را قرار دهیم. در این صورت نابرابری (۳.۲) به‌دست می‌آید.

□

۲.۲ نابرابری نمایی کلموگروف

از نابرابری‌های مهم دیگر می‌توان نابرابری نمایی کلموگروف[۳] را نام برد. در نظریه احتمال، نابرابری کلموگروف یک نابرابری ماکسیمال احتمالی است برای مجموع جزیی یک مجموعه متناهی از متغیرهای تصادفی مستقل. نابرابری مذکور در این پایان‌نامه برای مقایسه گشتاورهای متغیرهای تصادفی وابسته زیرجمعی منفی و متغیرهای تصادفی مستقل، به‌کار می‌رود.

کلموگروف (۱۹۶۳) برای متغیرهای تصادفی مستقل و هم‌توزیع $\{X_n; n \geq ۱\}$ از توزیع برنولی با پارامتر $۰ \leq p \leq ۱$، نابرابری ماکسیمال زیر را برای هر $\varepsilon > ۰$، به‌دست آورد:

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

که در آن $\bar{X}_k = \frac{۱}{k}\sum_{i=۱}^{k} X_i$. بانجیوایس (۱۹۸۴) نابرابری فوق را به شکل بهتری ارایه کرد که در آن به ازای هر $۰ \leq \varepsilon \leq ۱$ و $۰ \leq p \leq ۱$

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq ۲\exp\left\{-۲n\varepsilon^۲(۱-\frac{۲}{۳}\varepsilon)\right\}.$$

حال فرض کنید $\varepsilon \geq ۰$ باشد. اگر $\lambda > ۰$ی موجود باشد به‌طوری‌که $E[\exp\lambda(X_۱-\varepsilon)] \leq ۱$، به ازای $n \geq ۰$ نابرابری زیر برقرار است:

$$P\left(\sup_{k\geq n}\bar{X}_k \geq \varepsilon\right) \leq \left[E\exp\{\lambda(X_۱-\varepsilon)\}\right]^n.$$

در ادامه، یانگ و همکاران (۱۹۸۷) نابرابری ماکسیمال را دست‌خوش تغییر قرار دادند و نابرابری مذکور را به‌صورت زیر گزارش کردند. برای هر $۰ \leq \varepsilon \leq ۱$ و $۰ \leq p \leq ۱$

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

در حالت خاص برای توزیع دوجمله‌ای، یانگ، تورنر و همکاران (۱۹۸۸) نشان دادند که به ازای هر $۰ \leq p \leq \frac{۱}{۴}$ یا $\frac{۳}{۴} \leq p \leq \frac{۱}{۲}$ و همچنین $۰ \leq \varepsilon < \frac{۱}{۴}$ نابرابری کلموگروف به‌صورت زیر است:

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq ۲\exp\left\{-۲n\varepsilon^۲ - (\frac{۴}{۳})n\varepsilon^۴\right\}.$$

دیسچل و استروک (۱۹۸۹) برای دنباله‌ای از متغیرهای تصادفی $\{X_n; n \geq ۱\}$، نشان دادند به ازای هر $\varepsilon > ۰$،

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

که در آن C_b عددی ثابت است. تورنر و همکاران (۱۹۹۲) با تغییر شروط، نابرابری کلموگروف را برای توزیع دوجمله‌ای مورد بررسی قرار دادند. آن‌ها نشان دادند اگر $\{X_n; n \geq ۱\}$ یک دنباله‌ای از متغیرهای تصادفی برنولی هم‌توزیع و مستقل با میانگین p باشد، آن‌گاه برای هر $۰ \leq \varepsilon \leq min\{p, ۱-p\}$ که در آن $۰ < p < ۱$ و $n \geq ۰$، داریم:

$$P\left(\sup_{k\geq n}|\bar{X}_k - p| \geq \varepsilon\right) \leq ۲\exp\left\{-۲n\varepsilon^۲ - \frac{۴}{۹}n\varepsilon^۲\right\}.$$

---

[۳] Kolmogorov-type exponential inequality

مائوریکو (۲۰۰۷) نابرابری کلموگروف را برای میانگین نمونه‌ای متغیرهای تصادفی مستقل برنولی نه لزوما هم‌توزیع به‌دست آوردند. دنباله $\{X_i; 1 \leq i \leq n\}$ از متغیرهای تصادفی مستقل برنولی با $EX_i = p_i$، $i = 1, 2, \ldots, n$، را در نظر بگیرید. به ازای هر $\bar{p} + \varepsilon < \frac{1}{2}$ یا $\bar{p} < \frac{1}{2}$، $\varepsilon < 1$ می‌توان نوشت

$$P\left(\bar{X} - \bar{p} \geq \varepsilon\right) \leq \exp\left\{-n(2\varepsilon^2 - \frac{1}{3}\varepsilon^4 e^{\frac{\varepsilon^2}{4}})\right\}.$$

که در آن $\bar{p} = \frac{1}{n}\sum_{i=1}^{n} p_i$.

تا این‌جا نابرابری کلموگروف با شرط مستقل بودن متغیرهای تصادفی بیان شد. در ادامه به بیان این نابرابری‌ها با شرط وابسته بودن متغیرهای تصادفی می‌پردازیم.

ولیدین (۲۰۰۲) قضیه‌ای را بیان کرد که در آن نابرابری نمایی کلموگروف برای متغیرهای تصادفی پیوندی منفی به‌دست آمده است.

قضیه ۱.۲.۲. (ولیدین، ۲۰۱۱) متغیرهای تصادفی پیوندی منفی $\{X_i; 1 \leq i \leq n\}$ با میانگین صفر و واریانس متناهی را در نظر بگیرید. فرض کنید $B_n = \sum_{k=1}^{n} EX_k^2$ و به ازای هر $1 \leq k \leq n$ و $n \geq 1$، به‌طور تقریبا حتمی، $|X_k| \leq CS_n$ باشد. در این صورت برای هر $a > 0$، $n \geq 1$، $\varepsilon C \leq a$ و $0 < \alpha < \frac{a^3}{2(e^a - 1 - a - \frac{a^2}{2})}$، داریم

$$P\left\{\frac{S_n}{B_n} > \varepsilon\right\} \leq \exp\left\{-\frac{\varepsilon^2}{2}(1 - \frac{\varepsilon C}{\alpha})\right\}$$

که در آن $S_n = \sum_{k=1}^{n} X_k$.

ونگ و همکاران (۲۰۱۱) نابرابری نمایی را برای دنباله‌ای از متغیرهای وابسته متعامد منفی[۴] (NOD) به‌دست آوردند. فرض کنید $\{X_n, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی NOD باشد که به ازای هر $n \geq 1$، $EX_n = 0$. اگر دنباله‌ای از اعداد حقیقی مثبت مانند $\{C_n, n \geq 1\}$ موجود باشد به‌طوری که برای هر $1 \leq i \leq n$ و $n \geq 1$، $|X_i| \leq C_n$ باشد، برای هر $t > 0$ و هر عدد صحیح $n \geq 1$

$$E\exp\left\{t\sum_{i=1}^{n} X_i\right\} \leq \exp\left\{\frac{t^2}{2}e^{tC_n}\sum_{i=1}^{n} X_i^2\right\}. \tag{۴.۲}$$

اکنون اگر به ازای هر $1 \leq i \leq n$، $EX_i \neq 0$ باشد، در این صورت نابرابری (۴.۲) به شکل زیر تغییر می‌یابد:

$$E\exp\left\{t\sum_{i=1}^{n}(X_i - EX_i)\right\} \leq \exp\left\{\frac{t^2}{2}e^{2tC_n}\sum_{i=1}^{n} X_i^2\right\}.$$

ونگ و همکاران در ادامه با اعمال شروط ذکر شده در بالا و به ازای هر $0 < \varepsilon \leq \frac{eB_n}{(2C_n)}$ نابرابری نمایی زیر را ارایه کردند:

$$P\left(|\sum_{i=1}^{n}(X_i - EX_i)| \geq \varepsilon\right) \leq 2\exp\left\{-\frac{\varepsilon^2}{2eB_n}\right\}$$

[۴] Negatively orthant dependent variables

که در آن $B_n = \sum\limits_{i=۱}^{n} EX_i^۲$.

کیم (۲۰۱۱) نیز برخی نابرابری‌های نمایی را برای متغیرهای تصادفی وابسته ربعی منفی خطی توسعه‌یافته (ELNQD) ارایه کرد. فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی ELNQD با میانگین صفر و واریانس متناهی است. اگر برای هر $۱ \leq i \leq n$، و $n \geq ۱$ عدد مثبت C موجود باشد به‌طوری‌که $|X_i| \leq CB_n$ و $B_n = \sum\limits_{i=۱}^{n} EX_i^۲$، آن‌گاه برای هر $\varepsilon > ۰$ و $n \geq ۱$ ثابت $M > ۰$ وجود دارد به‌طوری‌که

$$P\left(\frac{S_n}{B_n} \geq \varepsilon\right) \leq \begin{cases} M\exp\left(-\frac{\varepsilon}{۲}(۱ - \frac{C\varepsilon}{۲})\right) & C\varepsilon \leq ۱ \\ M\exp\left(-\frac{\varepsilon}{۴C}\right) & C\varepsilon > ۱ \end{cases}$$

حال اگر در شرایط فوق واریانس متناهی نباشد و اگر برای هر $n \geq ۱$ ثابت مثبت b موجود باشد به‌طوری که $|X_n| \leq b$، آن‌گاه به ازای هر $\varepsilon > ۰$ و $M > ۰$

$$P\left(|S_n| \geq \varepsilon\right) \leq ۲M\exp\left\{-\frac{\varepsilon^۲}{۲(۲B_n^۲ + b\varepsilon)}\right\}.$$

در ادامه با تغییر ثابت مثبت به دنباله $\{C_n, n \geq ۱\}$ که دنباله‌ای از اعداد مثبت است، نابرابری را به شکل زیر خواهیم داشت که در آن به ازای هر $i \geq ۱$، $|X_i| \leq C_i$ است:

$$E\exp\left(tS_n\right) \leq \exp\left\{\frac{t^۲}{۲}\sum_{i=۱}^{n} e^{tC_i} EX_i^۲\right\}.$$

سرانجام در حالی که میانگین نیز مخالف صفر است، اگر به ازای هر $\varepsilon > ۰$، $\varepsilon \leq \frac{eB_n^۲}{۲eB_n^۲}$ باشد آن‌گاه

$$P\left(|S_n - ES_n| \geq \varepsilon\right) \leq ۲M\exp\left\{-\frac{\varepsilon^۲}{۲eB_n}\right\}.$$

شن و همکاران (۲۰۱۴) نشان دادند که اگر $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی ρ-آمیخته با $EX_n = ۰$ باشد و برای هر $n \geq ۱$، به‌طور تقریبا حتمی، $|X_n| \leq d < \infty$ را داشته باشیم، برای هر $\varepsilon > ۰$ و $n \geq ۱$ روابط زیر برقرار هستند:

$$P(S_n > \varepsilon) \leq C_۱\exp\left\{-\frac{\varepsilon^۲}{C_۲(۴B_n + nd\varepsilon)}\right\}$$

$$P(S_n < -\varepsilon) \leq C_۱\exp\left\{-\frac{\varepsilon^۲}{C_۲(۴B_n + nd\varepsilon)}\right\}$$

$$P(|S_n| > \varepsilon) \leq ۲C_۱\exp\left\{-\frac{\varepsilon^۲}{C_۲(۴B_n + nd\varepsilon)}\right\}$$

که در آن $C_۱ = \exp\left\{۱ + \rho(m+۱) + \frac{n-۴m}{۴m}\right\}$، $C_۲ = ۸\left(۱ + ۲\sum\limits_{k=۱}^{m}\rho(k)\right)$، و $۱ \leq m \leq n$ برخی اعداد صحیح مثبت هستند و به ازای هر $n \geq ۱$، $S_n = \sum\limits_{i=۱}^{n} X_i$ و $B_n = \sum\limits_{i=۱}^{n} EX_i^۲$ است.

حال در زیر نابرابری نمایی کلموگروف برای متغیرهای تصادفی NSD را از شائو (۲۰۰۰) بیان می‌کنیم و سپس به اثبات آن می‌پردازیم.

لم ۲.۲.۲. (نابرابری نمایی کلموگروف) فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی NSD با میانگین صفر و گشتاور دوم متناهی باشد. اگر برای هر $n \geq ۱$، $B_n = \sum_{i=۱}^{n} EX_i^۲$ باشد، آن‌گاه برای هر $x > ۰, y > ۰, ۰ < \eta < ۱$ و $n \geq ۱$، داریم

$$P\left(\max_{۱\leq k\leq n}\left|\sum_{i=۱}^{k} X_i\right| \geq x\right) \leq ۲P\left(\max_{۱\leq k\leq n}|X_k| > y\right) + \frac{۲}{۱-\eta}exp\left\{-\frac{x^۲\eta}{۲(xy+B_n)}\right\}.$$

برهان. به ازای هر $i = ۱, ۲, \ldots, n$ تعریف می‌کنیم

$$t = y^{-۱}\ln(۱+\frac{xy}{B_n}) \qquad \tilde{S}_i = \sum_{j=۱}^{i}\min(X_j, y) \qquad U_i = \sum_{j=۱}^{i} X_j^* \tag{۵.۲}$$

به‌طوری که $\{X_i^*; ۱ \leq i \leq n\}$ متغیرهای تصادفی مستقل هستند که در آن X_i^* و $\min(X_i, y)$ به ازای هر $i = ۱, ۲, \ldots, n$ هم توزیع‌اند. از آن‌جا که $\{\min(X_i, y), ۱ \leq i \leq n\}$ وابسته زبرجمعی منفی است، با استفاده از رابطه (۲.۱) می‌توان نوشت

$$\begin{aligned}
P\left(\max_{۱\leq k\leq n}\left|\sum_{i=۱}^{k} X_i\right| \geq x\right) &\leq ۲P\left(\max_{۱\leq k\leq n}|X_k| > y\right) + ۲P\left(\max_{۱\leq k\leq n}\tilde{S}_i \geq x\right) \\
&\leq ۲P\left(\max_{۱\leq k\leq n}|X_k| > y\right) + ۲e^{-t\eta x}E\exp\left\{t\eta\max_{۱\leq k\leq n}\tilde{S}_i\right\} \\
&\leq ۲P\left(\max_{۱\leq k\leq n}|X_k| > y\right) \\
&\quad +۲e^{-t\eta x}E\exp\left\{t\eta\max_{۱\leq k\leq n}U_i\right\}.
\end{aligned} \tag{۶.۲}$$

با قرار دادن

$$J_۱ = e^{-t\eta x}E\exp\left\{t\eta\max_{۱\leq k\leq n}U_i\right\}$$

و

$$T_i = \exp\left\{tU_i - \left(e^{ty} - ۱ - ty\right)y^{-۲}B_i\right\}$$

می‌توان گفت

$$\begin{aligned}
J_۱ &= e^{-t\eta x}E\exp\left\{t\eta\max_{۱\leq k\leq n}U_i\right\} \\
&= e^{-t\eta x}E\exp\left\{t\max_{۱\leq k\leq n}U_i\right\}^{\eta} \\
&= e^{-t\eta x}E\left\{\max_{۱\leq k\leq n}T_i e^{(e^{ty}-۱-ty)y^{-۲}B_i}\right\}^{\eta} \\
&\leq \exp\left\{-t\eta x + \eta(e^{ty}-۱-ty)y^{-۲}B_i\right\}E\left[\max_{۱\leq k\leq n}T_i\right]^{\eta}.
\end{aligned} \tag{۷.۲}$$

از طرفی $\frac{e^x-1-x}{x^2}$ تابعی ناکاهشی از x روی $\mathbb{R}$ است و $EX_i^*\leq 0$ و $tX_i^*\leq ty$. لذا

$$\begin{aligned} Ee^{tX_i^*} &= 1+EX_i^*+E\left[\frac{e^{tX_i^*}-1-tX_i^*}{(X_i^*)^2}(X_i^*)^2\right] \\ &\leq 1+\left(e^{ty}-1-ty\right)y^{-2}E(X_i^*)^2 \\ &= 1+\left(e^{ty}-1-ty\right)y^{-2}E(X_i)^2 \\ &\leq \exp\left\{(e^{ty}-1-ty)y^{-2}E(X_i)^2\right\}. \end{aligned} \tag{۸.۲}$$

عبارت آخر در نابرابری (۸.۲) از رابطه $1+x\leq e^x$ نتیجه شده است. نابرابری (۸.۲) نشان می‌دهد که $\{T_i, 1\leq i\leq n\}$ بر اساس تعریف آ.۳.۳، یک زبرمارتینگل است. لذا طبق لم آ.۳.۴ در پیوست آ می‌توان نوشت

$$J_1\leq \frac{1}{1-\eta}\exp\left\{-t\eta x+\eta(e^{ty}-1-ty)y^{-2}B_i\right\}. \tag{۹.۲}$$

حال با جای‌گذاری (۵.۲) در نابرابری (۹.۲)

$$\begin{aligned} J_1 &\leq \frac{1}{1-\eta}\exp\left\{-\eta\left[xy^{-1}\ln(1+\frac{xy}{B_n})-\left((1+\frac{xy}{B_n})-1-\ln(1+\frac{xy}{B_n})\right)y^{-2}B_n\right]\right\} \\ &= \frac{1}{1-\eta}\exp\left\{-\eta\left[xy^{-1}\ln(1+\frac{xy}{B_n})-\left(\frac{xy}{B_n}-\ln(1+\frac{xy}{B_n})\right)y^{-2}B_n\right]\right\} \\ &= \frac{1}{1-\eta}\exp\left\{-\eta\left[(xy+B_n)y^{-2}\ln(1+\frac{xy}{B_n})-\frac{x}{y}\right]\right\}. \end{aligned} \tag{۱۰.۲}$$

اکنون بنابر لم آ.۲.۱

$$\begin{aligned} (xy+B_n)y^{-2}\ln(1+\frac{xy}{B_n})-\frac{x}{y} &\geq (xy+B_n)y^2\left[\frac{xy}{xy+B_n}\right. \\ &\quad \left.+\frac{(xy)^2}{2(xy+B_n)}\left(1+\frac{2}{3}\ln(1+\frac{xy}{B_n})\right)\right]-\frac{x}{y} \\ &= \frac{(xy)^2}{2(xy+B_n)}\left(1+\frac{2}{3}\ln(1+\frac{xy}{B_n})\right) \\ &\geq \frac{(xy)^2}{2(xy+B_n)}. \end{aligned} \tag{۱۱.۲}$$

بنابراین با جای‌گذاری (۱۱.۲) در (۱۰.۲)، داریم

$$J_1\leq \frac{1}{1-\eta}\exp\left\{\frac{\eta x^2}{2(xy+B_n)}\right\}. \tag{۱۲.۲}$$

سرانجام با جای‌گذاری (۱۲.۲) در عبارت سمت راست (۶.۲) برهان کامل می‌شود و

$$P\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^k X_i\right|\geq x\right)\leq 2P\left(\max_{1\leq k\leq n}|X_k|>y\right)+\frac{2}{1-\eta}\exp\left\{\frac{\eta x^2}{2(xy+B_n)}\right\}.$$

□

فصل ۳

همگرایی کامل برای مجموع موزون متغیرهای تصادفی در دسته‌ای خاص از وابستگی

۱.۳ مقدمه

مفهوم آماری وابستگی سبب شده است تا رشد چشم‌گیری در آمار، احتمال و دیگر شاخه‌های علم آمار پدید آید. مفاهیم متنوع وابستگی در احتمال محققان را بر آن داشت تا در سه دهه اخیر کاربردهای آن را در زمینه‌های مختلف جستجو کنند. از جمله می‌توان به مفاهیم وابستگی منفی و مثبت که سبب فهم هر چه بهتر توزیع‌های چندمتغیره و همچنین مدل‌های چندمتغیره شده اند، اشاره نمود.

مفهوم متغیرهای تصادفی NSD برای اولین بار توسط هو (۲۰۰۰) بیان شد که مبتنی بر رده توابع زبرجمعی است. هو (۲۰۰۰) مثالی را در این راستا شرح داد که در زیر به بیان آن می‌پردازیم. فرض کنید $\mathbf{X}=(X_۱,X_۲,X_۳,X_۴)$ توزیعی گسسته باشد که تابع احتمال آن در جدول زیر آمده است (جاج و پروشان، ۱۹۸۳):

		(X_1, X_2)				
		$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	تابع احتمال کناری
(X_3, X_4)	$(0,0)$	۰/۰۵۷۷	۰/۰۶۲۳	۰/۰۶۲۳	۰/۰۵۷۷	۰/۲۴
	$(0,1)$	۰/۰۶۲۳	۰/۰۶۷۷	۰/۰۶۷۷	۰/۰۶۲۳	۰/۲۶
	$(1,0)$	۰/۰۶۲۳	۰/۰۶۷۷	۰/۰۶۷۷	۰/۰۶۳۴	۰/۲۶
	$(1,1)$	۰/۰۵۷۷	۰/۰۶۲۳	۰/۰۶۲۳	۰/۰۵۷۷	۰/۲۴
	تابع احتمال کناری	۰/۲۴	۰/۲۶	۰/۲۶	۰/۲۴	

برای این‌که نشان دهیم متغیرهای فوق متغیرهای تصادفی NSD هستند، فرض کنید

$$(X_1, X_2, X_3, X_4) \stackrel{d}{=} (X_2, X_1, X_3, X_4) \stackrel{d}{=} (X_2, X_1, X_4, X_3) \stackrel{d}{=} (X_3, X_4, X_1, X_2).$$

بدون کم شدن از کلیت مسأله، تابعی مانند ϕ را در نظر می‌گیریم به‌طوری‌که به ازای هر $\mathbf{x} = (x_1, x_2, x_3, x_4)$

$$\phi(x_1, x_2, x_3, x_4) = \phi(x_2, x_1, x_3, x_4) = \phi(x_2, x_1, x_4, x_3) = \phi(x_3, x_4, x_1, x_2).$$

طبق تعریف ۵.۳.۱ و با توجه به این‌که متغیرهای تصادفی $\mathbf{X}$ نسبت به جایگشت ثابت هستند، با استناد به تعریف تابع زبرجمعی که در آن

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

در نتیجه

$$2\phi_{0001} \leq \phi_{0000} + \phi_{0011}, \quad 2\phi_{0111} \leq \phi_{0011} + \phi_{1111}, \quad 2\phi_{0101} \leq \phi_{0001} + \phi_{0111},$$

که در آن $\phi_{ijkl} = \phi(i, j, k, l)$. به‌عنوان نمونه دو تابع ϕ_{0001} و ϕ_{0010} را در نظر بگیرید. برای حالت می‌نیمم مؤلفه‌ای در اولین مؤلفه به دلیل این‌که هردو صفر هستند، بنابراین مینیمم صفر خواهد بود. به‌همین ترتیب برای مؤلفه‌های بعدی و همچنین حالت ماکسیمم مؤلفه‌ای را به‌دست می‌آوریم. بنابراین با استدلالی مشابه و با استفاده از روابط بالا، داریم

$$26\phi_{0101} \quad \leq \quad 6\phi_{0000} + 12\phi_{0011} + 6\phi_{1111} + \phi_{0001} + \phi_{0111}.$$

حال با توجه به اعداد جدول، نابرابری فوق درست است. در این صورت $\mathbf{X}$، NSD خواهد بود. از طرفی با توجه به تعریف وابستگی پیوندی منفی باید

$$P(X_i = 1; i = 1, 2, 3, 4) < P(X_1 = X_2 = 1)P(X_3 = X_4 = 1).$$

این در حالی است که با جای‌گذاری اعداد جدول، ملاحظه می‌شود که نابرابری فوق برقرار نیست و در نتیجه $\mathbf{X}$، NA نخواهد بود. نتیجه این مثال بدان معناست که وابستگی زبرجمعی منفی تعمیمی از وابستگی پیوندی منفی می‌باشد. این نتیجه مهم می‌تواند در برخی موارد از جمله نابرابری‌های احتمالی و نابرابری‌های گشتاوری نقش بسزایی را ایفا نماید.

