

## مقایسه توزیع‌های دُم‌پهن با استفاده از ورنترویی

فرانک گودرزی

### چکیده

در این مقاله، اطلاع رنی از مرتبه  $\lambda$  برای یک متغیر تصادفی پیوسته و ارتباط آن با ورنترویی مطالعه می‌شود. همچنین با به‌کارگرفتن ورنترویی به‌عنوان اندازه‌شکل یک توزیع، می‌توان آن را برای مقایسه دُم‌ها و شکل‌های گوناگون چگالی‌هایی که غالباً استفاده می‌شوند، به‌کار برد در حالی که اندازه‌کشیدگی معمول قابل اجرا نیست. در میان این توزیع‌ها می‌توان به بسیاری از توزیع‌های دُم‌پهن مانند توزیع تی استودنت با هر درجه آزادی، کُشی، کرامر، لوی و بر اشاره کرد. همچنین در حالت چندمتغیره می‌توان به توزیع‌های نرمال، تی و دیریکله اشاره نمود.

### ۱. سرآغاز

اندازه آنتروپی شانون به‌عنوان زیربنای نظریه اطلاع، از جایگاه ویژه‌ای در علوم مختلف از قبیل فیزیک، آمار، احتمال، اقتصاد، بیمه، نظریه ارتباطات، مهندسی آب، زیست‌شناسی و زلزله‌شناسی برخوردار است. این اندازه، نخستین بار توسط کلود شانون<sup>۱</sup> مطرح شد و اندازه‌ای از عدم‌حتمیت یا متوسط مقدار اطلاع از یک متغیر تصادفی را نشان می‌داد. آنتروپی شانون برای متغیرهای تصادفی پیوسته به‌صورت زیر تعریف می‌شود که در این حالت، آن را آنتروپی تفاضلی می‌نامند:

---

عبارات و کلمات کلیدی. دُم‌های پهن؛ کشیدگی؛ اطلاع رنی؛ ورنترویی.

<sup>۱</sup>Claude Shannon

تعریف ۱.۱. اگر  $X$  یک متغیر تصادفی پیوسته با تابع چگالی احتمال  $f(x)$  باشد، آنگاه آنترپوی آن عبارت است از

$$h(X) = - \int_{-\infty}^{\infty} f(x) \ln f(x) dx = -E[\ln f(X)]. \quad (1.1)$$

بدیهی است که مقدار آنترپوی تفاضلی می‌تواند منفی نیز باشد.

اگر  $X$  طول عمر یک دستگاه باشد، آنگاه می‌توان از  $h(X)$  برای اندازه‌گیری عدم حتمیت آن استفاده کرد. یک مسئله متداول برای مطالعهٔ برآورد انحراف محتوای اطلاع از آنترپوی، از طریق واریانس آنترپوی است که به اختصار و آنترپوی<sup>۱</sup> خوانده می‌شود. در واقع، و آنترپوی میانگین توان دوم انحراف محتوای اطلاع  $\tilde{h}(x) = (-\ln f(x))$  از آنترپوی  $h(X)$  است.

تعریف ۲.۱. فرض کنیم  $X$  یک متغیر تصادفی با مقادیر حقیقی و تابع چگالی  $f$  باشد. در این صورت، و آنترپوی عبارت است از

$$\text{Var}[-\ln f(X)] = \int_{-\infty}^{\infty} f(x) (\ln f(x))^2 dx - \left( \int_{-\infty}^{\infty} f(x) \ln f(x) dx \right)^2. \quad (2.1)$$

بدیهی است که تعریف‌های بالا را می‌توان برای حالت چندمتغیره نیز تعمیم داد. توجه کنید همان‌طور که در تعریف‌های بالا ملاحظه می‌شود، آنترپوی و و آنترپوی به تحقق  $X$  وابسته نیستند و تنها به چگالی  $f$  بستگی دارند. بنابراین اگر  $\mathbf{X}$  یک بردار تصادفی با مقادیری در  $\mathbb{R}^n$  و با تابع چگالی  $f$  باشد، آنگاه

$$h(\mathbf{X}) = - \int_{\mathbb{R}^n} f \ln f, \quad V(\mathbf{X}) = \text{Var}(\ln f(\mathbf{X})) = \int_{\mathbb{R}^n} f (\ln f)^2 - \left( \int_{\mathbb{R}^n} f \ln f \right)^2.$$

بابکو و مادیمان [۳] نشان دادند که توزیع اختلاف  $\tilde{h}(\mathbf{X}) - h(\mathbf{X})$  تحت هر تبدیل آفین  $\mathbb{R}^n$  ناوردا است (یعنی  $\tilde{h}(\mathbf{TX}) - h(\mathbf{TX}) = \tilde{h}(\mathbf{X}) - h(\mathbf{X})$  برای هر نگاشت آفین وارون‌پذیر  $(\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n)$ ). بنابراین و آنترپوی  $V(\mathbf{X})$  آفین ناوردا است، حال آنکه آنترپوی تحت تبدیلات آفین ناوردا نخواهد بود. نوع دیگر ناوردایی برای هر دوی  $h(\mathbf{X})$  و  $V(\mathbf{X})$  از این حقیقت ناشی می‌شود که آنها تنها به توزیع  $\ln f(\mathbf{X})$  بستگی دارند به طوری که اگر  $f$  طوری تعدیل شود که مجموعه‌های زیرسطحی حجم یکسانی را حفظ کنند، این مقادیر بدون تغییر باقی می‌مانند. این نشان می‌دهد که اگر  $f^*$  تجدید آرایش کاهشی به‌طور کروی متقارن از  $f$  باشد و  $\mathbf{X}^*$  بر طبق چگالی  $f^*$  توزیع شود، آنگاه

$$h(\mathbf{X}) = h(\mathbf{X}^*), \quad V(\mathbf{X}) = V(\mathbf{X}^*).$$

<sup>۱</sup>varentropy

ناوردایی آنتروپی ناشی از تجدید آرایش، یک عنصر کلیدی در گسترش نامساوی‌های توانی آنتروپی نظریف‌شده در وانگ و مادیمان [۱۷] است.

قبل از بیان مطالب اصلی این مقاله، نیاز به تعریف برخی اصطلاحات داریم.

**تعریف ۳.۱.** فرض کنیم متغیر تصادفی  $X$  دارای تابع توزیع  $F$  با پارامتر  $\mu$ ،  $X \sim F(x; \mu)$  باشد. در این صورت،  $\mu$  یک پارامتر مکان است اگر  $Z = X - \mu \sim F_0(z)$  که در آن،  $F_0$  مستقل از  $\mu$  است؛ یعنی  $F(x; \mu) = F_0(x - \mu)$ .

**تعریف ۴.۱.** فرض کنیم  $X \sim F(x; \sigma)$ . در این صورت،  $\sigma$  یک پارامتر مقیاس است اگر

$$Z = X/\sigma \sim F_0(z)$$

که در آن،  $F_0$  مستقل از  $\sigma$  است؛ یعنی  $F(x; \sigma) = F_0(x/\sigma)$ .

**تعریف ۵.۱.** فرض کنیم  $X$  یک متغیر تصادفی با تابع توزیع  $F$  باشد. توزیع  $F$  را روی  $\mathbb{R}$  دم‌پهن<sup>۱</sup> گوئیم اگر برای هر  $t > 0$ ،  $\int_{-\infty}^{\infty} e^{tx} dF(x) = \infty$ .

فلز [۵] توزیع‌های پایدار را به صورت زیر معرفی کرد.

**تعریف ۶.۱.** اگر  $X, X_1, X_2, \dots$  دنباله‌ای از متغیرهای تصادفی مستقل با تابع توزیع  $F$  باشد و  $S_n = X_1 + \dots + X_n$ ، گوئیم توزیع ناتباهیده<sup>۲</sup>  $F$  پایدار (اکیداً پایدار) است اگر برای هر  $n \in \mathbb{N}$  عددهای  $c_n > 0$  و  $\gamma_n$  وجود داشته باشند به طوری که  $S_n \stackrel{d}{=} c_n X + \gamma_n$ .

علاوه بر تعریف بالا، تعریف‌های دیگری نیز برای توزیع پایدار ارائه شده است که از جمله آنها می‌توان به تعریف زیر که در نولان [۱۳] ذکر شده است، اشاره کرد.

**تعریف ۷.۱.** متغیر تصادفی  $X$  دارای توزیع پایدار است اگر و تنها اگر تابع مشخصه آن به صورت

$$E \exp(itX) = \begin{cases} \exp(-\sigma^\alpha |t|^\alpha [1 - i\beta \tan \frac{\pi\alpha}{2} (\text{sign}(t))]) + i\mu t & \alpha \neq 1 \\ \exp(-\sigma |t| [1 + i\beta \frac{1}{\pi} (\text{sign}(t)) \log |t|] + i\mu t) & \alpha = 1 \end{cases} \quad (3.1)$$

باشد که در آن،  $0 < \alpha \leq 2$  شاخص پایداری یا نمای مشخصه متغیر تصادفی  $X$ ،  $-1 \leq \beta \leq 1$  پارامتر چولگی،  $\sigma \in \mathbb{R}^+$  پارامتر مقیاس و  $\mu \in \mathbb{R}$  پارامتر مکان است.

