تعیین تابع توزیع احتمال توام دو متغیر بطور تصادفی مرتب شده

شهرام منصوری
دانشگاه صنعت آب و برق شهید عباسپور

در این مقاله روشی برای بدست آوردن تابع توزیع احتمال توام دو متغیره به طور تصادفی مرتب شده با معلوم بودن توزیع‌های حاشیه‌ای و ضریب همبستگی ارائه شده و نحوه اجرای آن در مثال‌هایی توضیح داده شده است.

واژه‌های کلیدی: متغیرهای به طور تصادفی مرتب شده، اصل ماکزیمم آنتروپی، نظریه اطلاع.

۱ مقدمه

بفرض X و Y متغیرهای تصادفی با توابع چگالی احتمال $G(x)$ و $H(y)$ باشند، گوییم X و Y متغیرهای بطور تصادفی مرتب شده هستند هرگاه $X \leq Y \quad a.s.$ باشد. بدیهی است که در اینصورت به ازای هر z, $G(z) \leq H(z)$ است.

می‌دانیم توابع چگالی احتمال توام مختلفی می‌توانند دارای توزیع‌های احتمال حاشیه‌ای مشابهی باشند. در این مقاله با رهیافت توزیع اصل ماکزیمم آنتروپی، بهترین توزیع توام را معین می‌کنیم. تابع توزیع ماکزیمم آنتروپی یک متغیره و چند متغیره تحت برخی قیود توسط بسیاری از محققان مانند جینز (۱۹۶۸) و کاپور (۱۹۸۳) مورد مطالعه قرار گرفته است. منصوری و همکاران (۲۰۰۵)، یک فرم کلی برای تابع توزیع احتمال چند متغیره ماکزیمم آنتروپی با شرط معلوم بودن توابع چگالی حاشیه‌ای و ماتریس کوواریانس متغیرهای تصادفی ارائه کردند. توزیع توام ماکزیمم آنتروپی برای متغیرهای بطور تصادفی مرتب شده توسط استراسن (۱۹۶۵) و کامایی و همکاران (۱۹۷۷) مورد مطالعه قرار گرفته است. کیفیر (۲۰۰۹)، روشی برای بدست آوردن توزیع ماکزیمم آنتروپی برای داده های بطور تصادفی مرتب شده ارائه داد. در بخش ۲ این مقاله با روشی دیگر به همان نتایج کیفیر (۲۰۰۹)، می‌رسیم. در بخش ۳ با ارائه دو مثال روش ارائه شده توضیح داده می‌شود. بحث و نتیجه‌گیری در بخش چهارم بیان شده است.

۲ توزیع توام احتمال ماکزیمم آنتروپی دو متغیره به طور تصادفی مرتب شده

در این بخش نتایج لازم برای تعیین تابع توزیع احتمال ماکزیمم آنتروپی بیان می‌شود.
لم ۱: اگر L^2 مجموعه توابع اندازه پذیر باشد که مربع آنها انتگرال پذیر است، آنگاه برای هر دو تابع h_1 و h_2 در L^2، $h_1 = h_2,\ a.e.$ است اگر و تنها اگر برای هر تابع دلخواه K تساوی زیر برقرار باشد:

$$\int_R K(x)h_1(x)dx = \int_R K(x)h_2(x)dx \qquad (1)$$

برهان: اگر $h_1 = h_2,\ a.e.$ باشد، بدیهی است که برای هر تابع دلخواه K (1) برقرار است. برعکس از (1) داریم:

$$\int_R K(x)(h_1(x) - h_2(x))dx = 0,$$

چون K دلخواه است، بازاء $K(x) = h_1(x) - h_2(x)$ داریم:

$$\int_R (h_1(x) - h_2(x))^2 dx = 0,$$

بنابراین

$$h_1 = h_2; \quad a.e.$$

تعریف ۱ آنتروپی دو متغیر تصادفی پیوسته (X,Y) با تابع چگالی احتمال $f(x,y)$ که با نماد $H(X,Y)$ نشان داده می‌شود، به صورت زیر تعریف شده است.

$$H(X,Y) = E(-\ln f(X,Y)) = -\int\int_{R^2} f(x,y)(\ln f(x,y))dxdy. \qquad (2)$$

قضیه ۱ اگر $g,\ h \in L^{2}$ توابع چگالی حاشیه‌ای یک کلاس از توزیع احتمال توام دو متغیر بطور تصادفی مرتب شده بصورت $X \leq Y \quad a.s.$ باشند، آنگاه تابع چگالی احتمال ماکزیمم آنتروپی دارای فرم کلی

$$f_M(x,y) = f_1(x) f_2(y) I(x \leq y) \tag{3}$$

است، که در آن f_1 و f_2 از حل دستگاه معادلات تابعی

$$\begin{cases} \int_R f_1(x) f_2(y) I(x \leq y) dy = g(x) \\ \int_R f_1(x) f_2(y) I(x \leq y) dx = h(y) \end{cases} \tag{4}$$

بدست می آیند.

