پایداری توزیعهای حدی

صفیه محمودی[1]

چکیده

خانوادهٔ توزیعهای پایدار از جمله توزیعهای مهم آماری هستند که خواص بسیار جالبی دارند. در این مقاله به معرفی این خانواده از توزیعها و توجه بـه آنها به عنوان توزیعهای حدی مجموعهای جزئی استاندارد شده، می‌پردازیم که بر اساس دنباله‌ای از متغیرهای مستقل و هم‌توزیع (iid) به دست آمده‌اند و چند توزیع مهم را که در خانواده این توزیعها قرار دارند؛ معرفی کرده و در انتها فرم تابع مشخصه و رفتار تـابع چگالی نظیر ایـن نـوع توزیعـها را بررسـی خواهیم کرد.

واژه‌های کلیدی: قضیه حد مرکزی، توزیع کشی، توزیع پایدار، تابع مشخصه، قضیه لوی.

۱. مقدمه

یکی از مشهورترین قضایا در آمار، قضیه حد مرکـزی اسـت کـه بـر اساس آن برای دنبالهٔ $\{X_i\}_{i=1}^{\infty}$ از متغیرهای مستقل و هم‌توزیع (iid) بـا میانگیـن و واریـانس مـتناهـی μ، σ^2 و $S_n = \sum_{i=1}^{n} X_i$، تـوزیـع متـغیر استاندارد شدهٔ $(S_n - n\mu)/\sigma\sqrt{n}$ با افزایش n، بـه سـمت توزیـع نرمال استاندارد میل می‌کند و آن را به صورت

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow N(0,1)$$

نشان می‌دهند. اثبات آن نیز با استفاده از تابع مشخصهٔ X_i ها بـه راحتـی ممکن است. در واقع در این مقاله به قضیـه فـوق از ایـن دیـدگـاه توجـه داریم کـه بـا متنـاهی بـودن واریـانس متغیرهـای iid دنبالـه‌ای از مقـادیر ثابت a_n و $b_n > 0$ وجود دارد که

$$\frac{S_n - a_n}{b_n} \Rightarrow N(0,1)$$

حال اگر واریانس متغیرها متناهی نباشد چه اتفاقی خواهد افتاد؟ آیـا در این صورت نیز دنباله‌ای از اعداد ثابت a_n $b_n > 0$ مناسبی می‌توان یـافت کـه $(S_n - a_n)/b_n$ دارای توزیـع حـدی باشـد؟ و آیـا توزیـع حدی باز هم نرمال خواهد بود؟

مثـال ۱: فـرض کنیـد X_i هـا دارای توزیـع کوشـی باشـند $(X_i \sim C(0,1))$ در نتیجـه تـابع مشـخصهٔ X_i هـا بـه فرم $\phi(t) = \exp(-|t|), t \in R$ خواهد بود، در نتیجه

$$\phi_{\frac{S_n}{n}}(t) = \phi_{S_n}\left(\frac{t}{n}\right) = \left(\phi\left(\frac{t}{n}\right)\right)^n = e^{-|t|} \quad , \ t \in R$$

---

[1] دانشجوی دکتری، بخش آمار، دانشگاه شیراز

و طبق قضیه پیوستگی توزیع حدی S_n/n همان کوشی $C(0,1)$ خواهد بود. البته با توجه به اینکه $E|X_i| = \infty$ ، قضیه حد مرکزی را نمی‌توان به کار برد ولی مقادیر a_n و $b_n > 0$ وجود دارند بطوری که $(S_n - a_n)/b_n$ در توزیع به توزیعی بجز نرمال میل کند.

حال موضوع جالب این است که در دنباله مستقل و هم‌توزیع $\{X_i\}_{i=1}^{\infty}$ ، اگر بدانیم برای مقادیر a_n و $b_n > 0$، $(S_n - a_n)/b_n \Rightarrow X$ ، انتظار داریم چه ویژگیهایی داشته باشد؟ البته فرض شده است که $b_n > 0$ و اگر برای برخی مقادیر n، $b_n < 0$ ، باشد می‌توان زیر دنباله‌ای از $\{b_n\}_{n=1}^{\infty}$ را در نظر گرفت که اعضای آن مثبت باشند و در صورتی که برای همه مقادیر n، b_n منفی باشد می‌توان مسأله را برای $-X_n$ بررسی کرد. به طور کلی در این مقاله به خواص چنین متغیرهایی که می‌توانند دارای توزیع حدی باشند می‌پردازیم.

