توزیع‌های لوگ- مقعر[۱] و تکنمایی[۲]

حمزه ترابی[۳]

چکیده

تابع های لوگ-مقعر و تکنمایی دارای خواص ریاضی جالبی هستند و در آمار واحتمال نقش به سزایی را بازی می کنند. در این مقاله نشان می‌دهیم که تابع چگالیهای نرمال چند متغیره، گاما و دیریکله تحت شرایط خاصی، لوگ- مقعر و تکنمایی هستند و سپس با توجه به ویژگی های این توابع، چند نامساوی جالب را اثبات می کنیم.

واژه‌های کلیدی: لوگ-مقعر، تکنمائی، نرمال چند متغیره، قضیه اندرسن.

۱. مروری بر مفاهیم مقدماتی

در این بخش، پس از تعریف مجموعه و تابع حـول مبـدا متقـارن، همچنین مجموعه و تابع محدب ، بـه بیـان دو قضیـه مـهم دربـاره توابـع محدب می‌پردازیم.

تعریف ۱.۱ : مجمـوعه A، $A \subseteq \Re^n$ را حـول مبـدأ متقـارن[٤] گوییم هر گاه برای هر $x \in A$ داشته باشیم $-x \in A$.

تعریف ۲.۱ : تـابع $f : \Re^n \to \Re$ را روی مجموعـه حـول مبـدأ متقـارن A، $A \subseteq \Re^n$ متقـارن حـول مبـدأ گوییـم هـرگـاه بـرای هـر $x \in A$ داشته باشیم $f(-x) = f(x)$.

تعریف ۳.۱ : مجموعه A، $A \subseteq \Re^n$ را محدب[٥] گوییم هرگاه

$$\forall x, y \in A, \forall a \in [0,1] \Rightarrow ax + (1-a)y \in A.$$

تعریف ٤.۱ : تابع $f : \Re^n \to \Re$ را روی مجموعه محـدب A، $A \subseteq \Re^n$، محدب (مقعر[٦]) گوییم هرگاه

$$\forall x, y \in B, \forall \alpha \in [0,1] \Rightarrow$$

$$f(\alpha x + (1-\alpha)y) \leq (\geq) \alpha f(x) + (1-\alpha) f(y).$$

اگر نامساوی‌های تعریف فوق به صورت اکید باشـد، تـابع را اکیـدا محدب(اکیدا مقعر) می گوییم.

بررسی محدب بودن توابع، معمولا کار ساده‌ای نیست؛ از این جهت از دو قضیه زیر که در این جا بدون اثبات بیان می‌شوند، می‌توان استفاده کرد.

قضیه ۱.۱: اگر $f : \Re^n \to \Re$ در A مشتق پذیرباشد؛ یعنی

---

[۱] Log – Concave

[۲] Unimodal

[۳] دانشجوی دکتری، بخش آمار، دانشگاه شیراز

[٤] Symmetric about Origin

[٥] Convex

[٦] Concave

گرادیان[۷] آن، $\nabla f = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})'$ درهرنقطه A موجود و متناهی باشد، آنگاه f در A محدب (مقعر) است اگر و تنها اگر

$$f(y) \geq (\leq) f(x) + \nabla f(x)'(y-x), \qquad \forall x, y \in A.$$

قضیه ۲.۱: اگر $f : \Re^n \to \Re$ در A دوبار مشتق پذیر باشد؛ یعنی ماتریس هسین[۸] آن، $\nabla^2 f = [\partial^2 f / (\partial x_i \partial x_j)]$ در هر نقطه A موجود و متناهی باشد، آنگاه f در A محدب (مقعر) است اگر و تنها اگر $\nabla^2 f$ نیم مثبت معین[۹] (نیم منفی معین [۱۰]) باشد.