۲.۳ همگرایی کامل متغیرهای تصادفی NSD

در این بخش همگرایی کامل برای مجموع موزون متغیرهای تصادفی وابسته زبرجمعی منفی مورد بررسی قرار خواهد گرفت. این مفهوم نقشی کلیدی در سازگاری کامل برآوردگرهای تبیینی آلوده به خطا در مدل رگرسیونی کمترین توان‌های دوم با خطاهای وابسته زبرجمعی منفی ایفا می‌کند که در فصل ۴ به طور کامل به آن می‌پردازیم.

اقبال و همکاران (۲۰۱۰) نابرابری‌های ماکسیمال برای شکل‌های درجه دوم متغیرهای تصادفی NSD را مورد مطالعه قرار دادند. سپس قانون قوی اعداد بزرگ و رتبه همگرایی را تحت شرایط ضعیفی با موجودیت گشتاور rام به‌دست آوردند.

فرض کنید $\{X_i; i \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع باشند. اکنون شکل درجه دوم زیر را در نظر بگیرید:

$$Q_n = \sum_{۱\leq i<j\leq n} a_{ij}X_iX_j \qquad n \geq ۲.$$

در حالت خاص اگر $a_{ij} = ۱$ و $i \neq j$ باشد، آن‌گاه

$$T_n = \sum_{۱\leq i<j\leq n} X_iX_j \qquad n \geq ۲.$$

با توجه به شکل درجه دوم معرفی‌شده T_n، فرض کنید $\{X_i; i \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی نامنفی وابسته زبرجمعی منفی با $EX_i^r < \infty$ برای هر $i \geq ۱$ و $r > ۱$ باشد. سپس به ازای هر $\varepsilon > ۰$

$$p\left[\max_{۲\leq k\leq n} T_k > \varepsilon\right] \leq C\left(\frac{r}{\varepsilon(r-۱)}\right)^r \left(\sum_{j=۲}^{n} E(X_j^r)\sum_{i=۱}^{j-۱} E(X_i^r)\right)$$

اکنون اگر شرایط فوق برقرار باشند و همچنین $\{a_{ij}; ۱ \leq i < j \leq n\}$ آرایه‌ای از اعداد حقیقی باشد، در این صورت برای شکل درجه دوم Q_n به ازای هر $\varepsilon > ۰$ و $r > ۱$ داریم

$$P\left[\max_{۲\leq k\leq n} Q_k > \varepsilon\right] \leq C\left(\frac{r}{\varepsilon(r-۱)}\right)^r \left(\sum_{j=۲}^{n} E(X_j^r)\sum_{i=۱}^{j-۱} a_{ij}E(X_i^r)\right).$$

با استفاده از دو نابرابری ماکسیمال فوق، همگرایی کلموگروف معیاری برای قانون قوی اعداد بزرگ با متغیرهای تصادفی NSD است. اقبال و همکاران (۲۰۱۰) با استفاده از این مطالب، به نتایج مهمی دست یافتند که در آن قانون قوی اعداد بزرگ و رتبه همگرایی قریب به یقین شکل‌های درجه دوم به‌دست می‌آیند. در ادامه به بیان مختصری از نتایج به‌دست آمده می‌پردازیم.

فرض کنید $\{X_i; i \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی زبرجمعی منفی و نامنفی باشد که در آن به ازای هر $j \geq ۱$ و $r > ۱$، $E(X_i^r) < \infty$ است. اگر به ازای هر $r > ۱$، $\sum_{j=۲}^{n} E(X_j^r)\sum_{i=۱}^{j-۱} E(X_i^r) < \infty$ آن‌گاه شکل درجه دوم $T_n = \sum_{۱\leq i<j\leq n} X_iX_j$ وقتی که $n \longrightarrow \infty$، به‌طور تقریبا حتمی همگراست. حال

اگر شرایط فوق برقرار باشند و $\sum_{j=2}^{n}\frac{1}{b_j^2}E(X_j^r)\sum_{i=1}^{j-1}E(X_i^r)<\infty$، آن‌گاه

$$\frac{1}{b_n}\sum_{1\le i<j\le n}X_iX_j\xrightarrow{a.s.}0$$

که در آن $\{b_n; n\ge 1\}$ دنباله‌ای از اعداد حقیقی صعودی مثبت است و زمانی‌که $n\longrightarrow\infty$، آن‌گاه b_n نیز به بی‌نهایت میل خواهد کرد.

اقبال و همکاران (۲۰۱۱) نیز برخی نابرابری‌های احتمالی کلموگروف برای شکل‌های درجه دوم و همچنین متغیرهای تصادفی به‌طور یکنواخت کراندار NSD را مورد مطالعه قرار دادند.

با استفاده از این نابرابری‌ها می‌توان برخی از نسخه‌های همگرایی کامل تصادفیده[1] شکل‌های درجه دوم را تحت شرایطی مناسب ارزیابی نمود. در زیر به‌طور مختصر، به بیان نتایج می‌پردازیم.

فرض کنید $\{X_i; i\ge 1\}$ دنباله‌ای از متغیرهای تصادفی NSD و نامنفی با $P(a\le X_i\le b)=1$ و $i=1,2,\dots,n$ باشد. بنابراین

۱. به ازای هر $x>0$

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

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

۲. به ازای هر $x>0$

$$P\left(T_n-E(T_n)\le -x\right)\le\exp\left\{-\frac{\left(\binom{n}{2}b^2-x+\sum\limits_{i=1}^{n}a_i-a^2\right)^2}{2\sum_{i=3}^{n}(\sigma_i^2+a_i^2)}\right\}$$

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

اکنون اگر $\{\gamma_n; n\ge 1\}$ دنباله‌ای ناکاهشی از اعداد حقیقی مثبت باشد و وقتی $n\longrightarrow\infty$، سپس $\gamma_n\longrightarrow\infty$ میل خواهد کرد. همچنین به ازای هر $x>0$، $\alpha>\frac{3}{2}$ و $0<C<\infty$ داشته باشیم

$$x\gamma_n-A_n=Cn^{\alpha}.$$

در نتیجه، زمانی که $n\longrightarrow\infty$، داریم

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

در حالت خاص، اگر $\gamma_n=\binom{n}{2}$ باشد، در این صورت

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

از طرفی اگر شکل درجه دوم به صورت Q_n باشد، شرایط در نظر گرفته‌شده برای شکل درجه دوم T_n را برای حالت Q_n نیز در نظر می‌گیریم. در این صورت اگر $\{a_{ij}; 1\le i\le n\}$ آرایه‌ای از اعداد حقیقی نامنفی باشد، آن‌گاه

[1] Randomized

۱. به ازای هر $x > 0$

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

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

۲. به ازای هر $x > 0$

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

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

وانگ و همکاران (۲۰۱۴) با استفاده از نابرابری ماکسیمال روزنتال و نابرابری نمایی کلموگروف همگرایی کامل برای آرایه‌ای از متغیرهای تصادفی NSD سطری را بیان کردند.

فرض کنید $\{X_{ni}; 1 \leq i \leq k_n, n \geq 1\}$ آرایه‌ای از متغیرهای تصادفی NSD سطری باشد که متغیر تصادفی X بر آن غلبه تصادفی دارد. همچنین فرض کنید $\{a_{ni}; 1 \leq i, n \geq 1\}$ آرایه‌ای از ثابت‌هاست در این صورت $r > 0$ی موجود است که

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

همچنین $0 < \theta < 2$ و αهایی موجودند که $\theta + \frac{\alpha}{r} < 2$. لذا داریم

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

در این صورت سه حالت زیر می‌توانند اتفاق بیافتند:

۱. اگر $1 + \alpha + \beta < 0$ و $E|X|^\theta < \infty$، آن‌گاه برای هر $\varepsilon > 0$

$$\sum_{n=1}^{\infty} n^\beta P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}\right| > \varepsilon\right) < \infty \tag{۱.۳}$$

و

$$\sum_{n=1}^{\infty} n^\beta P\left(\left|\sum_{i=1}^{\infty} a_{ni}X_{ni}\right| > \varepsilon\right) < \infty. \tag{۲.۳}$$

۲. اگر $1 + \alpha + \beta > 0$، $\beta > -1$ و $s = \theta + \dfrac{1+\alpha+\beta}{r}$، آن‌گاه با در نظر گرفتن $E|X|^s < \infty$، همچنین وقتی $s \geq 1$، $EX_{ni} = 0$، روابط (۱.۳) و (۲.۳) برقرار هستند.

۳. اگر $1 + \alpha + \beta = 0$، $E|X|^\theta \log|X| < \infty$ و وقتی که $1 \leq \theta < 2$، $EX_{ni} = 0$ باشد. در این صورت روابط (۱.۳) و (۲.۳) صدق می‌کند.

۱.۲.۳ کاربردی از نابرابری‌های مذکور

حال با استفاده از مطالب فوق، کاربرد نابرابری‌ها را در همگرایی کامل برآوردگر رگرسیون ناپارامتری که توسط ونگ و همکاران (۲۰۱۴) بیان شده است را به‌عنوان یک مثال، ارایه می‌کنیم. در ابتدا مدل رگرسیون ناپارامتری زیر را در نظر بگیرید:

$$y_{ni} = g(x_{ni}) + \varepsilon_{ni}$$

که در آن x_{ni} ثابت معلومی است که نقاط آن از A به‌دست آمده است و $A \subset \mathbb{R}^p$ مجموعه‌ای فشرده به ازای $p \geq ۱$ می‌باشد. همچنین $g(\cdot)$ تابع رگرسیون نامعلوم تعریف‌شده روی A و ε_{ni} نیز خطاهای تصادفی وابسته زبرجمعی منفی‌اند. یک برآوردگر برای $g(\cdot)$ توسط جرجیو (۱۹۸۵) به صورت زیر معرفی شده است:

$$g_n(x) = \sum_{i=۱}^{n} w_{ni}(x) y_{ni} \qquad x \in A \subset \mathbb{R}^p.$$

در این‌جا $w_{ni}(x) = \omega(x; x_{n۱}, x_{n۲}, \ldots, x_{nn})$ برای $i = ۱, ۲, \ldots, n$ توابع وزن هستند.

به ازای هر تابع $g(x)$، $c(g)$ را مجموعه نقاط پیوستگی تابع g در A تعریف می‌کنیم. برای هر نقطه ثابت $x \in A$ فرض می‌کنیم شرایط زیر برقرار باشند:

۱) اگر $n \longrightarrow \infty$، آن‌گاه

$$\sum_{i=۱}^{n} w_{ni} \longrightarrow ۱.$$

۲) به ازای هر n

$$\sum_{i=۱}^{n} |w_{ni}| \leq C < \infty$$

که C یک عدد حقیقی ثابت است.

۳) به ازای هر $a > ۰$ و زمانی که $n \longrightarrow \infty$

$$\sum_{i=۱}^{n} |w_{ni}|. \left|g(x_{ni}) - g(x)\right| I\left(||x_{ni} - x|| > a\right) \longrightarrow ۰$$

که در آن $||\cdot||$ تابع نرم (اقلیدسی) است.

بنابراین اگر به ازای هر $p \geq ۱$، $E|X|^{۲p} < \infty$ و

$$\sup_{i \geq ۱} |w_{ni}(x)| = O\left(n^{-\frac{۱}{p}}\right)$$

آن‌گاه برای هر $x \in c(g)$ و $n \longrightarrow \infty$، تابع $g_n(x)$ به‌طور کامل به $g(x)$ میل می‌کند (وانگ و همکاران، ۲۰۱۴).

۳.۳ سایر نتایج نظری لازم برای متغیرهای تصادفی NSD

در این بخش سایر نتایج نظری را که برای استفاده در فصل‌های بعدی نیاز داریم در مورد متغیرهای تصادفی NSD بیان می‌کنیم.

قضیه ۱.۳.۳. (وانگ و همکاران، ۲۰۱۵) فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی NSD باشد که متغیر تصادفی X بر آن‌ها غلبه تصادفی دارد به‌طوری‌که $p > ۰$ی موجود است به‌طوری که $E|X|^{۲p} < \infty$ و برای هر $n \geq ۱$، $E(X_n) = ۰$. همچنین، فرض کنید $\{a_{ni}, i \geq ۱, n \geq ۱\}$ آرایه‌ای از اعداد ثابت باشد در این صورت $\delta > ۰$ی موجود است که

$$\max_{۱\leq i\leq n} |a_{ni}| = O(n^{-\frac{۱}{p}}) \tag{۳.۳}$$

و

$$\begin{cases} \sum\limits_{i=۱}^{n} |a_{ni}|^p = O(n^{-\delta}), & ۰ \leq p \leq ۱ \\ \sum\limits_{i=۱}^{n} a_{ni}^{۲} = O((logn)^{-۱}), & p > ۱ \end{cases} \tag{۴.۳}$$

آن‌گاه برای هر $\varepsilon > ۰$

$$\sum_{n=۱}^{\infty} P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}X_i\right| > \varepsilon\right) < \infty. \tag{۵.۳}$$

برهان. بدون کم شدن از کلیت مسأله، فرض کنید $a_{ni} > ۰$ باشد (در غیر این‌صورت، به‌جای a_{ni} از a_{ni}^{+} و a_{ni}^{-} استفاده می‌شود که در آن $a_{ni} = a_{ni}^{+} - a_{ni}^{-}$). برای هر $\varepsilon > ۰$، $۱ \leq i \leq n$ و $n \geq ۱$، با انتخاب عدد صحیح و مثبت N و عدد ثابت مثبتی مانند q، به‌طوری که

$$\begin{cases} q < \dfrac{۱ - \frac{۱}{p}}{p} & \text{اگر } p > ۱ \\ q < \delta & \text{اگر } ۰ < p \leq ۱ \end{cases}$$

می‌توان افراز زیر را برای X_i در نظر گرفت:

$$\begin{aligned} X_{ni}(۱) &= X_i I(|a_{ni}X_i| \leq n^{-q}) - a_{ni}^{-۱}n^{-q}I(a_{ni}X_i < -n^{-q}) + a_{ni}^{-۱}n^{-q}I(a_{ni}X_i > n^{-q}) \\ X_{ni}(۲) &= (X_i - a_{ni}^{-۱}n^{-q})I(n^{-q} < a_{ni}X_i \leq \frac{\varepsilon}{N}) \\ X_{ni}(۳) &= (X_i + a_{ni}^{-۱}n^{-q})I(-\frac{\varepsilon}{N} \leq a_{ni}X_i < -n^{-q}) \\ X_{ni}(۴) &= (X_i + a_{ni}^{-۱}n^{-q})I(a_{ni}X_i < -\frac{\varepsilon}{N}) + (X_i - a_{ni}^{-۱}n^{-q})I(a_{ni}X_i > \frac{\varepsilon}{N}). \end{aligned} \tag{۶.۳}$$

زیرا

$$\begin{aligned}
X_{ni}(۱)+X_{ni}(۲)+X_{ni}(۳)+X_{ni}(۴) &= X_iI(|a_{ni}X_i|\leq n^{-q})-a_{ni}^{-۱}n^{-q}\\
&\quad I(a_{ni}X_i<-n^{-q})+a_{ni}^{-۱}n^{-q}I(a_{ni}X_i>n^{-q})\\
&\quad +(X_i-a_{ni}^{-۱}n^{-q})I(n^{-q}<a_{ni}X_i\leq\frac{\varepsilon}{N})\\
&\quad +(X_i+a_{ni}^{-۱}n^{-q})I(-\frac{\varepsilon}{N}\leq a_{ni}X_i<-n^{-q})\\
&\quad (X_i+a_{ni}^{-۱}n^{-q})I(a_{ni}X_i<-\frac{\varepsilon}{N})+\\
&\quad (X_i-a_{ni}^{-۱}n^{-q})I(a_{ni}X_i>\frac{\varepsilon}{N})\\
&= X_i\Big[I(|a_{ni}X_i|\leq n^{-q})+I(n^{-q}<a_{ni}X_i\leq\frac{\varepsilon}{N})\\
&\quad +I(-\frac{\varepsilon}{N}\leq a_{ni}X_i<-n^{-q})+I(a_{ni}X_i<-\frac{\varepsilon}{N})\\
&\quad +I(a_{ni}X_i>\frac{\varepsilon}{N})\Big]\\
&\quad +a_{ni}^{-۱}n^{-q}\Big[-I(a_{ni}X_i<-n^{-q})\\
&\quad +I(a_{ni}X_i>n^{-q})-I(n^{-q}<a_{ni}X_i\leq\frac{\varepsilon}{N})\\
&\quad +I(-\frac{\varepsilon}{N}\leq a_{ni}X_i<-n^{-q})\\
&\quad +I(a_{ni}X_i<-\frac{\varepsilon}{N})-I(a_{ni}X_i>\frac{\varepsilon}{N})\Big]. \qquad (۷.۳)
\end{aligned}$$

حال مجموعه‌ها را به‌صورت زیر در نظر بگیرید:

$$\begin{aligned}
&A_۱=(|a_{ni}X_i|\leq n^{-q});\quad A_۲=(n^{-q}<a_{ni}X_i\leq\tfrac{\varepsilon}{N});\quad A_۳=I(-\tfrac{\varepsilon}{N}\leq a_{ni}X_i<-n^{-q})\\
&A_۴=(a_{ni}X_i<-\tfrac{\varepsilon}{N});\quad A_۵=(a_{ni}X_i>\tfrac{\varepsilon}{N}).\\
&B_۱=(a_{ni}X_i<-n^{-q});\quad B_۲=(a_{ni}X_i>n^{-q});\quad B_۳=(n^{-q}<a_{ni}X_i\leq\tfrac{\varepsilon}{N})\\
&B_۴=(-\tfrac{\varepsilon}{N}\leq a_{ni}X_i<-n^{-q});\quad B_۵=(a_{ni}X_i<-\tfrac{\varepsilon}{N});\quad B_۶=(a_{ni}X_i>\tfrac{\varepsilon}{N}).
\end{aligned}$$

با در نظر گرفتن مجموعه‌های بالا، ملاحظه می‌شود که

$$\bigcup_{i=۱}^{۵}A_i=\mathbb{R}$$

در نتیجه (۷.۳) به‌صورت زیر تبدیل می‌شود:

$$\begin{aligned}
X_{ni}(۱)+X_{ni}(۲)+X_{ni}(۳)+X_{ni}(۴) &= X_i[I_{\{A_۱\}}+I_{\{A_۲\}}+I_{\{A_۳\}}+I_{\{A_۴\}}+I_{\{A_۵\}}]\\
&\quad +a_{ni}^{-۱}n^{-q}[-I_{\{B_۱\}}+I_{\{B_۲\}}-I_{\{B_۳\}}+I_{\{B_۴\}}\\
&\quad +I_{\{B_۵\}}-I_{\{B_۶\}}]\\
&= X_i[I_{\{\bigcup_{i=۱}^{۵}A_i\}}]+a_{ni}^{-۱}n^{-q}[I_{\{B_۲,B_۴,B_۵\}}-I_{\{B_۱,B_۳,B_۶\}}]\\
&= X_i[I_{\{\mathbb{R}\}}]\\
&= X_i. \qquad (۸.۳)
\end{aligned}$$

برای اثبات (۵.۳)، با استفاده از (۸.۳)، به‌صورت زیر عمل می‌کنیم:

$$\begin{aligned}\sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_i\right| > 4\varepsilon\right) &\le \sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(1)\right| > \varepsilon\right)\\ &+\sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(2)\right| > \varepsilon\right)\\ &+\sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(3)\right| > \varepsilon\right)\\ &+\sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(4)\right| > \varepsilon\right)\\ &= I_1 + I_2 + I_3 + I_4. \qquad (9.3)\end{aligned}$$

برای اثبات قضیه کافی است نشان دهیم که $I_1 < \infty$، $I_2 < \infty$، $I_3 < \infty$ و $I_4 < \infty$. برای این کار، ابتدا می‌توان گفت

$$\begin{aligned} I_1 &= \sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(1)\right| > \varepsilon\right)\\ &\le \sum_{n=1}^{\infty} P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}\left(X_{ni}(1) - EX_{ni}(1)\right)\right| > \varepsilon\right)\\ &< \infty. \qquad (10.3)\end{aligned}$$

وقتی $n\to\infty$، برابری عبارت دوم در (۱۰.۳) برقرار است، زیرا

$$\begin{aligned}\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}EX_{ni}(1)\right| &= \max_{1\le j\le n}\left\{\left|\sum_{i=1}^{1} a_{ni}EX_{ni}(1)\right|, \left|\sum_{i=1}^{2} a_{ni}EX_{ni}(1)\right|, \ldots,\right.\\ &\qquad \left.\left|\sum_{i=1}^{n-1} a_{ni}EX_{ni}(1)\right|, \left|\sum_{i=1}^{n} a_{ni}EX_{ni}(1)\right|\right\}\\ &\le \sum_{i=1}^{n}\left|a_{ni}EX_{ni}(1)\right| \qquad (11.3)\\ &= \sum_{i=1}^{n}\left|a_{ni}E\left[X_iI(|a_{ni}X_i|\le n^{-q}) - a_{ni}^{-1}n^{-q}I(a_{ni}X_i < -n^{-q})\right.\right.\\ &\qquad \left.\left.+a_{ni}^{-1}n^{-q}I(a_{ni}X_i > n^{-q})\right]\right|\\ &\le \sum_{i=1}^{n}\left|E\left[a_{ni}X_iI(|a_{ni}X_i|\le n^{-q}) + n^{-q}I(|a_{ni}X_i| > n^{-q})\right]\right|.\end{aligned}$$

از آن‌جایی که $E\left[I(|a_{ni}X_i| > n^{-q})\right] = P(|a_{ni}X_i| > n^{-q})$، آن‌گاه

$$\begin{aligned}\max_{1\le j\le n}\left|\sum_{i=1}^{j} a_{ni}EX_{ni}(1)\right| &\le \sum_{i=1}^{n}\left|E\left[a_{ni}X_iI(|a_{ni}X_i|\le n^{-q})\right]\right.\\ &\qquad \left.+n^{-q}P(|a_{ni}X_i| > n^{-q})\right|. \qquad (12.3)\end{aligned}$$

برای جمله آخر در سمت راست (۱۲.۳)، با توجه به نابرابری مارکف برای $\circ < p \leq ۱$، داریم

$$\begin{aligned}\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| &\leq \sum_{i=۱}^{n}\big|E\left[a_{ni}X_iI(|a_{ni}X_i|\leq n^{-q})\right] \\ &\quad +n^{-q}\frac{E|a_{ni}X_i|^p}{n^{-pq}}| \\ &= \sum_{i=۱}^{n}\big|E\left[a_{ni}X_iI(|a_{ni}X_i|\leq n^{-q})\right] \\ &\quad +n^{-(۱-p)q}E|a_{ni}X_i|^p| \\ &= \sum_{i=۱}^{n}\big|E|a_{ni}X_i|^pE|a_{ni}X_i|^{۱-p}I(|a_{ni}X_i|\leq n^{-q}) \\ &\quad +n^{-(۱-p)q}E|a_{ni}X_i|^p|.\end{aligned}$$

با توجه به این‌که $|a_{ni}X_i| \leq n^{-q}$، لذا

$$\begin{aligned}\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| &\leq \sum_{i=۱}^{n}\big|E|a_{ni}X_i|^p(n^{-q})^{۱-p}I(|a_{ni}X_i|\leq n^{-q}) \\ &\quad +n^{-(۱-p)q}E|a_{ni}X_i|^p| \\ &\leq \sum_{i=۱}^{n}\big|E|a_{ni}X_i|^p(n^{-q})^{۱-p}\big] \\ &\quad +n^{-(۱-p)q}E|a_{ni}X_i|^p| \\ &= ۲\sum_{i=۱}^{n}n^{-(۱-p)q}E|a_{ni}X_i|^p. \end{aligned} \qquad (۱۳.۳)$$

آن‌گاه در (۱۳.۳)، با اعمال لم ۴.۶.۱ می‌توان نوشت

$$\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| \leq C\sum_{i=۱}^{n}n^{-(۱-p)q}|a_{ni}|^pE|X|^p. \qquad (۱۴.۳)$$

سپس بنا به (۴.۳)، برای هر $\delta > \circ$ی موجود است به‌طوری که می‌توان نوشت

$$\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| \leq Cn^{-(۱-p)q}O(n^{-\delta}). \qquad (۱۵.۳)$$

اگر $n \to \infty$، آن‌گاه

$$Cn^{-(۱-p)q}O(n^{-\delta}) \longrightarrow \circ.$$

در نتیجه

$$\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| \xrightarrow{c.} \circ. \qquad (۱۶.۳)$$

حال نتیجه را وقتی $p > ۱$ باشد مورد بررسی قرار می‌دهیم.