۲. متغیرهای تصادفی پایدار

تعریف ۱: فرض کنید متغیرهای X_1, X_2 نمونه‌های مستقل از متغیر X باشند. اگر برای هر دو مقدار مثبت دلخواه C_1, C_2 مقادیر ثابت a و $b > 0$ (وابسته به C_1, C_2) وجود داشته باشند بطوری که

$$C_1X_1 + C_2X_2 \overset{D}{\equiv} bX + a \qquad (1)$$

در این صورت متغیر تصادفی X را دارای توزیع پایدار می‌نامند. در واقع در این صورت هر پیچش[3] (ترکیب) توزیع X بوسیلهٔ خود آن توزیع تولید می‌شود. یعنی

$$F\left(\frac{x}{c_1}\right)*F\left(\frac{x}{c_2}\right) = F\left(\frac{x-a}{b}\right) \quad ,\forall x\in R$$

بنابراین پایداری صفت مناسبی برای این مفهوم است و واضح است در صورتی که تابع مشخصهٔ X ، تابع $\phi(\cdot)$ باشد، رابطهٔ $\phi(c_1t)\phi(c_2t) = e^{ita}\phi(bt)$ به ازای هر $t\in R$ ، برقرار خواهد بود.

با توجه به رابطه (۱) به راحتی می‌توان نشان داد برای هر عدد طبیعی n ، مقادیر ثابت a_n و $b_n > 0$ وجود دارند بطوری که

[3] Convolution

$$S_n \overset{d}{\equiv} a_n + b_n X_1 \qquad (2)$$

یا

$$\phi^n(t) = e^{ita_n}\phi(b_n t) \qquad (3)$$

همچنین اگر X,Y دو متغیر پایدار مستقل باشند به سادگی پایداری هر ترکیب خطی آنها قابل اثبات است.

مثال ۲: به وضوح می‌توان دید توزیعهای تباهیده و همچنین توزیعهای نرمال پایدار هستند.

۳. پایداری توزیعهای حدی

مسأله‌ای که بر اساس قضیه حد مرکزی به آن پاسخ داده شد پیدا کردن شرایطی بود که مجموع جزئی متغیرهای *iid* دارای توزیع حدی نرمال باشد که مسأله‌ای بسیار قدیمی است. حال اگر مجموع‌های جزئی یک دنباله از متغیرهای *iid* بعد از استاندارد کردن دارای توزیع حدی باشد این توزیع چه خاصیتی دارد؟ این مسأله توسط لوی[4] در ۱۹۳۰ حل شد که در قضیه زیر آمده است. اثبات آن را می‌توانید در Rorat (1997) [۵] ملاحظه کنید.

قضیه ۱: فرض کنید $\{X_i\}_{i=1}^{\infty}$ دنباله‌ای از متغیرهای مستقل و هم‌توزیع باشد و $S_n = \sum_{i=1}^{n} X_i$ ، اگر دنباله‌ای از مقادیر ثابت a_n و $b_n > 0$ وجود داشته باشد بطوری که توزیع متغیرهای تصادفی $(S_n - a_n)/b_n$ به توزیع F میل کند آنگاه F یک توزیع پایدار خواهد بود و بالعکس، یعنی برای هر توزیع پایدار متغیرهای *iid* با شرایط فوق وجود دارد.

اگر X_i ها دارای توزیع G باشند اصطلاحا گوییم G جذب F شده است. به مجموعهٔ تمام توزیعهایی که می‌توانند جذب F شوند، دامنهٔ جذب[5] F گویند. با توجه به قضیه فوق تنها توابع توزیع پایدار دارای دامنهٔ جذب هستند.