با توجه به قضایای بالا به سادگی می توان ثابت کرد که،

۲- تابع e^{ax} در $\Re$ محدب است؛

۲- تابع x^a در $(0,\infty)$ در حالت $a \leq 0$ یا $a \geq 1$ محدب و در حالت $0 \leq a \leq 1$ مقعراست؛

۳- تابع $|x|^p$ در $\Re$ در حالت $p \geq 1$، محدب است؛

٤- تابع $\ln x$ در $(0,\infty)$ مقعر است؛

۵- تابع $x \ln x$ در $(0,\infty)$ محدب است؛

٦- هر نرمی در $\Re^n$ محدب است؛

۷- تابع $\max\{x_1, \dots, x_n\}$ در $\Re^n$ محدب است؛

۸- تابع x^2/y در حالت $x \in \Re$ و $y > 0$ محدب است؛

۹- تابع $\ln(\sum_{i=1}^n e^{x_i})$ در $\Re^n$ محدب است؛

۱۰- میانگین هندسی، یعنی $(\prod_{i=1}^n x_i)^{1/n}$ در $(0,\infty)^n$ مقعراست؛

۱۱- فرض کنید $g : \Re \to \Re$ و $h : \Re \to \Re$ و $f(x) = h(g(x))$ اگر

الف) g محدب و h محدب و غیرنزولی باشد، آنگاه f محدب است،

ب) g مقعر و h محدب و غیرصعودی باشد، آنگاه f محدب است،

پ) g مقعر و h مقعر و غیرنزولی باشد، آنگاه f مقعر است،

ت) g محدب و h مقعر و غیرصعودی باشد، آنگاه f مقعر است،

۲. لوگ- مقعری و قضیه پرکوپا[۱۱]

در این بخش در آغاز تابع‌های لوگ- محدب و لوگ-مقعر راتعریف، و درباره این توابع بدون اثبات نکاتی را بیان می کنیم سپس نشان می‌دهیم که تابع چگالی نرمال چند متغیره، گاما و دیریکله لوگ - مقعر هستند. در پایان به قضیه‌ای مهم که به قضیه پرکوپا معروف است، اشاره می کنیم.

تعریف۲.۱: تابع $f : \Re^n \to \Re$ را روی مجموعه محدب A، $A \subseteq \Re^n$ لوگ- محدب (لوگ- مقعر) گوییم هرگاه

$$\forall x, y \in B, \forall \alpha \in [0,1] \Rightarrow$$
$$f(\alpha x + (1-\alpha)y) \leq (\geq) [f(x)]^{\alpha} [f(y)]^{1-\alpha}$$

روشن است که اگر $f(x) > 0$، برای هر $x \in A$ تعریف۱.۲ هم‌ارز این است که تابع $\ln f(x)$ محدب (مقعر) باشد.

به سادگی می‌توان ثابت کرد که

۱. تابع لجستیک، یعنی $e^x/(1+e^x)$ روی $\Re$ لوگ-مقعر است؛

۲. لوگ - محدبی و لوگ- مقعری تحت ضرب اسکالر بسته است اما لزوما نسبت به جمع، تفریق و تقسیم بسته نیست؛

۳. اگر $f : \Re^n \to \Re$ لوگ- مقعر باشد، آنگاه به شرط $a \geq 0$، تابع $g = f - a$ روی مجموعه $\{x \in D_f \mid f(x) > a\}$ لوگ- مقعر است؛

٤. تابع گاما، $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ روی $x \geq 1$ لوگ- محدب است؛

۵. تابع $f(x) = \prod x_i / \sum x_i$ در $(0,\infty)^n$ لوگ- مقعر است؛

٦. تابع e^{ax} هم لوگ- محدب و هم لوگ- مقعر است؛

۷. تابع x^a در $(0,\infty)$ به شرط $a \leq 0$ لوگ-محدب و به شرط $a \geq 0$ لوگ-مقعر است؛

قضیه ۲.۱: تابع چگالی $N_n(\mu, \Sigma)$ به شرط $\Sigma > 0$ در $\Re^n$ لوگ - مقعر است.