برهان: چون توابع چگالی حاشیه‌ای $f(x,y)$ معلومند بنابراین برای توابع دلخواه N و M

$$\mu_1 = E(M(X)), \quad \mu_2 = E(N(Y)) \tag{5}$$

معلوم خواهند بود. با توجه به لم ۱ معلوم بودن h و g معادل معلوم بودن $E(N(Y))$ و $E(M(X))$ است. لذا برای تعیین تابع چگالی احتمال ماکزیمم آنتروپی تحت شرایط *(5)* می‌بایست *(2)* را ماکزیمم نماییم. حال با استفاده از روش اویلر- لاگرانژ (کاپور۱۹۸۹)، لاگرانژین بصورت

$$\begin{aligned} L &= \int\int_{R^2} f(-\ln f) dxdy \\ &+\lambda_1\left(\int\int_{R^2} M(x) f dxdy - \mu_1\right) + \lambda_2\left(\int\int_{R^2} N(y) f dxdy - \mu_2\right) \\ &= \int\int_{R^2} \{-f\ln f + (\lambda_1 M(x) + \lambda_2 N(y)) f\} dxdy - (\lambda_1\mu_1 + \lambda_2\mu_2). \end{aligned}$$

است. چون $(-f)\ln f + (\lambda_1 M(x) + \lambda_2 N(y))) f$ تابعی مقعر و پیوسته بر حسب f است جواب بهینه، بطور یکتا آنتروپی را ماکزیمم می کند. لذا تابع چگالی احتمال ماکزیمم آنتروپی بیزی جواب معادله

$$\frac{\partial}{\partial f}\{(-f)\ln f + (\lambda_1 M(x) + \lambda_2 N(y)) f\} = 0$$

است. بنابراین

$$f_M(x,y) = e^{\lambda_1 M(x)} e^{\lambda_2 N(y)} I(x \leq y)$$

واضح است که f_M حاصلضرب دو تابع یعنی یکی بر حسب x و دیگری بر حسب y است. لذا

$$f_M(x,y) = f_۱(x)f_۲(y)I(x \leq y)$$

از آنجایی که h و g توابع چگالی حاشیه‌ای می‌باشند، لذا $f_۱$ و $f_۲$ از *(4)* قابل حصول می‌باشند.

قضیه ۲ اگر $h \in L^۲$ و g توابع چگالی حاشیه‌ای یک کلاس از توزیع احتمال توام دو متغیر بطور تصادفی مرتب شده بصورت $X \leq Y \quad a.s.$ و با ضریب همبستگی ρ باشند آنگاه تابع چگالی احتمال ماکزیمم آنتروپی دارای فرم کلی

$$f_M(x,y) = f_۱(x)f_۲(y)e^{\lambda xy}I(x \leq y), \qquad (۶)$$

است که در آن λ، $f_۱$ و $f_۲$ از حل دستگاه معادلات

$$\begin{cases} \int_R f_۱(x)f_۲(y)e^{\lambda xy}I(x \leq y)dy = g(x) \\ \int_R f_۱(x)f_۲(y)e^{\lambda xy}I(x \leq y)dx = h(y) \\ \int\int_{R^۲} xyf_۱(x)f_۲(y)e^{\lambda xy}I(x \leq y)dxdy = \rho\sigma_X\sigma_Y + \mu_x\mu_Y \end{cases} \qquad (۷)$$

بدست می آیند.