با استفاده از قضیه فوق منطقی است که متغیرهای دارای توزیع پایدار را به صورت زیر نیز تعریف کنند که با تعریف ۱ معادل است.

[4] Levy

[5] Attraction domain

تعریف ۲: متغیر X را دارای توزیع پایدار گویند، اگر دارای دامنهٔ جذب باشد. یعنی اگر دنباله‌ای از متغیرهـای iid ، $\{X_i\}_{i=1}^{\infty}$ و دنبالـه‌ای از مقادیر ثابت a_n و $b_n > 0$ حقیقـی وجـود داشـته باشـد بطـوری کـه

اگر $S_n = \sum_{i=1}^{n} X_i$ آنگاه $(S_n - a_n)/b_n \Rightarrow X$.

با توجه به قضیهٔ حد مرکزی تمام توزیعـهایی کـه واریـانس متنـاهی دارند جزء دامنهٔ جذب توزیع نرمال هستند، بنابراین پایداری توزیع نرمال از این طریق نیز اثبات می‌شود. توزیـع نرمـال در ایـن خـانواده خاصیت جالب دیگری نیز دارد.

قضیهٔ ۲: توزیع نرمال تنها توزیع پایداری است کـه دارای واریـانس متناهی است.

اثبات: فـرض کنیـد $X_1, X_2, \ldots$ متغیـرهـای iid بـا توزیـع پایـداری بـاشند که دارای واریانس متناهی است، بـدون از دسـت دادن کلیـت مسـألـه می‌توان فـرض کرد $E(X_i) = 0$ و $V(X_i) = 1$. در واقـــع بـــا تغیـــیر متغـــیری بـــه صـــورت $X_i' = \left(X_i - E(X_i)\right)/\sqrt{V(X_i)}$ کـه توزیـع X_i' نـیز پـایدار خواهـد بود، کـار می‌کنیم. حال بـا توجـه بـه رابـطهٔ (۲) برای هـر n طبیـعی، مـقادیر ثـابت a_n و $b_n > 0$ وجـود دارنـد بـطوری کـه $S_n \overset{d}{\equiv} a_n + b_n X_1$ ، و با توجه به اینکـه امیـد ریـاضی و واریـانس دو طرف رابطه برابرند، $b_n^2 = n$ و $a_n = 0$ در نتیجـه $X_1 \overset{d}{\equiv} S_n/\sqrt{n}$. از طرفی با توجه به قضیهٔ حـد مرکـزی، $S_n/\sqrt{n}$ در توزیـع بـه نرمـال استاندارد میل می‌کند. بنابراین X_1 دارای توزیع نرمال استاندارد است.

۴. تابع مشخصه و رفتار چگالی احتمـال توزیعـهای پایدار

قبلا ذکر شد که یک متغیر تصادفی دارای توزیع پـایدار است اگر برای هر c_1, c_2 مثبـت، مقـادیر ثـابت a و $b > 0$ وجـود داشـته باشـد بطوری که

$$\phi(c_1 t)\phi(c_2 t) = e^{ita}\phi(bt).$$

ثابت شده است که تابع $\phi(t)$ دارای خاصیت فوق است اگر و فقط اگر پارامترهای $\mu, \sigma, \beta, \alpha$ وجود داشته باشند بطوری که

$$\phi(t) = e^{-\sigma^\alpha |t|^\alpha \Psi_\alpha(t) + i\mu t}$$

که $\mu \in R$ ، $\sigma \geq 0$ ، $\beta \in [-1,1]$ و $\Psi_\alpha(t)$ به فرم زیر است

$$\Psi_\alpha(t) = \begin{cases} 1 + i\,\mathrm{sgn}(t)\beta \frac{2}{\pi}\log|t| & , \alpha = 1 \\ 1 + i\,\mathrm{sgn}(t)\beta \tan\frac{\pi\alpha}{2} & , \alpha \neq 1 \end{cases}$$