اگر $p > ۱$ باشد، بنا به فرض قضیه $\frac{۲-\frac{۲}{p}}{۲p-۱} < \frac{۱-\frac{۱}{p}}{p} < q$ و $EX_i = ۰$ است و با توجه به نامساوی (۱۱.۳) داریم

$$\begin{aligned}\max_{۱\le j\le n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| &\le \sum_{i=۱}^{n}|a_{ni}EX_{ni}(۱)| \\ &= \sum_{i=۱}^{n}\Big|a_{ni}E\left[X_iI(|a_{ni}X_i|\le n^{-q})\right] - n^{-q}P(a_{ni}X_i < -n^{-q}) \\ &\quad + n^{-q}P(a_{ni}X_i > n^{-q})\Big| \\ &\le \sum_{i=۱}^{n}\Big|a_{ni}E\left[X_i\right] - n^{-q}P(a_{ni}X_i < -n^{-q}) \\ &\quad + n^{-q}P(a_{ni}X_i > n^{-q})\Big|. \end{aligned} \tag{۱۷.۳}$$

طبق فرض قضیه، $EX_i = ۰$ است. بنابراین نابرابری آخر در (۱۷.۳) برابر است با

$$\begin{aligned}\max_{۱\le j\le n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| &\le \sum_{i=۱}^{n}\left|n^{-q}P(a_{ni}X_i > n^{-q}) - n^{-q}P(a_{ni}X_i < -n^{-q})\right| \\ &= \sum_{i=۱}^{n}\left|n^{-q}P(a_{ni}X_i > n^{-q}) - n^{-q}\left(۱ - P(a_{ni}X_i > -n^{-q})\right)\right| \\ &= \sum_{i=۱}^{n}n^{-q}P(a_{ni}X_i > n^{-q}) + \sum_{i=۱}^{n}n^{-q}P(a_{ni}X_i > -n^{-q}) \\ &\quad -nn^{-q} \\ &\le \sum_{i=۱}^{n}n^{-q}P(a_{ni}X_i>n^{-q})+\sum_{i=۱}^{n}n^{-q}P(a_{ni}X_i>-n^{-q}). \end{aligned} \tag{۱۸.۳}$$

طبق نابرابری مارکف، به ازای هر $p > ۱$

$$P(a_{ni}X_i > n^{-q}) \le \frac{E|a_{ni}X_i|^{۲p}}{(n^{-q})^{۲p}}. \tag{۱۹.۳}$$

به‌طور مشابه

$$P(a_{ni}X_i > -n^{-q}) \le \frac{E|a_{ni}X_i|^{۲p}}{(-n^{-q})^{۲p}} \tag{۲۰.۳}$$

زیرا $EI(a_{ni}X_i > n^{-q}) = P(a_{ni}X_i > n^{-q})$ و $EI(a_{ni}X_i > -n^{-q}) = P(a_{ni}X_i > -n^{-q})$. آن‌گاه با جای‌گذاری (۱۹.۳) و (۲۰.۳) در (۱۸.۳)، خواهیم داشت

$$\begin{aligned}\max_{۱\le j\le n}\left|\sum_{i=۱}^{j} a_{ni}EX_{ni}(۱)\right| &\le \sum_{i=۱}^{n}n^{(۲p-۱)q}E|a_{ni}X_i|^{۲p} + \sum_{i=۱}^{n}n^{(۲p-۱)q}E|a_{ni}X_i|^{۲p} \\ &= ۲\sum_{i=۱}^{n}n^{(۲p-۱)q}E|a_{ni}X_i|^{۲p}. \end{aligned} \tag{۲۱.۳}$$

سپس طبق لم ۴.۶.۱ و با در نظر گرفتن ثابت مثبتی مانند $C_۱$، داریم

$$\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}EX_{ni}(1)\right| \leq \underbrace{2C_1}_{C_2}\sum_{i=1}^{n} n^{(2p-1)q}E|a_{ni}X|^{2p} = C_2\sum_{i=1}^{n} n^{(2p-1)q}E|a_{ni}X|^{2p}. \quad (۲۲.۳)$$

چون $E|X|<\infty$، لذا

$$\begin{aligned}\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}EX_{ni}(1)\right| &\leq Cn^{(2p-1)q}\sum_{i=1}^{n}|a_{ni}|^{2p}\\ &= Cn^{(2p-1)q}(\max_{1\leq i\leq n}|a_{ni}|^{2p-2})\sum_{i=1}^{n}|a_{ni}|^{2}\\ &\leq Cn^{(2p-1)q}\left(n^{-\frac{1}{p}}\right)^{2p-2}O\left((\log n)^{-1}\right)\\ &= Cn^{(2p-1)q-(2-\frac{2}{p})}O\left((\log n)^{-1}\right).\end{aligned} \quad (۲۳.۳)$$

عبارت سوم در نامعادله (۲۳.۳)، بر اساس (۴.۳) به دست آمده است. حال اگر $n\to\infty$،

$$Cn^{(2p-1)q-(2-\frac{2}{p})}O\left((\log n)^{-1}\right) \longrightarrow 0 \quad (۲۴.۳)$$

که بنا بر (۲۴.۳)، اگر $n\to\infty$

$$\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}EX_{ni}(1)\right| \xrightarrow{c.} 0 \quad (۲۵.۳)$$

و اثبات کامل می‌شود زیرا نشان می‌دهد عبارت دوم معادله (۱۰.۳) برقرار است. حال عبارت سوم در (۱۰.۳) را ثابت کنیم و نشان می‌دهیم $I_1<\infty$.

از آن‌جا که دنباله $\{X_n, n\geq 1\}$ NSD است، طبق ملاحظه ۱۱.۳.۱ دنباله $\{X_{ni}, n\geq 1, i\geq 1\}$، نیز NSD خواهد بود. در نتیجه بنابر لم ۶.۳.۱ هر تابعی از آن نیز NSD است. پس با توجه به لم ۹.۳.۱ دنباله $\{a_{ni}(X_{ni}(1)-EX_{ni}(1)), 1\leq i\leq n\}$ نیز NSD خواهد بود.

حال می‌خواهیم لم ۲.۲.۲ (نابرابری نمایی کلموگروف) را جهت اثبات، اعمال کنیم. در لم ۲.۲.۲ فرض کنید $x=\varepsilon$، $y=2n^{-q}$، و $\eta=\frac{1}{2}$. لذا داریم

$$\begin{aligned}
&\sum_{n=1}^{\infty} P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right|>\varepsilon\right)\\
&\leq 2\sum_{n=1}^{\infty} P\left(\max_{1\leq i\leq n}\left|a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right|>2n^{-q}\right)\\
&\quad +C\sum_{n=1}^{\infty}\exp\left\{-\frac{\frac{\varepsilon^2}{2}}{2(2\varepsilon n^{-q}+B_n)}\right\}\\
&= 2\sum_{n=1}^{\infty}\left(1-P\left(\max_{1\leq i\leq n}\left|a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right|\leq 2n^{-q}\right)\right)\\
&\quad +C\sum_{n=1}^{\infty}\exp\left\{-\frac{\frac{\varepsilon^2}{2}}{2(2\varepsilon n^{-q}+B_n)}\right\} \qquad (۲۶.۳)
\end{aligned}$$

که در آن $B_n=\sum_{i=1}^{n} E\left[a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right]^2$ است. برای جمله آخر سمت چپ در (۲۶.۳)، به ازای هر $n\geq 1$،

$$\begin{aligned}
\max_{1\leq i\leq n}\left|a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right| &\leq \max_{1\leq i\leq n}\left(\left|a_{ni}\right|\left|\left(X_{ni}(1)-EX_{ni}(1)\right)\right|\right)\\
&\leq \max_{1\leq i\leq n}\left(\left|a_{ni}\right|\left|X_{ni}(1)\right|\right)+\max_{1\leq j\leq n}\left(\left|a_{ni}\right|\left|EX_{ni}(1)\right|\right).
\end{aligned}$$

بر اساس (۲۵.۳) وقتی $n\longrightarrow\infty$، جمله سمت راست در عبارت آخر نابرابری فوق برابر صفر است. لذا

$$\begin{aligned}
\max_{1\leq i\leq n}\left|a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right| &\leq \max_{1\leq i\leq n}\left(\left|a_{ni}\right|\left|X_{ni}(1)\right|\right)\\
&= \max_{1\leq i\leq n}\Big(\left|a_{ni}\right|\Big|X_iI(|a_{ni}X_i|\leq n^{-q})\\
&\quad -a_{ni}^{-1}n^{-q}I(a_{ni}X_i<-n^{-q})\\
&\quad +a_{ni}^{-1}n^{-q}I(a_{ni}X_i>n^{-q})\Big|\Big)\\
&\leq \max_{1\leq i\leq n}\Big(|a_{ni}||X_iI(|a_{ni}X_i|\leq n^{-q})|\\
&\quad +|a_{ni}^{-1}n^{-q}I(a_{ni}X_i<-n^{-q})|\\
&\quad +|a_{ni}^{-1}n^{-q}I(a_{ni}X_i>n^{-q})|\Big)\\
&= \max_{1\leq i\leq n}\Big(\big(|a_{ni}X_iI(|a_{ni}X_i|\leq n^{-q})|\\
&\quad +|n^{-q}I(a_{ni}X_i<-n^{-q})|\\
&\quad +|n^{-q}I(a_{ni}X_i>n^{-q})|\Big)\\
&\leq 2n^{-q}. \qquad (۲۷.۳)
\end{aligned}$$

در (۲۷.۳) برای nهای به اندازه کافی بزرگ، داریم

$$P\left(\max_{1\leq i\leq n}\left|a_{ni}\left(X_{ni}(1)-EX_{ni}(1)\right)\right|\leq 2n^{-q}\right)=1. \qquad (۲۸.۳)$$

از طرفی اگر $۰ \leq p \leq ۱$ باشد، آن‌گاه

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq \sum_{i=۱}^{n} E\left(a_{ni}X_{ni}(۱)\right)^۲ \\ &= \sum_{i=۱}^{n} E\left[a_{ni}^۲\left[X_i I(|a_{ni}X_i| \leq n^{-q}) - a_{ni}^{-۱}n^{-q}I(a_{ni}X_i < -n^{-q})\right.\right. \\ &\quad \left.\left. + a_{ni}^{-۱}n^{-q}I(a_{ni}X_i > n^{-q})\right]^۲\right] \\ &= \sum_{i=۱}^{n} E\left[a_{ni}^۲\left[X_i^۲ I(|a_{ni}X_i| \leq n^{-q}) - a_{ni}^{-۲}n^{-۲q}I(a_{ni}X_i < -n^{-q})\right.\right. \\ &\quad \left.\left. + a_{ni}^{-۲}n^{-۲q}I(a_{ni}X_i > n^{-q})\right]\right] \\ &= \sum_{i=۱}^{n} E\left[a_{ni}^۲X_i^۲ I(|a_{ni}X_i| \leq n^{-q}) - n^{-۲q}I(|a_{ni}X_i| > n^{-q})\right]. \end{aligned} \tag{۲۹.۳}$$

عبارت دوم در نابرابری (۲۹.۳) از تعریف واریانس به‌دست آمده است. حال با استفاده از نابرابری مارکوف، جمله آخر (۲۹.۳) به صورت زیر به‌دست می‌آید:

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq \sum_{i=۱}^{n}\left(E\left[a_{ni}^۲X_i^۲ I(|a_{ni}X_i| \leq n^{-q})\right] + n^{-۲q}\frac{E|a_{ni}X_i|^p}{n^{-pq}}\right) \\ &= \sum_{i=۱}^{n}\left(E\left[a_{ni}^۲X_i^۲ I(|a_{ni}X_i| \leq n^{-q})\right] + n^{-(۲-p)q}E|a_{ni}X_i|^p\right). \end{aligned} \tag{۳۰.۳}$$

چون $۰ \leq p \leq ۱$ است، پس می‌توان نابرابری (۳۰.۳) را به صورت زیر تغییر داد:

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq \sum_{i=۱}^{n}\left(E\left[[a_{ni}X_i]^{۲-p}[a_{ni}X_i]^p I(|a_{ni}X_i| \leq n^{-q})\right] + n^{-(۲-p)q}E|a_{ni}X_i|^p\right) \\ &\leq \sum_{i=۱}^{n}\left(E\left[(n^{-q})^{۲-p}[a_{ni}X_i]^p I(|a_{ni}X_i| \leq n^{-q})\right] + n^{-(۲-p)q}E|a_{ni}X_i|^p\right) \\ &\leq \sum_{i=۱}^{n}\left(E\left[(n^{-q})^{۲-p}[a_{ni}X_i]^p\right] + n^{-(۲-p)q}E|a_{ni}X_i|^p\right) \\ &= ۲\sum_{i=۱}^{n} E\left[n^{-(۲-p)q}E[a_{ni}X_i]^p\right]. \end{aligned} \tag{۳۱.۳}$$

با توجه به لم ۴.۶.۱ و فرض قضیه که در آن متغیر تصادفی X بر دنباله $\{X_n, n \geq ۱\}$ غلبه تصادفی دارد، با انتخاب ثابت مثبتی مانند C، داریم:

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq Cn^{-(۲-p)q} \sum_{i=۱}^{n} |a_{ni}|^p E|X|^p. \end{aligned} \tag{۳۲.۳}$$

حال با استفاده از (۴.۳) در نابرابری (۳۲.۳)، می‌توان نوشت

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq Cn^{-(۲-p)q-\delta} \\ &= O\left((\log n)^{-۱}\right). \end{aligned} \tag{۳۳.۳}$$

از طرفی اگر $p > ۱$ باشد و با توجه به این‌که $|X_{ni}(۱)| \leq |X_i|$، زیرا $X_{ni}(۱)$ بریده شده X_i است. لذا

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱)\right)\right]^۲ \\ &= \sum_{i=۱}^{n} a_{ni}^۲ E\left[|X_{ni}(۱)|\right]^۲. \end{aligned} \tag{۳۴.۳}$$

حال با قرار دادن $|X_{ni}(۱)| \leq |X_i|$ در عبارت آخر نابرابری (۳۴.۳)، داریم

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq \sum_{i=۱}^{n} a_{ni}^۲ E\left[|X_i|\right]^۲. \end{aligned} \tag{۳۵.۳}$$

با در نظر گرفتن ثابتی مانند C و با استفاده از لم ۴.۶.۱، معادله (۴.۳) و همچنین $EX^۲ < \infty$ در نابرابری فوق، می‌توان نوشت

$$\begin{aligned} B_n &= \sum_{i=۱}^{n} E\left[a_{ni}\left(X_{ni}(۱) - EX_{ni}(۱)\right)\right]^۲ \\ &\leq C\sum_{i=۱}^{n} a_{ni}^۲ EX^۲ \\ &= O\left((\log n)^{-۱}\right). \end{aligned} \tag{۳۶.۳}$$

در نهایت با توجه به نتایج مشابه در حالت‌های $۰ < p \leq ۱$ و $p > ۱$ در (۳۳.۳) و (۳۶.۳)، با جای‌گذاری (۲۷.۳)، (۳۶.۳) و (۳۳.۳) در نابرابری (۲۶.۳)، خواهیم داشت

$$\sum_{n=۱}^{\infty} P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}\left(X_{ni}(۱)-EX_{ni}(۱)\right)\right|>\varepsilon\right) \leq C\sum_{n=۱}^{\infty}\exp\left\{-\frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon n^{-q}+O\left((\log n)^{-۱}\right))}\right\}. \quad (۳۷.۳)$$

می‌دانیم $n^{-q} < ۱$ است. بنابراین $۲\varepsilon n^{-q} < ۲\varepsilon$. پس می‌توان نوشت

$$۲\varepsilon n^{-q}+O\left((\log n)^{-۱}\right) < ۲\varepsilon + O\left((\log n)^{-۱}\right).$$

پس

$$۲\left(۲\varepsilon n^{-q}+O\left((\log n)^{-۱}\right)\right) < ۲\left(۲\varepsilon + O\left((\log n)^{-۱}\right)\right).$$

با معکوس کردن و ضرب $\frac{\varepsilon^۲}{۲}$، داریم

$$\frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon n^{-q}+O\left((\log n)^{-۱}\right))} > \frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon+O\left((\log n)^{-۱}\right))}.$$

در نتیجه

$$-\frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon n^{-q}+O\left((\log n)^{-۱}\right))} < -\frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon+O\left((\log n)^{-۱}\right))}.$$

از طرفی

$$\begin{aligned} -\frac{\frac{\varepsilon^۲}{۲}}{۲(۲\varepsilon+O\left((\log n)^{-۱}\right))} &= -\frac{\frac{\varepsilon}{۲}}{۴+۲\frac{O\left((\log n)^{-۱}\right)}{\varepsilon}} \\ &\leq -\frac{\frac{\varepsilon}{۲}}{۴+۲O\left((\log n)^{-۱}\right)} \\ &\leq -\frac{\frac{\varepsilon}{۲}}{O\left((\log n)^{-۱}\right)}. \end{aligned} \quad (۳۸.۳)$$

حال برای هر $\varepsilon > ۰$ و با جای‌گذاری (۳۸.۳) در نابرابری (۳۷.۳) داریم

$$\begin{aligned} \sum_{n=۱}^{\infty} P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}\left(X_{ni}(۱)-EX_{ni}(۱)\right)\right|>\varepsilon\right) &\leq C\sum_{n=۱}^{\infty}\exp\left\{-۲\log n\right\} \\ &\leq C\sum_{n=۱}^{\infty}\exp\left\{\log n^{-۲}\right\} \\ &= C\sum_{n=۱}^{\infty}\frac{۱}{n^۲} < \infty. \end{aligned} \quad (۳۹.۳)$$

این نشان می‌دهد که $I_۱ < \infty$ است.

برای اثبات $I_۲ < \infty$، مشابه آن‌چه در (۱۱.۳) بیان شد، داریم

$$\begin{aligned}\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}X_{ni}(۲)\right| &= \max\left\{\left|\sum_{i=۱}^{۱} a_{ni}X_{ni}(۲)\right|, \left|\sum_{i=۱}^{۲} a_{ni}X_{ni}(۲)\right|, \ldots,\right.\\ &\qquad \left.\left|\sum_{i=۱}^{n-۱} a_{ni}X_{ni}(۲)\right|, \left|\sum_{i=۱}^{n} a_{ni}X_{ni}(۲)\right|\right\}\\ &\leq \sum_{i=۱}^{n}|a_{ni}X_{ni}(۲)|\\ &= \sum_{i=۱}^{n} a_{ni}X_{ni}(۲).\end{aligned} \tag{۴۰.۳}$$

فرض کنید $N = N_۱$ باشد و چون $۰ < a_{ni}X_{ni}(۲) \leq \frac{\varepsilon}{N_۱}$ است، بنابراین

$$\begin{aligned}P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}X_{ni}(۲)\right| > \varepsilon\right) &\leq P\left(\sum_{i=۱}^{n} a_{ni}X_{ni}(۲) > \varepsilon\right)\\ &\leq \sum_{i=۱}^{n} P\left(a_{ni}X_{ni}(۲) > \frac{\varepsilon}{n}\right)\end{aligned} \tag{۴۱.۳}$$

رابطه (۴۱.۳) نشان می‌دهد حداقل یک عدد صحیح مانند $N_۱$ وجود دارد به‌طوری‌که $a_{ni}X_{ni}(۲) \neq ۰$. در نتیجه با تعریف زیربازه‌هایی در بازه $[۱, n]$ در (۴۱.۳)، داریم

$$\begin{aligned}P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}X_{ni}(۲)\right| > \varepsilon\right) &\leq \sum_{۱<i_۱<i_۲<\ldots<i_{N_۱}<n} P\left(a_{ni_۱}X_{ni_۱}(۲)\neq ۰, a_{ni_۲}X_{ni_۲}(۲)\neq ۰\right.\\ &\qquad \left., \ldots, a_{ni_{N_۱}}X_{ni_{N_۱}}(۲) \neq ۰\right).\end{aligned} \tag{۴۲.۳}$$

با توجه به (۶.۳) و بازه در نظر گرفته‌شده برای $X_{ni}(۲)$، در نابرابری (۴۲.۳)، داریم

$$\begin{aligned}P\left(\max_{۱\leq j\leq n}\left|\sum_{i=۱}^{j} a_{ni}X_{ni}(۲)\right| > \varepsilon\right) &\leq \sum_{۱<i_۱<i_۲<\ldots<i_{N_۱}<n} P\left(a_{ni_۱}X_{i_۱} > n^{-q}, a_{ni_۲}X_{i_۲} > n^{-q},\right.\\ &\qquad \left.\ldots, a_{ni_{N_۱}}X_{i_{N_۱}} > n^{-q}\right).\end{aligned} \tag{۴۳.۳}$$

حال با استفاده از قسمت سوم لم ۶.۳.۱ می‌توان نوشت

$$\begin{aligned}
P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(2)\right|>\varepsilon\right) &\le \sum_{1<i_1<i_2<\ldots<i_{N_1}<n} P\left(a_{ni_1}X_{i_1}>n^{-q}\right)\\
&\quad P\left(a_{ni_2}X_{i_2}>n^{-q}\right)\ldots P\left(a_{ni_{N_1}}X_{i_{N_1}}>n^{-q}\right)\\
&\le \sum_{1<i_1<i_2<\ldots<i_{N_1}<n}\left(\sum_{i=1}^{n}P\left(a_{ni}X_i>n^{-q}\right)\right.\\
&\quad \left.\sum_{i=1}^{n}P\left(a_{ni}X_i>n^{-q}\right)\ldots\sum_{i=1}^{n}P\left(a_{ni}X_i>n^{-q}\right)\right)\\
&\le \left(\sum_{i=1}^{n}P\left(a_{ni}X_i>n^{-q}\right)\right)^{N_1}\\
&\le \left(C\sum_{i=1}^{n}P\left(|a_{ni}X|>n^{-q}\right)\right)^{N_1}. \qquad (۴۴.۳)
\end{aligned}$$

عبارت آخر در نابرابری فوق با استفاده از لم ۴.۶.۱ و انتخاب ثابت مثبتی مانند C نتیجه شد. حال با توجه به نابرابری مارکف دو حالت اتفاق می‌افتد:

الف) $0<p\le 1$

ب) $p>1$

در ابتدا فرض می‌کنیم $0<p\le 1$ باشد. بنابر نابرابری مارکف، نابرابری (۴۴.۳) به صورت زیر تغییر می‌یابد:

$$\begin{aligned}
P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(2)\right|>\varepsilon\right) &\le C_1\left(\sum_{i=1}^{n}\frac{E|a_{ni}X|^p}{(n^{-q})^p}\right)^{N_1}\\
&= C_1\left(\sum_{i=1}^{n}n^{pq}E|a_{ni}X|^p\right)^{N_1}\\
&= \underbrace{C_1E|X|^{N_1p}}_{\le C}\left(\sum_{i=1}^{n}n^{pq}|a_{ni}|^p\right)^{N_1}\\
&\le C\left(\sum_{i=1}^{n}n^{pq}|a_{ni}|^p\right)^{N_1}. \qquad (۴۵.۳)
\end{aligned}$$

لذا با جای‌گذاری (۴.۳) درجمله آخر نابرابری (۴۵.۳) و برای $\delta>0$

$$\begin{aligned}
P\left(\max_{1\le j\le n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(2)\right|>\varepsilon\right) &\le C\left(n^{pq}O\left(n^{-\delta}\right)\right)^{N_1}\\
&\le Cn^{-(\delta-pq)N_1}. \qquad (۴۶.۳)
\end{aligned}$$

توجه داشته باشید اگر $1 \leq p < 0$ باشد آن‌گاه $q < \frac{\delta}{p}$ است. بنابراین با انتخاب عدد صحیح و بزرگ مانند N_1، به‌طوری که $(\delta - pq)N_1 > 1$ باشد، داریم

$$I_2 = P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(2)\right| > \varepsilon\right) < \infty. \tag{۴۷.۳}$$