<sup>۱</sup>heavy tailed

نولان [۱۳] بیان کرد که اگر  $X$  داری توزیع پایدار با پارامترهای  $\alpha, \beta, \mu$  و  $\sigma$  باشد به طوری که  $0 < \alpha < 2$  و  $0 < \beta \leq 1$ ، آن‌گاه

$$P(X > x) \sim \sigma^\alpha c_\alpha (1 + \beta)x^{-\alpha}, \quad x \rightarrow \infty$$

که در آن،  $c_\alpha = \sin(\frac{\pi\alpha}{2})\Gamma(\alpha)/\pi$ . همچنین او بیان کرد که برای  $1 > \beta \geq -1$ ،

$$P(X < -x) \sim \sigma^\alpha c_\alpha (1 - \beta)x^{-\alpha}, \quad x \rightarrow \infty.$$

## ۲. نتایج اصلی

اطلاع رنی از مرتبه  $\lambda$  برای یک متغیر تصادفی پیوسته با تابع چگالی  $f(x)$  به صورت

$$\mathcal{I}_R(\lambda) = (1 - \lambda)^{-1} \ln \left( \int_{\mathbb{R}} f^\lambda(x) dx \right)$$

برای  $\lambda \neq 1$  تعریف می‌شود، به طوری که

$$\lim_{\lambda \rightarrow 1} \mathcal{I}_R(\lambda) = - \int_{\mathbb{R}} f(x) \ln f(x) dx.$$

ویژگی‌های آن‌تروپی رنی توسط بسیاری از پژوهشگران از جمله رنی [۱۴]، مورالس و همکاران [۱۵]، سانگ [۱۵]، ناداراجا و زگرافوس [۱۱، ۱۲]، برچر [۱] و گلشنی و همکاران [۶] مطالعه شده‌اند. به عنوان تابعی از  $\lambda$ ،  $\mathcal{I}_R(\lambda)$  طیف اطلاع رنی<sup>۱</sup> نامیده می‌شود و با رسم نمودار آن نسبت به  $\lambda$  اطلاعات سودمندی به دست می‌آید. در کاربردها علاقه زیادی به استفاده از طیف متناظر با برخی مقادیر خاص  $\lambda$  وجود دارد. برای مثال، نسخه گسسته اطلاع رنی از مرتبه ۲ غالباً به عنوان اندازه تنوع اقتصادی<sup>۲</sup> [۸] مورد استفاده قرار می‌گیرد. این اندازه یکی از پُر کاربردترین اندازه‌های عدم حتمیت یا اطلاع در علوم است که از نظر آماری از دیدگاه لگاریتم درست‌نمایی قابل توجیه است، زیرا برابر با منفی مقدار مورد انتظار لگاریتم درست‌نمایی است. طیف، نه تنها اطلاعات سودمند را منتقل می‌سازد، بلکه با استفاده از گرادیان طیف،  $\mathcal{I}_R(\lambda)$  می‌توان لگاریتم درست‌نمایی و اطلاع رنی را به یکدیگر پیوند داد. در واقع، گرادیان طیف را در  $\lambda = 1$  مقدار یابی می‌کنیم؛ در صورتی که کمیت‌های مورد استفاده نظیر  $\int f^\lambda(x) dx$  خوش تعریف و عملگرهای مشتق مجاز باشند، آن‌گاه با مشتق‌گیری از  $\mathcal{I}_R(\lambda)$  نسبت به  $\lambda$  و سپس قرار دادن  $\lambda = 1$

<sup>۱</sup>spectrum <sup>۲</sup>economic diversity

به دست می‌آوریم

$$\begin{aligned}
\mathcal{J}_R(1) &= \lim_{\lambda \rightarrow 1} \frac{(\lambda - 1)(\int f^\lambda(x) dx)^{-1} \int f^\lambda(x) \ln f(x) dx + \ln \int f^\lambda(x) dx}{(\lambda - 1)^2} \\
&= \lim_{\lambda \rightarrow 1} \left( -\frac{\int f^\lambda(x) \ln f(x) dx}{\int f^\lambda(x) dx} \right. \\
&\quad \left. + (\lambda - 1) \frac{(\int f^\lambda(x) (\ln f(x))^2 dx)(\int f^\lambda(x) dx) - (\int f^\lambda(x) \ln f(x) dx)^2}{(\int f^\lambda(x) dx)^2} \right. \\
&\quad \left. + \frac{\int f^\lambda(x) \ln f(x) dx}{\int f^\lambda(x) dx} \right) / \{-2(\lambda - 1)\} \\
&= -\frac{1}{2} \lim_{\lambda \rightarrow 1} \left\{ \frac{\int f^\lambda(x) (\ln f(x))^2 dx}{\int f^\lambda(x) dx} - \left( \frac{\int f^\lambda(x) \ln f(x) dx}{\int f^\lambda(x) dx} \right)^2 \right\} \\
&= -\frac{1}{2} \text{Var}[\ln f(X)].
\end{aligned}$$

اگر آنتروپی را اندازه احتمال تمرکز یا پراکندگی بدانیم، آن‌گاه وراثتروپی  $\mathcal{J}_R(1) = -2\mathcal{L}_f$  شکل ذاتی یک توزیع را اندازه می‌گیرد [۲]. این موضوع با مشاهده اینکه  $\mathcal{L}_f$  تابعی مثبت است که به پارامترهای مکان و مقیاس وابسته نیست، قابل بررسی است؛ یعنی  $\mathcal{L}_f = \mathcal{L}_g$  که در آن،  $f(x) = \sigma^{-1}g((x - \mu)/\sigma)$

اگر گشتاور چهارم  $f(x)$  متناهی باشد، وراثتروپی در مقایسه شکل‌های چگالی‌های گوناگونی که غالباً استفاده می‌شوند، ضمن آنکه نقشی شبیه به اندازه کشیدگی ( $B_2 = \frac{\mu_4}{\sigma^4}$ ) که در آن،  $\mu_4$  چهارمین گشتاور مرکزی است) بازی می‌کند، بلکه سنگینی دم‌های توزیع‌ها را نیز اندازه می‌گیرد. پس وراثتروپی بیش از آنچه را که کشیدگی قادر به اندازه‌گیری آن باشد، اندازه می‌گیرد. یکی از مشکلات به‌کارگیری اندازه کشیدگی این است که در بسیاری از موقعیت‌ها وقتی دم‌های توزیع پهن هستند، گشتاور چهارم ممکن است موجود نباشد یا حتی بدتر، ممکن است میانگین موجود نباشد. یک مثال نوعی و جالب توجه، توزیع  $t$  با  $v \leq 4$  درجه آزادی است به طوری که اندازه کشیدگی که اغلب برای اندازه‌گیری سنگینی دم‌ها تصور می‌شود، در این موقعیت مناسب نیست. اندازه شکل  $\mathcal{L}_f$  تقریباً همواره در بیشتر شرایطی که در عمل رخ می‌دهد، مناسب است و محدودیت‌هایی از قبیل فرض تقارن را نیاز ندارد و می‌تواند برای توزیع‌هایی با دم‌های پهن نظیر کُشی، کرامر، لوی یا پارتو به‌کار رود. برای اینکه بتوانیم اندازه  $\mathcal{L}_f$  را برای یک خانواده پارامتری تعیین کنیم، لازم است که آنتروپی و گشتاور دوم لگاریتم درستنمایی را محاسبه کنیم. این کار را می‌توان با معرفی تابع مولد اطلاع  $\mathcal{S}(\lambda) = \int f^\lambda(x) dx$  و دو بار مشتق‌گیری از آن نسبت به  $\lambda$  و سپس تعیین حد وقتی  $\lambda \rightarrow 1$  انجام داد.

مثال ۱.۲. توزیع  $t$  با  $v > 0$  درجه آزادی را که دارای تابع چگالی احتمال

$$f(x) = \frac{1}{(v\pi)^{1/2}} \frac{\Gamma(v/2 + \frac{1}{2})}{\Gamma(v/2)} \left(1 + \frac{x^2}{v}\right)^{-(v+1)/2}, \quad -\infty < x < \infty$$

است، در نظر بگیرید. می‌دانیم که گشتاور  $n$ ام تنها برای  $v > n$  موجود است و برای  $v = 1$ ، توزیع  $t$  بر توزیع کُشی منطبق است به طوری که حتی میانگین وجود ندارد. برای  $\lambda > \frac{1}{1+v}$  تابع مولد اطلاع توزیع  $t$  برابر است با

$$\begin{aligned} \mathcal{J}(\lambda) &= \int_{-\infty}^{\infty} \frac{1}{(v\pi)^{\lambda/2}} \frac{\Gamma^\lambda(v/2 + \frac{1}{2})}{\Gamma^\lambda(v/2)} \left(1 + \frac{x^2}{v}\right)^{-(v+1)\lambda/2} dx \\ &= (\pi v)^{(1-\lambda)/2} \left\{ \frac{\Gamma(v/2 + \frac{1}{2})}{\Gamma(v/2)} \right\}^\lambda \frac{\Gamma(v\lambda/2 + \lambda/2 - \frac{1}{2})}{\Gamma(v\lambda/2 + \frac{1}{2})} \end{aligned}$$

و طیف اطلاع رنی مساوی است با

$$\mathcal{J}_R(\lambda) = \frac{1}{1-\lambda} \ln \left\{ \frac{B(v\lambda/2 + \lambda/2 - \frac{1}{2}, \frac{1}{2})}{B^\lambda(v/2, \frac{1}{2})} \right\} + \frac{1}{2} \ln(v)$$

که در آن،  $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$  تابع بتا است و مشتق طیف عبارت است از

$$\begin{aligned} \mathcal{J}'_R(\lambda) &= (\ln\{B(v\lambda/2 + \lambda/2 - \frac{1}{2}, \frac{1}{2})/B(v/2, \frac{1}{2})\}) \\ &\quad + \frac{1}{2}(1-\lambda)(v+1)\{\psi(v\lambda/2 + \lambda/2 - \frac{1}{2}) \\ &\quad - \psi(v\lambda/2 + \lambda/2)\}/(1-\lambda)^2 \end{aligned}$$

که در آن،  $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$  تابع دو گاما است. بنابراین

$$\begin{aligned} \mathcal{L}_f(v) &= \lim_{\lambda \rightarrow 1} \frac{1}{2}(v+1)^2 \left\{ \psi\left(\frac{v\lambda}{2} + \frac{\lambda}{2} - \frac{1}{2}\right) - \psi\left(\frac{v\lambda}{2} + \frac{\lambda}{2}\right) \right\} \\ &= \frac{1}{2}(v+1)^2 \left\{ \psi\left(\frac{v}{2}\right) - \psi\left(\frac{v+1}{2}\right) \right\} \end{aligned} \quad (1.2)$$

که در آن،  $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$  تابع سه گاما است. به ویژه برای توزیع کُشی داریم

$$\mathcal{L}_f(1) = \psi\left(\frac{1}{2}\right) - \psi(1) = 2\zeta(2) = \pi^2/3$$

که در آن،  $\zeta(s)$  تابع زتای ریمان است. برای مطالعه جزئیات بیشتر، به [۷] مراجعه کنید. به‌طور مشابه برای  $v = 2, 3, 4, 5$  به‌دست می‌آوریم

$$\mathcal{L}_f(2) = 9 - \frac{3}{4}\pi^2 \text{ (میانگین وجود دارد، اما واریانس موجود نیست)},$$

$$\mathcal{L}_f(3) = \frac{4}{3}\pi^2 - 12 \text{ (میانگین و واریانس موجودند)},$$

$$\mathcal{L}_f(4) = \frac{775}{36} - \frac{25}{12}\pi^2 \text{ (سه گشتاور نخست متناهی هستند)},$$

$$\mathcal{L}_f(5) = 3\pi^2 - \frac{115}{4} \text{ (چهارگشتاور نخست متناهی هستند)}.$$

جالب است که اندازه  $\mathcal{L}_f(v)$  با افزایش  $v$  کاهش می‌یابد:

$$\mathcal{L}_f(1) \approx 3,2899, \mathcal{L}_f(2) \approx 1,5978, \mathcal{L}_f(3) \approx 1,1595,$$

$$\mathcal{L}_f(4) \approx 0,9661, \mathcal{L}_f(5) \approx 0,8588$$

که منطقی است، زیرا وقتی  $v$  افزایش می‌یابد، دُم‌ها باریکتر می‌شوند. در واقع، با استفاده از فرمول‌های مجانبی برای تابع سه گاما می‌توان نشان داد که  $\lim_{v \rightarrow \infty} \mathcal{L}_f(v) = \frac{1}{4}$  که متناظر با اندازه  $\mathcal{L}_f$  برای توزیع نرمال است. توجه کنید که اگر اندازه کشیدگی را به‌کار ببریم، حتی نمی‌توان کُشی را با مثلاً توزیع لاپلاس مقایسه کرد، چون کشیدگی برای چگالی کُشی وجود ندارد. همچنین با استفاده از ملاک کشیدگی، باید ادعا کرد که توزیع لاپلاس و توزیع  $t$  استودنت با  $v = 6$  دارای سنگینی یکسانی برحسب دُم‌ها هستند، زیرا برای هر دو توزیع  $B_2 = 6$  اما

$$\mathcal{L}_f(6) = \frac{49}{4}\{\psi(3) - \psi(7/2)\} = \frac{147931}{3600} - \frac{49\pi^2}{12} \approx 0,79106.$$

پس از انجام محاسبات ساده جبری می‌توان نشان داد که برای توزیع لاپلاس  $\mathcal{L}_f = 1$  و می‌توان نتیجه گرفت که توزیع لاپلاس در مقایسه با توزیع  $t$  با درجه آزادی ۶ دارای دُم پهن‌تری است.