برهان: چون توابع چگالی حاشیه‌ای $f(x,y)$ معلومند پس برای توابع دلخواه N وM

$$\mu_۱ = E(M(X)),\ \mu_۲ = E(N(Y)) \qquad (۸)$$

معلوم هستند. با توجه به لم ۱ معلوم بودن h وg معادل معلوم بودن $E(N(Y))$ و $E(M(X))$ است. چون ضریب همبستگی مقدار معلوم ρ است پس

$$EXY = \rho\sigma_X\sigma_Y + \mu_X\mu_Y = C \qquad (۹)$$

معلوم می‌باشد. حال برای تعیین تابع چگالی احتمال ماکزیمم آنتروپی تحت شرایط *(8)* و *(9)* می‌بایست *(1)* را ماکزیمم نماییم. با استفاده از روش اویلر- لاگرانژ، لاگرانژین بصورت

$$\begin{aligned} L &= \int\int_{R^۲} f(-\ln f dxdy + \lambda(\int\int_{R^۲} xyf dxdy - C) \\ &+\lambda_۱(\int\int_{R^۲} M(x)f dxdy - \mu_۱) + \lambda_۲(\int\int_{R^۲} N(y)f dxdy - \mu_۲) \\ &= \int\int_{R^۲} \{-f\ln f + (\lambda xy + \lambda_۱ M(x) + \lambda_۲ N(y))f\}dxdy \\ &-(\lambda C + \lambda_۱\mu_۱ + \lambda_۲\mu_۲). \end{aligned}$$

است. حال چون $(-f)\ln f+(\lambda xy+\lambda_1 M(x)+\lambda_2 N(y)))f$ تابعی مقعر و پیوسته بر حسب f است جواب بهینه، بطور یکتا آنتروپی را ماکزیمم می‌کند. لذا تابع چگالی احتمال ماکزیمم آنتروپی بیزی جواب معادله

$$\frac{\partial}{\partial f}\{(-f)\ln f+(\lambda xy+\lambda_1 M(x)+\lambda_2 N(y))f\}=0,$$

است. بنابراین

$$f_M(x,y)=e^{\lambda_1 M(x)}e^{\lambda_2 N(y)}e^{\lambda xy},\quad x\le y.$$

در نتیجه فرم تابع چگالی احتمال ماکزیمم آنتروپی بیزی بصورت

$$f_M(x,y)=f_1(x)f_2(y)e^{\lambda xy}I(x\le y)$$

است. از آنجایی که توابع چگالی حاشیه‌ای و ضریب همبستگی معلوم هستند بدیهی است که λ،f_1 وf_2 از (۶) قابل حصول می‌باشند.

۳ مثال‌ها

مثال ۱ فرض کنیم X و Y متغیرها به طور تصادفی مرتب شده بصورت $Y\le X\ \ a.s.$ باشند و بترتیب دارای توابع توزیع

$$G(x)=\begin{cases}0 & x\le 0\\ x^2 & 0\le x\le 1,\\ 1 & 1\le x\end{cases}\qquad H(y)=\begin{cases}0 & y\le 0\\ y & 0\le y\le 1\\ 1 & 1\le y\end{cases}$$

باشند، در اینصورت با توجه به قضیه ۱ تابع چگالی احتمال توام(X,Y) دارای فرم کلی زیر است:

$$f_M(x,y)=f_1(x)\,f_2(y)\,I\,(0\le y\le x\le 1)$$

با توجه به مفروضات توابع چگالی X و Y متغیرها بترتیب برابرند با:

$$g(x)=2xI(0\le x\le 1),\quad h(y)=I(0\le y\le 1)$$

با استفاده از رابطه (۴) $f_۱$ و $f_۲$ از حل دستگاه معادلات زیر بدست می آید:

$$\begin{cases} g(x)=\int_R f(x,y)dy=\int_۰^x f_۱(x)f_۲(y)dy \\ h(y)=\int_R f(x,y)dx=\int_y^۱ f_۱(x)f_۲(y)dx \end{cases} \Rightarrow \begin{cases} ۲x=f_۱(x)\int_۰^x f_۲(y)dy \\ ۱=f_۲(y)\int_y^۱ f_۱(x)dx \end{cases}$$

تعریف می کنیم $A(x)=\int_۰^x f_۲(y)dy$ و $B(y)=\int_y^۱ f_۱(x)dx$ در نتیجه

$$A'(x)=f_۲(x),\quad B'(x)=-f_۱(x) \qquad (۱۰)$$

پس

$$\begin{cases} A(x)B'(x)=۲x \\ -A'(x)B(x)=۱ \end{cases} \qquad (۱۱)$$

لذا

$$[A(x)B(x)]'=۲x-۱\Rightarrow A(x)B(x)=x^۲-x \qquad (۱۲)$$

از روابط (۱۱) و (۱۲) نتیجه می گیریم:

$$\frac{B'(x)}{B(x)}=\frac{۲x}{x^۲-x}=\frac{۲}{x-۱}\Rightarrow B(x)=C_۱(۱-x)^۲\Rightarrow B'(x)=-۲C_۱(۱-x)$$

