ماکسیمم آنتروپی در توزیع‌های دومتغیره

زهره زمانی ده یعقوبی - غلامرضا محتشمی برزادران

دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

چکیده : بهینه‌سازی همواره مورد توجه علوم زیادی بوده است در نظریه اطلاع نیز این مبحث جایگاه خاص خودش را دارد. یکی از روش‌های بهینه‌سازی اصل ماکسیمم آنتروپی به روش لاگرانژ می‌باشد. در این مقاله ماکسیمم آنتروپی چندین توزیع دومتغیره را مورد بررسی قرار می‌دهیم.

واژه‌های کلیدی: آنتروپی شانون ، ماکسیمم آنتروپی ،

۱ مقدمه

اندازه آنتروپی شانون به عنوان زیربنای نظریه‌ی اطلاع نقش عمده ای در استنباط آماری دارد. این اندازه برای اولین بار توسط شانون در سال ۱۹۴۸ مطرح شد.همچنین جینز در سال ۱۹۵۷ اصل آنتروپی ماکسیمم را ارائه نموده است.

۱.۱ آنتروپی شانون متغیرهای تصادفی

در ابتدا آنتروپی شانون را برای متغیرهای تصادفی تعریف می‌کنیم.

تعریف ۱۶ اگر X یک متغیـــر تصادفی گسسته با تابع احتمال $p_X(x)$ باشد، آنگاه آنتروپی آن عبارت است از:

$$H(X) = -\sum_x p_X(x)\log p_X(x).$$

این معیار برای متغیرهای تصادفی پیوسته بطورمشابه به صورت زیر تعریف می‌شود:

تعریف ۱۷ اگر X یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f_X(x)$ باشد، آنگاه آنتروپی آن عبارت است از:

$$H(X) = -\int_{-\infty}^{\infty} f_X(x)\log f_X(x)dx.$$

در نظریه اطلاع بهینه‌سازی جایگاه خاصی دارد. یکی از روش‌های بهینه‌سازی اصل ماکسیمم آنتروپی به روش لاگرانژ می‌باشد. سوالی که مطرح می‌شود این است که چرا می‌خواهیم آنتروپی را بهینه کنیم؟ ما معمولاً می‌خواهیم عدم قطعیت را با بدست‌آوردن اطلاعات بیشتر وبیشتر کاهش دهیم.

اصل ماکسیمم آنتروپی: بین همه توزیع‌های احتمال سازگار با یک سری از محدودیت‌های داده شده، توزیعی که دارای ماکسیمم عدم قطعیت است را انتخاب می‌کنیم.

در بسیاری از مواقع به دنبال یافتن توزیعی که آنتروپی را ماکسیمم می‌کند هستیم. چگونگی یافتن آن با توجه به محدودیت‌هایی که علاوه بر تابع (احتمال) چگالی بودن لازم است توسط کاگان [1] و همکاران (۱۹۷۳)۱ و کاپور (۱۹۸۹ و ۱۹۹۲) بیان گردید. توزیع دارای آنتروپی ماکسیمم معمولا به فرم تابع نمایی از محدودیت‌ها بیان می‌شود. یکی از این حالت‌ها زمانی است که گشتاورهای مراتب مختلف مشخص باشد که به ازای دامنه تغییرات متغیر تصادفی توزیع‌های مشهوری را به عنوان توزیع آنتروپی ماکسیمم نتیجه می‌دهد.

در بسیاری از مواقع با توجه به دانستن مقادیر گشتاورهای مراتب مختلف یک توزیع علاقه مند به یافتن توزیعی که دارای آنتروپی ماکسیمم است می‌باشیم که کاربردهای دیگری نیز دارد. در این ارتباط کاگان وهمکاران (۱۹۷۳) قضیه زیر رابیان کرده‌اند.