مثال ۲.۲. فرض کنیم متغیر تصادفی  $X$  دارای توزیع کرامر با تابع چگالی زیر باشد:

$$f(x) = \frac{\theta}{2(1 + \theta|x|)^2}, \quad -\infty < x < \infty, \quad \theta > 0.$$

کشیدگی برای توزیع کرامر وجود ندارد. با استفاده از تقارن و تغییر متغیر، طیف اطلاع رنی برابر است با

$$\begin{aligned} \mathcal{I}_R(\lambda) &= \frac{1}{1-\lambda} \ln \left( \int_{-\infty}^{\infty} \frac{\theta^\lambda}{\sqrt{\lambda}(1+\theta|x|)^{\sqrt{\lambda}}} dx \right) \\ &= \frac{1}{1-\lambda} \ln \left( \int_0^{\infty} \frac{\theta^\lambda}{\sqrt{\lambda-1}(1+\theta y)^{\sqrt{\lambda}}} dy \right) \\ &= \frac{1}{1-\lambda} \ln \left\{ \frac{\theta^\lambda}{\sqrt{\lambda-1} \cdot \theta(\sqrt{\lambda-1})} \right\} \\ &= \ln \left( \frac{\sqrt{\lambda}}{\theta} \right) - \frac{\ln(\sqrt{\lambda-1})}{1-\lambda} \end{aligned}$$

و مشتق طیف به صورت زیر است:

$$\dot{\mathcal{I}}_R(\lambda) = \frac{\sqrt{\lambda-1}(\sqrt{\lambda-1})^{-1} - \ln(\sqrt{\lambda-1})}{(1-\lambda)^2}$$

به طوری که  $\mathcal{L}_f = \lim_{\lambda \rightarrow 1} \frac{\dot{\mathcal{I}}_R(\lambda)}{(\sqrt{\lambda-1})^2} = 4$  از این رو توزیع کرامر پهن تر از توزیع گشی است.

مثال سوم دربارهٔ توزیع لوی است که مساوی با توزیع اولین زمانی است که حرکت براونی استاندارد

به تراز  $a > 0$  می‌رسد.

**مثال ۳.۲.** فرض کنیم متغیر تصادفی  $X$  دارای توزیع لوی با تابع چگالی زیر باشد:

$$f(x) = \left( \frac{\sigma}{\sqrt{\pi}} \right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp \left\{ -\frac{\sigma}{\sqrt{\pi}(x-\mu)} \right\}$$

که در آن،  $x > \mu$  است. چون

$$\phi_X(t) = e^{i\mu t - \sqrt{-\pi}i\sigma t} = e^{i\mu t - |\sigma t|^{1/2}} (1 - i \operatorname{sign}(t)),$$

پس این یک توزیع پایدار با پارامتر مقیاس  $\sigma$ ، پارامتر مکان  $\mu$  و شاخص پایداری  $\frac{1}{2}$  است. توجه کنید که اولین و دومین گشتاور، نامتناهی هستند اما این توزیع برای مدل‌بندی تغییرپذیری بالا سودمند است، زیرا برای نوسانات بزرگ مجاز است. شبیه به توزیع گشی، دم توزیع لوی شبیه به یک تابع توانی کاهش می‌یابد اما این کاهش، نسبت به توزیع گشی خیلی ملایم‌تر است. اکنون با تغییر متغیر و استفاده از چگالی توزیع گاما، می‌توان نشان داد که برای  $\lambda > \frac{1}{2}$

$$\mathcal{I}_R(\lambda) = \ln \left( \frac{\sigma}{\sqrt{\pi}} \right) + \frac{1}{1-\lambda} \left\{ \ln \frac{\Gamma(\frac{1}{2}\lambda - 1)}{\Gamma(\frac{1}{2})} - \left( \frac{1}{2}\lambda - 1 \right) \ln \lambda \right\}$$



## جدول ۱. چگالی‌های پُر کاربرد

محدودیت‌ها	$f(x)$	نام
$\alpha > 0, \beta > 0, 0 < x < 1$	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$	بتا
$-\infty < \mu < \infty, \sigma > 0$	$\frac{1}{\pi\sigma\{1+(\frac{x-\mu}{\sigma})^2\}}$	کُشی
$-\infty < x < \infty, \theta > 0$	$\frac{\theta}{2(1+\theta x )^2}$	کرامر
$\alpha > 0, x > 0$	$\alpha \exp\{-\alpha x\}$	نمایی
$k > 0, m > 0, x > 0$	$\frac{(\frac{k}{m})^{k/2} x^{(k-2)/2}}{B(\frac{k}{2}, \frac{m}{2})(1+\frac{k}{m}x)^{(k+m)/2}}$	F
$\alpha > 0, \beta > 0, x > 0$	$\frac{\alpha^\beta x^{\beta-1} \exp\{-\alpha x\}}{\Gamma(\beta)}$	گاما
$-\infty < \alpha < \infty, \beta > 0$	$\beta^{-1} e^{-(x-\alpha)/\beta} \exp\{-e^{-(x-\alpha)/\beta}\}$	گامبل
$-\infty < x, \mu < \infty, \sigma > 0$	$\frac{1}{\sqrt{2\pi}\sigma} \exp\{-\sqrt{2} x-\mu /\sigma\}$	لاپلاس
$x > \mu$	$(\frac{\sigma}{\sqrt{\pi}})^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\{-\frac{\sigma}{2(x-\mu)}\}$	لوی
$-\infty < x, \mu < \infty, \sigma > 0$	$\frac{\pi}{\sigma^{3/2}} \frac{e^{-\pi(x-\mu)/(\sigma^{1/2})}}{(1+e^{-\pi(x-\mu)/(\sigma^{1/2})})^2}$	لجستیک
$\rho > 0, \tau > 0, x > 0$	$\frac{\tau x^{\tau-1} \rho^\tau}{(1+(\rho x)^\tau)^\tau}$	لگ لجستیک
$\rho > 0, \tau > 0, x > 0$	$\frac{1}{(\sqrt{\pi})^{1/2} \tau x} \exp\{-\frac{\log^\tau(\rho x)}{\tau}\}$	لگ نرمال
$-\infty < x, \mu < \infty, \sigma > 0$	$\frac{1}{(\sqrt{\pi})^{1/2} \sigma} \exp\{-\frac{(x-\mu)^2}{\sigma^2}\}$	نرمال
$ x  < \delta, \kappa \delta^\nu = \nu/\nu$	$\kappa(\delta^\nu - x^\nu)$	سه‌موی
$x > \beta, \alpha > 0, \beta > 0$	$\alpha\beta^\alpha x^{-(\alpha+1)}$	پارتو
$-\infty < x < \infty, \delta > 0$	$\frac{\delta}{\nu(1+\delta^{-1})\Gamma(1/\delta)} \exp\{-\frac{1}{\nu} x ^\delta\}$	نمایی توانی
$v > 0$	$\frac{1}{(v\pi)^{1/2}} \frac{\Gamma(v/2 + \frac{1}{2})}{\Gamma(v/2)} (1 + \frac{x^2}{v})^{-(v+1)/2}$	تی استوندت
$\alpha > 0$	$\begin{cases} 2x/\alpha & 0 \leq x \leq \alpha, \\ \frac{2(1-x)}{1-\alpha} & \alpha \leq x \leq 1, \end{cases}$	مثالی
$\alpha < x < \beta$	$\frac{1}{\beta-\alpha}$	یکنواخت
$1 + \beta x > 0$	$(1 + \beta x)^{-(1+\beta^{-1})} \exp\{-(1 + \beta x)^{-\beta^{-1}}\}$	فون میزس
$\alpha > 0, \beta > 0, x > 0$	$\beta\alpha^\beta x^{\beta-1} \exp\{-(\alpha x)^\beta\}$	وایبول

$$\mathcal{J}_R(\lambda) = \frac{\frac{2}{\nu}(1-\lambda)\psi(\frac{\nu}{2}\lambda - 1) + \ln\{\Gamma(\frac{\nu}{2}\lambda - 1)/(\lambda^{1/2}\Gamma(\frac{1}{2}))\}}{(1-\lambda)^\nu} + \frac{(3\lambda - 2)(\lambda - 1)}{(2\lambda)}$$

بنابراین

$$\begin{aligned} \mathcal{L}_f(\text{لوی}) &= -\frac{1}{2} \lim_{\lambda \rightarrow 1} \mathcal{J}_R(\lambda) = \lim_{\lambda \rightarrow 1} \left\{ \frac{9}{4} \psi\left(\frac{3}{2}\lambda - 1\right) - \frac{3\lambda + 2}{2\lambda^2} \right\} \\ &= \frac{9}{4} \psi\left(\frac{1}{2}\right) - \frac{5}{2} = \frac{9\pi^2}{8} - \frac{5}{2}. \end{aligned}$$