برای حالت $p > 1$ نیز به مانند حالت الف عمل می‌کنیم. می‌توان نوشت

$$\begin{aligned}
P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(2)\right| > \varepsilon\right) &\leq C_2\left(\sum_{i=1}^{n}\frac{E|a_{ni}X|^{2p}}{(n^{-q})^{2p}}\right)^{N_1}\\
&= C_2\left(\sum_{i=1}^{n} n^{2pq}E|a_{ni}X|^{2p}\right)^{N_1}\\
&= \underbrace{C_2E|X|^{2pN_1}}_{\leq C}\left(\sum_{i=1}^{n} n^{2pq}|a_{ni}|^{2p}\right)^{N_1}\\
&\leq C\left(n^{2pq}\sum_{i=1}^{n}|a_{ni}|^{2p-2}|a_{ni}|^{2}\right)^{N_1}\\
&\leq C\left(n^{2pq}\left(\max_{1\leq i\leq n}|a_{ni}|^{2p-2}\right)\sum_{i=1}^{n}|a_{ni}|^{2}\right)^{N_1}.
\end{aligned}$$

با جای‌گذاری (۳.۳) و (۴.۳) در عبارت آخر نابرابری فوق، داریم

$$\begin{aligned}
P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(2)\right| > \varepsilon\right) &\leq C\left(n^{2pq}\left(O\left(n^{-\frac{1}{p}}\right)\right)^{2p-2}\left(O\left((\log n)^{-1}\right)\right)\right)^{N_1}\\
&\leq Cn^{-2(1-\frac{1}{p}-pq)N_1}(\log n)^{N_1}.
\end{aligned} \tag{۴۸.۳}$$

با توجه به این‌که $p > 1$ است، لذا $q < \dfrac{1-\frac{1}{p}}{p}$. پس عدد صحیح و بزرگی مانند N_1 موجود است به‌طوری‌که $2(1 - \frac{1}{p} - pq)N_1 > 1$. در نتیجه

$$I_2 = P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(2)\right| > \varepsilon\right) < \infty. \tag{۴۹.۳}$$

برای این‌که نشان دهیم $I_3 < \infty$ ، فرض کنید $N = N_2$ باشد و چون

$$-\frac{\varepsilon}{N_2} \leq a_{ni}X_{ni}(3) < 0 \tag{۵۰.۳}$$

بنابراین مشابه (۴۰.۳) داریم

$$\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(3)\right| = -\sum_{i=1}^{n} a_{ni}X_{ni}(3). \tag{۵۱.۳}$$

این معادله نشان می‌دهد که عدد صحیح مثبتی مانند N_2 موجود است به‌طوری‌که $a_{ni}X_{ni}(3) \neq 0$. بنابراین مشابه آن‌چه برای $I_2 < \infty$ ثابت شد، می‌توان نشان داد

$$I_3 = \sum_{n=1}^{\infty} P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(3)\right| > \varepsilon\right) < \infty. \tag{۵۲.۳}$$

در آخر کافی است نشان دهیم که I_4 نیز متناهی است. مشابه آن‌چه در (۴۰.۳) و (۵۱.۳) اتفاق افتاد، برای عبارت ذیل نیز داریم:

$$\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(4)\right| = \sum_{i=1}^{n}|a_{ni}X_{ni}(4)|. \tag{۵۳.۳}$$

بنابراین با فرض کردن عدد صحیح و مثبت $N=\{N_1,N_2\}$، می‌توان نوشت

$$I_4 = \sum_{n=1}^{\infty}P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(4)\right|>\varepsilon\right) \leq \sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|a_{ni}X_{ni}(4)|>\frac{\varepsilon}{n}\right).$$

معادله فوق نشان می‌دهد که به ازای حداقل یک $i\geq 1$، $a_{ni}X_{ni}(4)\neq 0$ است. با توجه به این‌که $|a_{ni}X_i|>\frac{\varepsilon}{N}$، لذا

$$\sum_{n=1}^{\infty}P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(4)\right|>\varepsilon\right) \leq \sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|a_{ni}X_i|>\frac{\varepsilon}{N}\right). \tag{۵۴.۳}$$

حال با استفاده از لم ۴.۶.۱ و انتخاب ثابت مثبت C، داریم

$$\begin{aligned}
\sum_{n=1}^{\infty}P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j}a_{ni}X_{ni}(4)\right|>\varepsilon\right) &\leq C\sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|a_{ni}X|>\frac{\varepsilon}{N}\right)\\
&= C\sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|a_{ni}||X|>\frac{\varepsilon}{N}\right)\\
&= C\sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(O\left(n^{-\frac{1}{p}}\right)|X|>\frac{\varepsilon}{N}\right)\\
&\leq C\sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|X|>C_1\frac{\varepsilon}{N}n^{\frac{1}{p}}\right)\\
&= C\sum_{n=1}^{\infty}\sum_{i=1}^{n}P\left(|X|>C_2n^{\frac{1}{p}}\right)\\
&= C\sum_{n=1}^{\infty}nP\left(|X|>C_2n^{\frac{1}{p}}\right).
\end{aligned} \tag{۵۵.۳}$$

برابری دوم در (۵۵.۳) بنابر معادله (۳.۳) به‌دست آمده است. از آن‌جا که $|X|\in(Cn^{\frac{1}{p}},+\infty)$ است، می توانیم بازه مذکور را به صورت زیر افراز کنیم:

$$|X|\in\left\{(Cn^{\frac{1}{p}},C(n+1)^{\frac{1}{p}}],(C(n+1)^{\frac{1}{p}},C(n+2)^{\frac{1}{p}}],(C(n+2)^{\frac{1}{p}},C(n+3)^{\frac{1}{p}}],\ldots\right\}.$$

لذا داریم

$$P\left(|X|>C_2n^{\frac{1}{p}}\right)\leq\sum_{k=n}^{\infty}P\left(C_2k^{\frac{1}{p}}<|X|<C_2(k+1)^{\frac{1}{p}}\right). \tag{۵۶.۳}$$

حال با جای‌گذاری (۵۶.۳) در (۵۵.۳)، می‌توان نوشت

$$\begin{aligned}\sum_{n=1}^{\infty} P\left(\max_{1\leq j\leq n}\left|\sum_{i=1}^{j} a_{ni}X_{ni}(4)\right| > \varepsilon\right) &\leq \sum_{n=1}^{\infty} n \sum_{k=n}^{\infty} P\left(C_2 k^{\frac{1}{p}} < |X| < C_2(k+1)^{\frac{1}{p}}\right)\\ &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} nP\left(C_2 k^{\frac{1}{p}} < |X| < C_2(k+1)^{\frac{1}{p}}\right)\\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{k} kP\left(C_2 k^{\frac{1}{p}} < |X| < C_2(k+1)^{\frac{1}{p}}\right)\\ &= \sum_{k=1}^{\infty} k^2 P\left(C_2 k^{\frac{1}{p}} < |X| < C_2(k+1)^{\frac{1}{p}}\right)\\ &= E|X|^{2p} < \infty. \qquad (۵۷.۳)\end{aligned}$$

در پایان با جای‌گذاری (۳۹.۳)، (۴۹.۳)، (۵۲.۳) و (۵۷.۳) در (۹.۳) داریم:

$$\sum_{i=1}^{\infty} P\left(\max_{1\leq j\leq n}\left|\sum_{n=1}^{j} a_{ni}X_i\right| > 4\varepsilon\right) < \infty. \qquad (۵۸.۳)$$

برهان کامل می‌شود. □

ملاحظه ۲.۳.۳. (تروم، ۱۹۸۷) قضیه ۱.۳.۳ در همگرایی کامل، یک نتیجه کلی است. حال اگر در این قضیه، به ازای هر $p > 1$، $a_{ni} = \frac{b_{ni}}{n^{\frac{1}{p}}}$ باشد به‌طوری‌که اعداد حقیقی $\{b_{ni}, 1 \leq i \leq n, n \geq 1\}$ وجود داشته باشند که در برابری‌های

$$\max_{1\leq i\leq n} |b_{ni}| = O(1)$$

و

$$\sum_{i=1}^{n} b_{ni}^2 = O\left(n^{\frac{2}{p}} (\log n)^{-1}\right)$$

صدق کنند،

$$\frac{1}{n^{\frac{1}{p}}} \sum_{i=1}^{n} b_{ni}X_i \xrightarrow{c} 0.$$

یعنی $\frac{1}{n^{\frac{1}{p}}}\sum_{i=1}^{n} b_{ni}X_i$ به طور کامل به صفر همگرا خواهد بود.

فصل ۴

رگرسیون آلوده به خطا

۱.۴ مقدمه

فرضیه اساسی در تحلیل آماری این است که مشاهدات به درستی اندازه‌گیری شده‌اند. در مفهوم مدل‌های رگرسیونی یک‌متغیره یا چندمتغیره، فرض می‌شود که مشاهدات مورد مطالعه و همچنین متغیرهای تبیینی بدون هر خطایی به‌دست آیند. در بسیاری از موارد، این فرضیات اساسی می‌تواند نوعی تخطی به شمار آید. بنابراین، مقادیر درست متغیرها را نمی‌توان به دست آورد. یعنی آن‌ها با مقداری خطا مشاهده می‌شوند. تفاوت بین مقادیر مشاهده‌شده و مقادیر متغیرها، خطای اندازه‌گیری[۱] نامیده می‌شود و به این نوع متغیرها، متغیرهای آلوده به خطا[۲] می‌گویند.

دلایل زیادی برای به‌وجود آمدن خطاها وجود دارند. از رایج‌ترین این خطاها می‌توان به خطای ابزاری و خطای نمونه‌گیری اشاره کرد. در زیر مثال‌هایی را بیان می‌کنیم که نشان می‌دهند تنوع در مفاهیم خطای اندازه‌گیری می‌تواند نگران‌کننده باشد.

- **وضعیت بیماری:** در علم امراض مسری، متغیرهای پاسخ غالبا مبتلا شدن یا نشدن شخص به بیماری هستند. به‌عنوان مثال می‌توان به ایدز، سرطان و هپاتیت اشاره کرد که اغلب تشخیص ابتلا به‌واسطه یک مرحله تشخیصی ناقص انجام می‌شود. تکنیک‌هایی مانند آزمایش خون و تصویربرداری وجود دارند که می‌توان در مثبت یا منفی بودن وجود بیماری از آن‌ها بهره جست.

[۱] Measurement error

[۲] Errors-in -variables

- **داده‌های مربوط به بازماندگان بمب اتمی:** برخی محققان داده‌های مربوط به انفجار بمب اتمی در هیروشیما و ناگازاکی را مورد بررسی قرار دادند که در آن پاسخ، به دلیل تعداد عدم انطباق کروموزوم‌ها می‌تواند متنوع باشد. همچنین به علت تشعشعات اتمی روی متغیرهای پیش‌گو، این متغیرها را نمی‌توان به‌طور دقیق اندازه‌گیری کرد که با کمی خطا قابل اندازه‌گیری هستند.

- **عملکرد دستگاه تنفسی در کودکان:** در این‌جا متغیر پاسخ، وجود خس‌خس هنگام تنفس در کودکان می‌باشد که نشانه‌ای از عملکرد ریه است و می‌توان به آن مقادیر دودویی (صفر یا یک) نسبت داد. متغیر پیش‌گوی مورد علاقه، می‌تواند در معرض نیترید قرار گرفتن کودکان باشد. حال اگر متغیر پیش‌گو موجود باشد، به‌راحتی می‌توان ارتباط میان این متغیر با متغیر پاسخ را به‌دست آورد. اما متغیر پیش‌گو موجود نیست، زیرا ممکن است کودک در یک محل خاص نباشد. برای مثال در یک خانه، کودک می‌تواند در آشپزخانه یا اتاق‌خواب قرار گیرد که در آن‌ها میزان نیترید متفاوت است. در نتیجه اندازه‌گیری متغیر پیش‌گو با خطا همراه خواهد بود.

بسیاری از مجموعه داده‌ها توسط متغیرهای اندازه‌گیری‌شده آلوده می‌شوند. این بدان معنی است که مشکل خطای اندازه‌گیری مشکلی بنیادی در حالت تجربی به‌شمار می‌رود. وجود این خطاهای اندازه‌گیری سبب به وجود آمدن اریبی شده که نتایج را تحت تأثیر قرار می‌دهد و پارامتر به‌طور سازگار برآورد نمی‌شود. سرانجام این خطاها سبب می‌شود که تحلیل نتایج بر پایه نتایج نادقیقی به‌دست آمده و تنوع زیادی در تحلیل آن‌ها به وجود می‌آید.

برای رفع کردن مشکل خطاها در اندازه‌گیری‌ها، تکنیک‌هایی وجود دارند که می‌توان آن‌ها را در دو بعد رده‌بندی کرد. این تکنیک‌ها در مدل‌های رگرسیونی آلوده به خطای خطی و غیرخطی و همچنین در دو رده خطای اندازه‌گیری کلاسیک[۳] و غیرکلاسیک[۴] مطرح می‌شوند. در ابتدا به تعریف خطای اندازه‌گیری کلاسیک و غیرکلاسیک می‌پردازیم. مطالب این فصل از کارول و همکاران (۱۹۹۵) گرفته شده‌اند.

تعریف ۱.۱.۴. اگر خطای اندازه‌گیری مستقل از متغیرهای حقیقی نامعلوم باشد، در این صورت خطای اندازه‌گیری کلاسیک نامیده می‌شود.

تعریف ۲.۱.۴. اگر خطای اندازه‌گیری به متغیرهای حقیقی نامعلوم وابسته باشد، به آن خطای اندازه‌گیری غیرکلاسیک اطلاق می‌شود.

به عنوان مثالی از خطای اندازه‌گیری غیرکلاسیک، یک کارگر اتحادیه را در نظر بگیرید که ممکن است به اشتباه به عنوان یک کارگر خارج اتحادیه رده‌بندی شود. وقتی متغیر مورد نظر و خطای آن به‌صورت صفر و یک باشند، در این صورت خطای اندازه‌گیری نمی‌تواند مستقل از متغیر صحیح صفر و یک باشد.

۱.۱.۴ تفاوت بین خطای تصادفی و خطای اندازه‌گیری

خطای تصادفی در مدل رگرسیونی ناشی از عامل‌هایی مانند غیرقابل پیشگویی بودن عناصر به‌طور تصادفی، فقدان ارتباط قطعی بین متغیر پیشگو و پاسخ خواهد بود. این در حالی‌است که خطای اندازه‌گیری به سبب

[۳] Classical

[۴] Nonclassical

اندازه‌گیری ناقص از متغیرهای صحیح به وجود می‌آید.

۲.۱.۴ خطای اندازه‌گیری کوچک و بزرگ

اگر مقدار خطای اندازه‌گیری کوچک باشد، در این صورت می‌توان آن را با خطای تصادفی ترکیب کرد که تأثیر چندانی در استنباط آماری نخواهد داشت. در مقابل اگر مقدار این خطای اندازه‌گیری بزرگ باشد، در نظر نگرفتن آن به استنباط آماری غلطی منجر خواهد شد. برای مثال، در مدل رگرسیونی خطی برآوردگر کم‌ترین توان‌های دوم معمولی وقتی که خطای اندازه‌گیری وجود نداشته باشد، بهترین برآوردگر نااریب خطی ضرایب رگرسیونی خواهد بود. حال اگر خطای اندازه‌گیری در داده‌ها موجود باشد، همان برآوردگر کم‌ترین توان‌های دوم، اریب شده و برآوردگری ناسازگار برای ضرایب رگرسیونی است.

چون تاکنون روش‌های متنوعی برای مدل‌های آلوده به خطای خطی با خطای اندازه‌گیری کلاسیک معرفی شده‌اند، در این‌جا به توضیحی مختصر درباره مدل‌های رگرسیونی آلوده به خطای خطی با خطای اندازه‌گیری کلاسیک می‌پردازیم.

۲.۴ مدل رگرسیونی آلوده به خطای خطی یک‌متغیره با خطای اندازه‌گیری کلاسیک

فرضیات در خطای اندازه‌گیری بر این مدعی است که خطای اندازه‌گیری در هر متغیر تبیینی از مجموعه داده‌ها، مستقل از سایر متغیرهای تبیینی است. مفهوم این فرضیات در مدل رگرسیونی کم‌ترین توان‌های دوم زیر بسیار قابل فهم است. مدل خطی

$$y_i = \beta x_i^{*'} + \varepsilon_i, \qquad i = ۱, ۲, \dots, n$$

را در نظر بگیرید. تحت این مدل، خطای اندازه‌گیری در متغیر پاسخ به‌صورت $y_i = y_i^* + \nu_i$ منجر به تناقضی در برآورد ضرایب رگرسیونی نمی‌شود. زیرا با بازنویسی مدل در y_i^* می‌توان دید که

$$y_i = \beta x_i^{*'} + \varepsilon_i + \nu_i = \beta x_i^{*'} + \omega_i.$$

تنها اثر وجود خطاهای اندازه‌گیری در متغیرهای پاسخ، این است که خطاهای استاندارد در برآورد ضرایب رگرسیونی زیاد می‌شوند. در مقابل، خطاهای مستقل در متغیرهای تبیینی رگرسیونی یعنی در حالتی که $x_i = x_i^* + \eta_i$ است، به افزایش اریبی (منفی) در مدل رگرسیونی ساده منجر می‌شود و برآوردگرهای ضرایب رگرسیونی در حالت کلی ناسازگار هستند.

۱.۲.۴ کاهش اریبی

مدل رگرسیونی خطی کلاسیک یک‌متغیره زیر را در نظر بگیرید:

$$y = \alpha + \beta x^* + \varepsilon, \qquad E[X^* \varepsilon] = ۰. \tag{۱.۴}$$

در مدل (۱.۴)، x^* تنها مشاهده با یک افزوده و مستقل از خطای اندازه‌گیری $\eta \sim (\circ, \sigma_\eta^2)$ است به طوری که

$$x = x^* + \eta. \tag{۲.۴}$$

در نتیجه رگرسیون y روی x با وارد کردن (۲.۴) در مدل (۱.۴) به‌دست می‌آید. یعنی

$$y = \alpha + \beta x + u, \qquad u = \varepsilon - \beta\eta. \tag{۳.۴}$$

حال فرض کنید برای هر $i = 1, 2, \dots, n$، (y_i, x_i) n مشاهده از یک نمونه تصادفی روی (y, x) باشند. در این صورت، کم‌ترین توان‌های دوم با

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_n)y_i}{\sum_{j=1}^{n}(x_j - \bar{x}_j)^2} \tag{۴.۴}$$

به‌دست می‌آید. با وجود خطای اندازه‌گیری x و u به یک‌دیگر وابسته‌اند، زیرا

$$\begin{aligned} Cov(x, u) &= Cov(x^* + \eta, \varepsilon - \beta\eta) \\ &= Cov(\eta, -\beta\eta) \\ &= -\beta Cov(\eta, \eta) \\ &= -\beta\sigma_\eta^2 \neq \circ. \end{aligned}$$

بنابراین برآوردگر کم‌ترین توان‌های دوم ناسازگار است، زیرا طبق تعریف سازگاری، وقتی $n \longrightarrow \infty$ باید برآوردگر مورد نظر در احتمال به پارامتر مورد نظر همگرا شود. در نتیجه

$$\begin{aligned} \hat{\beta}_n &\approx \beta + \frac{Cov(x, u)}{Var(x)} \\ &\approx \beta - \frac{\beta\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2} \\ &\approx \beta\frac{\sigma_*^2}{\sigma_*^2 + \sigma_\eta^2} \end{aligned}$$

که در آن $\sigma_*^2 = Var(x^*)$ است. چون σ_*^2 و σ_η^2 هر دو مثبت اند، پس $\hat{\beta}$ برآوردی ناسازگار برای β با یک اریبی منفی است. علاوه بر این، با توجه به اریبی ضریب شیب در مدل رگرسیونی تک‌متغیره، برآورد α به صورت زیر به دست می‌آید:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}.$$

در این صورت برای بررسی سازگاری آن می‌توان نوشت

$$\begin{aligned} \hat{\alpha}_n &\approx E[\alpha + \beta x^* + \varepsilon] - \beta\frac{\sigma_*^2}{\sigma_*^2 + \sigma_\eta^2}E[x^* + \eta] \\ &\approx \alpha + \beta\frac{\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2}\mu^* \end{aligned}$$

که در آن $\mu^* = Ex^*$ است. پس $\hat{\alpha}$ نیز برآوردگری ناسازگار است. این نتایج را می‌توان به یک مدل رگرسیونی خطی چندمتغیره نیز تعمیم داد.

۳.۴ مدل رگرسیونی آلوده به خطای خطی چندمتغیره با خطای اندازه‌گیری کلاسیک

در ابتدا فرض کنید بین متغیرهای تبیینی و مورد مطالعه که به درستی مشاهده شده‌اند، ارتباط صحیحی به شکل

$$\tilde{y} = \tilde{X}\beta \tag{۵.۴}$$

وجو داشته باشد که در آن $\tilde{y}$ برداری $(n \times ۱)$ از مشاهدات صحیح از متغیر پاسخ، $\tilde{X}$ ماتریسی $(n \times k)$ از مشاهدات صحیح روی متغیرهای تبیینی و β برداری $(k \times ۱)$ از ضرایب رگرسیونی است.

مقادیر $\tilde{y}$ و $\tilde{X}$ از حضور خطای اندازه‌گیری که قابل مشاهده نیستند تأثیر می‌پذیرد. یعنی این مقادیر با خطای اندازه‌گیری افزوده به صورت زیر مشاهده می‌شوند:

$$\begin{aligned} y &= \tilde{y} + u \\ X &= \tilde{X} + V \end{aligned}$$

که در آن y برداری $(n \times ۱)$ از مقادیر مشاهده‌شده پاسخ است که با بردار $(n \times ۱)$ u، که شامل خطای اندازه‌گیری است مشاهده شده‌اند. به‌طور مشابه، X نیز یک ماتریس $(n \times k)$ از مقادیر مشاهدات متغیرهای تبیینی است که با ماتریس $(n \times k)$ V، از خطای اندازه‌گیری در X همراه هستند.

بدون کم شدن از کلیت مسأله، فرض می‌شود که خطای تصادفی در u بیان شود. بنابراین مدل رگرسیونی آلوده به خطا را می‌توان به‌صورت زیر بازنویسی کرد:

$$\begin{aligned} y &= \tilde{X}\beta + u \\ X &= \tilde{X} + V \end{aligned} \tag{۶.۴}$$

که در آن تنها X با خطای V آلوده شده است و u را می‌توان به‌عنوان خطای تصادفی معمولی در نظر گرفت. در حالتی که متغیرهای تبیینی بدون هر خطایی اندازه‌گیری می‌شوند، خطای V صفر در نظر گرفته می‌شود.
با توجه به مدل رگرسیونی (۶.۴)، فرض می‌شود که

$$\begin{aligned} &E[u] = ۰; \qquad E[uu'] = \sigma^۲ I; \\ &E[V] = ۰; \qquad E[VV'] = \Omega; \quad E[V'u] = ۰. \end{aligned}$$

حال مدل رگرسیونی (۶.۴) را با استفاده از روابط تعریف شده بالا به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned} y &= \tilde{y} + u \\ &= \tilde{X}\beta + u \\ &= (X - V)\beta + u \\ &= X\beta + (u - V\beta) \\ &= X\beta + \omega \end{aligned}$$

که در آن $\omega = u - V\beta$ به‌عنوان ترکیبی از خطای تصادفی نامیده می‌شود.

