فصل ۶

دنباله‌های تصادفی فشرده‌پذیر

۶-۱ مقدمه

همان‌طور که در فصول قبل اشاره شد، در مبحث نمونه‌برداری فشرده، نمونه‌برداری و بازسازی سیگنال‌های تنک و یا در حالت کلی‌تر، فشرده‌پذیر، مورد بررسی قرار می‌گیرد. در فصل‌های گذشته، جزئیات مربوط به نحوه‌ی نمونه‌برداری و روش‌های بازسازی مورد مطالعه قرار گرفتند. در این فصل، به مبحث پایه‌ای‌تر فشرده‌پذیری یک دنباله می‌پردازیم. در واقع، تمامی استدلال‌هایی که تاکنون ارائه شده، بر این پایه استوارندکه بردار نمونه‌های یک سیگنال پیوسته، در یک پایه خاص، نمایش تنک و یا فشرده‌پذیر دارند. بردار نمونه‌ها همواره شامل متناهی تعداد است حال آن که فشرده‌پذیری و یا تنک‌بودن مربوط به خواص ذاتی سیگنال است و در نتیجه، توقع می‌رود که بتوان فشرده‌پذیری را حتی برای دنباله‌های نامتناهی نیز بررسی کرد. از آنجا که در کاربردهای عملی، سیگنال‌های واقعی را به کمک مدل‌های آماری بیان می‌کنند، در این فصل، مدل‌سازی آماری دنباله‌های نامتناهی فشرده‌پذیر مورد بررسی قرار می‌گیرد. مطالب این فصل برگرفته از [۸] می‌باشند.

در گذشته روند رایج در بررسی فشرده‌پذیری یک دنباله نامتناهی، استفاده از فضاهای Besov بود. این فضاها زیر فضاهایی از دنباله‌های با نرم ℓ_p محدود هستند که خطای تقریبی آن‌ها به کمک بهترین k جمله، به صورت k^{-c} با افزایش k افت می‌کند (c یک ثابت است) [۲۶، ۳۳]. در تعاریف جدید، فشرده‌پذیری به کمک نرم ℓ_p ضعیف بیان می‌شود: دنباله $\{x_i\}_{i=1}^{\infty}$ فشرده‌پذیر با نرم ℓ_p نامیده می‌شود اگر $R > 0$ موجود باشد که:

$$|x_i| \leq R i^{-\frac{1}{p}} \tag{۶-۱}$$

با وجود نقاط قوت بسیار تعریف فوق، هنگامی که دنباله‌های تصادفی مورد نظر باشند، این تعریف کارآ نیست. برای روشن شدن این موضوع، فرض کنیم $\{x_i\}$ یک دنباله i.i.d. با توزیع احتمالی f_x است. در این صورت به غیر از چند حالت بدیهی، دنباله $\{x_i\}$ اصلاً میرا نخواهد بود و در نتیجه رابطه (۶-۱) صادق نیست. از طرف دیگر دنباله نمونه‌های تحقق‌های فرآیندهای تصادفی ایستان را می‌توان در حوزه نوآوری توسط دنباله‌های i.i.d. مدل کرد و در نتیجه فشرده‌پذیری سیگنال، منوط به فشرده‌پذیری دنباله نوآوری آن است. از آن‌جا که تعبیر آماری روش موفق Basis Pursuit (کمینه کردن نرم ℓ_1) به صورت تخمین گر MAP توزیع احتمال لاپلاس را برای ورودی i.i.d. پیشنهاد می‌کند، در بعضی موارد توزیع احتمال لاپلاس را به طور اشتباه یک توزیع فشرده‌پذیر قلمداد می‌کنند [۹۵، ۷۷]. این اشتباه در چندین مقاله مورد نقد قرار گرفته است [۲۷، ۲۸، ۹۹]. بررسی توزیع‌های فشرده‌پذیر، اولین بار در [۲۷] به صورت نسبتاً دقیق انجام گرفته‌است. در این مقاله به کمک تقریب‌های توزیع احتمال دنباله ترتیب‌ها[1] تعدادی از توزیع احتمال‌ها مثل Generalized Pareto و Student's t به عنوان توزیع فشرده‌پذیر معرفی شده‌اند. نقطه ضعف اصلی این مقاله، استفاده از تقریب در تعریف هاست که تعاریف را نادقیق می‌کند. علاوه بر این، نتایج به گونه‌ای است که بررسی فشرده‌پذیری توزیع‌های احتمال باید به صورت مورد به مورد انجام پذیرد؛ به عبارت بهتر، نمی‌توان فشرده‌پذیری را به صورت کامل برای یک دسته از توزیع‌ها نشان داد. در این فصل، به کمک نتایج حدی در مورد توزیع احتمال دنباله‌های مرتب، فشرده‌پذیری را به طور دقیق تعریف می‌کنیم و به کمک چند قضیه، دسته بزرگی از توزیع احتمال‌های فشرده‌پذیر را معرفی می‌کنیم.

۶-۲ دنباله‌های نامتناهی فشرده‌پذیر

برای بررسی فشرده‌پذیری دنباله‌های تصادفی، ابتدا از دنباله‌های یقینی شروع می‌کنیم. همانطور که قبلاً اشاره شد، می‌خواهیم مفهوم فشرده‌پذیری را بدون استفاده از میرایی دنباله تعریف کنیم تا بتوان این تعریف را به دنباله‌های تصادفی نیز تعمیم داد. به طور شهودی، یک دنباله متناهی فشرده‌پذیر است هرگاه درصد بالایی از انرژی کل دنباله در تعداد کمی از جملات متمرکز شده باشد. اکنون یک دنباله نامتناهی را فشرده‌پذیر می‌نامیم هرگاه تمام زیر دنباله‌های منقطع[2] آن فشرده‌پذیر باشد. برای تعریف دقیق این مفهوم، ابتدا نمادهای زیر را تعریف می‌کنیم:

تعریف ۶-۱ فرض کنید $\{a_i\}_{i=1}^n$ دنباله‌ای از اعداد حقیقی (مختلط) باشد و از لحاظ قدر مطلق، دنباله مرتب

[1] Order Statistics

[2] Truncated

شده آن $\{a_{ni}\}_{i=1}^{n}$ باشد ($|a_{n1}| \geq \cdots \geq |a_{nn}|$). برای $p > 0$ و $k \leq n$ تعریف می‌کنیم:

$$\sigma_p\big(k; \{a_i\}_{i=1}^{n}\big) = \Big(|a_{n1}|^p + \cdots + |a_{nk}|^p\Big)^{\frac{1}{p}} \tag{۶-۲}$$

به عبارت بهتر، $\sigma_p(k; \{a_i\}_{i=1}^{n})$ برابر است با نرم ℓ_p بهترین تقریب k جمله‌ای از دنباله.

تعریف ۶-۲ برای دنباله $\{a_i\}_{i=1}^{n}$ که جملات آن همگی صفر نیستند و $0 \leq r \leq 1$ و $p > 0$ مفروض، تعریف می‌کنیم:

$$\kappa_p\big(r; \{a_i\}_{i=1}^{n}\big) = \min\left\{k \;\middle|\; \frac{\sigma_p\big(k; \{a_i\}_{i=1}^{n}\big)}{\sigma_p\big(n; \{a_i\}_{i=1}^{n}\big)} \geq r\right\} \tag{۶-۳}$$

به بیان ساده‌تر $\kappa_p(.)$ نشان‌دهنده حداقل تعداد جمله مورد نیاز برای بیان کسر r از کل انرژی دنباله است.