جدول ۲. طیف اطلاع رنی

نام	$\mathcal{J}_R(\lambda)$
بتا	$(1 - \lambda)^{-1} \ln \frac{B(\lambda(\alpha-1)+1, \lambda(\beta-1)+1)}{B^\lambda(\alpha, \beta)}$
کُشی	$\ln \sigma - \frac{\lambda}{1-\lambda} \ln \pi + (1 - \lambda)^{-1} \ln B\left(\frac{1}{\tau}, \lambda - \frac{1}{\tau}\right)$
کرامر	$\ln\left(\frac{\lambda}{\theta}\right) - (1 - \lambda)^{-1} \ln(2\lambda - 1)$
نمایی	$-\ln(\alpha\lambda) - \frac{\lambda}{1-\lambda} \ln(\lambda)$
F	$(1 - \lambda)^{-1} \ln \frac{B(\frac{1}{\tau}\lambda(k-\tau)+1, \frac{1}{\tau}\lambda(m+\tau)-1)}{B^{\lambda(\frac{k}{\tau}, \frac{m}{\tau})}} - \ln\left(\frac{k}{m}\right)$
گاما	$(1 - \lambda)^{-1} \ln \frac{\Gamma(\lambda(\beta-1)+1)}{\Gamma^\lambda(\beta)} - \frac{\lambda\beta}{1-\lambda} \ln \lambda - \ln(\alpha\lambda)$
گامبل	$\ln(\beta) - \frac{\lambda}{1-\lambda} \ln(\lambda) + (1 - \lambda)^{-1} \ln(\Gamma(\lambda))$
لاپلاس	$\frac{1}{\tau} \ln(2\sigma^\tau) - (1 - \lambda)^{-1} \ln(\lambda)$
لوی	$\ln\left(\frac{\sigma}{\tau}\right) + \frac{1}{1-\lambda} \left\{ \ln \frac{\Gamma(\frac{\tau}{\lambda})}{\Gamma^\lambda(\frac{1}{\tau})} - \left(\frac{\tau}{\lambda} - 1\right) \ln \lambda \right\}$
لجستیک	$(1 - \lambda)^{-1} \ln B(\lambda, \lambda) - \frac{1}{\tau} \ln\left(\frac{\pi^\tau}{\tau\sigma^\tau}\right)$
لگ لجستک	$(1 - \lambda)^{-1} \ln B\left((\lambda - 1)\left(\frac{1}{\tau} + 1\right) + 1, \lambda\left(1 - \frac{1}{\tau}\right) + \frac{1}{\tau}\right) - \ln(\tau\rho)$
لگ نرمال	$\frac{(1+\lambda^\tau)\tau^\tau}{2\lambda(1-\lambda)} - \frac{1}{1-\lambda} \tau^\tau - \frac{1}{\tau(1-\lambda)} \ln \lambda + \frac{1}{\tau} \ln\left(\frac{\tau\pi\tau^\tau}{\rho^\tau}\right)$
نرمال	$\frac{1}{\tau} \ln(2\pi\sigma^\tau) - \frac{1}{\tau}(1 - \lambda)^{-1} \ln \lambda$
سه‌موی	$\ln(\delta) + \frac{\lambda}{1-\lambda} \ln\left(\frac{\tau}{\tau}\right) + (1 - \lambda)^{-1} \ln B(\lambda + 1, 1/\tau)$
پارتو	$\frac{\lambda}{1-\lambda} \ln \alpha - (1 - \lambda)^{-1} \ln\{(1 + \alpha)\lambda - 1\} + \ln \beta$
نمایی توانی	$\ln\{2^{(1+\delta^{-1})}\Gamma(1 + \delta^{-1})\} - \frac{\ln(\lambda)}{(1-\lambda)\delta}$
تی استوندت	$\frac{1}{1-\lambda} \ln\left\{\frac{B(v\lambda/\tau+\lambda/\tau-\frac{1}{\tau}, \frac{1}{\tau})}{B^\lambda(v/\tau, \frac{1}{\tau})}\right\} + \frac{1}{\tau} \ln(v)$
مثلی	$\frac{\lambda}{1-\lambda} \ln 2 - (1 - \lambda)^{-1} \ln(1 + \lambda)$
یکنواخت	$\ln(\beta - \alpha)$
فون میزس	$\beta \ln(\lambda) - \frac{\lambda}{1-\lambda} \ln(\lambda) + (1 - \lambda)^{-1} \ln(\Gamma((\lambda - 1)(\beta + 1) + 1))$
وایبول	$(1 - \lambda)^{-1} \ln \Gamma\left(\lambda - \frac{\lambda-1}{\beta}\right) - \frac{\lambda}{\lambda-1} \ln \lambda - \ln(\alpha\beta\lambda^\beta)$

همان‌طور که قبلاً ذکر شد، احتمالات دُمی برای توزیع لوی شبیه به  $x^{-\alpha}$  وقتی  $x \rightarrow \infty$  رفتار می‌کند که در اینجا  $\frac{1}{\alpha} = 0.33608$ . محاسبه بالا نتیجه می‌دهد  $\mathcal{L}_f(\text{لوی}) \approx 8.6033$  که در مقایسه با  $\mathcal{L}_f(\text{کشی}) \approx 3.2899$  و  $\mathcal{L}_f(\text{کرامر}) = 4$ ، نشان می‌دهد که دُم‌های توزیع لوی خیلی پهن‌تر از دُم‌های توزیع‌های کُشی و کرامر هستند. همچنین جالب است این را با  $\mathcal{L}_f(\text{پارتو}, \alpha) = (1 + 1/\alpha)^2$  متناظر با توزیع پارتو با پارامتر  $\alpha$  مقایسه کنیم که در آن، دُم‌ها همچنین شبیه به توزیع توانی  $x^{-\alpha}$  وقتی  $x \rightarrow \infty$  کاهش می‌یابد (جدول ۱ و ۳ را ببینید). مشاهده کنید که وقتی  $\alpha = \frac{1}{9}$ ،  $\mathcal{L}_f(\text{پارتو}, \frac{1}{9}) = 9$  که نشان می‌دهد تا حد زیادی،  $\mathcal{L}_f$  پهنی دُم‌های توزیع‌ها را اندازه می‌گیرد. در برخی کاربردهای جالب مشابه با این مثال، چگالی زیربنایی متقارن نیست و دارای تکیه‌گاهی مانند  $\mathbb{R}^+ = (0, \infty)$  است. در چنین موقعیت‌هایی، اگر بخواهیم بین چگالی‌های متقارن مقایسه مفیدتری انجام دهیم، می‌توان همواره چگالی  $f(x)$  را توسط

$$f^*(x) = \frac{1}{\sqrt{2}} \{f(x)I(x > 0) + f(-x)I(x < 0)\}$$

متقارن ساخت. بنابراین متقارن‌سازی نباید شکل اصلی توزیع زیربنایی را بجز عدم‌حتمیت تغییر دهد. در واقع، نشان می‌دهیم که متقارن‌سازی اندازه  $\mathcal{L}_f$  را تغییر نمی‌دهد؛ یعنی  $\mathcal{L}_f = \mathcal{L}_{f^*}$  در حالی که آنتروپی (عدم‌حتمیت)  $f$  را به مقدار  $\ln 2$  افزایش خواهد داد. در ابتدا می‌دانیم که

$$\mathcal{L}_{f^*} = \lim_{\lambda \rightarrow 1} \left\{ \frac{\int_{-\infty}^{\infty} (f^*(x))^\lambda (\ln f^*(x))^2 dx}{\int_{-\infty}^{\infty} (f^*(x))^\lambda dx} - \left( \frac{\int_{-\infty}^{\infty} (f^*(x))^\lambda \ln f^*(x) dx}{\int_{-\infty}^{\infty} (f^*(x))^\lambda dx} \right)^2 \right\}. \quad (2.2)$$

از طرفی،

$$\begin{aligned} \int_{-\infty}^{\infty} (f^*(x))^\lambda (\ln f^*(x))^2 dx &= \int_{-\infty}^0 \left(\frac{1}{\sqrt{2}} f(-x)\right)^\lambda (\ln \frac{1}{\sqrt{2}} f(-x))^2 dx \\ &+ \int_0^{\infty} \left(\frac{1}{\sqrt{2}} f(x)\right)^\lambda (\ln \frac{1}{\sqrt{2}} f(x))^2 dx \\ &= \left(\frac{1}{\sqrt{2}}\right)^{\lambda-1} \int_0^{\infty} (f(x))^\lambda [\ln f(x) - \ln 2]^2 dx. \end{aligned} \quad (3.2)$$

به‌طور مشابه می‌توان نشان داد که

$$\begin{aligned} \int_{-\infty}^{\infty} (f^*(x))^\lambda \ln f^*(x) dx &= \left(\frac{1}{\sqrt{2}}\right)^{\lambda-1} \int_0^{\infty} (f(x))^\lambda \ln \left(\frac{1}{\sqrt{2}} f(x)\right) dx \\ &= \left(\frac{1}{\sqrt{2}}\right)^{\lambda-1} \int_0^{\infty} (f(x))^\lambda (\ln f(x) - \ln 2) dx. \end{aligned} \quad (4.2)$$

جدول ۳. اندازه‌های  $\mathcal{L}_f$

اندازه‌های $\mathcal{L}_f(\lambda)$	آنتروپی	نام
$(\alpha - 1)^\gamma \psi(\alpha) + (\beta - 1)^\gamma \psi(\beta)$ $-(\alpha - 1 + \beta - 1)^\gamma \psi(\alpha + \beta)$ $\pi^\gamma / 3$ $\gamma$ $1$	$\ln B(\alpha, \beta) - (\alpha - 1)(\psi(\alpha) - \psi(\alpha + \beta))$ $-(\beta - 1)(\psi(\beta) - \psi(\alpha + \beta))$ $\ln(\sqrt[3]{\pi\sigma})$ $\gamma + \ln(\frac{\gamma}{\theta})$ $1 - \ln(\alpha)$	بنا کُشی کرامر نمای
$\frac{(k-\gamma)^\gamma}{\gamma} \psi(k/\gamma) + \frac{(m+\gamma)^\gamma}{\gamma} \psi(m/\gamma)$ $-\frac{(k+m)^\gamma}{\gamma} \psi(\frac{k+m}{\gamma})$ $(\beta - 1)^\gamma \psi(\beta) - \beta + \gamma$ $\pi^\gamma / 6 - 1$ $9\pi^\gamma / 8 - 5/\gamma$ $\gamma - \pi^\gamma / 3$ $\frac{\pi^\gamma}{\gamma} (\frac{1}{\gamma} - 1) + \gamma$ $\frac{1}{\gamma}$ $\frac{1}{\gamma}$ $31/9 - \frac{\pi^\gamma}{\gamma}$ $(1 + 1/\alpha)^\gamma$ $\frac{1}{\delta}$	$\ln[B(k/\gamma, m/\gamma)m/k] + (1 - k/\gamma)\psi(k/\gamma)$ $-(1 + m/\gamma)\psi(m/\gamma) + (k + m)/\gamma\psi(\frac{k+m}{\gamma})$ $\ln \frac{\Gamma(\beta)}{\alpha} - (\beta - 1)\psi(\beta) + \beta$ $\ln(\beta) - \psi(1) + 1$ $\frac{1}{\gamma} \ln(\sqrt[6]{\pi e \sigma^\gamma}) + \frac{\gamma}{\gamma} \gamma$ $\gamma - \ln \frac{\pi}{\sigma\sqrt{\gamma}}$ $\gamma - \ln(\tau\rho)$ $\frac{1}{\gamma} \ln(\sqrt[3]{\pi e \tau^\gamma}) - \ln(\rho)$ $\frac{1}{\gamma} \ln(\sqrt[3]{\pi e \sigma^\gamma})$ $\ln(\delta/\gamma) + 5/\gamma$ $\ln(\beta/\alpha) + 1/\alpha + 1$ $\ln\{\gamma^{1+\delta^{-1}}\Gamma(1 + \delta^{-1})\} + 1/\delta$	F گاما گامبل لوی لجستیک لُگ لجستیک لُگ نرمال نرمال سه‌موی پارتو
$\frac{(v+\gamma)^\gamma}{\gamma} \{\psi(\frac{v}{\gamma}) - \psi(\frac{v+\gamma}{\gamma})\}$ $1/\gamma$ $0$ $\frac{\pi^\gamma}{\gamma} (\beta + 1)^\gamma - 2\beta - 1$ $\frac{\pi^\gamma}{\gamma} (1 - \beta^{-1})^\gamma + 2\beta^{-1} - 1$	$\frac{v+\gamma}{\gamma} \{\psi(\frac{v}{\gamma} + \frac{1}{\gamma}) - \psi(\frac{v}{\gamma})\} + \frac{1}{\gamma} \ln\{vB^\gamma(v/\gamma, 1/\gamma)\}$ $1/\gamma - \ln \gamma$ $\ln(\beta - \alpha)$ $1 + \gamma(1 + \beta)$ $1 + \gamma(1 - \beta^{-1}) - \ln(\alpha\beta)$	نمای توانی تی استودنت مثالی یکتواخت فون میزس وایبول