قضیه ۱۴ هر گاه X متغیر تصادفی باشد که $f(x) > 0,\ \forall x \in D$ وهمچنین توابع, $h_i(x)$ $i = 1, 2, ...k$ طوری باشند که $E(h_i(X)) = \theta_i\ ,\ \ i = 1, 2, ...k$ و $\int_D f(x)dx = 1$ و θ_i ها نیز ثابت باشند، آنگاه توزیع دارای آنتروپی ماکسیمم به‌صورت $f(x) = e^{c_0 + c_1 h_1(x) + ... + c_k h_k(x)}$ است که $c_0, c_1, ..., c_k$ به کمک شرایط ذکرشده به‌دست می آیند.

از کارهای انجام شده در زمینه ماکسیمم آنتروپی می توان به جعفری (۱۹۹۱)، گلان (۱۹۹۸) اشاره‌کرد که ارتباط بین برآوردگر درستنمایی ماکسیمم و برآوردگر آنتروپی ماکسیمم را مورد بررسی قرار داه اند.

۲.۱ اصل ماکزیمم آنتروپی برای حالت پیوسته

در حالتی که متغیر تصادفی X پیوسته باشد، می خواهیم تحت محدودیت های

$$\int_a^b f(x)dx = 1, \qquad (1)$$

[1] $Kagan$

$$E(g_r(X)) = \int_a^b f(x) g_r(x) dx = a_r, \quad r = 1, \cdots, m \tag{2}$$

$f(x)$ را به دست آوریم.
تعداد نامتناهی از توابع چگالی احتمال وجود دارد که در محدودیت های داده شده صدق کند، بنابراین باید تابع چگالی احتمال را انتخاب کنیم که اندازه آنتروپی شانون در حالت پیوسته را ماکزیمم کند.

۱.۲.۱ تابع چگالی احتمال ماکزیمم آنتروپی متغیر پیوسته

قبل از به دست آوردن تابع چگالی احتمال ماکزیمم آنتروپی برای متغیر پیوسته لم زیر را درنظر بگیرید:

لم ۲ فرض کنید $I = \int_a^b F(x, f(x), f'(x))$ که F تابعی معلوم است. آنگاه با توجه به معادله اویلر- لاگرانژ تابع $f(x)$ که I را ماکزیمم می کند از معادله زیر به دست می آید.

$$\frac{\partial F}{\partial f(x)} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'(x)}\right) = 0.$$

حال می خواهیم تابع چگالی احتمال را به دست آوریم که هم اندازه آنتروپی شانون در حالت پیوسته را ماکزیمم کند و هم در محدودیت ها صدق کند. بنابراین تابع لاگرانژ را به صورت زیر تشکیل می دهیم.

$$\begin{aligned} L \ &= \ -\int_a^b f(x)\log f(x) dx - (\lambda_0 - 1)\left(\int_a^b f(x) dx - 1\right) \\ &\quad - \sum_{r=1}^{m} \lambda_r \left(\int_a^b f(x) g_r(x) dx - a_r\right), \end{aligned}$$

حال با استفاده از لم تابع چگالی احتمال ماکزیمم آنتروپی در حالت پیوسته به صورت زیر است:

$$\begin{aligned} \frac{\partial L}{\partial f(x)} \ &= \ -1 - \log f(x) - \lambda_0 + 1 - \sum_{r=1}^{m} \lambda_r g_r(x) = 0, \\ &\Longrightarrow f(\hat{x}) = e^{-\lambda_0 - \lambda_1 g_1(x) - \cdots - \lambda_m g_m(x)} \end{aligned}$$

۲ ماکسیمم آنتروپی

یکی از جالبترین کاربردهایی که از ماکزیمم آنتروپی به دست می‌آید، این است که با استفاده از اصل ماکزیمم آنتروپی، هنگامی که یکی یا بیشتر از گشتاورها معلوم باشند، می‌توان توزیع های احتمال را به دست آورد.