حال فشرده‌پذیری یک دنباله را چنین تعریف می‌کنیم:

تعریف ۶-۳ دنباله $\{a_i\}_{i\in\mathbb{N}}$ از اعداد حقیقی (مختلط) را ℓ_p-فشرده‌پذیر می‌نامیم هرگاه :

$$\forall\, 0 \leq r < 1 : \quad \lim_{n\to\infty} \frac{\kappa_p\big(r; \{a_i\}_{i=1}^{n}\big)}{n} = 0 \tag{۶-۴}$$

تعریف فوق بیان می‌دارد که در یک دنباله فشرده‌پذیر، برای دستیابی به هر کسری از کل انرژی، تنها تعداد اندکی از کل جملات مورد نیاز هستند. برای بررسی این تعریف، مثال $i \in \mathbb{N},\quad a_i = \frac{1}{\alpha^i}$ را در نظر می‌گیریم که $|\alpha| \neq 1$. برای $0 \leq r < 1$ دلخواه قرار می‌دهیم $k = \lceil -\frac{\ln(1-r^p)}{p|\ln|\alpha||} \rceil$. در این صورت داریم:

$$\begin{aligned} \frac{\sigma_p\big(k; \{a_i\}_{i=1}^{n}\big)}{\sigma_p\big(n; \{a_i\}_{i=1}^{n}\big)} &= \left(\frac{\sum_{i=1}^{k} e^{-ip\left|\ln|\alpha|\right|}}{\sum_{i=1}^{n} e^{-ip\left|\ln|\alpha|\right|}}\right)^{\frac{1}{p}} \geq \left(\frac{\sum_{i=1}^{k} e^{-ip\left|\ln|\alpha|\right|}}{\sum_{i=1}^{\infty} e^{-ip\left|\ln|\alpha|\right|}}\right)^{\frac{1}{p}} \\ &= \Big(1 - e^{-\left|\ln|\alpha|\right|kp}\Big)^{\frac{1}{p}} \geq r \end{aligned} \tag{۶-۵}$$

در نتیجه $\kappa_p\big(r; \{a_i\}_{i=1}^{n}\big) \leq k$. از آنجا که k فقط به r بستگی دارد و نه n، واضح است که:

$$\lim_{n\to\infty} \frac{\kappa_p\big(r; \{a_i\}_{i=1}^{n}\big)}{n} = 0 \tag{۶-۶}$$

و در نتیجه دنباله $\{a_i\}_{i\in\mathbb{N}}$ برای هر $p > 0$، ℓ_p-فشرده‌پذیر است. مشابه استدلال فوق، می‌توان نشان داد که دنباله‌های با نرم ℓ_p محدود، ℓ_p-فشرده‌پذیر هم هستند. نکته جالب توجه در مثال فوق این است که در حالت $|\alpha| > 1$، دنباله $\{a_i\}$ نه تنها میرا نیست، بلکه به صورت نمایی به سمت بی‌نهایت میل می‌کند.

حال به بررسی خواص دنباله‌های فشرده‌پذیر می‌پردازیم. یکی از مهم‌ترین خواص این دنباله‌ها، تودرتو بودن آن‌هاست؛ یعنی اگر دنباله‌ای ℓ_p–فشرده‌پذیر باشد، برای هر $q \geq p$ و ℓ_q–فشرده‌پذیر نیز هست. برای اثبات این موضوع از لم زیر کمک می‌گیریم:

لم ۶–۱ برای دنباله دلخواه $\{a_i\}$ و اعداد صحیح $k < n$، $\left(\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}\right)^p$ نسبت به p تابعی صعودی است.

اثبات: برای اندیس‌های $1 \leq i \leq j \leq n$ و اعداد $q \geq p > 0$ داریم:

$$\begin{aligned} |a_{ni}| \geq |a_{nj}| \quad &\Rightarrow \quad |a_{ni}|^{q-p} \geq |a_{nj}|^{q-p} \\ &\Rightarrow \quad |a_{ni}|^q |a_{nj}|^p \geq |a_{ni}|^p |a_{nj}|^q \end{aligned} \tag{۶–۷}$$

اکنون با جمع کردن تمام این نامساوی‌ها برای $1 \leq i \leq k$ و $k+1 \leq j \leq n$ داریم:

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=k+1}^{n} |a_{ni}|^q |a_{nj}|^p \quad &\geq \quad \sum_{i=1}^{k} \sum_{j=k+1}^{n} |a_{ni}|^p |a_{nj}|^q \\ \Rightarrow \frac{\sum_{i=1}^{k} |a_{ni}|^q}{\sum_{i=k+1}^{n} |a_{ni}|^q} \quad &\geq \quad \frac{\sum_{i=1}^{k} |a_{ni}|^p}{\sum_{i=k+1}^{n} |a_{ni}|^p} \\ \Rightarrow \frac{\sum_{i=1}^{k} |a_{ni}|^q}{\sum_{i=1}^{n} |a_{ni}|^q} \quad &\geq \quad \frac{\sum_{i=1}^{k} |a_{ni}|^p}{\sum_{i=1}^{n} |a_{ni}|^p} \end{aligned} \tag{۶–۸}$$

که اثبات را به پایان می‌رساند. ■

لم ۶–۱ نتیجه می‌دهد که $\kappa_p(r^{\frac{1}{p}}, \{a_i\}_{i=1}^n)$ تابعی نزولی نسبت به p است. پس اگر حد مورد نظر در تعریف ۶–۳ برای یک مقدار p به صفر میل کند، این حد برای تمامی مقادیر $q \geq p$ نیز به صفر میل خواهد کرد. پس هر دنباله نامتناهی یا فشرده ناپذیر است و یا برای p از جایی به بعد ℓ_p–فشرده‌پذیر است. کوچکترین مقدار p که برای آن دنباله ℓ_p–فشرده‌پذیر است، معیاری از فشرده‌پذیری ارائه می‌دهد. یعنی دنباله ℓ_p– فشرده‌پذیر با p کمتر، فشرده‌پذیرتر تلقی می‌شود.

۶–۳ فشرده‌پذیری دنباله‌های تصادفی

برای تعمیم مفهوم فشرده‌پذیری به دنباله‌های تصادفی، لازم است که احتمال پیشامدهای مختلف نیز در تعریف گنجانده شود. به طور شهودی، یک دنباله تصادفی فشرده‌پذیر است اگر تحقق‌های آن با احتمال به دلخواه زیاد (نزدیک به یک) فشرده‌پذیر باشند. برای بیان دقیق این تعریف، ابتدا تعریف ۶–۲ را مجدداً به کمک توزیع احتمال بازنویسی می‌کنیم:

تعریف ۶-۴ فرض کنید $\{x_i\}_{i\in\mathbb{N}}$ یک دنباله $i.i.d.$ از متغیرهای تصادفی با توزیع احتمال f_x باشد. همچنین $0 \le r, \delta \le 1$ و $p \ge 0$ اعدادی دلخواه هستند. تعریف می‌کنیم:

$$\tilde{\kappa}_p(r,\delta,n;f_x) = \min\left\{k \;\middle|\; P\Big(\frac{\sigma_p\big(k;\{x_i\}_{i=1}^n\big)}{\sigma_p\big(n;\{x_i\}_{i=1}^n\big)} \ge r\Big) \ge \delta\right\} \tag{۶-۹}$$