این مدل، مشابه مدل رگرسیونی خطی معمولی می‌باشد. یک فرضیه اساسی در مدل رگرسیونی خطی این است که در آن متغیرهای تبیینی و خطای تصادفی ناهمبسته هستند. حال این فرضیات را در مدل $y = X\beta + \omega$ مورد بررسی قرار می‌دهیم. برای محاسبه این همبستگی، داریم

$$\begin{aligned} E\left[(X - E[X])'(\omega - E[\omega])\right] &= E\left[V'(u - V\beta)\right] \\ &= E[V'u] - E[V'V]\beta \\ &= ۰ - \Omega\beta \\ &= -\Omega\beta \\ &\neq ۰. \end{aligned}$$

بنابراین X و ω همبسته‌اند. حال فرض کنید خطای اندازه‌گیری نادیده گرفته شود. در این صورت برآوردگر کم‌ترین توان‌های دوم را به دست می‌آوریم. توجه کنید که نادیده گرفتن خطای اندازه‌گیری در داده‌ها بدان معنا نیست که این خطاها وجود ندارند. برآوردگر کم‌ترین توان‌های دوم به‌صورت زیر حاصل می‌شود:

$$b = (X'X)^{-۱}X'y.$$

برای محاسبه اریبی این برآوردگر ابتدا مقدار $b - \beta$ را به‌دست می‌آوردیم:

$$\begin{aligned} b - \beta &= (X'X)^{-۱}X'(X\beta + \omega) - \beta \\ &= (X'X)^{-۱}X'\omega. \end{aligned}$$

بنابراین می‌توان نوشت

$$\begin{aligned} E[b - \beta] &= E[(X'X)^{-۱}X'\omega] \\ &= (X'X)^{-۱}X'E[\omega] \\ &\neq ۰ \end{aligned}$$

که در آن X یک ماتریس تصادفی است که با ω همبسته است. بنابراین b یک برآوردگر اریب برای β خواهد بود.

اکنون سازگاری برآوردگر کمترین توان‌های دوم را مورد بررسی قرار می‌دهیم. به همین منظور فرض کنید وقتی $n \longrightarrow \infty$

$$\frac{۱}{n}\tilde{X}'\tilde{X} \xrightarrow{p} \Sigma_{xx}$$

$$\frac{۱}{n}V'V \xrightarrow{p} \Sigma_{vv}$$

$$\frac{1}{n}\tilde{X}'V \xrightarrow{p} 0$$
$$\frac{1}{n}V'u \xrightarrow{p} 0.$$

با استفاده از روابط بالا، خواهیم داشت

$$b-\beta = (\frac{X'X}{n})^{-1}(\frac{X'\omega}{n}). \tag{۷.۴}$$

از طرفی

$$\begin{aligned}\frac{1}{n}X'X &= \frac{1}{n}(\tilde{X}+V)'(\tilde{X}+V)\\ &= \frac{1}{n}\tilde{X}'\tilde{X}+\frac{1}{n}\tilde{X}'V+\frac{1}{n}V'\tilde{X}+\frac{1}{n}V'V.\end{aligned}$$

بنابراین

$$\frac{1}{n}X'X \xrightarrow{p} \Sigma_{xx}+0+0+\Sigma_{VV}=\Sigma_{xx}+\Sigma_{VV}. \tag{۸.۴}$$

برای محاسبه قسمت دوم سمت راست رابطه (۷.۴) به صورت زیر عمل می‌کنیم:

$$\begin{aligned}\frac{1}{n}X'\omega &= \frac{1}{n}\tilde{X}'\omega+\frac{1}{n}V'\omega\\ &= \frac{1}{n}\tilde{X}'(u-V\beta)+\frac{1}{n}V'(u-V\beta)\\ &= \frac{1}{n}\tilde{X}'u-\frac{1}{n}\tilde{X}'V\beta+\frac{1}{n}V'u-\frac{1}{n}V'V.\end{aligned} \tag{۹.۴}$$

در نتیجه

$$\frac{1}{n}X'\omega \xrightarrow{p} 0-0+0-\Sigma_{VV}\beta=-\Sigma_{VV}\beta. \tag{۱۰.۴}$$

سرانجام با جای‌گذاری (۸.۴) و (۱۰.۴) در (۷.۴) خواهیم داشت

$$b-\beta \xrightarrow{p} -(\Sigma_{xx}+\Sigma_{VV})^{-1}\Sigma_{VV}\beta\neq 0. \tag{۱۱.۴}$$

بنابراین ملاحظه می‌شود که b یک برآوردگر ناسازگار برای پارامتر β خواهد بود، به‌طوری‌که این ناسازگاری اساسا ناشی از ارتباط بین X و ω است.

نتیجه ۱.۳.۴. نتایج نباید این‌طور تعبیر شوند که برآوردگر کم‌ترین توان‌های دوم معمولی $b=(X'X)^{-1}X'y$ با می‌نیمم کردن $S=\omega'\omega=(y-X\beta)'(y-X\beta)$ در مدل $y=X\beta+\omega$ به‌دست می‌آید. در حقیقت $\omega'\omega$ را نمی‌توان در حالت رگرسیون خطی معمولی می‌نیمم کرد زیرا برخلاف خطای تصادفی در مدل رگرسیون خطی معمولی، $\omega=u-V\beta$ خودش تابعی از β است.

حال به جهت دیدن ماهیت سازگاری، مدل رگرسیون خطی ساده با خطای اندازه‌گیری را به صورت زیر در نظر بگیرید:

$$\begin{aligned} \tilde{y}_i &= \beta_۰ + \beta_۱ \tilde{x}_i \quad ; \quad i = ۱, ۲, \ldots, n \\ y_i &= \tilde{y}_i + u_i \\ x_i &= \tilde{x}_i + V_i \end{aligned} \tag{۱۲.۴}$$

که در آن ماتریس‌های X، $\tilde{X}$ و V به صورت زیر تعریف می‌شوند:

$$X = \begin{pmatrix} ۱ & x_۱ \\ ۱ & x_۲ \\ \vdots & \vdots \\ ۱ & x_n \end{pmatrix} \qquad \tilde{X} = \begin{pmatrix} ۱ & \tilde{x}_۱ \\ ۱ & \tilde{x}_۲ \\ \vdots & \vdots \\ ۱ & \tilde{x}_n \end{pmatrix} \qquad V = \begin{pmatrix} ۰ & V_۱ \\ ۰ & V_۲ \\ \vdots & \vdots \\ ۰ & V_n \end{pmatrix}.$$

همچنین فرض کنید

$$\frac{۱}{n} \sum_{i=۱}^{n} \tilde{x}_i \xrightarrow{p} \mu \tag{۱۳.۴}$$

و

$$\frac{۱}{n} \sum_{i=۱}^{n} (\tilde{x}_i - \mu)^۲ \xrightarrow{p} \sigma_x^۲. \tag{۱۴.۴}$$

در این صورت با توجه به (۱۰.۴) و (۱۳.۴)، داریم

$$\frac{۱}{n} \tilde{X}' \tilde{X} = \begin{pmatrix} ۱ & \frac{۱}{n} \sum_{i=۱}^{n} \tilde{x}_i \\ \frac{۱}{n} \sum_{i=۱}^{n} \tilde{x}_i & \frac{۱}{n} \sum_{i=۱}^{n} \tilde{x}_i^۲ \end{pmatrix} \xrightarrow{p} \begin{pmatrix} ۱ & \mu \\ \mu & \mu^۲ + \sigma^۲ \end{pmatrix} = \Sigma_{xx}. \tag{۱۵.۴}$$

همچنین

$$\frac{۱}{n} V' V \xrightarrow{p} \begin{pmatrix} ۰ & ۰ \\ ۰ & \sigma^۲ \end{pmatrix} = \Sigma_{VV}. \tag{۱۶.۴}$$

با جای‌گذاری روابط (۱۵.۴) و (۱۶.۴)، در معادله (۱۱.۴)، خواهیم داشت

$$\begin{aligned} \begin{pmatrix} b_۰ - \beta_۰ \\ b_۱ - \beta_۱ \end{pmatrix} &\xrightarrow{p} -\begin{pmatrix} ۱ & \mu \\ \mu & \mu^۲ + \sigma_x^۲ + \sigma_v^۲ \end{pmatrix} \begin{pmatrix} ۰ & ۰ \\ ۰ & \sigma_V^۲ \end{pmatrix} \begin{pmatrix} \beta_۰ \\ \beta_۱ \end{pmatrix} \\ &= -\frac{۱}{\sigma_V^۲ + \sigma_x^۲} \begin{pmatrix} \sigma_x^۲ + \mu^۲ + \sigma_V^۲ & -\mu \\ -\mu & ۱ \end{pmatrix} \begin{pmatrix} ۰ \\ \beta \sigma_V^۲ \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\sigma_V^۲}{\sigma_V^۲ + \sigma_x^۲} \mu \beta \\ -\dfrac{\sigma_V^۲}{\sigma_V^۲ + \sigma_x^۲} \beta \end{pmatrix}. \end{aligned} \tag{۱۷.۴}$$

ملاحظه می‌شود که برآوردگرهای کم‌ترین توان‌های دوم معمولی $\beta_۰$ و $\beta_۱$ اریب و ناسازگارند. بنابراین اگر یک متغیر در معرض خطای اندازه‌گیری قرار داشته باشد، نه‌تنها برآورد پارامتر روی خودش اثر می‌گذارد، بلکه بر برآوردگرهای سایر پارامترها که با این متغیر مرتبط هستند و بدون هر خطایی اندازه‌گیری شده‌اند نیز اثر می‌گذارد که این اتفاق در مدل رگرسیونی ساده نیز می‌تواند روی دهد.

فصل ۵

سازگاری کامل برآوردگرهای کمترین توان‌های دوم در مدل رگرسیونی آلوده به خطا

۱.۵ مقدمه

در فصل ۳، همگرایی کامل آرایه‌ای از متغیرهای تصادفی NSD مورد بررسی قرار دادیم. در این فصل سازگاری برآوردگرهای کم‌ترین توان‌های دوم[۱] (LS) در مدل رگرسیونی تبیینی آلوده به خطا که به اختصار آن را EV می‌نامیم، با خطای NSD مورد مطالعه قرار می‌دهیم. مطالب نظری این فصل از میائو و همکاران (۲۰۱۳) گرفته شده‌اند.

مدل رگرسیونی EV، در ابتدا توسط دیتون (۱۹۸۵) بیان شد. او نشان داد که خطای نمونه‌گیری می‌تواند تأثیر بسیار زیادی در نتایج پیش‌گویی در رگرسیون داشته باشد و مدل رگرسیونی EV می‌تواند عملی‌تر از مدل رگرسیونی LS باشد؛ زیرا در این مدل خطای نمونه‌گیری لحاظ می‌شود و نتایج دقیق‌تری را ارایه می‌کند. برای اطلاعات بیشتر می‌توانید به فولر و همکاران (۱۹۸۷)، فوزک و فوزوکاوا (۱۹۸۹)، میتگ (۱۹۸۹)، کارول و همکاران (۱۹۹۵) و هسالو و همکاران (۱۹۹۷) مراجعه کنید.

تعریف ۱.۱.۵. مدل رگرسیونی EV با پارامترهای θ و β به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} \eta_i &= \theta + \beta x_i + \varepsilon_i \\ \xi &= x_i + \delta_i \qquad ۱ \leq i \leq n \end{aligned} \tag{۱.۵}$$

[۱] Least square

که در آن $\theta, \beta, x_1, x_2, \ldots$ پارامترهای نامعلوم، $(\varepsilon_1, \delta_1), (\varepsilon_2, \delta_2), \ldots$ بردارهای تصادفی و برای هر $i = 1, 2, \ldots, n$، ξ_i و η_i مشاهدات به ترتیب متغیرهای تبیینی آلوده و پاسخ هستند.

شکل دیگری از مدل (۱.۵) به صورت زیر قابل ارایه است:

$$\begin{aligned} \eta_i &= \theta + \beta\xi_i + \upsilon_i \\ \upsilon_i &= \varepsilon_i - \beta\delta_i \qquad 1 \leq i \leq n. \end{aligned} \tag{۲.۵}$$

حال با در نظر گرفتن مدل (۲.۵) به‌عنوان مدل رگرسیونی معمولی η_i روی ξ_i، برآوردگرهای LS پارامترهای θ و β عبارتند از:

$$\begin{aligned} \hat{\beta}_n &= \frac{\sum\limits_{i=1}^{n}(\xi_i - \bar{\xi}_n)(\eta_i - \bar{\eta}_n)}{\sum\limits_{i=1}^{n}(\xi_i - \bar{\xi}_n)^2} \\ \hat{\theta}_n &= \bar{\eta}_n - \hat{\beta}_n\bar{\xi}_n \end{aligned} \tag{۳.۵}$$

که در آن $\bar{\xi}_n = \sum\limits_{i=1}^{n}\xi_i$، $\bar{\eta}_n = \sum\limits_{i=1}^{n}\eta_i$ و به‌طور مشابه $\bar{\delta}_n$ و $\bar{x}_n$ تعریف می‌شوند.

دیتون (۱۹۸۵) رفتار حدی برآوردگرهای LS برای پارامترهای θ و β را در مدل رگرسیونی EV مورد بررسی قرار داد. در همین راستا محققان دیگری نیز مطالعاتی را انجام داده‌اند. برای مثال، می‌توان به سوی (۱۹۹۷) اشاره کرد که نرمال مجانبی بودن M-برآوردگر[2] پارامترها را در مدل رگرسیونی EV بیان نمود. لیو و چن (۲۰۰۵) نیز سازگاری قوی و ضعیف برآوردگرهای LS را برای مدل رگرسیون EV خطی گزارش کرد و نشان داد شرط لازم و کافی برای این‌که برآوردگر $\hat{\beta}_n$ سازگار قوی و ضعیف برای β باشد عبارت است

$$\lim_{n\to\infty}\frac{S_n}{n} = \infty$$

که در آن $S_n = \sum\limits_{i=1}^{n}(x_i - \bar{x}_n)^2$ است.

میائو و همکاران (۲۰۰۷) قضیه حد مرکزی را در مدل رگرسیونی EV تحت مدل (۱.۵) با شرطهای $E\delta_1 = E\varepsilon_1 = 0$، $E\delta_1^2 = \sigma_1^2$، $E\varepsilon_1^2 = \sigma_2^2$، $\sigma_1^2 > 0$، $\sigma_2^2 < \infty$ و $\delta_i \overset{iid}{\sim} \varepsilon_i$ برای پارامترهای θ و β به صورت زیر نشان دادند.

قضیه ۲.۱.۵. تحت مدل (۱.۵) برای پارامتر β، فرض کنید $S_n = \sum\limits_{i=1}^{n}(x_i - \bar{x}_n)^2$ و همچنین

$$\lim_{n\to\infty}\frac{n}{S_n} = 0$$

باشد. در این صورت ثابتی مانند $\alpha > 0$ وجود دارد به‌طوری‌که $E|\varepsilon_i|^{2+\alpha} < \infty$، $E|\delta_i|^{2+\alpha} < \infty$ و $\lim\limits_{n\to\infty}\max\limits_{1\leq i\leq n}\frac{|x_i - \bar{x}_n|}{S_n^{\frac{1}{2}}} = 0$ است. آن‌گاه

$$\frac{\sqrt{S_n}(\hat{\beta}_n - \beta)}{\sqrt{Var(\varepsilon_1 - \beta\delta_1)}} \xrightarrow{d} N(0, 1)$$

که در آن $N(0, 1)$، توزیع نرمال استاندارد است.

[2] M-estimator

قضیه ۳.۱.۵. تحت مدل (۱.۵) برای پارامتر θ، فرض کنید که شرط‌های قضیه ۲.۱.۵ برقرار باشند. افزون بر این، فرض کنید

$$\frac{S_n}{n\bar{x}_n^2} \longrightarrow \infty.$$

آن‌گاه $\hat{\theta}_n - \theta$ به‌طور مجانبی به توزیع نرمال استاندارد میل می‌کند. به عبارتی

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sqrt{Var(\varepsilon_1 - \beta\delta_1)}} \xrightarrow{d} N(0, 1).$$

میائو و همکاران (۲۰۱۱) نیز برخی رفتارهای حدی برآوردگرهای LS را برای مدل رگرسیونی EV خطی ساده مورد مطالعه قرار دادند که در آن خطاها متغیرهای هم‌توزیع و مستقل هستند. فرض کنید تحت مدل (۱.۵)، به ازای $p \geq 2$، $E|\varepsilon_1|^p < \infty$، $E|\delta_1|^p < \infty$ و $\lim_{n\longrightarrow\infty} \frac{S_n}{n^{2-\frac{2}{p}}} = \infty$، آن‌گاه

$$\frac{\sqrt{S_n}}{n^{\frac{1}{p}}}(\hat{\beta}_n - \beta) \xrightarrow{a.s.} 0.$$

حال اگر شرط‌های فوق برقرار باشند و برای $\alpha \in (\frac{1}{2}, 1]$، $\frac{n^{1-\alpha+\frac{1}{p}}}{\sqrt{S_n}}|\bar{x}_n| = O(1)$، آن‌گاه

$$n^{1-\alpha}(\hat{\theta}_n - \theta) \xrightarrow{a.s.} 0.$$

همچنین، تحت شرط‌های گشتاوری ضعیف‌تر $E\delta_1^2 < \infty$، $E\delta_1^2 < \infty$ و همچنین

$$\lim_{n\longrightarrow\infty} \frac{n}{\sqrt{S_n}} = 0, \qquad \lim_{n\longrightarrow\infty} \max_{1\leq i\leq n} \frac{|x_i - \bar{x}_n|}{\sqrt{S_n}} = 0$$

$\frac{\sqrt{S_n}(\hat{\beta}_n-\beta)}{\sqrt{Var(\varepsilon_1-\beta\delta_1)}}$ و $\frac{\sqrt{S_n}(\hat{\theta}_n-\theta)}{\sqrt{Var(\varepsilon_1-\beta\delta_1)}}$ به توزیع نرمال استاندارد میل می‌کنند.

تا به این‌جا نتایج با فرض مستقل بودن دنباله‌ای از متغیرهای تصادفی برای جمله خطا مورد بررسی قرار گرفتند. در ادامه خطاهایی که دنباله‌ای از متغیرهای تصادفی وابسته هستند، را مورد مطالعه قرار می‌دهیم.

فازیکاس و کوکاش (۱۹۹۷) ویژگی‌های مجانبی برآوردگرها را در مدل رگرسیونی EV در توابع غیرخطی با شرط خطای α-آمیخته بررسی کردند. به همین دلیل لازم است در ابتدا α-آمیخته را از فان و همکاران (۲۰۱۰) بیان کنیم.

تعریف ۴.۱.۵. دنباله $\{\xi_i; i \geq 1\}$ را α-آمیخته[۳] گویند اگر

$$\alpha(m) := \sup \sup \left\{|P(AB) - P(A)P(B)| : A \in \mathcal{F}_{m+k}^{\infty}, B \in \mathcal{F}_1^k\right\}$$

وقتی که $m \longrightarrow \infty$، به صفر همگرا باشد. در این‌جا $\mathcal{F}_l^m = \sigma\left\{\xi_l, \xi_{l+1}, \ldots, \xi_{l+m}\right\}$ یک σ-جبر تولیدشده با $\xi_l, \xi_{l+1}, \ldots, \xi_{l+m}$ و $l \leq m$ است.

حال مدل زیر را در نظر بگیرید:

$$\begin{aligned} y_i &= g(\xi_i, \beta_0) + \delta_i \\ x_i &= \xi_i + \varepsilon_i \qquad i = 1, 2, \ldots \end{aligned} \tag{۴.۵}$$

[۳] α-mixing

که در آن $\xi_۱, \xi_۲, \ldots$ غیرتصادفی و نامعلوم، y_i و x_i مشاهدات و ε_i و δ_i خطاهای تصادفی نامعلومند. $\beta_۰$ نیز مقداری صحیح از پارامتر نامعلوم β است.

فرض کنید $\beta_۰$ عضو مجموعه‌ای p-بعدی و $\beta_۰ \in \Theta \subset \mathbb{R}^p$، x_i، ξ_i و ε_i بردارهای q-بعدی، δ_i و y_i یک‌بعدی و $g : \mathbb{R}^q \times \Theta \longrightarrow \mathbb{R}$ تابعی معلوم است. حال فرضیات زیر را نیز در نظر بگیرید:

۱) مجموعه $\{\delta_i; i = ۱, ۲, \ldots\}$ مستقل از $\{\varepsilon_i; i = ۱, ۲, \ldots\}$ هستند و برای $i = ۱, ۲, \ldots$ ، $E\varepsilon_i = ۰$، $E\delta_i = ۰$.

۲) $\varepsilon_۱, \varepsilon_۲, \ldots$ متغیرهای تصادفی مستقل و هم‌توزیع‌اند.

۳) مجموعه باز $U \supset \Theta$ و تابع $f \in C(\mathbb{R}^q \times U)$ موجود است به‌طوری‌که به ازای هر $\xi \in \mathbb{R}^q$ و $\beta \in \Theta$ داریم

$$Ef(\xi + \varepsilon_۱, \beta) = g(\xi, \beta).$$

۴) تابعی مانند $h \in C(\mathbb{R}^q \times U)$ وجود دارد به‌طوری‌که به ازای هر $\beta \in \Theta$ و $\xi \in \mathbb{R}^q$، داریم

$$Eh(\xi + \varepsilon_۱, \beta) = g^۲(\xi, \beta).$$

۵) برای هر $d > ۰$، $l > ۰$ موجود است به‌طوری‌که به ازای هر $s \in \mathbb{R}^q$ اگر $||\beta_۱ - \beta_۲|| < l$ آن‌گاه

$$||g(s, \beta_۱) - g(s, \beta_۲)|| < d$$

۶) g تابعی کراندار است.

۷) مجموعه Θ فشرده[۴] و محدب[۵] است و همچنین

$$\Psi_n(\beta_۱, \beta_۲) = \frac{۱}{n} \sum_{i=۱}^{n} \left(g(\xi_i, \beta_۱) - g(\xi_i, \beta_۲)\right)^۲$$

که در آن $\beta_۱, \beta_۲ \in \Theta$.

۸) $۰ < k_۱ \leq k_۲ < \infty$ وجود دارند به‌طوری‌که به ازای هر $\beta_۱, \beta_۲ \in \Theta$ و $n \in N$

$$k_۱||\beta_۱ - \beta_۲||^۲ \leq \Psi_n(\beta_۱, \beta_۲) \leq k_۲||\beta_۱ - \beta_۲||^۲.$$

۹) $E\sup_{\beta \in \Theta}||\frac{df(\xi + \varepsilon_۱, \beta)}{d\beta}||^۲$ تابعی کراندار از ξ است.

۱۰) برای هر $\beta \in \Theta$

$$\sup_n E|h(x_n, \beta)|^t < \infty.$$

۱۱) $\sup_n E|\delta_n|^t < \infty$.

---

[۴] Compact

[۵] Convex

۱۲) با تغییراتی در مدل رگرسیونی کم‌ترین توان‌های دوم و این‌که $\hat{\beta} = \hat{\beta}_n$ نقطه می‌نیمم تابع

$$Q_n(\beta) = \frac{۱}{n} \sum_{i=۱}^{n} \left\{ (y_i - f(x_i, \beta))^۲ + h(x_i, \beta) - f^۲(x_i, \beta) \right\}$$

باشد، باید از شرط‌های آمیختگی برای وابستگی $\{\delta_i\}$ استفاده گردد، که در آن به ازای $d > ۰$

$$a(\varphi, t) = \sum_{k=۱}^{\infty} \varphi^{\frac{۱}{j(t)}}(k)(k+۱)^{j(t)-۲} < \infty \tag{۵.۵}$$

و

$$b(\alpha, t, d) = \sum_{k=۱}^{\infty} \alpha^{\frac{d}{j(t)+d}}(k)(k+۱)^{j(t)-۲} < \infty. \tag{۶.۵}$$

قضیه ۵.۱.۵. (فازیکاس و کوکاش، ۱۹۹۷) فرض کنید تحت مدل (۴.۵) فرضیات ۱ تا ۱۲ برقرار باشند. علاوه بر آن فرض کنید $r \geq ۲$ و $p < r$، به‌طوری‌که (۵.۵) و موارد ۷، ۸ و ۹ با $t = r$ برقرار باشند. بنابراین $c > ۰$ موجود است به‌طوری که به ازای هر $\varrho > ۰$ و $n \in N$

$$P\left(\sqrt{n}||\hat{\beta}_n - \beta_۰|| > \varrho\right) \leq \frac{c}{\varrho^r}.$$

فان و همکاران (۲۰۱۰) نیز سازگاری قوی و نرمال مجانبی برآوردگرهای LS را در مدل رگرسیونی EV گزارش کردند. این در حالی بود که شکل خطاها، دنباله‌ای از متغیرهای α-آمیخته ایستاست. میائو و همکاران (۲۰۱۳) به‌طور مجانبی نرمال بودن و سازگاری قوی برآوردگرهای LS در مدل رگرسیونی EV را با خطای پیوندی منفی به‌دست آوردند. در ادامه نتایج را به‌طور مختصر و در قالب چند قضیه بیان می‌کنیم. در ابتدا به‌طور مجانبی نرمال بودن برآوردگرهای LS را برای برآوردگرهای پارامترهای نامعلوم β و θ نشان می‌دهیم.