همچنین

$$\int_{-\infty}^{\infty} (f^*(x))^\lambda dx = (\frac{1}{\gamma})^{\lambda-1} \int_0^{\infty} (f(x))^\lambda dx. \quad (5.2)$$

با جایگذاری سمت راست (۳.۲)، (۴.۲) و (۵.۲) در (۲.۲)، به دست می‌آوریم

$$\begin{aligned} \mathcal{L}_{f^*} &= \lim_{\lambda \rightarrow 1} \frac{1}{\int_0^{\infty} f^\lambda(x) dx} \left\{ \int_0^{\infty} (f(x))^\lambda (\ln f(x))^\gamma dx \right. \\ &\quad \left. + (\ln \gamma)^\gamma \int_0^{\infty} (f(x))^\lambda dx - \gamma \ln \gamma \int_0^{\infty} (f(x))^\lambda \ln f(x) dx \right\} \\ &\quad - \lim_{\lambda \rightarrow 1} \frac{1}{(\int_0^{\infty} f^\lambda(x) dx)^\gamma} \left( \int_0^{\infty} (f(x))^\lambda (\ln f(x) - \ln \gamma) dx \right)^\gamma = \mathcal{L}_f. \end{aligned}$$

برای به دست آوردن آنتروپی  $f^*$  با استفاده از (۴.۲)، با قرار دادن  $\lambda = 1$ ، نتیجه می‌گیریم که

$$\begin{aligned} - \int_{-\infty}^{\infty} f^*(x) \ln f^*(x) dx &= - \int_{-\infty}^{\infty} f(x) (\ln f(x) - \ln 2) \\ &= - \int_{-\infty}^{\infty} f(x) \ln f(x) dx + \ln 2. \end{aligned}$$

این ویژگی مطلوب است، زیرا  $\mathcal{L}_f$  به‌عنوان یک اندازه‌شکلی، نباید مقیاس را اندازه بگیرد. این شاخص را باید با اندازه کشیدگی متداول  $\beta_2$  با در نظر گرفتن مثلاً توزیع گاما با چگالی  $f(x)$  در جدول ۱، مقایسه کرد. چگالی متقارن شده به‌صورت

$$f^*(x) = \frac{\alpha^\beta |x|^{\beta-1} \exp\{-\alpha|x|\}}{\Gamma(\beta)}$$

است. پس از مقداری محاسبات جبری، به دست می‌آوریم

$$\mathcal{L}_{f^*} = \mathcal{L}_f = (\beta - 1)^2 \psi(\beta) - \beta + 2.$$

به‌ویژه وقتی  $\beta = 1$ ،  $f(x)$  و  $f^*(x)$  به‌ترتیب، متناظر با توزیع‌های نمایی و لاپلاس (نمایی دوگانه) هستند. در این حالت،  $\mathcal{L}_{f^*} = \mathcal{L}_f = 1$  در حالی که کشیدگی برای توزیع لاپلاس،  $\beta_2 = 6$  و برای توزیع نمایی،  $\beta_2 = 9$  است. این اختلاف در مقادیر برای اندازه کشیدگی ممکن است کاملاً گمراه‌کننده باشد، زیرا توزیع نمایی را می‌توان توزیع لاپلاس تا شده حول  $x = 0$  در نظر گرفت. برای شرح بیشتر درباره عدم کفایت استفاده از کشیدگی به‌عنوان نشانگر شکل توزیعی، مثال نوعی توزیع لگ نرمال را در نظر بگیرید. برای این توزیع، کشیدگی برابر است با

$$\beta_2 = (e^{6\tau^2} + 6e^{\tau^2} - 4e^{3\tau^2} - 3)/(e^{\tau^2} - 1)^2.$$

به‌ویژه وقتی  $\tau^2 = 1$ ،  $\beta_2 \approx 113/9346$ . اما از جدول ۳، مشاهده می‌کنیم که برای توزیع لگ نرمال  $\mathcal{L}_f(\tau^2) = \tau^2 + \frac{1}{4}$  و  $\mathcal{L}_f(1) = \frac{3}{4}$  که شبیه به توزیع وایبول با پارامتر شکلی

$$\beta = 2\pi^2 / \{(144 + 12\pi^2)^{1/2} + 2\pi^2 - 12\} \approx 0.82456$$

رفتار می‌کند که برای آن،  $\mathcal{L}_f$  (وایبول) در حالی که برای توزیع وایبول با مقدار  $\beta = 0.82456$ ، کشیدگی برابر با  $\beta_2 \approx 14.47$  است. این اختلاف زیاد بین  $\beta_2 \approx 113/9346$  و  $\beta_2 \approx 14.47$  به‌روشنی محدودیت کشیدگی را آشکار می‌کند در حالی که اندازه  $\mathcal{L}_f$  به ما ایده‌ای درباره شکل (دم‌های) چگالی زیربنایی می‌دهد.

در ادامه مثالی خواهیم داشت که به ما بینشی برای روش جدید برآورد پارامتر شکلی، بر اساس گشتاورهای لگ درستنمایی ارائه می‌دهد. در این مثال، برآوردیابی با استفاده از روش‌های دیگر دشوار است.

مثال ۴.۲. توزیع توانی نمایی با تابع چگالی زیر را در نظر بگیرید:

$$f(x) = \frac{\delta}{\sqrt[2]{(1+\frac{1}{\delta})}\Gamma(1/\delta)} \exp\{-\frac{1}{\sqrt[2]{\delta}}|x|^\delta\}, \quad -\infty < x < \infty.$$

توجه کنید که اگر  $\delta = 1$ ، آن‌گاه این توزیع، توزیع نمایی دوگانه است،  $\delta = 2$  متناظر با توزیع نرمال و  $\delta = p$  (یک عدد زوج مثبت معین) توزیع توانی  $p$ ام است. همچنین اگر  $\delta \rightarrow \infty$ ، این توزیع به توزیع یکنواخت منجر می‌شود. بنابراین پارامتر شکلی  $\delta$  بیانگر میزان «غیر نرمال» بودن جامعه اولیه است. برای این توزیع، طیف اطلاع رنی عبارت است از

$$\begin{aligned} \mathcal{I}_R(\lambda) &= \frac{1}{1-\lambda} \ln \int_{-\infty}^{\infty} \frac{\delta^\lambda}{\sqrt[2]{\lambda(1+\frac{1}{\delta})}[\Gamma(\frac{1}{\delta})]^\lambda} \exp\{-\frac{1}{\sqrt[2]{\delta}}|\lambda^{\frac{1}{\delta}}x|^\delta\} dx \\ &= \frac{1}{1-\lambda} \ln \left\{ \frac{\delta^\lambda}{\sqrt[2]{\lambda(1+\frac{1}{\delta})}[\Gamma(\frac{1}{\delta})]^\lambda} \cdot \frac{\sqrt[2]{1+\frac{1}{\delta}}\Gamma(\frac{1}{\delta})\lambda^{-\frac{1}{\delta}}}{\delta} \right\} \\ &= \ln \left\{ \sqrt[2]{1+\frac{1}{\delta}}\Gamma(1+\frac{1}{\delta}) \right\} - \frac{\ln(\lambda)}{(1-\lambda)\delta} \end{aligned}$$

و مشتق طیف برابر است با

$$\dot{\mathcal{I}}_R(\lambda) = \frac{1 - \lambda^{-1} - \ln(\lambda)}{(1-\lambda)^2 \delta}.$$

بنابراین

$$\mathcal{L}_f = -2 \lim_{\lambda \rightarrow 1} \frac{1 - \lambda^{-1} - \ln(\lambda)}{(1-\lambda)^2 \delta} = \lim_{\lambda \rightarrow 1} \frac{\lambda^{-2} - \lambda^{-1}}{(1-\lambda)\delta} = \lim_{\lambda \rightarrow 1} \frac{1}{\lambda^2 \delta} = \frac{1}{\delta}.$$

این عبارت بسیار ساده، در مقایسه با کشیدگی  $\beta_2 = \frac{\Gamma(5/\delta)\Gamma(1/\delta)}{\Gamma^2(3/\delta)}$  به ما یاری می‌دهد تا پارامتر شکلی  $\delta$  را به‌طور تجربی با حل معادله

$$\frac{1}{\sqrt[2]{\delta}} \left\{ \frac{1}{n} \sum_{i=1}^n |X_i|^{\sqrt[2]{\delta}} - \left( \frac{1}{n} \sum_{i=1}^n |X_i|^\delta \right)^{\sqrt[2]{\delta}} \right\} = \frac{1}{\delta}$$