در حقیقت $\tilde{\kappa}_p(.)$ همان نقش $\kappa_p(.)$ در دنباله‌های یقینی را ایفا می‌کند با این تفاوت که به جای نامساوی قطعی، نامساوی با احتمال δ برقرار است. به این ترتیب در تعریف $\tilde{\kappa}_p$ دو کران پایین دخالت دارند. یکی بیان‌گر حداقل درصد انرژی لازم (r) و دیگری حداقل احتمال وقوع پیشامد (δ) است. از آن‌جا که تمام خواص یک دنباله i.i.d. براساس توزیع احتمال آن تعیین می‌شود، فشرده‌پذیری یک دنباله تصادفی را به جای دنباله، به خود توزیع احتمال نسبت می‌دهیم. تعریف زیر مشابه حالت غیرتصادفی، فشرده‌پذیری تصادفی (برای توزیع احتمال) را براساس رفتار حدی $\tilde{\kappa}_p$ بیان می‌کند:

تعریف ۶-۵ توزیع احتمال f_x را ℓ_p-فشرده‌پذیر می‌نامیم هرگاه:

$$\forall\, 0 \le r, \delta < 1, \quad \lim_{n\to\infty} \frac{\tilde{\kappa}_p(r,\delta,n;f_x)}{n} = 0 \tag{۶-۱۰}$$

دقت کنید که حالات $\delta = 1, r = 1$ در رابطه فوق درنظر گرفته نشده‌اند. در بردارهای با طول متناهی، براساس این که بهترین تقریب k-جمله از لحاظ انرژی نمایش دقیق و یا تقریبی از بردار ایجاد می‌کند، آن را تنک یا فشرده‌پذیر می‌نامند. در این‌جا نیز می‌توان دنباله تصادفی تنک را به طور مشابهی تعریف کرد:

تعریف ۶-۶ توزیع احتمال f_x را ℓ_p-تنک می‌نامیم هرگاه ℓ_p-فشرده‌پذیر باشد و در هر تحقق از دنباله‌های $i.i.d.$ با این توزیع احتمال، کل انرژی دنباله در کسری از جملات متمرکز باشد:

$$\forall\, p,\ 0 \le \delta < 1, \quad \lim_{n\to\infty} \frac{\tilde{\kappa}_p(1,\delta,n;f_x)}{n} < 1 \tag{۶-۱۱}$$

به زبان ساده‌تر، تنها فرق فشرده‌پذیری و تنک‌بودن، رفتار دنباله در حالت $r = 1$ است. در یک دنباله تنک توقع داریم همواره کسری از جملات صفر باشند و در انرژی کل مشارکت نکنند. در واقع (۶-۱۱) بیان می‌کند که $P(x = 0) = \pi_0 > 0$ و یا به عبارت دیگر توزیع احتمال f_x در نقطه $x = 0$ شامل یک تابع ضربه است. از سوی دیگر اگر $P(x = 0) = \pi_0 > 0$، با احتمال بسیار بالا کل انرژی یک دنباله nتایی در $n(1 - \pi_0)$ جمله قرار می‌گیرد. با وجود شرط اضافه در تعریف تنک‌بودن، این توزیع‌ها به راحتی براساس توزیع‌های فشرده‌پذیر

مشخص می‌شوند. فرض کنید f_x یک توزیع فشرده‌پذیر باشد، برای تولید یک دنباله تنک می‌توان ابتدا به کمک f_x یک دنباله فشرده‌پذیر ساخت و سپس با احتمال π_0 هریک از این جملات را صفر کرد و یا ثابت نگه داشت. به راحتی می‌توان نشان داد که دنباله‌ی حاصل، تنک است. شایان ذکر است که توزیع‌های فشرده‌پذیر و یا تنک، ساختاری تودرتو مشابه آن چه در لم ۶-۱ برای ساختارهای غیرتصادفی نشان داده شد، دارند: یک توزیع ℓ_p-فشرده‌پذیر (ℓ_p-تنک) برای هر $q \geq p$ ℓ_q-فشرده‌پذیر (ℓ_q-تنک) نیز هست. همچنین لم زیر رابطه بین توزیع‌های ℓ_p-فشرده‌پذیر و ℓ_1-فشرده‌پذیر را نشان می‌دهد:

لم ۶-۲ توزیع احتمال متغیر تصادفی x از نوع ℓ_p-فشرده‌پذیر است اگر و تنها اگر، توزیع احتمال متغیر $y = |x|^p$ ℓ_1-فشرده‌پذیر باشد.

اثبات: با توجه به این نکته که نرم ℓ_p دنباله $\{x_i\}$ با نرم ℓ_1 دنباله $\{y_i\}$ معادل است، لم ثابت می‌شود.

۶-۴ تشخیص توزیع احتمال‌های فشرده‌پذیر

در ادامه توسط دو قضیه، دسته بزرگی از توزیع‌های فشرده‌پذیر را مشخص می‌کنیم. ابتدا در قضیه اول نشان می‌دهیم که توزیع‌هایی که به صورت نمایی میرا می‌شوند فشرده ناپذیرند. در قضیه دوم نیز نشان می‌دهیم که اگر میرایی به صورت چندجمله‌ای باشد، توزیع فشرده‌پذیر است. با توجه به لم ۶-۲ برای شناسایی توزیع‌های فشرده‌پذیر، فقط لازم است توزیع‌های ℓ_1-فشرده‌پذیر را شناسایی کنیم. از این رو در ادامه تنها بر روی دنباله‌های ℓ_1-فشرده‌پذیر متمرکز می‌شویم.

قضیه ۶-۱ اگر توزیع احتمال f_x به گونه‌ای باشد که برای یک $\gamma > 0$، $\mathscr{E}\{e^{\gamma|x|}\}$ محدود باشد، این توزیع ℓ_1-فشرده‌پذیر نیست.

اثبات: فرض کنید $\{x_i\}$ دنباله‌ای i.i.d. با توزیع احتمال f_x باشد و قرار دهید $y_i = |x_i|$. به وضوح دنباله $\{y_i\}$ نیز i.i.d. خواهد بود اما با توزیع $f_y(y) = \big(f_x(y) + f_x(-y)\big)u(y)$. همچنین فرض کنید μ_y و σ_y^2 به ترتیب میانگین و واریانس توزیع f_y باشند (از آن‌جا که $\mathscr{E}_x\{e^{\gamma|x|}\}$ محدود است، μ_y و σ_y نیز محدودند). برای $0 < r < 1$ دلخواه، فرض کنید m عدد مثبت به اندازه کافی بزرگ است که $\mathscr{E}_y\{e^{\gamma y}\} < e^{m\gamma r\mu_y}(3m)^{-1}$ (γ همان مقدار مثبتی است که برای آن $\mathscr{E}\{e^{\gamma y}\}$ محدود است). برای بررسی فشرده‌پذیری پیشامدهای زیر را تعریف

می‌کنیم:

$$\begin{cases} EV_1: & \sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\sum_{i=1}^{mk} y_i \\ EV_2: & \sum_{i=1}^{mk} y_i \geq mk\mu_y - (mk)^{0.75}\sigma_y \\ EV_3: & \sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y \end{cases} \tag{۶-۱۲}$$

در نتیجه:

$$\begin{aligned} P(EV_1) &= P(EV_1|EV_2)P(EV_2) + P(EV_1|EV_3)P(EV_3) \leq P(EV_1|EV_2) + P(EV_3) \\ &\leq P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) + P\Big(\sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y\Big) \end{aligned} \tag{۶-۱۳}$$

از آن‌جا که y_iها مستقل‌اند (i.i.d.) داریم:

$$\begin{cases} \mathscr{E}_y\{\sum_{i=1}^{mk} y_i\} = mk\mu_y \\ \mathrm{Var}\{\sum_{i=1}^{mk} y_i\} = mk\sigma_y^2 \end{cases} \tag{۶-۱۴}$$

بنابراین با کمک نامساوی چبیشف[۳] بدست می‌آید:

$$P\Big(\sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y\Big) \leq \frac{1}{\sqrt{mk}} \tag{۶-۱۵}$$

علاوه بر این، $\sigma_1(k;\{y_i\}_{i=1}^{mk})$ توسط یکی از $\binom{mk}{k}$ حالت انتخاب k جمله از کل mk جمله بدست آمده است.