همان طور که انتظار می‌رفت یک تابع توزیع پایدار به چهار پـارامتر بستگی دارد و به همـین دلیـل پایدار بودن توزیع X با این چهار پـارامتر را با نماد $X \sim S_\alpha(\sigma, \beta, \mu)$ نشـان می‌دهنـد. بـا توجـه بـه فـرم تـابع مشـخصـه مـی‌تـوان دیـد μ پـارامتر مکـان[6] اسـت. بدیـن معنـا کـه $X - \mu \sim S_\alpha(\sigma, \beta, 0)$ اسـت و بـرای حالـت $\alpha \neq 1$ ، σ پـارامتر مقیـــاس اســـت. بدیـــن دلیـــل کـــه در ایـــن حـــالت $(X - \mu)/\sigma \sim S_\alpha(1, \beta, 0)$ ، β را پارامتر چولگی می‌نامند. البتـه در اکثر موارد نمی‌توان فرم دقیـق تـابع چگـالی را بدسـت آورد ولـی بـا شبیه‌سازی می‌توان تأثیر این پارامتر را بر چولگی تابع چگالی به وضـوح دید. پـارامتر α را نـیز پـارامتر پایـداری می‌نـامند و متغیر X را $\alpha-$ پایدار[7] می‌نامند.

مثال ۳: در صورتی که $\alpha = 2$ باشد فرم تابع مشخصه به صورت

$$\phi(t) = e^{-\sigma^2 t^2 + i\mu t}$$

خواهد بود. بنابراین توزیع پایدار با $\alpha = 2$ همان توزیع نرمال است و بـا توجه به اینکه چولگی نرمال صفـر اسـت β برابـر صفـر در نظـر گرفتـه می‌شود یعنی $S_2(\mu, 0, \sigma^2) = N(\mu, 2\sigma^2)$.

مثال ۴: در صورتی که $\alpha = 1$ و $\beta = 0$ آنگاه

$$\phi(t) = e^{-\sigma^\alpha |t| + i\mu t} \quad , t \in R$$

که تابع مشخصهٔ توزیع کوشی با پـارامتر مکـان μ و پـارامتر مقیـاس σ است. یعنی $X \sim C(\mu, \sigma)$ با تابع چگالی