برآورد کنیم حال آنکه با استفاده از روش‌های دیگر، برآورد آن ساده نیست، زیرا برآورد بیشینه درست‌نمایی  $\delta$  شامل تابع دو گاما است و نیز در روش گشتاورها، عبارت‌های پیچیده تابع گاما ظاهر می‌شوند. برای مثال، گشتاورهای مرکزی  $m$  برای  $m$  زوج عبارت‌اند از

$$\mu_m = \frac{2^{m/\delta} \Gamma(m/\delta + 1/\delta)}{\Gamma(1/\delta)}.$$

آخرین مثال مربوط به توزیع‌های مقدار کرانگین است. فرض کنیم  $X_1, \dots, X_n$  متغیرهای تصادفی مستقل با تابع توزیع مشترک  $F$  باشند و  $Y = \max\{X_1, \dots, X_n\}$  در این صورت،  $P\{Y \leq t\} = F^n(t)$  اگر برای یک انتخاب مناسب از ثابت‌های  $a_n$  و  $b_n$ ، تابع توزیع ماکسیمم به یک تابع توزیع حدی پیوسته همگرا باشد، یعنی

$$F^n(a_n x + b_n) \rightarrow G(x), \quad n \rightarrow \infty,$$

آن‌گاه  $G$  برابر یکی از سه تابع توزیع زیر است:

$G_0(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}$	الف) گامبل
$G_{1,\alpha}(x) = \exp(-x^{-\alpha}), \quad x \geq 0, \alpha > 0$	ب) فرشه
$G_{2,\alpha}(x) = \exp(-(-x)^{-\alpha}), \quad x \leq 0, \alpha < 0$	پ) وایبول

علاوه بر این، تابع توزیع  $F$  متعلق به دامنه ریاضی توزیع مقدار کرانگین  $G$  است. در ادامه با پارامتری‌سازی مجدد  $\beta = 1/\alpha$  و پارامترهای مقیاس و مکان مناسب، می‌توان این مدل‌ها را در پارامتری‌سازی فون‌میزس به صورت زیر متحد ساخت:

$$G_\beta(x) = \begin{cases} \exp(-(1 + \beta x)^{-1/\beta}) & 1 + \beta x > 0, \beta \neq 0 \\ \exp(-e^{-x}) & x \in \mathbb{R}, \beta = 0 \end{cases}$$

توجه کنید که برای  $\beta \rightarrow 0, G_\beta(x) \rightarrow G_0(x)$  و نیز تابع چگالی توزیع مقدار کرانگین برابر خواهد شد با

$$f(x) = \begin{cases} (1 + \beta x)^{-(1+\beta^{-1})} \exp\{-(1 + \beta x)^{-\beta^{-1}}\} & 1 + \beta x > 0, \beta \neq 0 \\ \exp\{-x - e^{-x}\}, & x \in \mathbb{R}, \beta = 0 \end{cases}$$

برای محاسبه طیف اطلاع رنی، با تغییر متغیرها و استفاده از چگالی توزیع گاما، داریم

$$\begin{aligned}\mathcal{I}_R(\lambda) &= \frac{1}{1-\lambda} \ln \int_{-1/\beta}^{\infty} (1+\beta x)^{-\lambda(1+1/\beta)} \exp\{-\lambda(1+\beta x)^{-1/\beta}\} dx \\ &= \frac{1}{1-\lambda} \ln \int_0^{\infty} y^{(\beta+1)(\lambda-1)} e^{-\lambda y} dy \\ &= \frac{1}{1-\lambda} \ln\left\{(\lambda/\beta)^{(\beta+1)(\lambda-1)+1} \Gamma((\beta+1)(\lambda-1)+1)\right\} \\ &= \frac{\{\beta(1-\lambda) - \lambda\} \ln \lambda + \ln \Gamma((\lambda-1)(\beta+1)+1)}{1-\lambda}\end{aligned}$$

و

$$\begin{aligned}\dot{\mathcal{I}}_R(\lambda) &= \frac{1}{(1-\lambda)^2} \left[ \left\{ (-\beta-1) \ln \lambda + \frac{1}{\lambda} (\beta(1-\lambda) - \lambda) \right. \right. \\ &\quad \left. \left. + (\beta+1) \psi((\lambda-1)(\beta+1)+1) \right\} (1-\lambda) \right. \\ &\quad \left. + (\beta(1-\lambda) - \lambda) \ln \lambda + \ln \Gamma((\lambda-1)(\beta+1)+1) \right] \\ &= \left( (1-\lambda)(\beta+1) \psi((\lambda-1)(\beta+1)+1) \right. \\ &\quad \left. + \ln \left\{ \lambda^{-1} \Gamma((\lambda-1)(\beta+1)+1) \right\} \right. \\ &\quad \left. + (\beta(1-\lambda))^2 - \lambda(1-\lambda) \right) / \lambda / (1-\lambda)^2.\end{aligned}$$

بنابراین

$$\begin{aligned}\mathcal{L}_f(\text{فون میزس}, \beta) &= -2 \lim_{\lambda \rightarrow 1} \dot{\mathcal{I}}_R(\lambda) \\ &= \lim_{\lambda \rightarrow 1} \left\{ (\beta+1)^2 \psi((\lambda-1)(\beta+1)+1) - \frac{\beta(1+\lambda) + \lambda}{\lambda^2} \right\} \\ &= (\beta+1)^2 \psi(1) - 2\beta - 1 \\ &= \frac{\pi^2}{6} (\beta+1)^2 - 2\beta - 1.\end{aligned}$$

از این، بلافاصله می‌توان نتیجه گرفت که وقتی  $\beta = 0$ ، متناظر با توزیع مقدار کرانگین گامبل،

$$\mathcal{L}_f(\text{فون میزس}, 0) = \mathcal{L}_f(\text{گامبل}) = \frac{\pi^2}{6} - 1.$$



در ادامه برخی از مهم‌ترین و پُر کاربردترین توزیع‌ها در آمار را در نظر می‌گیریم و طیف اطلاع رنی آنها را همراه با ورنترویی متناظر  $\mathcal{L}_f$  به دست می‌آوریم. توابع چگالی در جدول ۱ فهرست شده‌اند. طیف متناظر در جدول ۲ داده می‌شوند. جدول ۳ شامل اندازه  $\mathcal{L}_f$  همراه با آنتروپی‌های آنها است که در آن،  $\lambda$  نشان‌دهنده ثابت اویلر ( $\gamma = 0.57721\dots$ ) است. توجه کنید که از جدول ۳ رابطه بین آنتروپی بیان شده توسط تابع دو گاما و واریانس لگاریتم درست‌نمایی نشان داده شده توسط تابع سه گاما، رابطه میان طیف اطلاع رنی و مشتق طیف را منعکس می‌کند. همچنین از جدول ۳ مشاهده می‌کنیم که ورنترویی  $\mathcal{L}_f$ ، به میزان زیادی، پهنی دُم‌ها را اندازه می‌گیرد. برای مثال، برای توزیع لگ لجستیک مذکور در جدول ۱،  $\mathcal{L}_f(\tau^2) = \frac{\pi^2}{3}(\frac{1}{\tau^2} - 1) + 4$ ، وقتی  $\tau^2 = 1$ ،  $\mathcal{L}_f(1) = 4$  همان اندازه برای توزیع کرامر را نتیجه می‌دهد و وقتی  $\tau^2 \rightarrow \infty$ ،  $\mathcal{L}_f(\tau^2) \rightarrow 4 - \frac{\pi^2}{3} = \mathcal{L}_f$  (لجستیک). اگرچه در این مقاله، ما بررسی مقایسه‌ای ورنترویی  $\mathcal{L}_f$  را در نظر نمی‌گیریم، مایلیم خاطر نشان کنیم که اندازه  $\mathcal{L}_f$  را می‌توان برای القای یک ترتیب جزئی مورد استفاده قرار داد. به عبارت دیگر، اگر نماد  $f \leq g$  به این معنی باشد که  $\mathcal{L}_f \leq \mathcal{L}_g$ ؛ یعنی  $g$  پهن‌تر از  $f$  است، آن‌گاه می‌توانیم این نماد پهنی را برای ترتیب دُم‌های توزیع‌ها به کار ببریم. برای مثال،

لجستیک  $\leq$  مقدار کرانگین گامبل  $\leq$  نرمال  $\leq$  مثلثی  $\leq$  سهمی گون  $\leq$  یکنواخت

$\leq$  لاپلاس  $\leq$  (با برای مثال  $5 \leq v \leq 4$ ) تی استودنت

لوی  $\leq$  کرامر  $\leq$  گُشی  $\leq$  (با برای مثال  $3 \leq v \leq 1$ ) تی استودنت

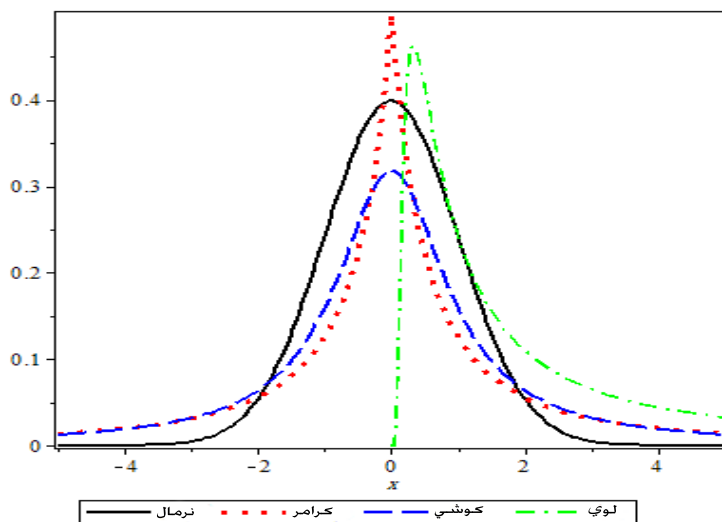
شکل ۱ نمودار چهار تابع چگالی نرمال  $N(0, 1)$ ، تابع چگالی گُشی  $C(0, 1)$ ، تابع چگالی لوی با پارامتر مکان ۰ و پارامتر مقیاس ۱ و تابع چگالی کرامر با پارامتر  $\theta = 1$  را نشان می‌دهد. مقایسه کردن ترتیب حاصل از اندازه  $\mathcal{L}_f$ ، با ترتیب جزئی حاصل از کشیدگی برای توزیع‌های در نظر گرفته شده در [۱۶]، آموزنده است.

مثال ۵.۲. فرض کنیم  $X$  دارای توزیع بر نوع ۱۲ با تابع چگالی زیر باشد:

$$f(x) = \frac{kcx^{c-1}}{(1+x^c)^{k+1}}, \quad x > 0, c > 0, k > 0 \quad (6.2)$$

در ابتدا داریم

$$\int_0^\infty \frac{k^\lambda c^\lambda x^{\lambda(c-1)}}{(1+x^c)^{\lambda(k+1)}} = \frac{k^\lambda c^{\lambda-1} \Gamma(\frac{\lambda ck + \lambda - 1}{c}) \Gamma(\frac{\lambda c - \lambda + 1}{c})}{\Gamma(\lambda(k+1))}$$



شکل ۱. نمودار برخی از توابع چگالی.