پس:

$$P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) \leq \binom{mk}{k} P\Big(\sum_{i=1}^{k} y_i \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) \tag{۶-۱۶}$$

از آن‌جا که $\mathscr{E}\{\sum_{i=1}^{k} y_i\} = k\mu_y$، احتمال برقراری $\sum_{i=1}^{k} y_i \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)$ جزء احتمالات انتهایی[۴] طبقه‌بندی می‌شود. اکنون نشان می‌دهیم این احتمال نسبت به k به صورت نهایی میرا می‌شود:

$$P\big(\sum_{i=1}^{k} y_i \geq T\big) = P\big(e^{\gamma\sum_{i=1}^{k} y_i} \geq e^{\gamma T}\big) \leq \frac{\mathscr{E}\{e^{\gamma\sum_{i=1}^{k} y_i}\}}{e^{\gamma T}} = \big(\mathscr{E}\{e^{\gamma y}\}e^{-\gamma\frac{T}{k}}\big)^k < \Big(\frac{e^{\gamma(rm\mu_y - \frac{T}{k})}}{3m}\Big)^k \tag{۶-۱۷}$$

که اولین نامساوی در رابطه فوق توسط نامساوی مارکف[۵] بدست آمده است. حال به استناد $\binom{mk}{k} < (em)^k$ و نتایج بدست آمده در (۶-۱۶) و (۶-۱۷) داریم:

$$P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) < \Big(\frac{e}{3}e^{\gamma r\sigma_y m^{0.75}k^{-0.25}}\Big)^k \tag{۶-۱۸}$$

[۳] Chebychev

[۴] Tail Probability

[۵] Markov

که با در نظر گرفتن (۶-۱۳) و نتیجه می‌دهد:

$$P(EV_1) = P\left(\frac{\sigma_1(k;\{x_i\}_{i=1}^{mk})}{\sigma_1(mk;\{x_i\}_{i=1}^{mk})} \geq r\right) < \frac{1}{\sqrt{mk}} + \left(\frac{e}{3}e^{\gamma r \sigma_y m^{0.75} k^{-0.25}}\right)^k \qquad (۶-۱۹)$$

به راحتی می‌توان بررسی کرد که برای یک m ثابت، هنگامی که $k \to \infty$ کران بالایی فوق و در نتیجه احتمال $P(EV_1)$ به صفر میل می‌کنند. از این رو برای این که احتمال $P(EV_1)$ از یک کران پایین دلخواه مانند δ بزرگتر باشد، به بیش از k جمله از کل mk جمله دنباله نیازمندیم. پس:

$$\lim_{k\to\infty} \frac{\tilde{\kappa}_p(r,\delta,km;f_x)}{km} \geq \frac{1}{m} \qquad (۶-۲۰)$$

که m تنها به r بستگی دارد و نه k. در نتیجه، توزیع احتمال f_x، ℓ_1-فشرده‌پذیر نیست. ■

قضیه ۶-۱ نشان می‌دهد که توزیع‌هایی نظیر لاپلاس، گاما و گوسی که میرایی نمایی دارند، ℓ_1-فشرده‌پذیر نیستند. در ادامه نشان می‌دهیم که در صورت میرایی از مرتبه چندجمله‌ای، توزیع احتمال فشرده‌پذیر است. پیش از بیان این نتیجه، ابتدا به بیان پیش نیازهایی در نظریه احتمال می‌پردازیم:

براساس قضیه معروف حد مرکزی[۶] اگر $\{x_i\}_{i\in\mathbb{N}}$ دنباله‌ای i.i.d. از متغیرهای تصادفی با میانگین (μ_x) و واریانس (σ_x^2) محدود باشد، زیرجمع‌های نرمالیزه شده آن‌ها به فرم $n^{-0.5}\sum_{i=1}^{n}(x_i-\mu_x)$ در حد، به سمت متغیرتصادفی گوسی با میانگین صفر و واریانس σ_x^2 میل می‌کند. علاوه بر این، جمع دو متغیرتصادفی مستقل گوسی مجدداً متغیر گوسی است؛ به بیان دیگر، توزیع گوسی نسبت به عمل جمع پایدار[۷] است. نکته جالب آن است که مجموعه توزیع‌های پایدار، تنها محدود به توزیع‌های گوسی نمی شود. دسته‌ای از توزیع‌ها معروف به α-پایدار موجودند که توسط پارامتر $0 < \alpha \leq 2$ اندیس‌گذاری می‌شوند و همگی نسبت به عمل جمع پایدارند؛ حالت خاص $\alpha = 2$ مربوط به توزیع گوسی است. به غیر از حالت $\alpha = 2$ (توزیع گوسی)، میرایی این توزیع‌ها به صورت $|t|^{-(\alpha+1)}$ است و در نتیجه ممان‌های آن‌ها تنها تا مرتبه α متناهی است. مثلاً هیچ کدام از آن‌ها واریانس متناهی ندارند [۸۲]. مشابه حالت گوسی، توزیع حدی زیرجمع‌های نرمالیزه شده (با $n^{-\frac{1}{\alpha}}$) دسته خاصی از توزیع احتمال‌ها به صورت α-پایدار خواهد بود؛ به مجموعه این توزیع احتمال‌ها حوزه جذب[۸] توزیع α-پایدار می‌گویند.

---

[۶] Central Limit Theorem

[۷] Stable

[۸] Domain of Attraction

برای شناسایی حوزه جذب در حالت $0 < \alpha < 2$، فرض کنید x یک متغیرتصادفی باشد و تعریف کنید $y = |x|$ و $G(t) = P(y > t)$. طبق نتیجه‌ای در نظریه احتمال [۴۵] می‌دانیم که توزیع احتمال x در حوزه جذب یک توزیع α-پایدار است اگر و فقط اگر:

۱. تابع $h(t) = t^{\alpha}G(t)$ تغییرات کندی در $t \to \infty$ داشته باشد، به بیان دیگر، برای هر $c > 0$ حد $\frac{h(ct)}{h(t)}$ وقتی t به سمت بی‌نهایت میل می‌کند، برابر یک باشد؛

۲. حد $\lim_{t\to\infty} \frac{P(x>t)}{G(t)}$ موجود باشد.