قضیه ۶.۱.۵. (میائو و همکاران، ۲۰۱۳) برای برآورد کردن پارامتر β تحت مدل (۱.۵)، فرض کنید $S_n = \sum_{i=۱}^{n}(x_i - \bar{x}_n)^۲$ و شرط‌های زیر برقرار باشند:

$$\lim_{n \longrightarrow \infty} \frac{n}{\sqrt{S_n}} = ۰$$

$$\lim_{n \longrightarrow \infty} r_n = ۰, \qquad r_n = \max_{۱ \leq i \leq n} \frac{|x_i - \bar{x}_n|}{\sqrt{S_n}}$$

و برای هر ثابت $c > ۰$ و هر $۱ \leq i \leq n$ و $۱ \leq j \leq n$

$$|x_i - x_j| < c|i - j|.$$

به‌علاوه فرض کنید برای $i \geq ۱$، $X_i = \varepsilon_i - \beta\delta_i$ به‌طوری که $E\varepsilon_۱^۲ < \infty$، و $E\delta_۱^۲ < \infty$. همچنین

$$\sigma^۲ = EX_۱^۲ + ۲\sum_{i=۲}^{\infty} E(X_۱ X_i) > ۰$$

و

$$\sum_{i=2}^{n}(j-1)|E(X_1X_i)| = O\left(\frac{1}{l_n}\right)$$

که در آن $l_n = r_n max\{r_n, (\frac{n}{\sqrt{S_n}})\}$ است. در این صورت $\frac{\sqrt{S_n}}{\sigma}(\hat{\beta}_n - \beta)$ به توزیع نرمال استاندارد همگراست. یعنی

$$\frac{\sqrt{S_n}}{\sigma}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0,1).$$

قضیه ۷.۱.۵. (میائو و همکاران، ۲۰۱۳) با استفاده از شرط‌های قضیه ۶.۱.۵ و همچنین فرض‌های

$$\frac{S_n}{n\bar{x}_n^2} \longrightarrow \infty$$

و

$$\sum_{i=2}^{\infty}|Cov(X_1, X_j)| < \infty$$

برای برآوردگر پارامتر θ داریم

$$\frac{\sqrt{n}}{\sigma}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0,1).$$

میائو و همکاران (۲۰۱۳) برای به‌دست آوردن سازگاری قوی برآوردگرهای LS در مدل رگرسیونی EV با استفاده از خطاهای وابسته پیوندی منفی به نتایج زیر دست یافتند.

قضیه ۸.۱.۵. تحت مدل (۱.۵) فرض کنید $\{\varepsilon_i; i \geq 1\}$ و $\{\delta_i; i \geq 1\}$ دنباله‌ای از متغیرهای تصادفی NA باشند که بر یک‌دیگر غلبه تصادفی دارند، به‌طوری‌که $E\varepsilon_i = 0$، $E\delta_i = 0$ و به ازای هر $p \geq 2$، $E|\varepsilon_1|^p < \infty$ و $E|\delta_1|^p < \infty$. فرض کنید $\tau > 0$ و

$$\max_{1\leq i\leq n}\frac{|x_i - \bar{x}_n|}{\sqrt{S_n}n^{\frac{\tau-1}{p}}} = O(1), \qquad \frac{n^{\tau}}{\sqrt{S_n}} = O(1) \tag{۷.۵}$$

که در آن $S_n = \sum_{i=1}^{n}(x_i - \bar{x}_n)^2$. برای حالت $p = 2$، فرض کنید به ازای هر $\gamma > 0$

$$\frac{\sqrt{S_n}}{n^{1-\tau+\gamma}} \longrightarrow \infty. \tag{۸.۵}$$

در نتیجه $\frac{\sqrt{S_n}}{n^{\tau}}(\hat{\beta}_n - \beta)$ به طور تقریبا حتمی به صفر همگراست. به عبارتی

$$\frac{\sqrt{S_n}}{n^{\tau}}(\hat{\beta}_n - \beta) \xrightarrow{a.s.} 0. \tag{۹.۵}$$

حال اگر $p > 2$ و

$$\frac{n^{1-\tau}}{\sqrt{S_n}} \longrightarrow 0 \tag{۱۰.۵}$$

آن‌گاه (۹.۵) نتیجه می‌شود.

در ابتدا به لم زیر توجه کنید.

لم ۹.۱.۵. (جینگ و لیانگ، ۲۰۰۸) فرض کنید $\{X_n, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی وابسته پیوندی منفی و هم‌توزیع باشند. همچنین برای هر $p > 0$، $E|X_1|^p < \infty$ و به ازای هر $p > 1$، $EX_1 = 0$. آرایه $\{a_{k,n}, 1 \leq k \leq n\}$ که آرایه‌ای از اعداد حقیقی است را در نظر بگیرید به‌طوری‌که در معادله زیر صدق می‌کند:

$$\max_{1 \leq k \leq n} |a_{k,n}| = O\left(n^{-\frac{1}{p}}\right).$$

در نتیجه

۱. اگر $p > 2$ و $\sum_{k=1}^{n} a_{k,n}^2 = O\left((\log n)^{-1}\right)$ آن‌گاه به‌طور تقریبا حتمی

$$\sum_{k=1}^{n} a_{k,n} X_k = O(1).$$

۲. اگر $0 < p \leq 2$ و برای هر $\gamma > 0$، $\sum_{k=1}^{n} |a_{k,n}|^p = O\left(n^{-\gamma}\right)$ باشد، آن‌گاه به‌طور تقریبا حتمی

$$\sum_{k=1}^{n} a_{k,n} X_k = O(1).$$

میائو و همکاران (۲۰۱۳) برای اثبات قضیه، با استفاده از قسمت دوم لم ۹.۱.۵ و با قرار دادن $X_k = \delta_k^2 I(\delta_k \geq 0) - E\delta_k^2 I(\delta_k \geq 0)$ و $X_k = \delta_k^2 I(\delta_k < 0) - E\delta_k^2 I(\delta_k < 0)$ نشان دادند که اگر $p = 2$ باشد، آن‌گاه

$$I_n = \frac{1}{\sqrt{S_n} n^{\tau}} \sum_{k=1}^{n} \left(\delta_k^2 I(\delta_k \geq 0) - E\delta_k^2 I(\delta_k \geq 0)\right) \xrightarrow{a.s.} 0$$

و

$$II_n = \frac{1}{\sqrt{S_n} n^{\tau}} \sum_{k=1}^{n} \left(\delta_k^2 I(\delta_k < 0) - E\delta_k^2 I(\delta_k < 0)\right) \xrightarrow{a.s.} 0$$

برقرار هستند.

ملاحظه ۱۰.۱.۵. در قضیه ۸.۱.۵، شرط‌های متناهی بودن گشتاور مرتبه pام ε_1 و δ_1 ($E|\varepsilon_1|^p < \infty$ و $E|\delta_1|^p < \infty$) با شرط‌های $E|\delta_1|^{2p} < \infty$ و $E|\varepsilon_1|^{2p} < \infty$ به جهت استفاده از قضیه ۱.۳.۳ جایگزین می‌گردند.

۲.۵ کاربرد همگرایی کامل مجموع موزون متغیرهای تصادفی NSD در مدل رگرسیونی EV

در این بخش، سازگاری کامل برآوردگرهای کم‌ترین توان‌های دوم در مدل رگرسیونی تبیینی آلوده به خطا با خطاهای وابسته زبرجمعی منفی را در تناظر با قضیه ۸.۱.۵ قرار می‌دهیم که توسط میائو و همکاران (۲۰۱۳) بیان شده است. ابتدا مدل ارایه‌شده در (۱.۵) را به صورت دقیق‌تری در زیر بیان می‌کنیم:

$$\begin{aligned} \eta_i = \theta + \beta x_i + \varepsilon_i, \qquad & \xi_i = x_i + \delta_i \\ E\varepsilon_i = E\delta_i = ۰ \qquad & ۱ \le i \le n \end{aligned} \tag{۱۱.۵}$$

در این مدل (ξ_i, η_i) برای $۱ \le i \le n$ مشاهدات هستند و x_i، $۱ \le i \le n$، و β و θ پارامترهای نامعلومند. در ادامه فرض کنید که $\{\varepsilon_i; i \ge ۱\}$ و $\{\delta_i; i \ge ۱\}$ دنباله‌ای از متغیرهای تصادفی هستند که بر دنباله‌ای از متغیرهای تصادفی NSD غلبه تصادفی دارند، به‌طوری که از دیگر خطاها مستقل هستند. در ابتدا سازگاری کامل برآوردگرهای LS را برای پارامتر نامعلوم β در مدل رگرسیونی EV مورد بررسی قرار می‌دهیم.

قضیه ۱.۲.۵. (وانگ و همکاران، ۲۰۱۵) تحت مدل (۱۱.۵)، فرض کنید به ازای هر $p > ۱$، $E|\varepsilon_۱|^{۴p} < \infty$ و $E|\delta_۱|^{۴p} < \infty$. اگر $\tau > ۰$ موجود باشد به‌طوری‌که

$$\max_{۱\le i\le n} \frac{|x_i - \bar{x}_n|}{\sqrt{S_n} n^{\frac{\tau-۱}{p}}} = O(۱) \tag{۱۲.۵}$$

$$\frac{n^{\tau}}{\sqrt{S_n}} = O(۱) \tag{۱۳.۵}$$

و

$$\frac{n^{۱-\tau}}{\sqrt{S_n}} \longrightarrow ۰ \tag{۱۴.۵}$$

آن‌گاه

$$\frac{\sqrt{S_n}}{n^{\tau}}(\hat{\beta}_n - \beta) \xrightarrow{c} ۰ \tag{۱۵.۵}$$

که در آن $S_n = \sum_{i=۱}^{n}(x_i - \bar{x}_n)^۲$ است.

برهان. با توجه به معادلات (۳.۵) در بخش قبل، می‌توان نوشت

$$\hat{\beta}_n - \beta = \frac{\sum_{i=۱}^{n}(\xi_i - \bar{\xi}_n)(\eta_i - \bar{\eta}_n)}{\sum_{i=۱}^{n}(\xi_i - \bar{\xi}_n)^۲} - \beta \tag{۱۶.۵}$$

و با توجه به این‌که

$$\eta_i - \bar{\eta}_n = \theta + \beta x_i + \varepsilon_i - \theta - \beta \bar{x}_n - \bar{\varepsilon}_n \tag{۱۷.۵}$$

و

$$\xi_i - \bar{\xi}_n = (x_i - \bar{x}_n) + (\delta_i - \bar{\delta}_n) \tag{۱۸.۵}$$

با جای‌گذاری (۱۷.۵) و (۱۸.۵) در معادله (۱۶.۵) می‌توان نوشت

$$\hat{\beta}_n - \beta = \frac{\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)\varepsilon_i - \sum_{i=1}^{n}(x_i - \bar{x}_n)(\varepsilon_i - \beta\delta_i) - \beta\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)^2}{\sum_{i=1}^{n}(\xi_i - \bar{\xi}_n)^2}.$$

همچنین می‌توان نشان داد که

$$\hat{\theta}_n - \theta = (\beta - \hat{\beta}_n)\bar{x}_n + (\beta - \hat{\beta}_n)\bar{\delta}_n + \bar{\varepsilon}_n - \beta\bar{\delta}_n.$$

بنابراین کافی است نشان دهیم که به ازای $\varepsilon > 0$

$$P\left(\frac{\sqrt{S_n}}{n^{\tau}}\left|\hat{\beta}_n - \beta\right| > \varepsilon\right) =$$
$$P\left(\frac{\frac{1}{\sqrt{S_n}n^{\tau}}\left|\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)\varepsilon_i - \sum_{i=1}^{n}(x_i - \bar{x}_n)(\varepsilon_i - \beta\delta_i) - \beta\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)^2\right|}{\frac{1}{S_n}\sum_{i=1}^{n}(\xi_i - \bar{\xi}_n)^2} > \varepsilon\right) < \infty$$

با توجه به تساوی فوق، ملاحظه می‌شود که $\hat{\beta}_n - \beta$ از چهار قسمت تشکیل شده است که این چهار قسمت عبارتند از

- $\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)\varepsilon_i$
- $\sum_{i=1}^{n}(x_i - \bar{x}_n)(\varepsilon_i - \beta\delta_i)$
- $\beta\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)^2$
- $\sum_{i=1}^{n}(\xi_i - \bar{\xi}_n)^2$

برای ثابت نمودن (۱۵.۵) لازم است نشان دهیم که سه عبارت اول فوق به‌طور کامل و عبارت چهارم به‌طور تقریبا حتمی همگرا هستند. به عبارتی

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)^2 \xrightarrow{c} 0 \tag{۱۹.۵}$$

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i - \bar{\delta}_n)\varepsilon_i \xrightarrow{c} 0 \tag{۲۰.۵}$$

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(x_i - \bar{x}_n)(\varepsilon_i - \beta\delta_i) \xrightarrow{c} 0 \tag{۲۱.۵}$$

و

$$\frac{1}{S_n}\sum_{i=1}^{n}(\xi_i - \bar{\xi}_n)^2 \xrightarrow{a.s.} 1 \tag{۲۲.۵}$$

در ابتدا نشان می‌دهیم که (۱۹.۵) برقرار است. لذا داریم

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \quad\leq\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\delta_i^2.$$

با اضافه و کم کردن $\frac{1}{\sqrt{S_n}n^{\tau}}\sum\limits_{i=1}^{n}E\delta_i^2$ به عبارت سمت راست فوق، می‌توان نوشت

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \quad\leq\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i^2-E\delta_i^2)+\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}E\delta_i^2. \tag{۲۳.۵}$$

حال عبارت داخل پرانتز جمله سمت راست در نابرابری (۲۳.۵) را به صورت زیر افراز می‌کنیم:

$$\begin{aligned}
\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \quad&\leq\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\big(\delta_i^2I(\delta_i\geq 0)-E\delta_i^2I(\delta_i\geq 0)\\
&\quad +\delta_i^2I(\delta_i<0)-E\delta_i^2I(\delta_i<0)\big)\\
&\quad +\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}E\delta_i^2\\
&=\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\big(\delta_i^2I(\delta_i\geq 0)-E\delta_i^2I(\delta_i\geq 0)\big)\\
&\quad +\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\big(\delta_i^2I(\delta_i<0)-E\delta_i^2I(\delta_i<0)\big)\\
&\quad +\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}E\delta_i^2\\
&=\quad J_{1n}+J_{2n}+\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}E\delta_i^2
\end{aligned} \tag{۲۴.۵}$$

که در آن

$$J_{1n} \quad=\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\big(\delta_i^2I(\delta_i\geq 0)-E\delta_i^2I(\delta_i\geq 0)\big) \tag{۲۵.۵}$$

و

$$J_{2n} \quad=\quad \frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}\big(\delta_i^2I(\delta_i<0)-E\delta_i^2I(\delta_i<0)\big)\,. \tag{۲۶.۵}$$

از آن‌جا که دنباله $\{\delta_i; i\geq 1\}$، NSD است، طبق لم ۶.۳.۱، $\delta_i^2I(\delta_i\geq 0)$، $-E\delta_i^2I(\delta_i\geq 0)$، $\delta_i^2I(\delta_i<0)$ و $-E\delta_i^2I(\delta_i<0)$ نیز NSD هستند. بنابراین با توجه به لم ۹.۳.۱ دو دنباله $\{\delta_i^2I(\delta_i\geq 0)-E\delta_i^2I(\delta_i\geq 0); i\geq 1\}$ و $\{\delta_i^2I(\delta_i<0)-E\delta_i^2I(\delta_i<0); i\geq 1\}$، NSD خواهند بود.

حال با در نظر گرفتن $X_i=\delta_i^2I(\delta_i\geq 0)-E\delta_i^2I(\delta_i\geq 0)$ و $a_{ni}=\frac{1}{\sqrt{S_n}n^{\tau}}$، همچنین (۱۴.۵)، ملاحظه می‌شود

.۱

$$\begin{aligned}
E|X_1|^{2p} &= E|\delta_i^2 I(\delta_1 \geq 0) - E\delta_i^2 I(\delta_1 \geq 0)|^{2p} \\
&\leq E|\delta_1^2 I(\delta_1 \geq 0)|^{2p} + E|\delta_1^2 I(\delta_1 \geq 0)|^{2p} \\
&= 2E|\delta_1^2 I(\delta_1 \geq 0)|^{2p} \\
&\leq CE|\delta_1^2|^{2p} \\
&= CE|\delta_1|^{4p} < \infty.
\end{aligned}$$

.۲

$$\begin{aligned}
\underset{1\leq i\leq n}{max}\, a_{ni} &= \frac{1}{\sqrt{S_n} n^{\tau}} \\
&\leq \frac{1}{\sqrt{S_n} n^{\tau-p}} . \frac{n}{n} \\
&= \frac{n^{1-\tau}}{\sqrt{S_n}} . \frac{1}{\frac{n}{n^p}} \\
&= \frac{n^{1-\tau}}{\sqrt{S_n}} . \frac{1}{n^{\frac{1}{p}}} \\
&\leq Cn^{-\frac{1}{p}} \\
&= O\left(n^{-\frac{1}{p}}\right).
\end{aligned}$$

.۳

$$\sum_{i=1}^{n} a_{ni}^2 = \frac{n^{2-2\tau}}{\sqrt{S_n}} . \frac{1}{n}. \tag{۲۷.۵}$$

از آن‌جا که $1 + n \leq e^n$ است، بنابراین داریم

$$\log n \leq \log(1+n) \leq n \leq n^{2\tau}.$$

با توجه به نابرابری فوق

$$\frac{\frac{1}{n^{2\tau}}}{\frac{1}{\log n}} \leq C.$$

بنابراین طبق تعریف ترتیب‌های تصادفی

$$\frac{1}{n^{2\tau}} = O\left((\log n)^{-1}\right). \tag{۲۸.۵}$$

در نتیجه با جای‌گذاری (۲۸.۵) در (۲۷.۵) داریم

$$\sum_{i=1}^{n} a_{ni}^2 = O\left((\log n)^{-1}\right).$$

لذا با توجه به موارد فوق و بنابر قضیه ۱.۳.۳ در فصل ۳، نتیجه می‌شود که J_{1n} به‌طور کامل به صفر همگراست. به عبارتی

$$J_{1n} = \frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}\left(\delta_i^2 I(\delta_i \geq 0) - E\delta_i^2 I(\delta_i \geq 0)\right) \xrightarrow{c} 0. \tag{۲۹.۵}$$

مشابه (۲۹.۵)، معادله (۲۶.۵) نیز از قضیه ۱.۳.۳ ثابت می‌شود. لذا داریم

$$J_{2n} = \frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}\left(\delta_i^2 I(\delta_i < 0) - E\delta_i^2 I(\delta_i < 0)\right) \xrightarrow{c} 0. \tag{۳۰.۵}$$

بنابراین با جای‌گذاری (۲۹.۵) و (۳۰.۵) در (۲۵.۵)، می‌توان نوشت

$$\begin{aligned}
\frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 &\leq \frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}E\delta_i^2 \\
&= \frac{1}{\sqrt{S_n}n^\tau}\left(E\delta_1^2 + E\delta_2^2 + \ldots + E\delta_n^2\right) \\
&\leq \frac{1}{\sqrt{S_n}n^\tau}\left(E\delta_1^2 + E\delta_1^2 + \ldots + E\delta_1^2\right) \\
&= \frac{n^{1-\tau}}{\sqrt{S_n}}\left(E\delta_1^2\right).
\end{aligned} \tag{۳۱.۵}$$

با توجه به رابطه (۱۴.۵) و $E|\delta_1|^{4p} < \infty$، نابرابری فوق به صفر همگراست. در نتیجه برای nهای به اندازه کافی بزرگ، به ازای هر $\varepsilon > 0$

$$\frac{n^{1-\tau}}{\sqrt{S_n}}\left(E\delta_1^2\right) < \frac{\varepsilon}{2}.$$

بنابراین با استفاده از (۳۱.۵)

$$\sum_{n=1}^{\infty}P\left(\frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 > \varepsilon\right) < \infty$$

و به تبع آن (۱۹.۵) نتیجه می‌شود.

مشابه اثبات (۱۹.۵)، می‌توان نشان داد که

$$\frac{1}{\sqrt{S_n}n^\tau}\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon}_n)^2 \xrightarrow{c} 0. \tag{۳۲.۵}$$

است. حال به دو یادآوری زیر توجه کنید.

یادآوری ۲.۲.۵. با توجه به اتحاد مربع و $(a-b)^2 \geq 0$، داریم

$$\frac{1}{2}(a^2+b^2) \geq ab.$$

برای اثبات (۲۰.۵)، بنابر آن‌چه در یادآوری فوق بیان شد، داریم

$$\begin{aligned}
\frac{1}{\sqrt{S_n n^{\tau}}}\left|\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)\varepsilon_i\right| &= \frac{1}{\sqrt{S_n n^{\tau}}}\left|\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)(\varepsilon_i-\bar{\varepsilon}_n)\right| \\
&\leq \frac{1}{\sqrt{S_n n^{\tau}}}.\frac{1}{2}\left|\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)+\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon}_n)\right| \\
&\leq \frac{1}{2}\left|\frac{1}{\sqrt{S_n n^{\tau}}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)\right|+\frac{1}{2}\left|\frac{1}{\sqrt{S_n n^{\tau}}}\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon}_n)\right|.
\end{aligned}$$

حال با جای‌گذاری (۳۲.۵) و (۱۹.۵) در جمله آخر عبارت فوق، می‌توان نتیجه گرفت عبارت

$$\frac{1}{\sqrt{S_n n^{\tau}}}\left|\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)\varepsilon_i\right|$$

به‌طور کامل به صفر میل می‌کند. لذا (۲۰.۵) به‌دست می‌آید.