به طور مشابه از (۱۱) و (۱۲) نتیجه می گیریم:

$$\frac{A'(x)}{A(x)}=-\frac{۱}{x^۲-x}\Rightarrow A(x)=C_۲\frac{x}{۱-x}\Rightarrow A'(x)=C_۲\frac{۱}{(۱-x)^۲}$$

با توجه به رابطه (۱۰) داریم:

$$\begin{aligned} f_۱(x)=-B'(x)&\Rightarrow f_۱(x)=-۲C_۱(۱-x) \\ f_۲(x)=A'(x)&\Rightarrow f_۲(x)=-C_۲\frac{۱}{(۱-x)^۲} \end{aligned}$$

لذا تابع چگالی احتمال توام دو متغیره ماکزیمم آنتروپی برابر است با:

$$\begin{aligned} f_M(x,y)&=f_۱(x)f_۲(y)I(۰\le y\le x\le ۱) \\ &=۲C_۱C_۲\frac{۱-x}{(۱-y)^۲}I(۰\le y\le x\le ۱) \\ &=C\frac{۱-x}{(۱-y)^۲}I(۰\le y\le x\le ۱) \end{aligned}$$

با توجه به آنکه $f_M(x,y)$ تابع چگالی احتمال است C از حل معادله $\int\int_{R^2} f_M(x,y)dxdy = 1$ ، بدست می آید. داریم:

$$C\int_0^1 dx \int_0^x \frac{1-x}{(1-y)^2}dy = 1 \Rightarrow C = 2$$

لذا تابع چگالی احتمال توام (X,Y) بصورت

$$f_M(x,y) = 2\frac{1-x}{(1-y)^2}I(0 \le y \le x \le 1)$$

بدست می آید. همچنین آنتروپی توام این دو متغیر بطور تصادفی مرتب شده برابر است با:

$$\begin{aligned} H(X,Y) &= E(-\ln f_M(X,Y)) = -E(\ln 2 + \ln(1-X) - 2\ln(1-Y)) \\ &= -\ln 2 - \int_0^1 2x\ln(1-x)dx + \int_0^1 2\ln(1-y)dy \\ &= -\ln 2 - 2\int_0^1 (1-x)\ln x dx + \int_0^1 2\ln x dx \\ &= -\ln 2 - 2\int_0^1 x\ln x dx = -\frac{1}{2} - \ln 2 \end{aligned}$$

مثال ۲ فرض کنیم X و Y متغیرها بطور تصادفی مرتب شده و بصورت $Y \le X \quad a.s.$ باشند و بترتیب دارای توزیع بتای $(3,3)$ و بتای $(1,3)$ باشند، در اینصورت فرم تابع چگالی احتمال توام (X,Y) بصورت

$$f_M(x,y) = f_1(x)f_2(y)I(0 \le y \le x \le 1)$$

است لذا توابع چگالی احتمال حاشیه‌ای بصورت

$$\begin{cases} g(x) = \int_R f_M(x,y)dy = \int_0^x f_1(x)f_2(y)dy \\ h(y) = \int_R f_M(x,y)dx = \int_y^1 f_1(x)f_2(y)dx \end{cases}$$

هستند، از طرفی با توجه به مفروضات مثال توابع چگالی احتمال برابرند با:

$$g(x) = 30x^2(1-x)^2, \quad 0 < x < 1; \qquad h(y) = 3(1-y)^2, \quad 0 < y < 1$$

لذا

$$\begin{cases} ۳۰x^۲(۱-x)^۲ = f_۱(x)\int_۰^x f_۲(y)dy \\ ۳(۱-y)^۲ = f_۲(y)\int_y^۱ f_۱(x)dx \end{cases}$$

تعریف می‌کنیم $A(x)=\int_۰^x f_۲(y)dy$ و $B(y)=\int_y^۱ f_۱(x)dx$ از این رو $A'(x)=f_۲(x)$ و $B'(x)=-f_۱(x)$ پس

$$\begin{cases} A(x)B'(x) = -۳۰x^۲(۱-x)^۲ \\ A'(x)B(x) = ۳(۱-x)^۲ \end{cases} \tag{۱۳}$$