طیف اطلاع رنی برابر است با

$$\mathcal{I}_R(\lambda) = -\ln(kc) + \frac{1}{1-\lambda} \ln \left[ kB \left( \frac{\lambda c - \lambda + 1}{c}, \frac{\lambda kc + \lambda - 1}{c} \right) \right]$$

و بنابراین

$$\begin{aligned} \mathcal{I}'_R(\lambda) &= \frac{c-1}{c(1-\lambda)} \psi \left( \frac{\lambda c - \lambda + 1}{c} \right) + \frac{kc+1}{c(1-\lambda)} \psi \left( \frac{\lambda kc + \lambda - 1}{c} \right) \\ &\quad - \frac{1+k}{1-\lambda} \psi(\lambda + \lambda k) + \frac{1}{(1-\lambda)^2} \ln \left[ kB \left( \frac{\lambda c - \lambda + 1}{c}, \frac{\lambda kc + \lambda - 1}{c} \right) \right]. \end{aligned}$$

از این رو وراثتروپی توزیع بر خواهد شد

$$\mathcal{L}_R(\lambda) = \left(1 - \frac{1}{c}\right)^2 \psi'(1) + \left(k + \frac{1}{c}\right)^2 \psi'(k) - (k+1)^2 \psi'(k+1).$$

### ۳. تعمیم چندمتغیره

در این بخش می‌خواهیم اندازه  $\mathcal{L}_f$  را به حالت چندمتغیره تعمیم دهیم. فرض کنیم  $\mathbf{x}$  یک بردار  $d$  بُعدی  $\mathbf{x} = (x_1, \dots, x_d) \in R^d$  باشد. اطلاع رنی از مرتبه  $\lambda$  برای یک بردار تصادفی با تابع چگالی  $f(\mathbf{x})$  به صورت  $\mathcal{I}_R(\lambda) = (1-\lambda)^{-1} \ln(\int f^\lambda(\mathbf{x}) d\mathbf{x})$  با تفسیر  $\int f^\lambda(\mathbf{x}) d\mathbf{x}$  به عنوان

انتگرال چندگانه  $\int \dots \int f^\lambda(x_1, \dots, x_d) dx_1 \dots dx_d$  تعریف می‌شود. طیف اطلاع رنی و اندازه  $\mathcal{L}_f$  به‌طور مشابه تعریف می‌شوند. این مفاهیم را طی چند مثال شرح می‌دهیم.

مثال ۱.۳. توزیع  $t$  چندمتغیره زیر را با  $v$  درجه آزادی در نظر بگیرید:

$$f(\mathbf{x}) = \frac{\Gamma((d+v)/2)}{(v\pi)^{d/2} \Gamma(v/2) |\Sigma|^{1/2}} \left(1 + \frac{1}{v} \mathbf{x} \Sigma^{-1} \mathbf{x}^T\right)^{-(d+v)/2}.$$

با تعویض متغیر  $\mathbf{y} = \left(\frac{\lambda(d+v)-d}{v}\right)^{d/2} \mathbf{x}$  و با توجه به اینکه ژاکوبی تبدیل برابر  $\left\{\frac{v}{\lambda(d+v)-d}\right\}^{d/2}$  است، طیف اطلاع رنی برای  $\lambda > d/(d+v)$  به‌صورت زیر محاسبه می‌شود:

$$\begin{aligned} \mathcal{J}(\lambda) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{[\Gamma(\frac{v+d}{v})]^\lambda}{[\Gamma(\frac{v}{v})]^\lambda (v\pi)^{\frac{\lambda d}{2}} |\Sigma|^{\lambda/2}} \left(1 + \frac{1}{v} \mathbf{x} \Sigma^{-1} \mathbf{x}^T\right)^{-(\frac{d+v}{v})\lambda} dx_1 dx_2 \dots dx_d \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{[\Gamma(\frac{v+d}{v})]^\lambda}{[\Gamma(\frac{v}{v})]^\lambda (v\pi)^{\frac{\lambda d}{2}} |\Sigma|^{\lambda/2}} \left(1 + \frac{1}{\lambda(d+v)-d} \mathbf{y} \Sigma^{-1} \mathbf{y}^T\right)^{-(\frac{d+v}{v})\lambda} \\ &\quad \times \left(\frac{v}{\lambda(d+v)-d}\right)^{\frac{d}{2}} dy_1 \dots dy_d \\ &= \frac{[\Gamma(\frac{v+d}{v})]^\lambda}{[\Gamma(\frac{v}{v})]^\lambda (v\pi)^{\frac{\lambda d}{2}} |\Sigma|^{\lambda/2}} \left(\frac{v}{\lambda(d+v)-d}\right)^{\frac{d}{2}} \frac{\Gamma(\frac{\lambda(d+v)-d}{v})}{\Gamma(\frac{\lambda(d+v)}{v})} ((\lambda(d+v)-d)\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}} \\ &= \frac{B(\frac{\lambda(d+v)-d}{v}, \frac{d}{v})}{[B(\frac{v}{v}, \frac{d}{v})]^\lambda} \frac{1}{|\Sigma|^{\frac{\lambda-1}{2}} (v\pi)^{\frac{d}{2}(\lambda-1)}} [\Gamma(\frac{d}{v})]^{\lambda-1} \end{aligned}$$

و بنابراین

$$\mathcal{J}_R(\lambda) = \frac{1}{1-\lambda} \ln \frac{B(\frac{\lambda(d+v)-d}{v}, \frac{d}{v})}{[B(\frac{v}{v}, \frac{d}{v})]^\lambda} + \frac{1}{v} \ln\{(v\pi)^d |\Sigma|\} - \ln \Gamma(\frac{d}{v}).$$

همچنین مشتق طیف عبارت است از

$$\begin{aligned} \dot{\mathcal{J}}_R(\lambda) &= \left( \ln \left\{ \frac{B((\lambda(d+v)-d)/2, d/2)}{B(v/2, d/2)} \right\} \right. \\ &\quad \left. + \frac{1}{v} (1-\lambda)(d+v) \{ \psi((\lambda(d+v)-d)/2) \right. \\ &\quad \left. - \psi((\lambda(d+v))/2) \} \right) / (1-\lambda)^2, \end{aligned}$$

در نتیجه

$$\begin{aligned}\mathcal{L}_f &= \lim_{\lambda \rightarrow 1} \frac{(d+v)^\nu}{\nu} \left\{ \psi \left( \frac{\lambda(d+v) - d}{\nu} \right) - \psi \left( \frac{\lambda(d+v)}{\nu} \right) \right\} \\ &= \frac{(d+v)^\nu}{\nu} \left\{ \psi \left( \frac{v}{\nu} \right) - \psi \left( \frac{d+v}{\nu} \right) \right\}.\end{aligned}$$

به‌ویژه وقتی  $d = 1$ ، این به سمت راست (۱.۲) تبدیل می‌شود.

مثال ۲.۳. در این مثال، توزیع نرمال چندمتغیره با تابع چگالی زیر را در نظر می‌گیریم:

$$f(x) = (\nu\pi)^{-d/\nu} |\Sigma|^{-1/\nu} \exp \left\{ -\frac{1}{\nu} (\mathbf{x} - \mu) \Sigma^{-1} (\mathbf{x} - \mu)^T \right\}.$$

تابع مولد اطلاع به‌صورت زیر محاسبه می‌شود:

$$\begin{aligned}\mathcal{J}(\lambda) &= \int \dots \int (\nu\pi)^{-d\lambda/\nu} |\Sigma|^{-\lambda/\nu} \exp \left\{ -\frac{\lambda}{\nu} (\mathbf{x} - \mu) \Sigma^{-1} (\mathbf{x} - \mu)^T \right\} dx_1 \dots dx_d \\ &= \int \dots \int (\nu\pi)^{-d\lambda/\nu} |\Sigma|^{-\lambda/\nu} \exp \left\{ -\frac{1}{\nu} (\mathbf{x} - \mu) \left( \frac{\Sigma}{\lambda} \right)^{-1} (\mathbf{x} - \mu)^T \right\} dx_1 \dots dx_d \\ &= (\nu\pi)^{-d\lambda/\nu} |\Sigma|^{-\lambda/\nu} (\nu\pi)^{-d/\nu} \left| \frac{\Sigma}{\lambda} \right|^{-1} = (\nu\pi)^{-d(\lambda-1)/\nu} |\Sigma|^{-(\lambda-1)/\nu} \lambda^{-d/\nu}\end{aligned}$$

و بنابراین طیف اطلاع رنی برابر است با

$$\mathcal{J}_R(\lambda) = \ln \{ (\nu\pi)^{d/\nu} |\Sigma|^{1/\nu} \} - \frac{d}{\nu(1-\lambda)} \ln(\lambda).$$

همچنین مشتق طیف عبارت است از

$$\dot{\mathcal{J}}_R(\lambda) = -\frac{d}{\nu(1-\lambda)^2} \{ \ln(\lambda) + \lambda^{-1} - 1 \},$$

به‌طوری که وراثت‌روپی برای این توزیع خواهد شد:

$$\begin{aligned}\mathcal{L}_f &= d \lim_{\lambda \rightarrow 1} \frac{\ln \lambda + \lambda^{-1} - 1}{(1-\lambda)^2} = -d \lim_{\lambda \rightarrow 1} \frac{1/\lambda - 1/\lambda^2}{\nu(1-\lambda)} \\ &= \frac{d}{\nu} \lim_{\lambda \rightarrow 1} (\nu/\lambda^2 - 1/\lambda^2) = \frac{d}{\nu}.\end{aligned}$$

مقایسه این تابع خطی ساده از  $d$ ، با کشیدگی نرمال چندمتغیره ماردیا [۹]،  $\beta_{\nu,d} = d(d+\nu)$ ، جالب است.