یادآوری ۳.۲.۵. به ازای هر $r>0$، چون $(\frac{a}{\sqrt{r}}-b\sqrt{r})^2\geq 0$. بنابراین

$$\frac{a^2}{2r}+\frac{r}{2}b^2\geq ab$$

برای اثبات (۲۲.۵) با توجه به رابطه ۳.۲.۵، داریم

$$\begin{aligned}
\sum_{i=1}^{n}|x_i-\bar{x}_n|.\left|\delta_i-\bar{\delta}_n\right| &\leq \sqrt{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2}.\sqrt{\sum_{i=1}^{n}\left(\delta_i-\bar{\delta}_n\right)^2} \\
&\leq \frac{r}{2}\sum_{i=1}^{n}(x_i-\bar{x}_n)^2+\frac{1}{2r}\sum_{i=1}^{n}\left(\delta_i-\bar{\delta}_n\right)^2 \\
&= \frac{r}{2}S_n+\frac{1}{2r}\sum_{i=1}^{n}\left(\delta_i-\bar{\delta}_n\right)^2. \qquad (۳۳.۵)
\end{aligned}$$

از طرفی می‌توان ثابت نمود:

$$\begin{aligned}\left|\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-S_n\right| &= \left|\sum_{i=1}^{n}\left[(x_i-\bar{x}_n)+(\delta_i-\bar{\delta}_n)\right]^2-S_n\right| \\ &= \left|\sum_{i=1}^{n}(x_i-\bar{x}_n)^2+\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2\right. \\ &\quad \left.+2\sum_{i=1}^{n}(x_i-\bar{x}_n)(\delta_i-\bar{\delta}_n)-S_n\right| \\ &= \left|2\sum_{i=1}^{n}(x_i-\bar{x}_n)(\delta_i-\bar{\delta}_n)+\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2\right| \\ &\leq 2\left|\sum_{i=1}^{n}(x_i-\bar{x}_n)(\delta_i-\bar{\delta}_n)\right|+\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \\ &\leq 2\left(\frac{r}{2}S_n+\frac{1}{2r}\sum_{i=1}^{n}\left(\delta_i-\bar{\delta}_n\right)^2\right)+\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \qquad (۳۴.۵)\end{aligned}$$

نابرابری آخر در (۳۴.۵) از رابطه ۳.۲.۵ و (۳۳.۵) به‌دست آمده است.

$$\left|\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-S_n\right| = rS_n+\frac{1+r}{r}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2.$$

حال با استفاده از (۱۹.۵) و طبق لم‌های برل-کانتلی در فصل اول داریم

$$\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2 \xrightarrow{a.s.} 0. \qquad (۳۵.۵)$$

لذا در نابرابری

$$\left|\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-S_n\right| \leq rS_n+\frac{1+r}{r}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2$$

با ضرب نمودن دوطرف عبارت در $\frac{1}{S_n}$ و همچنین ضرب و تقسیم کردن n^{τ} در جمله دوم عبارت سمت راست، می‌توان نوشت

$$\begin{aligned}\left|\frac{1}{S_n}\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-1\right| &\leq r+\frac{1+r}{rS_n}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2.\frac{n^{\tau}}{n^{\tau}} \\ &= r+\frac{1+r}{r}.\frac{n^{\tau}}{\sqrt{S_n}}.\frac{1}{\sqrt{S_n}n^{\tau}}\sum_{i=1}^{n}(\delta_i-\bar{\delta}_n)^2.\end{aligned}$$

بنابراین با توجه به (۱۳.۵) و (۳۵.۵) عبارت $\left|\frac{1}{S_n}\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-1\right|$ به‌طور تقریبا حتمی به $r>0$ همگراست، یعنی

$$\left|\frac{1}{S_n}\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2-1\right| \xrightarrow{a.s.} r.$$

لذا چون $r > ۰$ است، آن‌گاه (۲۲.۵) نتیجه می‌شود.

در پایان می‌خواهیم (۲۱.۵) را ثابت کنیم. برای اثبات فرض کنید که

$$a_{ni} = \frac{x_i - \bar{x}_n}{\sqrt{S_n} n^{\tau}}, \quad X_i = \varepsilon_i - \beta\delta_i.$$

بنابراین با استفاده از لم ۹.۳.۱و با توجه به این‌که دو دنباله $\{\delta_i; i \geq ۱\}$ و $\{\varepsilon_i; i \geq ۱\}$ متغیرهای تصادفی NSD هستند، ملاحظه می‌شود که دنباله‌ی $\{X_i, i \geq ۱\}$ نیز متغیر تصادفی NSD است. به‌علاوه، به ازای $p > ۰$ و با توجه به $E|\delta_۱|^{۴p} < \infty$ ،$E|\varepsilon_۱|^{۴p} < \infty$

.۱

$$\begin{aligned} E|X_۱|^{۲p} &= E|\varepsilon_۱ - \beta\delta_۱|^{۲p} \\ &\leq CE|\varepsilon_۱|^{۲p} + CE|\delta_۱|^{۲p} \\ &< \infty. \end{aligned}$$

۲. با توجه به (۱۲.۵)

$$\begin{aligned} \underset{۱\leq i\leq n}{max}|a_{ni}| &= \underset{۱\leq i\leq n}{max}\left(\frac{|x_i - \bar{x}_n|}{\sqrt{S_n} n^{\tau}} . \frac{n^{-\frac{۱}{p}}}{n^{-\frac{۱}{p}}}\right) \\ &= \underset{۱\leq i\leq n}{max}\left(\frac{|x_i - \bar{x}_n|}{\sqrt{S_n} n^{\tau - \frac{۱}{p}}} . n^{-\frac{۱}{p}}\right) \\ &= O(۱).n^{-\frac{۱}{p}} \\ &= O\left(n^{-\frac{۱}{p}}\right). \end{aligned}$$

۳. بر اساس آن‌چه در (۲۸.۵) بیان شد، داریم

$$\begin{aligned} \sum_{i=۱}^{n} a_{ni}^{۲} &= \sum_{i=۱}^{n}\left(\frac{x_i - \bar{x}_n}{\sqrt{S_n} n^{\tau}}\right)^{۲} \\ &= \sum_{i=۱}^{n}\frac{(x_i - \bar{x}_n)^{۲}}{S_n n^{۲\tau}} \\ &= \frac{۱}{n^{۲\tau}} \\ &= O\left((\log n)^{-۱}\right). \end{aligned}$$

موارد فوق نشان می‌دهند که شرط‌های قضیه ۱.۳.۳ برقرار هستند. بنابراین همگرایی کامل (۲۱.۵) با استفاده از قضیه ۱.۳.۳ ثابت می‌شود و حاکی از آن است که (۱۵.۵) به‌دست آمده است. لذا برهان قضیه را کامل می‌کند. □

ملاحظه ۴.۲.۵. اگر $۰ < \tau \leq \frac{۱}{۲}$ باشد، آن‌گاه (۱۴.۵) ، معادله (۱۳.۵) و همچنین اگر $\tau > \frac{۱}{۲}$ باشد سپس (۱۳.۵)، (۱۴.۵) را نتیجه می‌دهد. در این صورت می‌توان نتایج زیر را با استفاده از قضیه ۱.۲.۵ به‌دست آورد.

نتیجه ۵.۲.۵. تحت مدل (۴۴.۵)، فرض کنید برای هر $p > ۱$، $E|\delta_۱|^{۴p} < \infty$ و $E|\varepsilon_۱|^{۴p} < \infty$. اگر $۰ < \tau \leq \frac{۱}{۲}$ موجود باشد به‌طوری که (۱۲.۵) و (۱۴.۵) برقرار باشند، آن‌گاه (۱۵.۵) نتیجه می‌شود.

نتیجه ۶.۲.۵. تحت مدل (۴۴.۵)، فرض کنید برای هر $p > ۱$، $E|\delta_۱|^{۴p} < \infty$ و $E|\varepsilon_۱|^{۴p} < \infty$. اگر $\tau > \frac{۱}{۲}$ موجود باشد به‌طوری که (۱۲.۵) و (۱۳.۵) برقرار باشند. آن‌گاه (۱۵.۵) نتیجه می‌شود.

در انتها سازگاری کامل برآوردگرهای LS را برای پارامتر نامعلوم θ در مدل رگرسیونی EV مورد بررسی قرار می‌دهیم.

قضیه ۷.۲.۵. (وانگ و همکاران، ۲۰۱۵) تحت فرضیات قضیه ۱.۲.۵، اگر $۰ < \nu < min(۱-\frac{۱}{p}, \frac{۱}{۲})$ موجود باشد به‌طوری‌که

$$\frac{n^{\tau+\nu}}{\sqrt{S_n}}|\bar{x}_n| = O(۱) \tag{۳۶.۵}$$

آن‌گاه

$$n^{\nu}\left(\hat{\theta}_n - \theta\right) \xrightarrow{c} ۰. \tag{۳۷.۵}$$

برهان. بر طبق رابطه (۱۷.۵)، مشاهده می‌شود که رابطه (۳۷.۵) از سه قسمت به صورت زیر تشکیل شده است و کافی است نشان دهیم

$$n^{\nu}\left(\bar{\varepsilon}_n - \beta\bar{\delta}_n\right) \xrightarrow{c} ۰ \tag{۳۸.۵}$$

$$n^{\nu}\left(\beta - \hat{\beta}_n\right)\bar{x}_n \xrightarrow{c} ۰ \tag{۳۹.۵}$$

$$n^{\nu}\left(\beta - \hat{\beta}_n\right)\bar{\delta}_n \xrightarrow{c} ۰. \tag{۴۰.۵}$$

در ابتدا برای اثبات (۳۸.۵) مشاهده می‌شود که

$$n^{\nu}\left(\bar{\varepsilon}_n - \beta\bar{\delta}_n\right) = \frac{۱}{n^{۱-\nu}}\sum_{i=۱}^{n}(\varepsilon_i - \beta\delta_i). \tag{۴۱.۵}$$

فرض کنید $a_{ni} = \frac{۱}{n^{۱-\nu}}$ و $X_i = \varepsilon_i - \beta\delta_i$ باشد که در آن $۰ < \nu < min(۱-\frac{۱}{p}, \frac{۱}{۲})$ و $p > ۱$ است. ملاحظه می‌شود که شرط‌های قضیه ۱.۳.۳ برقرار هستند، زیرا:

۱.

$$\begin{aligned} E|X_۱|^{۲p} &= E|\varepsilon_۱ - \beta\delta_۱|^{۲p} \\ &\leq CE|\varepsilon_۱|^{۲p} + CE|\delta_۱|^{۲p} \\ &< \infty. \end{aligned}$$

.۲

$$\begin{aligned}\max_{1\leq i\leq n}|a_{ni}| &= \max_{1\leq i\leq n}\left(\frac{1}{n^{1-\nu}}\right)\\ &= \frac{1}{n^{1-\nu}}\\ &= O\left(n^{-\frac{1}{p}}\right).\end{aligned}$$

.۳ با توجه به آن‌چه در (۲۸.۵) بررسی شد، خواهیم داشت

$$\begin{aligned}\sum_{i=1}^{n}a_{ni}^{2} &= \sum_{i=1}^{n}\left(\frac{1}{n^{1-\nu}}\right)^{2}\\ &= \frac{1}{n^{2(1-\nu)}}\\ &= O\left((\log n)^{-1}\right).\end{aligned}$$

بنابراین رابطه (۳۸.۵) از قضیه ۱.۳.۳ نتیجه می‌شود.

حال همگرایی کامل رابطه (۳۹.۵) را به ازای هر $\varepsilon > 0$ به صورت زیر نشان می‌دهیم:

$$\begin{aligned}\sum_{n=1}^{\infty}P\left(\left|n^{\nu}\left(\beta-\hat{\beta}_n\right)\bar{x}_n\right|>\varepsilon\right) &= \sum_{n=1}^{\infty}P\left(\left|\frac{\sqrt{S_n}}{n^{\tau}}\left(\beta-\hat{\beta}_n\right).\frac{n^{\nu+\tau}}{\sqrt{S_n}}\bar{x}_n\right|>\varepsilon\right)\\ &\leq \sum_{n=1}^{\infty}P\left(\frac{\sqrt{S_n}}{n^{\tau}}\left|\beta-\hat{\beta}_n\right|>C\right)\\ &< \infty. \end{aligned} \qquad (۴۲.۵)$$

عبارت آخر در رابطه فوق از قضیه ۱.۲.۵ نتیجه شده است و این اثبات (۳۹.۵) را کامل می‌کند. سرانجام، برای اثبات رابطه (۴۰.۵) می‌توان نوشت

$$\begin{aligned}\sum_{n=1}^{\infty}P\left(\left|n^{\nu}\left(\beta-\hat{\beta}_n\right)\bar{\delta}_n\right|>\varepsilon\right) &= \sum_{n=1}^{\infty}P\left(\left|\frac{\sqrt{S_n}}{n^{\tau}}\left(\beta-\hat{\beta}_n\right).\frac{n^{\nu}}{\sqrt{S_n}}.\frac{1}{n^{1-\nu}}\sum_{i=1}^{n}\delta_i\right|>\varepsilon\right)\\ &\leq \sum_{n=1}^{\infty}P\left(\frac{\sqrt{S_n}}{n^{\tau}}\left|\beta-\hat{\beta}_n\right|>C\right)\\ &\quad+\sum_{n=1}^{\infty}P\left(\left|\frac{1}{n^{1-\nu}}\sum_{i=1}^{n}\delta_i\right|>C\right). \end{aligned} \qquad (۴۳.۵)$$

همگرایی کامل جمله اول در عبارت آخر نابرابری فوق، از رابطه (۴۲.۵) به‌دست می‌آید. برای جمله دوم عبارت مذکور، مشابه روند اثبات (۳۸.۵) عمل می‌کنیم. در این‌جا با در نظر گرفتن $a_{ni}=\frac{1}{n^{1-\nu}}$ و $X_i=\sum_{i=1}^{n}\delta_i$، شرایط قضیه ۱.۳.۳ را روی آن اعمال می‌کنیم که مطابق نتیجه قضیه همگرایی کامل $\frac{1}{n^{1-\nu}}\sum_{i=1}^{n}\delta_i$ به صفر به‌دست می‌آید.

بنابراین با توجه به مطالب فوق همگرایی کامل (۴۳.۵) نتیجه می‌شود و

$$\sum_{n=1}^{\infty} P\left(\left|n^{\nu}\left(\beta-\hat{\beta}_n\right)\bar{\delta}_n\right|>\varepsilon\right)<\infty.$$

لذا برهان کامل می‌شود. □

ملاحظه ۸.۲.۵. ترکیب دو قضیه ۸.۱.۵ و ۱.۲.۵ نتایجی را در پی خواهد داشت که در زیر به بیان آن می‌پردازیم:

۱. $p \geq 2$ در قضیه ۸.۱.۵ را می‌توان به $p > 1$ تغییر داد.

۲. در قضیه ۸.۱.۵ خطای NA را می‌توان به خطاهای NSD تبدیل کرد.

۳. همگرایی تقریبا حتمی در قضیه ۸.۱.۵ به همگرایی کامل تغییر می‌یابد.

۴. برای $p = 2$ شرط (۸.۵) در قضیه ۸.۱.۵ با شرط (۱۴.۵) ضعیف‌تر خواهد شد.

۳.۵ مطالعه شبیه‌سازی

شبیه‌سازی[۶] یک ابزار قدرتمند برای طراحی و تحلیل سیستم‌ها یا فرآیندهای پیچیده مهندسی است که به استفاده‌کنندگان آن، امکان انجام آزمایش با سیستم‌هایی را می‌دهد که در عمل غیرممکن یا پرهزینه هستند. به‌عنوان مثال، قبل از پرتاب یک ماهواره عملیات پرتاب در شرایط مختلف بارها شبیه‌سازی می‌شود تا هزینه‌ها به حداقل برسد.

شبیه‌سازی در آمار از قرن بیستم مطابق با شروع انتشار رادیو و تلویزیون آغاز شد. همان‌طور که استفاده از این وسایل در زندگی هر کسی معمول و متداول شده، استفاده از شبیه‌سازی نیز در بسیاری از شاخه‌های علم آمار گسترش یافته است.

از کاربردهای شبیه‌سازی در آمار می‌توان به تولید داده‌های تصادفی از توزیع‌های گسسته یا پیوسته، تولید اعداد شبه‌تصادفی (تصادفی)، محاسبه (تقریب) انتگرال‌ها، محاسبه (تقریب) احتمالات پیچیده و محاسبه (تقریب) امید ریاضی توابعی از متغیرهای تصادفی اشاره کرد. برای پاسخ به سوالاتی که نظریه آمار برای آن جوابی ندارد نیز می‌توان از شبیه‌سازی استفاده کرد و جوابی تجربی به‌دست آورد. در این پایان‌نامه برای ارزیابی سازگاری برآوردگرهای معرفی شده با استفاده از یک مطالعه شبیه‌سازی، مدل زیر را در نظر گرفتیم:

$$\boldsymbol{\eta} = -1 + 2\mathrm{x} + \varepsilon \qquad \boldsymbol{\xi} = \mathrm{x} + \delta$$

[۶] Simulation

که در آن $E\delta = ۰$ و $E\varepsilon = ۰$ است. به‌طور دقیق‌تر قرار می‌دهیم

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_۱ \\ \eta_۲ \\ \vdots \\ \eta_n \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} ۱ & x_۱ \\ ۱ & x_۲ \\ \vdots & \vdots \\ ۱ & x_n \end{pmatrix} \qquad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_۱ \\ \varepsilon_۲ \\ \vdots \\ \varepsilon_n \end{pmatrix} \qquad \boldsymbol{\xi} = \begin{pmatrix} \xi_۱ \\ \xi_۲ \\ \vdots \\ \xi_n \end{pmatrix} \qquad \boldsymbol{\delta} = \begin{pmatrix} \delta_۱ \\ \delta_۲ \\ \vdots \\ \delta_n \end{pmatrix}.$$

مطالعه شبیه‌سازی در محیط نرم‌افزار R اجرا شده است و کدهای لازم برای بازتولید نتایج شبیه‌سازی در پیوست ب آمده‌اند. داده‌ها را از یک مدل خطی ساده آلوده به خطا که در آن خطای اندازه‌گیری از نوع کلاسیک است، تولید کردیم. با توجه به آن‌چه هو (۲۰۰۰) گزارش کرده است، توزیع نرمال چندمتغیره در صورتی NSD خواهد بود که عناصر غیر قطر اصلی ماتریس کوواریانس آن منفی باشند و از آن‌جا که ε و δ باید NSD باشند، از نتیجه هو (۲۰۰۰) استفاده کرده و ε و δ را از توزیع نرمال چندمتغیره با میانگین صفر و ماتریس کوواریانس Σ تولید کردیم که در آن $\Sigma = (\sigma_{ij})$، ماتریسی $n * n$ با درایه‌های زیر است:

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

برای ارزیابی برآوردگرهای کمترین توان‌های دوم در حضور خطای اندازه‌گیری،سه مدل مختلف را در شبیه‌سازی در نظر گرفته و بر روی داده‌های شبیه‌سازی شده برازش دادیم:

الف) **مدل** $g_۱$: رگرسیون کم‌ترین توان‌های دوم معمولی خطی ساده $\boldsymbol{\eta}$ روی متغیر تبیینی $\boldsymbol{X}$.

ب) **مدل** $g_۲$: رگرسیون کم‌ترین توان‌های دوم معمولی $\boldsymbol{\eta}$ روی $\boldsymbol{\xi}$.

ج) **مدل** $g_۳$: رگرسیون آلوده به خطای خطی با خطای NSD.

متغیر تبیینی $\boldsymbol{X}$ را نیز از یک توزیع نرمال استاندارد تولید کردیم.

تذکر: از آن‌جا که $\boldsymbol{X}$ دارای توزیع نرمال است، لذا شرایط دیگر قضیه برای آن صدق می‌کند. اما اگر توزیع $\boldsymbol{X}$ را به توزیعی مانند کوشی تغییر دهیم، آن‌گاه شرط‌های (۱۲.۵)، (۱۳.۵) و (۱۴.۵) برقرار نخواهند بود.

برای بررسی سازگاری برآوردگرها، در شبیه‌سازی حجم‌های نمونه $n = ۱۰۰, ۲۰۰, ۳۰۰$ را در نظر گرفتیم. برای ارزیابی عملکرد، معیارهای اریبی و میانگین توان‌های دوم خطای[۷] (MSE) برآوردگرها را بر اساس ۲۰۰ مجموعه داده شبیه‌سازی‌شده در هر سه مدل محاسبه کردیم.

نتایج در جدول‌های ۱.۵ و ۲.۵ گزارش شده‌اند.

[۷] Mean squared errors

جدول ۱.۵: معیار MSE برآوردگرهای پارامترها در سه مدل مختلف $g_۱$، $g_۲$ و $g_۳$

حجم نمونه	پارامتر	MSE		
		$g_۱$	$g_۲$	$g_۳$
۱۰۰	θ	۰/۰۰۰۹	۰/۰۲۱۱	۰/۰۲۱۱
	β	۰/۰۰۹۹	۱/۰۳۷۸	۰/۱۹۳۹
۲۰۰	θ	۰/۰۰۰۳	۰/۰۱۱۱	۰/۰۱۱۱
	β	۰/۰۰۵۵	۱/۰۱۲۴	۰/۱۱۵۸
۳۰۰	θ	۰/۰۰۰۱	۰/۰۰۶۷	۰/۰۰۶۷
	β	۰/۰۰۳۶	۱/۰۰۷۴	۰/۰۵۱۵

جدول ۲.۵: معیار اریبی برآوردگرهای پارامترها در سه مدل مختلف $g_۱$، $g_۲$ و $g_۳$

حجم نمونه	پارامتر	اریبی		
		$g_۱$	$g_۲$	$g_۳$
۱۰۰	θ	۰/۰۰۰۵	۰/۰۰۶۲	۰/۰۰۶۲
	β	−۰/۰۰۰۶	−۱/۰۱۱۲	۰/۰۸۱۱
۲۰۰	θ	۰/۰۰۰۶	−۰/۰۰۸۱	−۰/۰۰۸۱
	β	−۰/۰۰۱۱	−۱/۰۰۱۹	۰/۰۶۵۹
۳۰۰	θ	۰/۰۰۰۴	−۰/۰۰۸۳	−۰/۰۰۸۳
	β	−۰/۰۰۴۷	−۱/۰۰۱۵	۰/۰۲۴۳

۱.۳.۵ نتایج شبیه‌سازی

همان‌طور که در جدول‌های ۱.۵ و ۲.۵ ملاحظه می‌کنید، دقت برآورد دو پارامتر θ و β در مدل رگرسیونی تبیینی آلوده به خطا با خطای NSD همانند دقت برآورد در مدل رگرسیونی کمترین توان‌های دوم است. از این رو در برخی از کاربردها می‌توان از این رهیافت مهم استفاده نمود.

لذا نتایج به‌دست آمده بیان‌گر آن است که اگر خطاها دارای نوعی وابستگی (وابستگی زبرجمعی منفی) باشند، با استفاده از رگرسیون آلوده به خطا، می‌توان پارامترها را به مانند روش رگرسیون کم‌ترین توان‌های دوم برآورد کرد. این دقت برآورد را می‌توان در جدول ۱.۵ و در ستون‌های $g_۱$ و $g_۳$ نشان داد که این نزدیک بودن دقت برای هر دو پارامتر θ در حجم نمونه ۳۰۰ بسیار چشم‌گیر است. اما دقت برآورد برای پارامتر β در مدل رگرسیونی آلوده به خطا بیشتر است. در نتیجه می‌توان این موضوع را متذکر شد که با افزایش حجم نمونه دقت برآورد نیز افزایش می‌یابد.

در همین راستا، با توجه به جدول ۲.۵ میزان اریبی نیز در این دو نوع مدل رگرسیونی را می‌توان مشاهده کرد. میزان اریبی در مدل رگرسیونی تبیینی آلوده به خطا با داده‌های x و خطای NSD همانند میزان اریبی در مدل رگرسیونی کم‌ترین توان‌های دوم معمولی با داده‌های ξ است. ملاحظه می‌شود که این اریبی و همچنین

دقت برآورد در دو مدل و در ستون‌های $g_۱$ و $g_۳$ اختلاف فاحشی وجود ندارد. این نتیجه، اهمیت موضوع وابستگی زبرجمعی منفی در خطاها به‌خصوص در رگرسیون را نشان می‌دهد که نقشی اساسی در سازگاری دو مدل رگرسیونی ایفا می‌کند.

۴.۵ نتیجه‌گیری و آینده تحقیق

NSD یک رده بسیار گسترده از دنباله‌های وابسته است که در حالت خاص دنباله‌های مستقل و همچنین دنباله‌های NA را شامل می‌شود. در این پایان‌نامه، برخی ویژگی‌های اساسی برای متغیرهای تصادفی NSD را ارایه کردیم که برای ساختن همگرایی کامل مجموع موزون متغیرهای تصادفی NSD مورد استفاده قرار گرفته اند. به‌عنوان یک کاربرد مهم می‌توان به سازگاری کامل برآوردگرهای LS در مدل رگرسیونی EV، با خطاهای NSD و تحت شرط‌های ضعیفی اشاره کرد.

با توجه به مطالب نظری و شبیه‌سازی می‌توان موارد زیر را به‌عنوان آینده تحقیق خروجی این پایان‌نامه به‌صورت زیر بیان نمود:

۱) می‌توان با تغییر نوع همگرایی، نتیجه را برای وابسته زبرجمعی منفی مورد بررسی قرار داد.

۲) نتایج این پایان‌نامه بر اساس نابرابری نمایی کلموگروف شکل گرفته است که می‌توان با تغییر نابرابری، نتایج را بررسی نمود.

۳) در این تحقیق، همگرایی کامل در رده خاصی از وابستگی نشان داده شد که می‌توان همگرایی کامل را با استفاده از رده دیگری از وابستگی‌ها نشان داد.