---

[۷] Gradient

[۸] Hessian

[۹] Positive Semi - definite

[۱۰] Negative Semi- definite

[۱۱] Prekopa

اثبات: می‌دانیم که تابع چگالی $N_n(\mu,\Sigma)$ به صورت زیر است

$$f(x;\mu,\Sigma)=(2\pi)^{-\frac{n}{2}}\,|\Sigma|^{-\frac{1}{2}}\exp\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\}$$

چون $f(x)>0$ برای هر $x\in\Re^n$ کافی است ثابت کنیم برای هر x و y دلخواه در $\Re^n$ و هر مقدار دلخواه $\alpha\in[0,1]$ داریم،

$$\ln f(\alpha x+(1-\alpha)y;\mu,\Sigma)-[\alpha\ln f(x;\mu,\Sigma)+(1-\alpha)\ln f(y;\mu,\Sigma)]\geq 0$$

اما

$$\ln f(\alpha x+(1-\alpha)y;\mu,\Sigma)-[\alpha\ln f(x;\mu,\Sigma)+(1-\alpha)\ln f(y;\mu,\Sigma)]=\frac{1}{2}\alpha(1-\alpha)[(x-\mu)'\Sigma^{-1}(x-\mu)+(y-\mu)'\Sigma^{-1}(y-\mu)-2(x-\mu)'\Sigma^{-1}(y-\mu)]=\frac{1}{2}\alpha(1-\alpha)[(x-y)'\Sigma^{-1}(x-y)]\geq 0$$

نابرابری اخیر به دلیل نامنفی بودن $\alpha(1-\alpha)/2$ و مثبت معین بودن Σ^{-1} نتیجه شده است.

قضیه ۲.۲: تابع چگالی گاما با پارامترهای α و β $(G(\alpha,\beta))$ به شرط $\alpha\geq 1$ در $(0,\infty)$ لوگ-مقعر است.

اثبات: داریم

$$\ln f(x;a,\beta)=-\frac{x}{\beta}+(\alpha-1)\ln x-\ln(\beta^{\alpha}\Gamma(\alpha))$$

بنابراین

$$\frac{\partial\ln f(x;a,\beta)}{\partial x}=-\frac{1}{\beta}+\frac{\alpha-1}{x}\Rightarrow$$

$$\frac{\partial^2\ln f(x;a,\beta)}{\partial x^2}=-\frac{\alpha-1}{x^2}\leq 0$$

با توجه به قضیه ۱.۲ با در نظر گرفتن $n=2$، حکم ثابت می شود.

قضیه ۳.۲: تابع چگالی توزیع دیریکله[14] (بتای چند متغیره) با پارامترهای $\lambda_1,\ldots,\lambda_{n+1}$ $(D(\lambda_1,\ldots,\lambda_{n+1}))$ که به صورت

$$f(x)=\frac{\Gamma(1'\lambda)}{\Gamma(\lambda_1)\ldots\Gamma(\lambda_{n+1})}x_1^{\lambda_1-1}\ldots x_n^{\lambda_n-1}(1-\sum_{i=1}^{n}x_i)^{\lambda_{n+1}-1}$$

$i=1,\ldots,n,x_i>0$ ، $1'x\leq 1$ ، $i=1,\ldots,n+1$ ، $\lambda_i>0$ تعریف می‌شود، به شرط $\lambda_i\geq 1$ ، $i=1,\ldots,n+1$ روی مجموعه $\{x\in(0,\infty)^n\mid 1'x\leq 1\}$ لوگ-مقعر است.