همانطور که اشاره شد، به دلیل میرایی چندجمله‌ای در توزیع‌های متعلق به حوزه جذب توزیع‌های α-پایدار، این توزیع‌ها از دیدگاه فشرده‌پذیری گزینه‌های مناسبی به شمار می‌روند. برای اثبات این موضوع ابتدا قضیه زیر را که در [۶۵] اثبات شده است، بیان می‌کنیم:

قضیه ۶-۲ فرض کنید $\{\eta_i\}$ دنباله‌ای *i.i.d.* از متغیرهای تصادفی با توزیع احتمال نمایی استاندارد (میانگین واحد) باشند و فرض کنید $\Gamma_i = \sum_{j=1}^{i} \eta_j$. اگر توزیع احتمال دنباله $\{x_i\}$ متعلق به حوزه جذب توزیع α-پایدار باشد داریم:

$$\lim_{n\to\infty} a_n^{-1}\left(y_{n1}, \ldots, y_{nn}, 0, 0, \ldots\right) =_d \left(\Gamma_1^{-\frac{1}{\alpha}}, \Gamma_2^{-\frac{1}{\alpha}}, \ldots\right) \qquad (۶-۲۱)$$

که در آن $=_d$ به معنای همگرایی در توزیع است و $\{y_{ni}\}_{i=1}^{n}$ دنباله مرتب شده (نزولی) $\{|x_i|\}_{i=1}^{n}$ را نشان می‌دهید. همچنین:

$$a_n = \lim_{t\to\infty} \frac{G^{-1}\left(\frac{1}{nt^{\alpha}}\right)}{t} \qquad (۶-۲۲)$$

نکته در خور توجه قضیه ۶-۲ این است که توزیع‌های حدی دنباله مرتب شده دسته‌ای از توزیع‌ها را می‌توان تنها با بررسی توزیع نمایی استاندارد بدست آورد. اکنون ابزار لازم برای نشان دادن فشرده‌پذیری دسته بزرگی از توزیع‌ها را در اختیار داریم:

قضیه ۶-۳ اگر توزیع احتمال متغیر x که با f_x نمایش داده می‌شود متعلق به حوزه جذب یک توزیع α-پایدار با $\alpha < 1$ باشد، در این صورت f_x ℓ_1-فشرده‌پذیر است.

اثبات: برای اثبات فشرده‌پذیری از احتمال بیان شده در تعریف ۶-۴ شروع می‌کنیم. از آن‌جا که $\sigma_1\big(k;\{x_i\}_{i=1}^n\big)=\sigma_1\big(k;\{y_i\}_{i=1}^n\big)$ (که $y_i=|x_i|$) داریم:

$$P\Big(\frac{\sigma_1\big(k;\{x_i\}_{i=1}^n\big)}{\sigma_1\big(n;\{x_i\}_{i=1}^n\big)}\geq r\Big)=P\Big(\frac{a_n^{-1}\big(y_{n(k+1)}+\cdots+y_{nn}\big)}{a_n^{-1}\big(y_{n1}+\cdots+y_{nn}\big)}\leq 1-r\Big) \tag{۶-۲۳}$$

در تعریف فشرده‌پذیری ما به دنبال رفتار حدی احتمال فوق هستیم. طبق قضیه ۶-۲ در حالت حدی می‌توانیم $a_n^{-1}y_{ni}$ را با $\Gamma_i^{-\frac{1}{\alpha}}$ جایگذاری کنیم (این موضوع را به طور دقیق‌تر در انتهای اثبات نشان می‌دهیم). در نتیجه داریم:

$$P\Big(\frac{\sigma_1\big(k;\{x_i\}_{i=1}^n\big)}{\sigma_1\big(n;\{x_i\}_{i=1}^n\big)}\geq r\Big)\sim P\Big(\frac{\Gamma_{k+1}^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}{\Gamma_1^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}\leq 1-r\Big) \tag{۶-۲۴}$$

دقت کنید :

$$\begin{aligned}\frac{\Gamma_{k+1}^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}{\Gamma_1^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}} &= \frac{k^{2-\alpha}\big(\frac{\Gamma_1}{\Gamma_{k+1}}\big)^{\frac{1}{\alpha}}}{1+\big(\frac{\Gamma_1}{\Gamma_2}\big)^{\frac{1}{\alpha}}+\cdots+\big(\frac{\Gamma_1}{\Gamma_n}\big)^{\frac{1}{\alpha}}}\times\frac{1+\big(\frac{\Gamma_{k+1}}{\Gamma_{k+2}}\big)^{\frac{1}{\alpha}}+\cdots+\big(\frac{\Gamma_{k+1}}{\Gamma_n}\big)^{\frac{1}{\alpha}}}{k^{2-\alpha}}\\ &\leq \underbrace{k^{2-\alpha}\big(\frac{\Gamma_1}{\Gamma_{k+1}}\big)^{\frac{1}{\alpha}}}_{A_k}\underbrace{\frac{1+\big(\frac{\Gamma_{k+1}}{\Gamma_{k+2}}\big)^{\frac{1}{\alpha}}+\cdots+\big(\frac{\Gamma_{k+1}}{\Gamma_n}\big)^{\frac{1}{\alpha}}}{k^{2-\alpha}}}_{B_{k,n}}\end{aligned} \tag{۶-۲۵}$$

که نتیجه می‌دهد:

$$\begin{aligned}P\Big(\frac{\Gamma_{k+1}^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}{\Gamma_1^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}\leq 1-r\Big) &\geq 1-P\big(A_kB_{k,n}\geq 1-r\big)\geq 1-P\big(A_k>1\big)-P\big(B_{k,n}>1-r\big)\\ &\geq 1-\mathscr{E}\{A_k\}-\frac{1}{1-r}\mathscr{E}\{B_{k,n}\}\end{aligned} \tag{۶-۲۶}$$

آخرین نامساوی با استفاده از نامساوی مارکف حاصل شده است. در ادامه اثبات نشان می‌دهیم هر دوی $\mathscr{E}\{A_k\}$ و $\mathscr{E}\{B_{k,n}\}$ در حالت حدی $k\to\infty$ و $n>k+2$ به صفر میل می‌کنند.

برای بدست آوردن حد $\mathscr{E}\{A_k\}$ وقتی $k\to\infty$، ابتدا لم زیر را اثبات می‌کنیم:

لم ۶-۳ فرض کنید $\{\eta_i\}_{i\in\mathbb{N}}$ دنباله‌ای از متغیرهای تصادفی *i.i.d.* با توزیع احتمال نمایی استاندارد باشد و θ,β اعداد مثبت دلخواهی باشند. داریم:

$$0\leq\mathscr{E}_{\{\eta_i\}_i}\left\{\frac{1}{\big(1+\frac{1}{\theta}\sum_{i=1}^k\eta_i\big)^{\beta}}\right\}<0/83^k+\Big(\frac{2\theta}{k}\Big)^{\beta} \tag{۶-۲۷}$$

اثبات: از آن‌جا که η_iها و θ نامنفی‌اند، نامساوی سمت چپ بدیهی است. برای سادگی در ادامه اثبات تعریف می‌کنیم $\Gamma_k = \sum_{i=1}^{k} \eta_i$. در نتیجه:

$$\begin{aligned}\mathscr{E}_{\{\eta_i\}}\Big\{\frac{1}{(1+\frac{\Gamma_k}{\theta})^{\beta}}\Big\} &= \int_0^{\infty} \frac{1}{(1+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma = \int_0^{\frac{k}{2}} \frac{1}{(1+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma + \int_{\frac{k}{2}}^{\infty} \frac{1}{(1+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma \\ &\leq \int_0^{\frac{k}{2}} f_{\Gamma_k}(\gamma)d\gamma + \int_{\frac{k}{2}}^{\infty} \Big(\frac{2\theta}{k}\Big)^{\beta} f_{\Gamma_k}(\gamma)d\gamma \leq P\Big(\Gamma_k \leq \frac{k}{2}\Big) + \Big(\frac{2\theta}{k}\Big)^{\beta} \qquad (۲۸-۶)\end{aligned}$$