مراجع

[۱] اسحاقی ا، (۱۳۹۲)، پایان‌نامه ارشد: ”مدل های نیمه پارامتری تحلیل بقا برای داده های بازگشتی با روش هسته“، دانشکده ریاضی، دانشگاه صنعتی شاهرود

[2] Asadian, N., Amini, V., and bozorgnia, A. (2006). Rosenthal's type inequalities for negatively orthant dependent random variabeles, *JIRSS*, **5**, 69-75.

[3] Benjevic, D. (1984). On a Kolmogorov inequality, *Theor Probab Appl*, **29**, 391-394.

[4] Baum, L. E., and Katz, M. (1963). Convergence rates in the law of large numbers, *Bullet Amer MATH Soc*, **69**(6), 771-772.

[5] Bolck, H. W., and Sampson, A. R. (1988). Conditionally orderd distributions, *J Multi Anal*, **27**, 91-104.

[6] Carroll, R. J., Ruppert D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*, CRC press. Newyork.

[7] Christofides, T. C., and Vaggelatou, E. (2004). A connection between supermodular ordering and positive/negative association, *J Multivar anal*, **88**(1), 138-151.

[8] Cui, H. (1997). Asymptotic normality of M-estimates in the EV model, *Sys Sci and Math Sci*, **10**(3), 225-236.

[9] Deaton, A. (1985). Panel data from a time series of cross-sections, *J Econ*, **30**, 109–126.

[10] Deuschel, J. D., and Strook, D. W. (1989). Large deviations, *Academic Press San Diego*.

[11] Eghbal, N., Amini, M., and Bozorgnia, A. (2010). Some maximal inequalities for quadratic forms of negative superadditive dependence random variables, *Stat Probab Lett*, **80**, 587–591.

[12] Eghbal, N., Amini M., and Bozorgnia, A. (2011). On the Kolmogorov inequalities for quadratic forms of dependent uniformly bounded random variables, *Stat Probab Lett* **81**, 1112–1120.

[13] Erdos, P. (1949). On a theorem of Hsu and Robbins, *Ann Stat*, **20**, 286–291.

[14] Efron, D. (1965). Increasing properteies of polya frequency function, *Ann Math Statist*, **36**, 272-279.

[15] Fan, G., Liang, HY., Wang, J., and Xu, H. (2010). Asymptotic properties for LS estimators in EV regression model with dependent errors, *AStA Adv Stat Anal*, **94**, 89–103.

[16] Fazekas, I., and Kukush, A. (1997). Asymptotic properties of an estimator in nonlinear functional errors-invariables models with dependent error terms, *Comput Math Appl*, **34**, 23–39.

[17] Fuller, W. (1987). *Measurement Error Models*, Wiley, New York.

[18] Fusek, I., and Fuskova, L. (1989). A combined estimator in the simple errors-invariables model, *Stat Pap*, **30**, 1–15.

[19] Fakoor, V., and Azarnoosh, H. A. (2005). Probability ineqalities for sums of negatively dependent random variables, *Pak J Stat*, **21**(3), 257-264.

[20] Georgiev, A. A. (1983). Local properties of function fitting estimates with applications to system identification. Mathematical Statistics and Applications. *Proc 4th Pan Symp Math Statist*, 4–10.

[21] Gut, A. (1994). *Probability: A Graduate Course*, Spring T Stat, Sweden.

[22] Hu, T. (2000). Negatively superadditive dependence of random variables with applications, *Chin J Appl Probab Stat*, **16**, 133–144.

[23] Hu, T., and Xie, C., Ruan, LY (2005). Dependence structures ofmultivariate Bernoulli random vectors, *J Multivar Anal*, **94**, 172–195.

[24] Hslao, C., Wang, L., and Wang, Q. (1997). Estimation of nonlinear errors-invariables models: an approximate solution, *Stat Pap*, **38**, 1–25.

[25] Hsu, P., and Robbins, H. (1947). Complete convergence and the law of large numbers, *Proc Natl Acad Sci USA*, **33**, 25–31.

[26] Joag-Dev, K., and Proschan, F. (1983). Negative association of random variables with applications, *Ann Stat*, **11**, 286–295.

[27] Johnson, W. B., Schechtman, G., and Zinn, J. (1985). Best Constants in Moment Inequalities for Linear Combinations of Independent and Exchangeable Random Variables, *Ann Probab*, **13**, 234-253.

[28] Jing, B. Y., and Liang, H. Y. (2008). Strong limit theorems for weighted sums of negatively associated random variables, *J Theor Probab*, **21**(4), 890–909.

[29] Kemperman, J. H. B., (1977). On the FKG-inequalities for measures on a partially ordered space, *Nederl Akad Wetensch Proc Ser A*, **80**, 313–331.

[30] Kim, H. C. (2011). Exponential inequalities for ELNQD random variables with applications, *J Chung Math Soc*, **24**(4), 783-793.

[31] Kolmogorov, A. N. (1963). On tables of random numbers, *Sankh: Indian J Stat Ser A*, 369-376.

[32] Liu, J. X., and Chen, X. R. (2005). Consistency of LS estimator in simple linear EV regression models, *Acta Math Sci*, **25**, 50–58.

[33] Levy, H. (2006). *Stochastic Dominance: Investment Decision Making Under Uncertainty*, 2nd ed, Spring Sci, New York.

[34] Miao, Y., Yang, G. Y., and Shen, L. M. (2007). The central limit theorem for LS estimator in simple linear EV regression models, *Com Stat Theor Meth*, **36**, 2263–2272.

[35] Miao, Y., Wang, K., and Zhao, F. F. (2011). Some limit behaviors for the LS estimator in simple linear EV regression models, *Stat Probab Lett*, **81**, 92–102.

[36] Miao, Y., Zhao, F. F., Wang, K., and Chen, Y.P. (2013). Asymptotic normality and strong consistency of LS estimators in the EV regression model with NA errors, *Stat Pap*, **54**, 193–206.

[37] Mittag, H. J. (1989). Estimating parameters in a simple errors-in-variables model: a new approach based on finite sample distribution theory, *Stat Pap*, **30**, 133–140.

[38] Mavrikiou, P. M. (2007). Kolmogorov inequalities for the partial sum of independent Bernoulli random variables, *Stat prob lett*, **77**(11), 1117-1122.

[39] Mari, D. D., and Kotz, S. (2001). *Correlation And Dependence*, Imp Coll Press, London.

[40] Petrov, V. V. (1995). Limit Theorems of Probability Theory: Sequences of Independent Random Variables, *Oxf: Claren Press*.

[41] Peligrad, M., Utev, S., and Wu, W. B. (2007). A Maximal Lp- Inequality for Stationary Sequences and its Applications, *Proc Amer Math Soc*, **135**, 541-550.

[42] Pinelis, I. F., and Utev, S. A. (1984). Estimates of Moments of Sums of Independent Random Variables, *Theor Veroya I Prim*, **29**, 554-557.

[43] Rosenthal, H. P. (1970). On the Subspaces of $L_p(p > ۲)$ Spanned by Sequences of Independent Random Variables, Israel J Math, **8**, 273-303.

[44] Shao, Q. M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables, *J Theor Probab*, **13**, 343–355.

[45] Shen, A. T. (2013). On the strong convergence rate for weighted sums of arrays of rowwise negatively orthant dependent random variables, *RACSAM*, **107**, 257–271.

[46] Shen, A.T., and Wu, R. C. (2013). Strong and weak convergence for asymptotically almost negatively associated random variables, *Discret Dyn Nat Soc*, Article ID 235012, doi:10.1155/2013/235012.

[47] Shen, A. T., Wu R. C., Wang, X. H., and Shen, Y. (2013a). Complete convergence for weighted sums of arrays of rowwise $\tilde{\rho}-$ mixing random variables, *J Inequal Appl*, Article ID 356, doi:10.1186/1029-242X-2013-356.

[48] Shen, A. T., Wu, R. C., Chen, Y., and Zhou, Y. (2013b). Complete convergence of the maximum partial sums for arrays of rowwise of AANA random variables, *Discret Dyn Nat Soc*, Article ID 741901, doi:10.1155/ 2013/741901.

[49] Spitzer, F. L. (1956). A combinatorial lemma and its application to probability theory, *T Am Math Soc*, **82**, 323–339.

[50] Sung, S. H. (2010). Complete convergence for weighted sums of ρ^*-mixing random variables, *Discret Dyn Nat Soc*, Article ID 630608, doi:10.1155/2010/630608.

[51] Sung, S. H. (2012). A note on the complete convergence for weighted sums of negatively dependent random variables, *J Inequal Appl*, Article ID 158, doi:10.1186/1029-242X-2012-158.

[52] Shao, Q. M. (1995). Maximal Inequalities for Partial Sums of -mixing Sequences, *Ann Probab*, **23**, 948-965.

[53] Shen, A., Huayan, Z., and Ying, Z. (2014). Exponential Inequality for $\tilde{\rho}$-mixing Sequences and its Applications, *Fac Sci Math* **28**(4), 859-870.

[54] Shen, A., and Wu, R. (2013). Strong and weak convergence for asymptotically almost negatively associated random variables, *Discret Dyna Nat Soc*, Vol 2013, Article ID 235012, p 7.

[55] Thrum, R. (1987). A remark on almost sure convergence of weighted sums, *Probab Theory Rel*, **75**, 425–430.

[56] Turner, D. W., Young, D. M., and Seaman, J. W. (1992). Improved Kolmogorov inequalities for the binomial distribution, *Stat Prob lett*, **13**(3), 223-227.

[57] Volodin, A. (2002). On the Kolmogorov exponential inequality for negatively dependent random variables, *pak J Stat Ser*, **18**(2), 249-254.

[58] Van der Vaart, A. W. (2000). *Asymptotic Statistics*, Cambridge university press, Newyork.

[59] Wang, X. J., Li, X. Q., Yang, W. Z., and Hu, S. H. (2012a). On complete convergence for arrays of rowwise weakly dependent random variables, *Appl Math Lett*, **25**, 1916–1920.

[60] Wang, X. J., Hu, S. H., and Yang, W. Z. (2012b). Complete convergence for arrays of rowwise negatively orthant dependent random variables, *RACSAM*, **106**, 235–245.

[61] Wu, Q. Y. (2010). Complete convergence for negatively dependent sequences of random variables, *J Inequal Appl*, **2010**(1), 1-10.

[62] Wu, Q. J. (2012a). A complete convergence theorem for weighted sums of arrays of rowwise negatively dependent random variables, *Inequal Appl*, doi:10.1186/1029-242X-2012-50.

[63] Wu, Q. J. (2012). Sufficient and necessary conditions of complete convergence for weighted sums of PNQD random variables, *J Appl Math*, doi:10.1155/2012/104390.

[64] Wang, X., Deng, X., Zheng, L., and Hu, S. (2014). Complete convergence for arrays of rowwise negatively superadditive-dependent random variables and its applications, *J Theor Appl Stat*, **48**(4), 834-850.

[65] Wang, X. J., Hu, S. H., Li, X. Q., and Yang, W.Z. (2011). Maximal inequalities and strong law of large numbers for AANA sequences, *Com Kore Math Soc*, **26**(1), 151-161.

[66] Wu, Q. Y. (2006). Probability limit throey for mixing sequences, *Sci Press China*, Beijing.

[67] Wang, X., Shen, A., Chen, Z., and Hu, S. (2015). Complete convergence for weighted sums of NSD random variables and its application in the EV regression model, *TEST*, **24**, 166–184.

[68] Wolfstetter, E. (1993). Stochastic dominance: Theory and applications, *Humbold Univ*, Wirtschafswiss Fak.

[69] Young, D.M., Seaman, W.J., and Marco, R.V., (1987). A note on a Kolmogorov inequality, *Stat Probab Lett*, **5**, 217–218.

[70] Young, D. M., Turner, D. W., and Seaman, J. W. (1988). A note on a Kolmogorov inequality for the binomial distribution, *Theor Prob Appl*, **33**(4), 747-749.

پیوست آ

قضایا و لم‌های مورد نیاز

آ.۱ قضایایی در اثبات نامساوی ماکسیمال روزنتال

لم آ.۱.۱. (فکور و آذرنوش، ۲۰۰۵) فرض کنید $0 < t \leq 1$ باشد. به ازای هر $h > 0$، $x > 0$ و $y > 0$

$$P(|S_n| \geq x) \leq \sum_{k=1}^{n} P(|X_k| \geq y) + 2\exp\left\{\frac{x}{y} - \frac{x}{y}\log(1 + \frac{xy^{t-1}}{M_{t,n}})\right\} \tag{آ.۱}$$

که در آن $M_{t,n} = \sum_{k=1}^{n} EX_k^t$ و $S_n = \sum_{k=1}^{n} EX_k$.

حال با استفاده از تغییراتی در لم آ.۱.۱، لم زیر به دست می‌آید.

لم آ.۱.۲. فرض کنید $1 \leq t \leq 2$ و برای $EX_k = 0, k = 1, 2, \ldots, n$، باشد. به ازای هر $h > 0$، $x > 0$ و $y > 0$

$$P(|S_n| \geq x) \leq \sum_{k=1}^{n} P(|X_k| \geq y) + 2\exp\left\{\frac{e^{hy} - 1 - hy}{y^t} M_{t,n} - hx\right\}. \tag{آ.۲}$$

قضیه آ.۱.۳. (اسدیان و همکاران، ۲۰۰۶) فرض کنید $1 \leq t \leq 2$ و برای $k = 1, 2, \ldots, n$، $EX_k = 0$ باشد. همچنین $g(x)$ تابعی نامنفی و صعودی در $[0, \infty)$ است که شرط‌های $g(0) = 0$ و $Eg(X_k) < \infty$ در آن صدق می‌کنند. در نتیجه برای هر $r > 0$

$$Eg(S_n) \leq \sum_{k=1}^{n} Eg(rX_k) + 2e^r \int_0^\infty (1 + \frac{x^t}{r^{t-1} M_{t,n}})^{-r} dg(x).$$

برهان. در ابتدا $h = \frac{۱}{y}\log(۱+\frac{xy^{t-۱}}{M_{t,n}})$ را در نظر بگیرید. حال با قرار دادن h در سمت راست (آ.۲) و با توجه به $\frac{M_{t,n}}{y^t}\log(۱+\frac{xy^{t-۱}}{M_{t,n}}) \geq ۰$، داریم

$$P(|S_n| \geq x) \leq \sum_{k=۱}^{n} P(|X_k| \geq y) + ۲\exp\left\{\frac{x}{y} - \frac{x}{y}\log(۱+\frac{xy^{t-۱}}{M_{t,n}})\right\}. \qquad (\text{آ.۳})$$

با قرار دادن $\frac{x}{y} = r$ در (آ.۳) می‌توان نوشت

$$P(|S_n| \geq x) \leq \sum_{k=۱}^{n} P(|X_k| \geq \frac{x}{r}) + ۲e^r(۱+\frac{x^t}{r^{t-۱}M_{t,n}})^{-r} \qquad (\text{آ.۴})$$

سپس با انتگرال‌گیری عبارت (آ.۴)، داریم:

$$\int_۰^\infty P(|S_n| \geq x)dg(x) \leq \sum_{k=۱}^{n}\int_۰^\infty P(|X_k| \geq \frac{x}{r})dg(x) + ۲e^r\int_۰^\infty (۱+\frac{x^t}{r^{t-۱}M_{t,n}})^{-r}dg(x).$$

سرانجام با استفاده از لم ۲.۴ در پتروف (۱۹۹۵)، برهان کامل می‌شود. □

نتیجه آ.۱.۴. فرض کنید $۱ \leq t \leq ۲$، $p \geq t$ و به ازای $k = ۱, ۲, \ldots, n$، $EX_k = ۰$ باشد. آن‌گاه

$$E|S_n|^p \leq C(p,t)\left(M_{p,n} + M_{t,n}^{\frac{p}{t}}\right).$$

برهان. در قضیه آ.۱.۳ فرض کنید $p \geq t$ و همچنین $g(x) = |x|^p$ باشد. بنابراین

$$|S_n|^p \leq \sum_{k=۱}^{n} E|X_k|^p + ۲pe^r\int_۰^\infty x^{p-۱}(۱+\frac{x^t}{r^{t-۱}M_{t,n}})^{-r}dx.$$

حال فرض کنید $I = \int_۰^\infty x^{p-۱}(۱+\frac{x^t}{r^{t-۱}M_{t,n}})^{-r}dx$. با توجه به این‌که توزیع بتا با پارامترهای $\alpha > ۰, \beta > ۰$ به شکل زیر است:

$$B(\alpha,\beta) = \int_۰^۱ x^{\alpha-۱}(۱-x)^{\beta-۱}dx$$

در نتیجه I به صورت زیر تغییر می‌یابد:

$$I = \frac{B(\frac{p}{p}, r-\frac{p}{t})}{t} r^{\frac{t-۱}{t}p} M_{t,p}^{\frac{p}{t}}.$$

بنابراین برای $r > \frac{p}{t}$ و همچنین با انتخاب $C(p,t) = max\left\{r^p, ۲pe^r\frac{B(\frac{p}{p}, r-\frac{p}{t})}{t}r^{\frac{t-۱}{t}p}M_{t,p}^{\frac{p}{t}}\right\}$ در نامساوی زیر داریم:

$$E|S_n|^p \leq r^p\sum_{k=۱}^{n}E|X_k|^p + ۲pe^r\frac{B(\frac{p}{p}, r-\frac{p}{t})}{t}r^{\frac{t-۱}{t}p}\left(\sum_{k=۱}^{n}E|X_k|^t\right)^{\frac{p}{t}}.$$

و برهان کامل می‌شود.

□

۲.آ تعاریف و لم‌هایی در اثبات نامساوی نمایی کلموگروف

در این بخش برخی تعاریف و لم‌هایی را که برای اثبات نابرابری کلموگروف لازم است، به‌طور مختصر، از گات (۱۹۹۴) بیان می‌کنیم.

لم ۱.۲.آ. برای هر $x \geq 0$

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

۳.آ مارتینگل

تعریف ۱.۳.آ. اگر $\{X_n; n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی در فضای احتمال $(\Omega, \mathcal{F}, P)$ و $\mathcal{F}_1$، $\mathcal{F}_2$، ... 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) $\mathbb{E}(|X_n|) < \infty$

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

برقرار باشند. دنباله $\{(X_n, \mathcal{F}_n), n \geq 0\}$ یک مارتینگل[1] 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 این صورت

$$\int_A X_{n+1} dP = \int_A X_n dP. \tag{۵.آ}$$

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

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

با جایگذاری $A = \Omega$ در رابطه (۵.آ)، تساوی زیر برقرار می‌شود:

$$\mathbb{E}(X_1) = \mathbb{E}(X_2) = \dots$$

تعریف ۲.۳.آ. اگر متغیر تصادفی X_n در شرایط زیر صدق کند

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

[1] Martingale

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

۳) $\mathbb{E}(|X_n|) < \infty$

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

آن‌گاه X_n زیرمارتینگل[۲] 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 $h > 1$

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

و

$$\mathbb{E}(X_1) \leq \mathbb{E}(X_2) \leq ...$$

برقرار می‌باشند.

تعریف آ.۳.۳. اگر متغیر تصادفی X_n در شرایط

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

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

۳) $\mathbb{E}(|X_n|) < \infty$،

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

صدق کند. آن‌گاه X_n زبرمارتینگل[۳] متناظر با $\mathcal{F}_n$ است.

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

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

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

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

و

$$\mathbb{E}(X_1) \geq \mathbb{E}(X_2) \geq ...$$

برقرار می‌باشند.

برای مارتینگل، زیرمارتینگل و زبرمارتینگل، تعریف‌هایی معادل نیز به صورت زیر وجود دارند.

لم آ.۴.۳. اگر $\{T_i, 1 \leq i \leq n\}$ زبرمارتینگل نامنفی باشد، برای هر $0 \leq \alpha \leq 1$ داریم

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

---

[۲] Submartingale

[۳] Supermartingale

پیوست ب

کد مربوط به شبیه‌سازی

جهت انجام شبیه‌سازی، کد زیر را در نرم‌افزار R انجام دهید.

```
   rm(list=ls())
library(MASS)
library(mvtnorm)
library(sem)
library(semdiag)

   set.seed(741392)

   ### generating data from EV linear model and classical linear model

   gen.data <- function(theta, beta, sigma, n){|
epsilon <- mvrnorm(1, mu=rep(0,n), Sigma=sigma)
delta <- mvrnorm(1,mu=rep(0,n), Sigma=diag(n))
#epsilon <- as.vector(rmvt(1, sigma=sigma, df=3))
real.cov <- rnorm(n)
y <- theta + beta*real.cov + epsilon
```

```
error.cov <- real.cov + delta
data <- data.frame(cbind(real.cov, error.cov, y))
return(data)
}

   ######################
### To set real values for parameters

   theta <- -1
beta <- 2
real.coef <- c(theta, beta)
n <- 100 # 200, 300
a <- -1/(n-1)
sigma <- a*matrix(rep(1,n*n), ncol=n)
diag(sigma) <- rep(1,n)
# data <- gen.data(theta=theta, beta=beta, sigma=sigma, n=n)

  ######################################
#### The main function to compute mse and bias for the estimators
MSE <- function(theta, beta, sigma, n, B){|
est.ev <- est.ignor <- est.real <- matrix(NA, ncol=2, nrow=B)

   for (i in 1:B){
data <- gen.data(theta=theta, beta=beta, sigma=sigma, n=n)
x <- data$real.cov
xi <- data$error.cov
y <- data$y
###
fit.real.cov <- lm(y~x)
est.real[i,] <- fit.real.cov$coeff
###
fit.ignor.error <- lm(y~xi)
est.ignor[i,] <- fit.ignor.error$coeff
###
```

```
data.ev <- data.frame(xi,y)
eqns <- specifyEquations(covs="x", text="
y = theta*Intercept + beta*x
xi = 1*x
V(y) = sig
V(xi) = 1
V(x) = phi
")
#
model.EV <- sem(eqns, data=data.ev, raw=TRUE,
fixed.x="Intercept")
est.ev[i,] <- model.EV$coeff[1:2]
}

   ################################
### Extracting the results
B <- 200 ## number of iteration to generate datasets for computing MSE
res <- MSE(theta=theta, beta=beta, sigma=sigma, n=n, B=B)
## Computing the Bias of regression parameters
Bias.real.coef <- apply(res$est.real,2,mean)-real.coef
Bias.ignor.coef <- apply(res$est.ignor,2,mean)-real.coef
Bias.ev.coef <- apply(res$est.ev,2,mean)-real.coef

   Bias.real.coef
Bias.ignor.coef
Bias.ev.coef

   ## Computing the MSE of regression parameters

   MSE.real.coef <- apply((res$est.real-matrix(real.coef, nrow=B,
ncol=2, byrow=TRUE))^2,2,mean)
MSE.ignor.coef <- apply((res$est.ignor-matrix(real.coef, nrow=B,
ncol=2, byrow=TRUE))^2,2,mean)
MSE.ev.coef <- apply((res$est.ev-matrix(real.coef, nrow=B,
```

```
ncol=2, byrow=TRUE))^2,2,mean)

   MSE.real.coef
MSE.ignor.coef
MSE.ev.coef
```

Aabstract

Some basic properties for negatively superadditive dependent (NSD) random variables are presented, such as the Rosenthal-type inequality and the Kolmogorov-type exponential inequality. Using these properties, we further study the complete convergence for weighted sums of NSD random variables, which generalizes and improves some corresponding ones for independent random variables and negatively associated random variables. Some sufficient conditions to prove the complete convergence for weighted sums of NSD random variables are provided. As an application, the complete consistency of LS estimators in the EV regression model with NSD errors is investigated under mild conditions, which generalizes and improves the corresponding one for negatively associated random variables. Finally, we will evaluate theoretical results with the simulation study.

Keywords: Negatively superadditive dependent random variables, Complete convergence, Complete consistency, EV regression model, Rosenthal-type inequality, Kolmogorov-type exponential inequality

Shahrood University of Technology

Faculty of Mathematical Sciences

Complete convergence for weighted sums of NSD random variables and its application in the EV regression model

By: Samaneh Habibi Komani

Supervisors

Dr. Negar Eghbal
Dr. Hossein Baghishani

January 2017