اثبات: داریم

$$\ln f(x)=\ln\left[\frac{\Gamma(1'\lambda)}{\Gamma(\lambda_1)\ldots\Gamma(\lambda_{n+1})}\right]+\sum(\lambda_i-1)\ln x_i+(\lambda_{n+1}-1)\ln(1-\sum x_i)$$

فرض کنید

$$g_1(x)=\sum(\lambda_i-1)\ln x_i$$

و

$$g_2(x)=(\lambda_{n+1}-1)\ln(1-\sum x_i)$$

با توجه به مقعر بودن تابع ln و نامنفی بودن λ_i-1 به سادگی دیده می‌شود که هر g_1 و g_2 هردو مقعر هستند. در نتیجه بنابر بسته بودن توابع مقعر تحت جمع، حکم ثابت می شود.

ترکیب محدب چند مجموعه، به صورت زیر تعریف می شود

تعریف ۲.۲: فرض کنید $A_1,\ldots,A_k$ زیر مجموعه‌های دلخواهی از $\Re^n$ باشند. ترکیب محدب این k مجموعه را با وزن های $\alpha_1,\ldots,\alpha_k$ ، با شرط $\alpha_i\geq 0$ و $\sum_{i=1}^{n}\alpha_i=1$ به صورت

$$\sum_{i=1}^{k}a_iA_i=\left\{z\mid z\in\Re^n,z=\sum_{i=1}^{k}\alpha_ix_i,\exists x_i\in A_i,i=1,\ldots,k\right\}$$

تعریف می کنیم.

پرکوپا (۱۹۷۱) نشان داده است که

قضیه ۴.۲: فرض کنید $X=(X_1,\ldots,X_n)'$ دارای تابع چگالی احتمال $f(x)$ باشد و $A_1,\ldots,A_k$ زیرمجموعه‌های بورل دلخواهی از $\Re^n$ باشند و $\alpha_i\geq 0$ و $\sum_{i=1}^{k}\alpha_i=1$ آنگاه

---

[14] Dirichlet

$$P(X \in \sum_{i=1}^{k} \alpha_i A_i) \geq \prod_{i=1}^{k} \{P(X \in A_i)\}^{\alpha_i} \quad (۲.۱)$$

نتیجه ۲.۱: تابع چگالی نرمال چند متغیره، به شرط $\Sigma > 0$، تابع چگالی $G(\alpha, \beta)$ به شرط $\alpha \geq 1$ و تابع چگالی $D(\lambda_1, \ldots, \lambda_{n+1})$ به شرط $\lambda_i \geq 1$، $i = 1, \ldots, n+1$، در رابطه (۲.۱) صدق می‌کنند.

اثبات: با توجه به قضیه های ۲. ۱، ۲. ۲، ۲. ۳، ۲. ۴، حکم نتیجه می‌شود.

برای روشن شدن موضوع به مثال‌های زیر توجه کنید.

مثال ۲.۱: فرض کنید متغیر تصادفی X دارای توزیع نرمال با پارامترهای μ و σ^2 باشد. همچنین فرض کنید a, b, c, d مقادیر حقیقی دلخواهی باشند. در این صورت داریم

$$p(\frac{a+c}{2} < X < \frac{b+d}{2}) \geq \sqrt{p(a < X < b) \times p(c < X < d)}. \quad (۲.۲)$$

اثبات: تعریف می‌کنیم $A_1 = (a,b)$ و $A_2 = (c,d)$ داریم

$$\frac{1}{2}(A_1 + A_2) = (\frac{a+c}{2}, \frac{b+d}{2}).$$

بنابر نتیجه ۲.۱ با $k = 2$ رابطه به دست می‌آید.

رابطه (۲. ۲) مثال فوق، برای توزیع $G(\alpha, \beta)$ به شرط $\alpha \geq 1$ و $a, c > 0$ همچنین برای توزیع $Beta(\alpha, \beta)$ به شرط $\alpha, \beta \geq 1$ و $0 < a < b < 1$ و $0 < c < d < 1$ بر قرار است.