همانطور که اشاره شد، Γ_k مجموع k متغیرتصادفی مستقل با میانگین واحد است که ایجاب می‌کند $\mathscr{E}\{\Gamma_k\} = k$. بنابراین $P(\Gamma_k \leq \frac{k}{2})$ یک احتمال انتهایی محسوب می‌شود. روش زیر راهی معروف برای نشان دادن میرایی نمایی این‌گونه احتمالات است:

$$\begin{aligned}P\Big(\Gamma_k \leq \frac{k}{2}\Big) &= P\Big(-\sum_{i=1}^{k}\eta_i \geq -\frac{k}{2}\Big) = P\Big(e^{-\sum_{i=1}^{k}\eta_i} \geq e^{-\frac{k}{2}}\Big) \leq \frac{\mathscr{E}_{\{\eta_i\}_i}\{e^{-\sum_{i=1}^{k}\eta_i}\}}{e^{-\frac{k}{2}}} \\ &= e^{\frac{k}{2}}\big(\mathscr{E}_{\eta}\{e^{-\eta}\}\big)^k = \Big(\frac{\sqrt{e}}{2}\Big)^k < 0/83^k \qquad (۲۹-۶)\end{aligned}$$

نامساوی اول به کمک نامساوی کرافت معتبر است. ■

حال به سراغ $\mathscr{E}\{A_k\}$ می‌رویم. طبق تعریف داریم:

$$\mathscr{E}\{A_k\} = k^{2-\alpha}\mathscr{E}_{\{\eta_i\}_i}\Big\{\big(\frac{\Gamma_1}{\Gamma_{k+1}}\big)^{\frac{1}{\alpha}}\Big\} = k^{2-\alpha}\mathscr{E}_{\eta_1}\Big\{\mathscr{E}_{\{\eta_i\}_{i>1}}\Big\{\big(1+\frac{1}{\eta_1}\sum_{i=2}^{k+1}\eta_i\big)^{-\frac{1}{\alpha}}\Big\}\Big\} \qquad (۳۰-۶)$$

اکنون با استفاده از لم ۶-۳ کران بالایی زیر حاصل می‌شود:

$$\mathscr{E}_{\{\eta_i\}_{i>1}}\Big\{\big(1+\frac{1}{\eta_1}\sum_{i=2}^{k+1}\eta_i\big)^{-\frac{1}{\alpha}}\Big\} < 0/83^k + \big(\frac{2\eta_1}{k}\big)^{\frac{1}{\alpha}} \qquad (۳۱-۶)$$

که در نهایت منجر به رابطه زیر می‌شود:

$$\mathscr{E}\{A_k\} \leq k^{2-\alpha}0/83^k + k^{2-\alpha-\frac{1}{\alpha}}\mathscr{E}_{\eta}\Big\{(2\eta)^{\frac{1}{\alpha}}\Big\} \qquad (۳۲-۶)$$

با توجه به $2 < \alpha + \frac{1}{\alpha}$ واضح است که $\lim_{k\to\infty}\mathscr{E}\{A_k\} = 0$. مشابهاً برای بدست آوردن حد $\mathscr{E}\{B_{k,n}\}$ هنگامی که $k \to \infty$، به بیان لم زیر می‌پردازیم:

لم ۶-۴ فرض کنید $\{\eta_i\}_{i\in\mathbb{N}}$ دنباله‌ای از متغیرهای تصادفی *i.i.d.* با توزیع احتمال نمایی استاندارد باشد و $1 < \beta$, $0 < \theta$ اعداد دلخواهی باشند. داریم:

$$\mathscr{E}_{\{\eta_i\}_i}\Big\{1+\sum_{k=1}^{\infty}\frac{1}{(1+\frac{1}{\theta}\sum_{i=1}^{k}\eta_i)^{\beta}}\Big\} = 1+\frac{\theta}{\beta-1} \qquad (۳۳-۶)$$

اثبات: مشابه قبل، فرض کنید $\Gamma_k = \sum_{i=1}^{k} \eta_i$. به دلیل .i.i.d بودن متغیرهای تصادفی، توزیع احتمال Γ_k از k مرتبه کانولوشن توزیع نمایی با خودش حاصل می‌شود:

$$f_{\Gamma_k}(\gamma) = \big(\underbrace{f_\eta * \cdots * f_\eta}_{k-\text{times}}\big)(\gamma) = \mathcal{F}_\omega^{-1}\Big\{(1+j\omega)^{-k}\Big\}(\gamma) \tag{۶-۳۴}$$

در نتیجه با استفاده از نمایش حوزه فرکانس داریم:

$$\mathscr{E}_{\{\eta_i\}}\Big\{\frac{1}{\big(1+\frac{\Gamma_k}{\theta}\big)^\beta}\Big\} = \int_0^\infty \int_{\mathbb{R}} \frac{1}{\big(1+\frac{\gamma}{\theta}\big)^\beta} \frac{e^{j\omega\gamma}}{(1+j\omega)^k} \frac{d\omega}{2\pi}\, d\gamma \tag{۶-۳۵}$$

که نتیجه می‌دهد:

$$\begin{aligned}
\mathscr{E}_{\{\eta_i\}_i}\Big\{1+\sum_{k=1}^{\infty} \frac{1}{\big(1+\frac{\Gamma_k}{\theta}\big)^\beta}\Big\} &= \int_0^\infty \int_{\mathbb{R}} \frac{e^{j\omega\gamma}}{\big(1+\frac{\gamma}{\theta}\big)^\beta} \sum_{k=0}^{\infty} \frac{1}{(1+j\omega)^k} \frac{d\omega}{2\pi}\, d\gamma \\
&= \int_0^\infty \int_{\mathbb{R}} \frac{e^{j\omega\gamma}}{\big(1+\frac{\gamma}{\theta}\big)^\beta} \big(1+\frac{1}{j\omega}\big) \frac{d\omega}{2\pi}\, d\gamma \\
&= \int_{\mathbb{R}} \frac{\delta(\gamma)+1}{\big(1+\frac{\gamma}{\theta}\big)^\beta} d\gamma = 1 + \frac{\theta}{\beta-1} < \infty \quad \blacksquare
\end{aligned} \tag{۶-۳۶}$$

نکته جالب توجه در لم فوق ارتباط آن با تابع زتای ریمان[۹] است: در لم ۶-۴ هنگامی که $\theta = 1$، ما به دنبال یافتن امید ریاضی عبارتی به صورت $1 + \sum_{k=1}^{\infty} \Lambda_k^{-\beta}$ هستیم که $\{\Lambda_k\}_k$ دنباله‌ای از متغیرهای تصادفی با شرط $\Lambda_k \leq \Lambda_{k+1}$ و $\mathscr{E}\{\Lambda_k\} = k+1$ است. در حالت حدی، هنگامی که واریانس این متغیرهای تصادفی به سمت صفر میل کند (عبور از حالت تصادفی به یقینی) این امید ریاضی به سمت $1 + \sum_{k=2}^{\infty} k^{-\beta}$ میل می‌کند که همان تابع زتای ریمان در نقطه β است. بنابراین امید ریاضی مورد نظر را می‌توان در حالت کلی به صورت تعمیم تصادفی تابع ریمان تفسیر کرد. نکته جالب توجه این است که $\beta = 1$ در هر دو حالت یک قطب ساده است.