۱.۲ توزیع های ماکزیمم آنتروپی پیوسته

در این بخش می‌خواهیم توزیع‌های دو متغیره مختلف آنتروپی را تحت محدودیت‌های مختلف روی دامنه‌های متفاوت به دست آوریم. همانطور که خواهیم دید خوشبختانه اغلب توزیع های مفید آماری را می توانیم با اصل ماکزیمم آنتروپی به دست آوریم. حال به بررسی این توزیع ها می پردازیم.

۱.۱.۲ توزیع های ماکزیمم آنتروپی روی بازه‌های خاص

-- اگر $E(\log X) = \psi(p) - \psi(p+q) + \log a$ ، $E(\log(Y_X) = \psi(q) - \psi(p+q) + \log a$ ، $E(\log Y) = \log a$ و $E(Y^c) = ba^c$ مقادیری معلوم باشند، آنگاه تابع چگالی ماکزیمم آنتروپی برابر است با

$$f(x,y) = \frac{|c|}{a^{bc}\Gamma(bc)B(p,q)} x^{p-1}(y-x)^{q-1} y^{bc-p-q} exp\left[-\left(\frac{y}{a}\right)^{c}\right], \ 0 < x < y, \quad a,b,c,p,q > 0.$$

که به توزیع بتا‌ـاس تی سی [2] معروف است و در آن تابع ψ تابع دیاگاما است یعنی $\psi(x) = \frac{\partial \Gamma(x)}{\partial x} = \frac{\Gamma'(x)}{\Gamma(x)}$

-- با معلوم بودن مقادیر $E(\log X) = \psi(p) - \log(a)$، $E(\log(Y_X) = \psi(q) - \log(a)$ و $E(Y) = \frac{p+q}{a}$، توزیع دارای ماکزیمم توزیع گامای دومتغیره مکی [3] با تابع چگالی زیر است:

$$f(x,y) = \frac{a^{p+q}}{\Gamma(p)\Gamma(q)} x^{p-1}(y-x)^{q-1} e^{-ay}, \ \ y > x > 0, \ \ a,p,q > 0.$$

---

۲ $Beta - Stacy\ Distribution$

۳ $Mckay^s\ Bivarite\ Gamma\ Distribution$

-- اگر $E(\log(X)) = \psi(\theta_1) - \psi(\theta_1+\theta_2+\theta_3)$، $E(\log(Y)) = \psi(\theta_2) - \psi(\theta_1 + \theta_2 + \theta_3)$ و $E(\log(X)) = \psi(\theta_3) - \psi(\theta_1+\theta_2+\theta_3)$ را داشته باشیم آنگاه توزیع

$$f(x,y) = \frac{\Gamma(\theta_1+\theta_2+\theta_3)}{\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(\theta_3)} x^{\theta_1-1} y^{\theta_2-1} (1-x-y)^{\theta_3-1}, \quad x,y \geq 0,\ x+y \leq 1.$$

ماکسیمم آنتروپی را دارد که به عنوان توزیع بتای دومتغیره [4] یا توزیع دریخله دومتغیره [5] شناخته می شود.

-- هنگامی که‌رابطه $E\left(\log\left(1 - \frac{1-X^2-2\rho XY+Y^2}{1-\rho^2}\right)\right) = -\frac{1}{\nu+1}$ معلوم باشد، توزیع پیرسون نوع II[6]، توزیع ماکزیمم آنتروپی است که دارای تابع احتمال

$$f(x,y) = \frac{\nu+1}{\pi\sqrt{1-\rho^2}} \left[1 - \frac{1-x^2-2\rho xy+y^2}{1-\rho^2}\right]^{\nu},\ \nu > 1, -1 < \rho < 1,$$

می باشد که (x,y) داخل بیضی گون $x^2 - 2\rho xy + y^2 = 1 - \rho^2$ که داخل مربع واحد است، قرار می گیرند.