$$f(x) = \frac{\sigma}{\pi(\sigma^2 + (x - \mu)^2)} \quad , x \in R$$

---

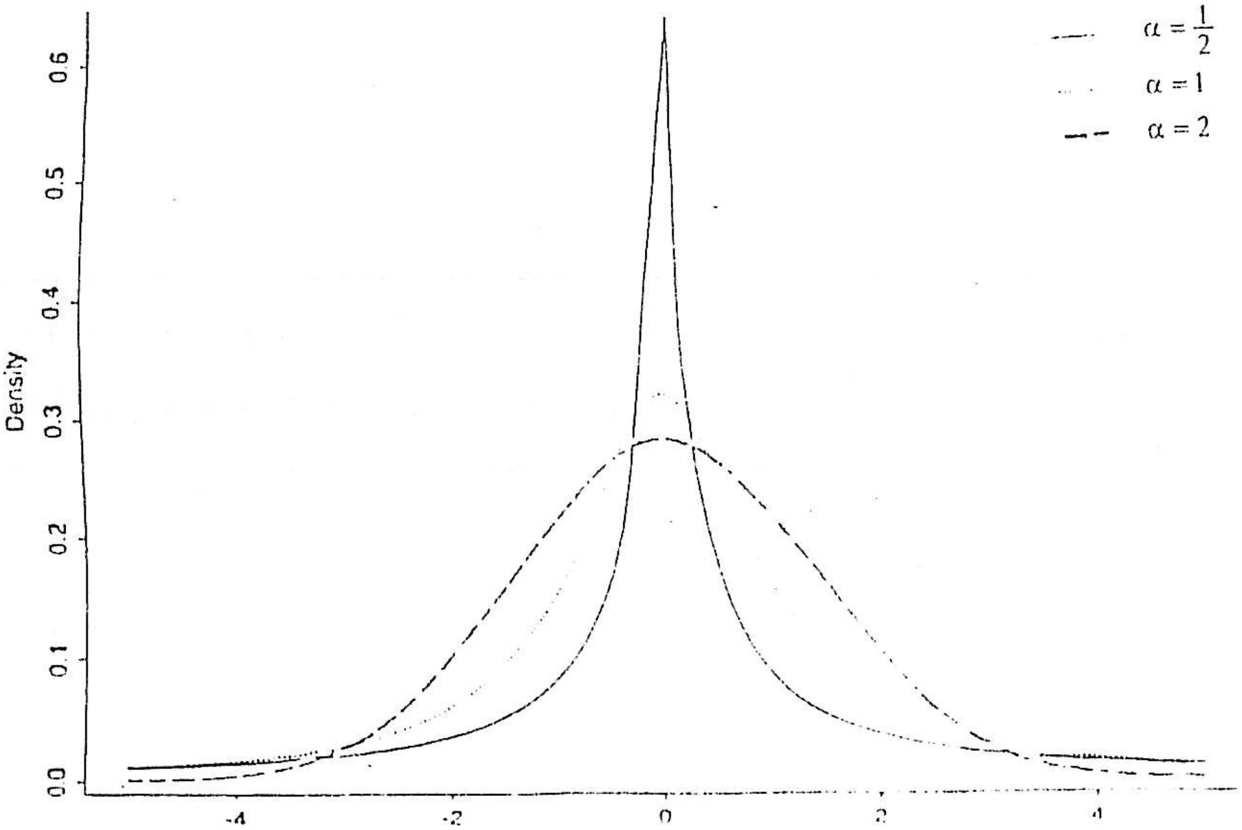
[6] shift parameter

[7] α - stable

$$f(x) = \frac{\sigma}{\pi(\sigma^{2}+(x-\mu)^{2})} \quad ، x \in R$$

بنابراین کوشی و نرمال جزو خانوادهٔ توزیعهای پایدار هستند. به طور کلی چون تابع مشخصهٔ متغیرهای پایدار مطلقا انتگرال‌پذیر است، طبق فرمول انعکاس[8] توابع، نظیر آنها دارای چگالیهای پیوسته هستند. همچنین چگالیهای پایدار همگی به صورت تک مدی هستند و دمهای چگالیهای پایدار برای $\alpha < 2$ به صورت تابع توانی با توان $-\alpha$ نزول می‌کند و البته نسبت به توزیع نرمال احتمال بیشتری در دمها قرار دارداین بدان معناست که میزان تغییرات از نرمال بیشتر است و احتمال اینکه متغیرها از میانه دور باشند بیشتر خواهد بود. تکیه‌گاه این نوع متغیرها تمام اعداد حقیقی یا نیم‌خط $x \geq \alpha$ یا $x \leq \alpha$ است. مثلا تکیه‌گاه متغیری با توزیع پایدار با پارامترهای $\alpha = 1/2$ ، $\beta = 1$ و $\mu = 0$ نیم‌خط $x \geq 0$ است.

نمودار چند چگالی پایدار با $\beta = 0$ و چند α متفاوت



[8] Inversion formula

مراجع

[1] Ash, R.B., 1972. *Real Analysis and Probability*, Academic Press, New York, pp.344-348.

[2] Billingsley, P., 1979. *Probability and Measures,* John Wiley & Sons, New York, pp.308-309.

[3] Petrov, V.V., 1975. *Sums of Independent Random Variables,* Translated from Russian by Brown, A.A., Springer-verlag, New York, pp.82-103.

[4] Prohorov, Y.V. and Rozanov, Y.A., 1969. *Probability Theory,* Translated by Kricheberg, K. and Urmitzer, H., pp.186-190.

[5] Rotar, V., 1997. *Probability Theory,* World Scientific, London, pp.318-314.

[6] Samorodinisky, G. and Taqqu, M.S., 1994. *Stable Non-Gaussian Random Processes,* Chapman & Hall, New York, pp.1-50.

آمار، مطالعه لذت‌بخشی است در باب این موضوع که چگونه می‌توان جهان ناشناخته را با گشودن دریچه بر روی آن توصیف کرد.