اکنون به بررسی حد $\mathscr{E}\{B_{k,n}\}$ بر می‌گردیم:

$$\mathscr{E}\{B_{k,n}\} = k^{\alpha-2}\mathscr{E}\Big\{1+\sum_{t=k+2}^{n} \big(\frac{\Gamma_{k+1}}{\Gamma_t}\big)^{\frac{1}{\alpha}}\Big\} \leq k^{\alpha-2}\mathscr{E}\Big\{1+\sum_{t=k+2}^{\infty} \big(\frac{\Gamma_{k+1}}{\Gamma_t}\big)^{\frac{1}{\alpha}}\Big\} \tag{۶-۳۷}$$

با استفاده از لم ۶-۴ داریم:

$$\begin{aligned}
\mathscr{E}_{\{\eta_i\}_{i>k+1}}\Big\{1+\sum_{t=k+2}^{\infty} \big(\frac{\Gamma_{k+1}}{\Gamma_t}\big)^{\frac{1}{\alpha}}\Big\} &= \mathscr{E}_{\{\eta_i\}_{i>k+1}}\Big\{1+\sum_{t=k+2}^{\infty} \frac{1}{\Big(1+\frac{\sum_{i=k+2}^{t}\eta_i}{\Gamma_{k+1}}\Big)^{\frac{1}{\alpha}}}\Big\} \\
&= 1 + \frac{\alpha}{1-\alpha}\Gamma_{k+1}
\end{aligned} \tag{۶-۳۸}$$

[۹] Reiman Zeta Function

جدول ۶-۱: تعدادی از توزیع‌های فشرده‌پذیر با میرایی مرتبه $|t|^{-(q+1)}$

$f_x(t)$	Distribution
$\frac{q}{2\lambda}\left(1+\frac{\lvert t\rvert}{\lambda}\right)^{-(q+1)}$	Generalized Pareto
$\frac{\Gamma((q+1)/2)}{\Gamma(q/2)\sqrt{2\pi\lambda}}\left(1+\frac{t^2}{2\lambda}\right)^{-(q+1)/2}$	Student's t
$\frac{q}{2\lambda}e^{-(1+\frac{\lvert t\rvert}{\lambda})^{-q}}\left(1+\frac{\lvert t\rvert}{\lambda}\right)^{-(q+1)}$	Extreme Value
$\frac{(q/\lambda)(t/\lambda)^{q-1}}{\left(1+(t/\lambda)^q\right)^2}$	Log-Logistic

بنابراین:

$$\mathscr{E}\{B_{k,n}\} \quad \leq \quad k^{\alpha-2}\mathscr{E}_{\Gamma_{k+1}}\left\{1+\frac{\alpha}{1-\alpha}\Gamma_{k+1}\right\} = k^{\alpha-2}\left(1+\frac{\alpha(k+1)}{1-\alpha}\right) \tag{۶-۳۹}$$

مجدداً از آن‌جا که $\alpha < 1$، کران بالای فوق در حالت حدی $k \to \infty$ به سمت صفر میل می‌کند. با استفاده از (۶-۲۶) و این موضوع که $\mathscr{E}\{A_k\} = 0$ و $\mathscr{E}\{B_{k,n}\} = 0$ در حالت حدی صفر هستند، می‌دانیم که:

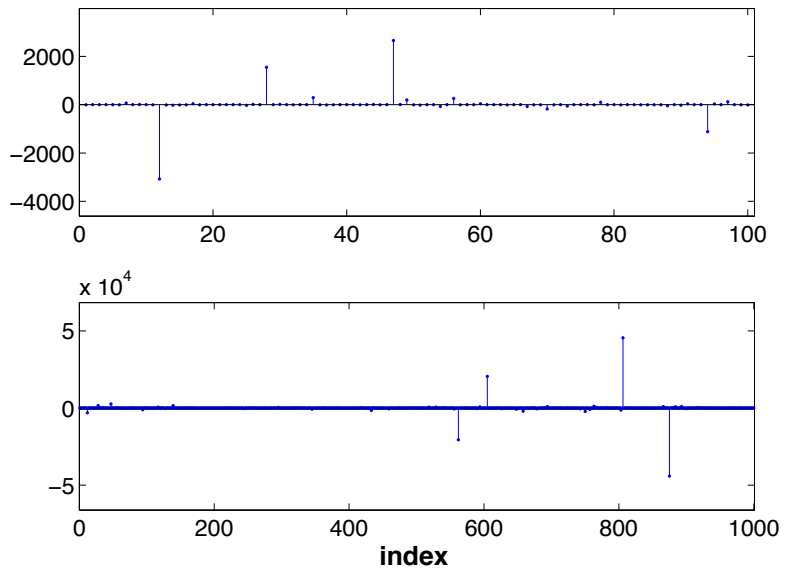
$$\lim_{k\to\infty} P\left(\frac{\Gamma_{k+1}^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}}{\Gamma_1^{-\frac{1}{\alpha}}+\cdots+\Gamma_n^{-\frac{1}{\alpha}}} \leq 1-r\right) = 1 \tag{۶-۴۰}$$

دقت کنید که کران بالای اثبات شده برای $\mathscr{E}\{A_k\}$ و $\mathscr{E}\{B_{k,n}\}$ در (۶-۳۲) و (۶-۳۹)، تنها تابعی از k و مستقل از n هستند. پس برای این که احتمال طرح شده در (۶-۴۰) برای $0 < \epsilon < 1$ دلخواه از $1 - \frac{\epsilon}{2}$ بزرگتر باشد، تنها کافی است k از حدی بزرگتر اختیار شود ($k > k_0$). از طرف دیگر، قضیه ۶-۲ نشان می‌دهد که برای این مقدار k_0 (که تنها به ϵ بستگی دارد)، n به اندازه کافی بزرگ، اختلاف بین دو احتمال در (۶-۲۴) از $\frac{\epsilon}{2}$ کوچکتر خواهد بود. بنابراین برای این مقدار k_0 و nهای به اندازه کافی بزرگ داریم:

$$P\left(\frac{\sigma_1\left(k_0;\{x_i\}_{i=1}^n\right)}{\sigma_1\left(n;\{x_i\}_{i=1}^n\right)} \geq r\right) \geq 1-\epsilon \tag{۶-۴۱}$$

به وضوح هنگامی که n رشد می‌کند، $\frac{k_0}{n} \to 0$، که اثبات را به پایان می‌برد. ■

قضیه ۶-۳ تقریباً بیان می‌کند که توزیع احتمال $f_x(t)$ با میرایی چندجمله‌ای از مرتبه $|t|^{-(\alpha+1)}$ که $\alpha < 1$، ℓ_1-فشرده‌پذیر است. حال فرض کنید میرایی توزیع احتمال از مرتبه $|t|^{-(q+1)}$ باشد که q لزوماً کمتر از یک نیست. در این صورت توزیع احتمال $y = |x|^p$ که برابر است با $f_y(t) = \frac{f_x(t^{1/p})+f_x(-t^{1/p})}{pt^{1-1/p}}$ از مرتبه $|t|^{-(q/p+1)}$ میرا خواهد بود. لذا با توجه به لم ۶-۲، f_x برای هر $p > q$، ℓ_p-فشرده‌پذیر خواهد بود. جدول ۶-۱ تعدادی از