در نتیجه $[A(x)B(x)]' = -۳۰x^۲(۱-x)^۲ + ۳(۱-x)^۲$ و از آنجا

$$A(x)B(x) = -۶x^۵ + ۱۵x^۴ - ۹x^۳ - ۳x^۲ + ۳x \tag{۱۴}$$

از (۱۳) و (۱۴) نتیجه می‌شود:

$$\frac{A'(x)}{A(x)} = \frac{۳(۱-x)^۲}{-۶x^۵ + ۱۵x^۴ - ۹x^۳ - ۳x^۲ + ۳x}$$

پس از انتگرال گیری داریم: $\ln A(x) = -\frac{۱}{۳}\ln(x-۱) + \ln x - \frac{۲}{۳}\ln(۲x+۱) + C'_۱$. لذا

$$A(x) = C_۱\frac{۱}{(۱-x)^{\frac{۱}{۳}}(۲x+۱)^{\frac{۲}{۳}}}$$

به طور مشابه از (۱۳) و (۱۴) نتیجه می‌شود:

$$\frac{B'(x)}{B(x)} = \frac{۳۰x^۲(۱-x)^۲}{-۶x^۵ + ۱۵x^۴ - ۹x^۳ - ۳x^۲ + ۳x}$$

پس از اندکی محاسبات داریم: $B(x) = C_۲(۱-x)^{\frac{۱}{۳}}(۲x+۱)^{\frac{۵}{۳}}$ حال مشتقات $A(x)$ و $B(x)$ را بدست می آوریم:

$$A'(x) = C_۱\frac{-۱}{(۱-x)^{\frac{۴}{۳}}(۲x+۱)^{\frac{۵}{۳}}}, \quad B'(x) = ۱۰C_۲x(۱-x)^{\frac{۷}{۳}}(۲x+۱)^{\frac{۲}{۳}}$$

لذا تابع چگالی احتمال توام (X,Y) بصورت

$$\begin{aligned} f_M(x,y) &= f_۱(x)f_۲(y)I(۰\le y\le x\le ۱) = -A'(y)B'(x)I(۰\le y\le x\le ۱) \\ &= C\frac{x(۱-x)^{\frac{۷}{۳}}(۲x+۱)^{\frac{۲}{۳}}}{(۱-y)^{\frac{۴}{۳}}(۲y+۱)^{\frac{۵}{۳}}}I(۰\le y\le x\le ۱) \end{aligned}$$

بدست می آید از آنجا که f_M یک تابع چگالی احتمال توام است $\int\int_{R^2} f_M(x,y)dxdy =$ ۱ و از آن $C =$ ۱/۹۱ بدست می آید. بنابراین

$$f_M(x,y) = 1/91 \frac{x(1-x)^{\frac{7}{3}}(2x+1)^{\frac{2}{3}}}{(1-y)^{\frac{4}{3}}(2y+1)^{\frac{5}{3}}} I(0 \leq y \leq x \leq 1)$$

۴ بحث و نتیجه‌گیری

قضیه‌های ۱ و ۲ برای توزیع‌های گسسته و توزیع‌های چند متغیره گسسته و پیوسته قابل تعمیم است. برای تعیین تابع ماکزیمم آنتروپی دو متغیره به روش عددی ممکن است از سایر روشها نظیر بسط به توابع متعامد مانند هرمیت، بسل، ... و همچنین سری فوریه، موجک‌ها و ... نیز استفاده نمود که نیاز به مطالعه بیشتر دارد. برای تعیین تابع توزیع ماکزیمم آنتروپی می‌توان از اندازه‌های دیگر آنتروپی نیز استفاده نمود.

مراجع

[1] Jaynes, E. T. (1968), "Prior Probabilities". *IEEE Trans. S.S.C.* 4, 227-248

[2] Kamae, T., U. Krengel, and G.L. O'Brien (1977), "Stochastic Inequalities on Partially Ordered Spaces" *The Annals of Probability*, 5(6),899-912.

[3] Kapur, J. N. (1983), "Bayesian Entropy and Its Application to Entropy Maximization on Models". *IIT/K Res. Rep.*

[4] Strassen, V. (1965), "The Existence of Probability Measures with Given Marginals," *Annals of Mathematical Statistics,* 36(2),423-439.

[5] Mansoury, S. Pasha, E. and Mohammadzadeh M. (2005), "Determination of Maximum Bayesian

Entropy Probability Distribution" *J. of Sciences Islamic Republic of Iran 16(4)*: 339-345

[6] Kiefer, N. (2009), "The Maximum Entropy Distribution for Stochastically Ordered Random Variables with Fixed Marginals" *Internet. CAE Working paper* 09-01.