مثال ۳.۳. در این مثال، توزیع دیریکله (بتای چندمتغیره) را مورد بحث قرار می‌دهیم:

$$f(x_1, \dots, x_{d-1}) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_d)} x_1^{\alpha_1-1} \dots x_{d-1}^{\alpha_{d-1}-1} (1 - x_1 - \dots - x_{d-1})^{\alpha_d-1},$$

که در آن،  $\alpha_0 = \sum_{i=1}^d \alpha_i$ ، برای محاسبه طیف اطلاع رنی، داریم

$$\begin{aligned} & \int \cdots \int \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} \right)^\lambda x_1^{\lambda(\alpha_1-1)} \cdots x_{d-1}^{\lambda(\alpha_{d-1}-1)} (1-x_1-\cdots-x_{d-1})^{\lambda(\alpha_d-1)} \\ & \times dx_1 \cdots dx_d \\ & = \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} \right)^\lambda \frac{\prod_{i=1}^d \Gamma(\lambda(\alpha_i-1)+1)}{\Gamma(\lambda(\alpha_0-d)+d)} \end{aligned}$$

و بنابراین طیف اطلاع رنی برای توزیع دیریکله برابر است با

$$\begin{aligned} \mathcal{J}_R(\lambda) &= \frac{\lambda}{1-\lambda} \left\{ \ln \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} \right\} + \frac{1}{1-\lambda} \sum_{i=1}^d \ln \Gamma(\lambda(\alpha_i-1)+1) \\ & - \frac{1}{1-\lambda} \ln \Gamma(\lambda(\alpha_0-d)+d) \end{aligned}$$

و در نتیجه مشتق طیف عبارت است از

$$\begin{aligned} \dot{\mathcal{J}}_R(\lambda) &= \frac{1}{(1-\lambda)^2} \ln \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} + \frac{1}{(1-\lambda)^2} \sum_{i=1}^d \ln \Gamma(\lambda(\alpha_i-1)+1) \\ & + \frac{1}{1-\lambda} \sum_{i=1}^d (\alpha_i-1) \psi(\lambda(\alpha_i-1)+1) - \frac{1}{(1-\lambda)^2} \ln \Gamma(\lambda(\alpha_0-d)+d) \\ & - \frac{1}{1-\lambda} (\alpha_0-d) \psi(\lambda(\alpha_0-d)+d) \\ & = \frac{1}{1-\lambda} \left\{ \ln \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} + \sum_{i=1}^d (\alpha_i-1) \psi(\lambda(\alpha_i-1)+1) \right. \\ & \left. - (\alpha_0-d) \psi(\lambda(\alpha_0-d)+d) \right\} + \frac{1}{1-\lambda} \mathcal{J}_R(\lambda). \end{aligned}$$

از طرف دیگر،

$$\begin{aligned} \lim_{\lambda \rightarrow 1} \dot{\mathcal{J}}_R(\lambda) &= - \lim_{\lambda \rightarrow 1} \left\{ \sum_{i=1}^d (\alpha_i-1) \psi(\lambda(\alpha_i-1)+1) \right. \\ & \left. - (\alpha_0-d) \psi(\lambda(\alpha_0-d)+d) \right\} - \lim_{\lambda \rightarrow 1} \mathcal{J}_R(\lambda), \end{aligned}$$

و لذا وراثتروپی خواهد شد:

$$\begin{aligned}\mathcal{L}_f &= \lim_{\lambda \rightarrow 1} \left\{ \sum_{i=1}^d (\alpha_i - 1)^\lambda \psi(\lambda(\alpha_i - 1) + 1) - (\alpha_0 - d)^\lambda \psi(\lambda(\alpha_0 - d) + d) \right\} \\ &= \sum_{i=1}^d (\alpha_i - 1)^\lambda \psi(\alpha_i) - (\alpha_0 - d)^\lambda \psi(\alpha_0).\end{aligned}$$

به ویژه وقتی  $d = 2$ ، عبارت فوق وراثتروپی توزیع بتا است که در قبل بحث شد.

این مثال‌ها به روشنی توان و سودمندی اندازه در توصیف شکل تابعی و رفتار دُمی را شرح می‌دهند. به طور مشابه می‌توان اندازه  $\mathcal{L}_f$  را برای چندین توزیع چندمتغیره به دست آورد که چون تمرکز این مقاله روی حالت تک‌متغیره بود، آنها را بسط نمی‌دهیم.

#### ۴. نتیجه‌گیری

در این مقاله، اطلاع رنی از مرتبه  $\lambda$  را برای یک متغیر تصادفی پیوسته در نظر گرفتیم و رابطه آن را با لگاریتم درست‌نمایی ثابت کردیم. همچنین یک اندازه توزیع بر پایه لگاریتم تابع درست‌نمایی بدون نیاز به فرض‌های گشتاور متناهی یا هموار بودن بیان گردید. برخلاف اندازه کشیدگی سنتی، اندازه  $\mathcal{L}_f$  تقریباً برای همه چگالی‌هایی که مقایسه‌های معنی داری میان آنها مجاز است، پیشنهاد می‌شود. این اندازه می‌تواند در مطالعه شکل‌های توزیعی به ویژه برای توزیع‌های دُم پهن مورد استفاده قرار گیرد؛ در حالی که اندازه کشیدگی برای بسیاری از این توزیع‌ها وجود ندارد. علاوه بر این، وراثتروپی در بسیاری از کاربردها نظیر ساختن ابزارهای تشخیصی بر اساس باقیمانده‌ها در تحلیل رگرسیون خیلی مفید خواهد بود. همچنین دارای کاربردهای بالقوه‌ای در نظریه مقدار کرانگین است.

#### مراجع

- [1] Bercher, J. F., On some entropy functionals derived from Rényi information divergence, *Inform. Sci.*, **178** (2008), 2489–2506.
- [2] Bickel, P. J., and Lehmann, E. L., Descriptive statistics for nonparametric models: I. Introduction, *Ann. Statist.*, **3** (1975), 1038–1044.
- [3] Bobkov, S., Madiman, M., Concentration of the information in data with log-concave distributions, *Annal. Prob.*, **39** (2011), 1528–1543.
- [4] Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.
- [5] Feller, W., *An introduction to Probability Theory and Its Applications*, vol. 2, 2nd edn., John Wiley & Sons, New York, 1971.

- [6] Golshani, L., Pasha, E., Yari, G., Some properties of Rényi entropy and Rényi entropy rate, *Inform. Sci.*, **179** (2009), 2426–2433.
- [7] Gradshteyn, I. S., Ryzhik, I. M., *Table of Integrals, Series, and Products*, 7th edn. San Diego, Academic Press, 2007.
- [8] Hart, P. E., Moment distributions in economics: an exposition, *J. Roy. Statist. Soc. Ser. A*, **138** (1975), 423–434.
- [9] Mardia, K. V., Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57** (1970), 519–530.
- [10] Morales, D., Pardo, L., Vajda, I., Some new statistics for testing hypotheses in parametric models, *J. Multivar. Anal.*, **62** (1997), 137–168.
- [11] Nadarjah, S., Zografos, K., Formulas for Rényi information and related measures for univariate distributions, *Inform. Sci.*, **155** (2003), 119–138.
- [12] Nadarjah, S., Zografos, K., Expressions for Rényi and Shannon entropies for bivariate distributions, *Inform. Sci.*, **170** (2005), 173–189.
- [13] Nolan, J. P., *Stable Distributions: Models for Heavy Tailed Data*, Birkhauser, Boston, 2013.
- [14] Rényi, A., On Measures of entropy and information, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. I, University of California Press, 1961.
- [15] Song, K. S., Rényi information, loglikelihood and an intrinsic distribution measure, *J. Statist. Plann. Inference*, **93** (2001), 51–69.
- [16] van Zwet, W. R., *Convex Transformations of Random Variables*, Math. Centrum. Amsterdam, 1964.
- [17] Wang, L., Madiman, M., Beyond the entropy power inequality, via rearrangements, *IEEE Trans. Inform. Theory*, **60** (2014), no. 9, 5116–5137.

---

تاریخ ارسال: ۱۳۹۷/۷/۲۵؛ تاریخ بازنگری: ۱۳۹۸/۷/۲۹؛ تاریخ پذیرش: ۱۳۹۸/۷/۳۰

فرانک گودرزئی: کاشان، دانشگاه کاشان، دانشکده علوم ریاضی، گروه آمار

تارنما: <https://faculty.kashanu.ac.ir/goodarzi>

رایانامه: [f-goodarzi@kashanu.ac.ir](mailto:f-goodarzi@kashanu.ac.ir)

## Varentropy and Comparing Heavy Tailed Distributions

F. Goodarzi

Department of Statistics, Faculty of Mathematics, University of Kashan, Iran

**Abstract.** In this article, the Rényi information of order for a continuous random variable and its relationship with varentropy are studied. Also, varentropy is used as a measure of the shape of a distribution, allowing for comparison of different tails and shapes of densities. This is particularly useful when the usual skewness measure is not applicable. Among these distributions, we can mention many broad-tailed distributions such as T-Student distribution with any degree of freedom, Cauchy, Kramer, Levy and Burr. Also, in multivariate mode, we can refer to normal, t and Dirichlet distributions.

---

*Keywords:* heavy tailed, kurtosis, Rényi information, varentropy

*Article history:* Received 17 October 2018; Accepted 22 October 2019

*Article type:* review

---

## References

- [1] Bercher, J. F., On some entropy functionals derived from Rényi information divergence, *Inform. Sci.*, **178** (2008), 2489–2506.
- [2] Bickel, P. J., and Lehmann, E. L., Descriptive statistics for nonparametric models: I. Introduction, *Ann. Statist.*, **3** (1975), 1038–1044.
- [3] Bobkov, S., Madiman, M., Concentration of the information in data with log-concave distributions, *Annal. Probab.*, **39** (2011), 1528–1543.
- [4] Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.
- [5] Feller, W., *An introduction to Probability Theory and Its Applications*, vol. 2, 2nd edn., John Wiley Sons, New York, 1971.
- [6] Golshani, L., Pasha, E., Yari, G., Some properties of Rényi entropy and Rényi entropy rate, *Inform. Sci.*, **179** (2009), 2426–2433.

---

<sup>1</sup>fgoodarzi@kashanu.ac.ir



- [7] Gradshteyn, I. S., Ryzhik, I. M., *Table of Integrals, Series, and Products*, 7th edn. San Diego, Academic Press, 2007.
- [8] Hart, P. E., Moment distributions in economics: an exposition, *J. Roy. Statist. Soc. Ser. A*, **138** (1975), 423–434.
- [9] Mardia, K. V., Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57** (1970), 519–530.
- [10] Morales, D., Pardo, L., Vajda, I., Some new statistics for testing hypotheses in parametric models, *J. Multivar. Anal.*, **62** (1997), 137–168.
- [11] Nadarjah, S., Zografos, K., Formulas for Rényi information and related measures for univariate distributions, *Inform. Sci.*, **155** (2003), 119–138.
- [12] Nadarjah, S., Zografos, K., Expressions for Rényi and Shannon entropies for bivariate distributions, *Inform. Sci.*, **170** (2005), 173–189..
- [13] Nolan, J. P., *Stable Distributions: Models for Heavy Tailed Data*, Birkhauser, Boston, 2013.
- [14] Rényi, A., On Measures of entropy and information, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. I, University of California Press, 1961.
- [15] Song, K. S., Rényi information, loglikelihood and an intrinsic distribution measure, *J. Statist. Plann. Inference*, **93** (2001), 51–69.
- [16] van Zwet, W. R., *Convex Transformations of Random Variables*, Math. Centrum. Amsterdam, 1964.
- [17] Wang, L., Madiman, M., Beyond the entropy power inequality, via rearrangements, *IEEE Trans. Inform. Theory*, **60** (2014), no. 9, 5116–5137.