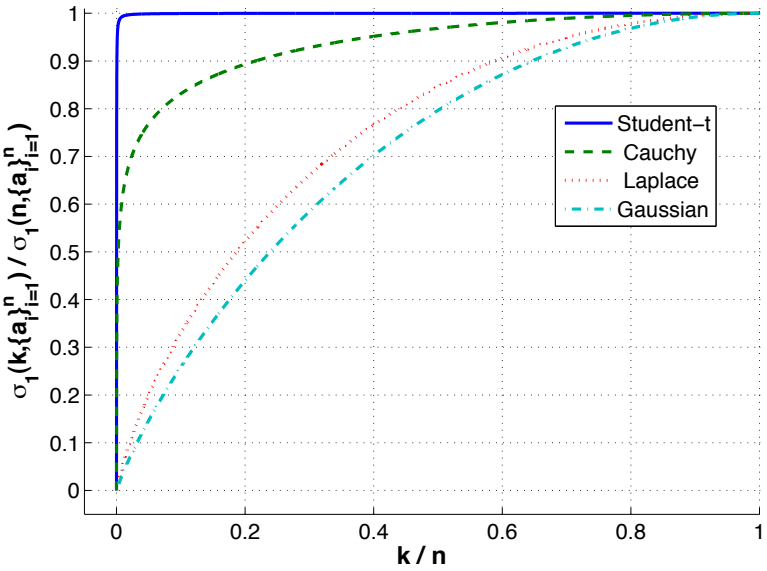
شکل ۶-۱: یک تحقق زمانی از دنباله تصادفی و .i.i.d با توزیع احتمال Student's t و پارامتر $q = ۰/۵$.

توزیع احتمال‌های میرا از مرتبه $|t|^{-(q+۱)}$ را نشان می‌دهد که مطابق آن چه گفته شد، فشرده‌پذیر نیز هستند. جالب آن که این توزیع‌ها با استدلال‌های مطرح شده در [۲۷] مطابقت دارند. علاوه بر این توزیع‌ها، تمامی توزیع‌های α-پایدار برای $۰ < \alpha < ۲$ که شامل توزیع کوشی و Lévy نیز می‌شود، برای $p > \alpha$، ℓ_p-فشرده‌پذیر هستند.

۶-۵ نتایج عددی

تعریف فشرده‌پذیری در دنباله‌های تصادفی بیان می‌کند که تحقق‌های این دنباله‌ها با احتمال بالا فشرده‌پذیرند. برای بررسی این مطلب، در شکل ۶-۱ یک تحقق از دنباله .i.i.d با توزیع Student's t با پارامتر $q = ۰/۵$ (میرایی از مرتبه $|t|^{-۰/۵}$) رسم شده است. در این شکل یک بار ۱۰۰ جمله اول و بار دیگر ۱۰۰۰ جمله اول که شامل ۱۰۰ جمله قبلی هم می‌شوند، دیده می‌شود. در هر دو شکل مشاهده می‌شود که تعداد جملاتی که مقدار قابل توجهی دارند، اندک است. اما نکته جالب‌تر از آن این است که تمام جملات پر اهمیت در ۱۰۰ جمله اول، جزء جملات بسیار کوچک در پنجره ۱۰۰۰ تایی بشمار می‌آیند. این امر نشان می‌دهد که با جلو رفتن در دنباله‌های .i.i.d فشرده‌پذیر، همواره شاهد رشد جملات اصلی هستیم به نحوی که باعث کم اهمیت شدن جملات قبلی می‌شوند.

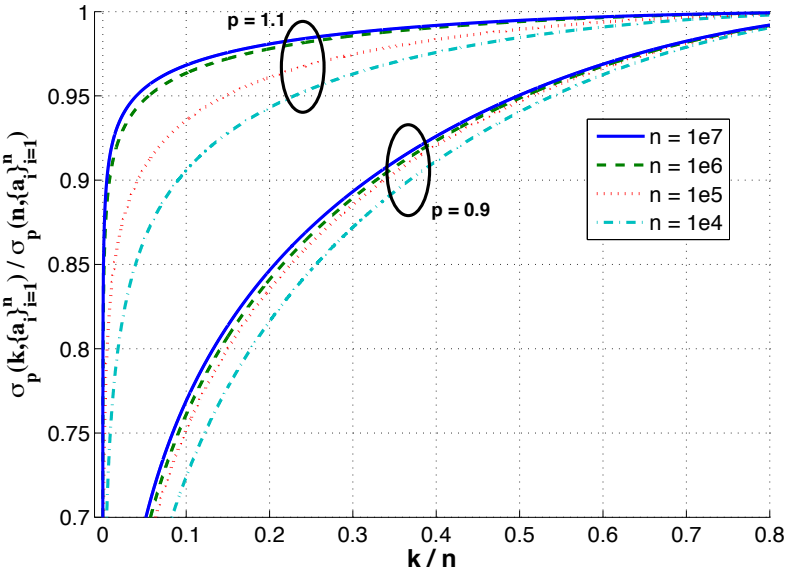
برای به تصویر کشیدن مفهوم فشرده‌پذیری در دنباله‌های تصادفی، مقادیر متوسط $\frac{\sigma_۱(k;\{a_i\}_{i=۱}^n)}{\sigma_۱(n;\{a_i\}_{i=۱}^n)}$ در ۵۰۰ تحقق متفاوت را بر حسب k و برای توزیع احتمال‌های مختلف در شکل ۶-۲ رسم کرده‌ایم. برای این



شکل ۶-۲: نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}$ بر حسب $\frac{k}{n}$ برای تحقق‌های i.i.d. از توزیع احتمال‌های گوسی، لاپلاس، کوشی و Student's t ($q = 0/5$) به ازای $p = 1$ و $n = 10^4$.

نتایج، طول دنباله $n = 10^4$ قرارداده شده است. قضیه ۶-۳ بیان می‌کند که توزیع Student's t با $q = 0/5$ از نوع ℓ_1-فشرده‌پذیر است و همان‌گونه که در شکل ۶-۲ دیده می‌شود، نمودار در مقادیر خیلی کوچک $\frac{k}{n}$ اشباع می‌شود. فشرده‌پذیری دنباله ایجاب می‌کند که در nهای بزرگ این نمودار به سمت تابع پله میل می‌کند. در نقطه مقابل، توزیع‌های گوسی و لاپلاس فشرده ناپذیرند و در نتیجه نمودار حدی آن‌ها هنگامی که $n \to \infty$ فرق چندانی با نمودارهای نمایش داده شده در شکل ۶-۲ ندارند. در این مثال توزیع کوشی در مرز ℓ_1-فشرده‌پذیری قرار دارد؛ از آن‌جا که میرایی این توزیع از مرتبه t^{-2} است، برای هر $p > 1$، ℓ_1-فشرده‌پذیر است اما در مورد $p = 1$ با استفاده از قضایای مطرح شده، نتیجه‌ای نمی‌توان گرفت.

برای بررسی بیشتر اثر p در ℓ_p-فشرده‌پذیری، مقدار متوسط کسر انرژی را برای توزیع کوشی با دو نرم $\ell_{0/9}$ و $\ell_{1/1}$ در شکل ۶-۳ رسم کرده‌ایم. همانطور که قضیه ۶-۳ پیش‌بینی می‌کند، نمودارهای مربوط به نرم $\ell_{1/1}$ با افزایش n به سمت تابع پله میل می‌کنند ولی در حالت $\ell_{0/9}$ چنین به نظر نمی‌رسد.



شکل ۶-۳: نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}$ بر حسب $\frac{k}{n}$ در nهای مختلف برای تحقق‌های i.i.d. از توزیع احتمال کوشی هنگامی که $p = 0/9$ و $p = 1/1$ مورد نظر باشند.