مثال ۲. ۲: فرض کنید $X = (X_1, X_2)'$ دارای توزیع $N_2(\mu, \Sigma)$ باشد به طوری که $\mu_1 = \mu_2$، $\sigma_1^2 = \sigma_2^2$ و $|\rho| < 1$ در این صورت برای هر $(a_1, a_2)' \in \Re^2$ تابع توزیع X در رابطه زیر صدق می‌کند.

$$F_{X_1,X_2}(\overline{a}, \overline{a}) \geq F_{X_1,X_2}(a_1, a_2), \quad \overline{a} = (a_1 + a_2)/2$$

اثبات: تعریف می‌کنیم $A_1 = \{(u_1, u_2) \mid u_1 \leq a_1, u_2 \leq a_2\}$ و $A_2 = \{(u_1, u_2) \mid u_1 \leq a_2, u_2 \leq a_1\}$ چون $F_{X_1,X_2}(a_1, a_2) = F_{X_1,X_2}(a_2, a_1)$ (به این ویژگی، متقارن - جایگشت بودن X گفته می‌شود) و همچنین

$$\frac{1}{2}A_1 + \frac{1}{2}A_2 = \{z = (z_1, z_2) \mid (z_1, z_2) = \frac{1}{2}(x_1, x_2) + \frac{1}{2}(y_1, y_2), (x_1, x_2) \in A_1, (y_1, y_2) \in A_2\} = \{(z_1, z_2) \mid z_1 \leq \overline{a}, z_2 \leq \overline{a}\},$$

با توجه به نتیجه ۲.۱ رابطه ثابت می شود.

می‌توان مثال‌های فوق را به صورت زیر تعمیم داد.

مثال ۲.۳: فرض کنید X دارای توزیع نرمال $N_n(\mu, \Sigma)$ باشد به طوری که متقارن- جایگشت باشد؛ یعنی

$$\rho_{ij} = \rho \in (-\frac{1}{n-1}, 1), \sigma_1^2 = \ldots = \sigma_n^2, \mu_1 = \ldots = \mu_n$$

(محدودیت ρ مربوط به شرط $\Sigma > 0$ است) داریم

$$F_{X_1,\ldots,X_n}(\overline{a}, \ldots, \overline{a}) \geq F_{X_1,\ldots,X_n}(a_1, \ldots, a_n) \quad \forall a_i \in \Re, i = 1, \ldots, n.$$

۳. تکنمایی بودن و قضیه اندرسون[15]

در این بخش خواهیم دید که ویژگی تکنمایی بودن تابع چگالی با مفهوم تحدب در ارتباط است. تعریف زیر را از توابع تکنمایی، اندرسون (۱۹۹۵) ارائه داده است.

تعریف۳.۱: هرتابع چگالی $f : \Re^n \to [0, \infty)$ را تکنمایی (شبه- مقعر[16])گوییم هرگاه مجموعه $A_\lambda = \{x \mid f(x) \geq \lambda\}$، برای هر $\lambda > 0$، محدب باشد.

قضیه ۳.۱: هر تابع لوگ- مقعر، تکنمایی است.

اثبات: فرض کنید $\lambda > 0$ همچنین x ,y در A_λ دلخواه هستند. برای هر $\alpha \in [0,1]$ داریم

$$f(\alpha x + (1-\alpha)y) \geq [f(x)]^\alpha [f(y)]^{1-\alpha} \geq \lambda^\alpha \lambda^{1-\alpha} = \lambda.$$

بنابراین $\alpha x + (1-\alpha)y \in A_\lambda$.

باتوجه به قضیه‌های ۲. ۱، ۳. ۱، داریم

[15] Anderson

[16] quasi-concave

نتیجه ۱.۳: تابع چگالی $N_n(\mu,\Sigma)$ با شرط $\Sigma > 0$ تابع چگالی $G(\alpha,\beta)$ به شرط $\alpha \ge 1$ و تابع چگالی $D(\lambda_1,\ldots,\lambda_{n+1})$ به شرط $\lambda_i \ge 1$، $i = 1,\ldots,n+1$ تکنمایی هستند.

قضیه زیر که منسوب به اندرسون است، در رابطه با ویژگی‌های یکنوایی انتگرال احتمال چگالی‌های متقارن و تکنمایی روی مجموعه‌های متقارن حول صفر و محدب است. اثبات این قضیه در اندرسون (۱۹۸۸) یافت می‌شود.

قضیه ۲.۳: فرض کنید که X دارای تابع چگالی $f(x)$ باشد. برای هر $B \subseteq \Re^n$ و هر مقدار دلخواه ثابت و غیر صفر $u \in \Re^n$ تعریف می‌کنیم

$$B + cu = \{x \mid x = y + cu, \exists y \in B\}, \qquad c \in \Re.$$

اگر f حول صفرمتقارن تکنمایی باشد و B نیز حول مبدأ متقارن و محدب باشد، آنگاه انتگرال احتمال $\int_{B+cu} f(x)dx$ در $|c|$ غیر صعودی است. علاوه بر این اکیداً نزولی است مگر این که یک $\nu \in \Re^n$ موجود باشد به طوری که

$$(B+u) \cap A_\lambda = (B \cap A_\lambda) + \nu, \qquad \forall \lambda > 0$$

با توجه به قضیه های ۳. ۱ و ۳. ۲ داریم

قضیه ۳.۳: اگر X دارای توزیع $N_n(0,\Sigma)$، $\Sigma > 0$ باشد و $B \subseteq \Re^n$ حول مبدأ متقارن و محدب باشد، آنگاه $P(X \in B + cu)$ برای هر $u \in \Re^2$ دلخواه، یک تابع غیر صعودی بر حسب $|c|$ است.

نتیجه ۲.۳: اگر X و Y به ترتیب دارای توزیع‌های $N_n(0,\Sigma_1)$ و $N_n(0,\Sigma_2)$ باشند که $\Sigma_i > 0$، $i = 1,2$ و $\Sigma_2 - \Sigma_1$ مثبت معین و $B \subseteq \Re^n$ حول مبدا متقارن و محدب باشد، آنگاه

$$P(X \in B) \ge P(Y \in B).$$

اثبات: فرض کنید U دارای توزیع $N_n(0,\Sigma_2 - \Sigma_1)$ و مستقل از X باشد. Y و X+U هم توزیعند. بنابراین با توجه به تعریف تابع نشانگر و رابطه امید ریاضی مکرر داریم

$$\begin{aligned} P(Y \in B) &= P(X + U \in B) = P(X \in B - U) \\ &= E[I(X \in B - U)] = E[E[I(X \in B - U) \mid U]] \end{aligned}$$

اما با استفاده از قضیه ۳. ۲ خواهیم داشت

$$\begin{aligned} E[I(X \in B - U) \mid U = u] &= E[I(X \in B - u) \mid U = u\} \\ &= E[I(X \in B - u)] = P(X \in B - u) \\ &\le P(X \in B + 0u) = P(X \in B). \end{aligned}$$

مراجع

[1] Anderson, T.W., 1955. *The Integral of a Symmetric Unimodal Function Over a Symmetric Convex Set and Some Probability Inequalities,* Proc. Amer. Math. Soc., Vol.6, pp. 170- 176.

[2] Anderson, T.W., 1988. Personal Communication.

[3] Hollander, M., Proschan, F., and Sethurman, J., 1977. *Functions Decreasing in Transformation and Their Applications in Ranking Problem,* Ann. Stat., Vol.63, pp. 722-733.

[4] Marshal, A.W., and Olkin, I., 1974. *Majorization multivariate distributions,* Ann. Stat., Vol.2, pp. 1189- 1200.

[5] Marshal, A.W., and Olkin, I., 1979. *Inequalities: Theory of Majorization and its Applications,* Academic Press, New York.

[6] Prekopa, A., 1971. *Logarithmic Concave Measure With Applications,* Acta. Sci. Math., Vol.32, pp. 302- 316.

[7] Tong, Y.L., 1990. *The Multivariate Normal Distribution,* Springer-Verlag.