-- یک مورد خاص از توزیع پیرسون نوع II، به ازای $\rho = 0$ و $b = \nu + \frac{3}{2}$ به صورت زیر است:

$$f(x,y) = \frac{\Gamma(b+\frac{1}{2})}{\Gamma(b-\frac{1}{2})} (1-x^2-y^2)^{b-\frac{3}{2}},\ b > \frac{1}{2}$$

که هر یک از توزیع های حاشیه ای یک توزیع یک متغیره بتا است که متقارن بوده و چگالی آن به صورت $f(x) = \frac{1}{B(b,\frac{1}{2})}(1-x^2)^{b-1},\ -1 < x < 1$ می باشد. حال توزیع دومتغیره مزبور با داشتن مقدار $E(\log(1-X^2-Y^2)) = \psi(b-\frac{1}{2}) - \psi(b+\frac{1}{2})$، توزیع ماکزیمم آنتروپی است.

-- زمانی که

۱) $E(\log X) = \psi(\alpha_1) - \psi(\alpha_3)$

۲) $E(\log(Y) = \psi(\alpha_2) - \psi(\alpha_3)$

۳) $E(\log(1+XY) = \psi(\alpha_1+\alpha_2+\alpha_3)$

---

[4] *Bivariate Beta Distribution*

[5] *Bivariate Dirichlet Distribution*

[6] *Pearson Type II Distribution*

آنگاه توزیع ماکسیمم آنتروپی، توزیع دومتغیره بتای معکوس [۷] (توزیع دومتغیره دریخله معکوس [۸] با تابع چگالی احتمال زیر است:

$$f(x,y)=\frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(alpha_2)\Gamma(\alpha_3)}\frac{x^{\alpha_1-1}y^{\alpha_2-1}}{(1+x+y)^{\alpha_1+\alpha_2+\alpha_3}}$$

هنگامی که $E(\log(Y-X))=\frac{2\gamma+1+(\log(\xi-\eta))\gamma(\gamma+1)}{\gamma(\gamma+1)}$، آنگاه تابع چگالی احتمال

$$f(x,y)=\gamma(\gamma+1)(\xi-\eta)^{\gamma}(y-x)^{-(\gamma+2)},\quad x<\eta<\xi<y$$

که به توزیع پاره تو دومتغیره [۹] معروف است، توزیع ماکسیمم آنتروپی است.

۲.۱.۲ توزیع های ماکزیمم آنتروپی روی بازه $(-\infty,\infty)\times(-\infty,\infty)$

-- اگر $E(X)=\psi(\nu)-\psi(\alpha)$ ، $E(Y)=\psi(\nu)$ و $E(\log(1+e^{-X}+e^{-Y})=\psi(\nu+\alpha+\beta)-\psi(\nu)$ معلوم باشند آنگاه توزیع ماکزیمم آنتروپی توزیع دومتغیره Z [۱۰] با تابع چگالی احتمال زیر است:

$$f(x,y)=\frac{\Gamma(\nu+\alpha+\beta)}{\Gamma(\alpha)\Gamma(\nu)}\frac{e^{-\alpha x-\beta y}}{(1+e^{-x}+e^{-y})^{\alpha+\beta+\nu}}$$

-- هنگامی که $\frac{1}{N-1}=E(\log\left(1+\frac{X^2+Y^2-2\rho XY}{m(1-\rho^2)}\right))$ معلوم باشد تابع چگالی ماکسیمم آنتروپی برابر است با

$$f(x,y)=\frac{N-1}{m\pi\sqrt{1-\rho^2}}\left\{1+\frac{x^2+y^2-2\rho xy}{m(1-\rho^2)}\right\}^{-N}\quad ,N>1,m>0.$$

که به عنوان توزیع پیرسون نوع VII [۱۱] شناخته می شود.

-- اگر داشته باشیم

---

۷ *Bivariate Inverted Beta Distribution*

۸ *Bivariate Inverted Dirichlet Distribution*

۹ *Bilateral Bivariate Pareto Distribution*

۱۰ *Bivariate Z Distribution*

۱۱ *Bivariate Pearson Type VII Distribution*

۱) $E(\log(X^2 - 2\rho XY + Y^2)) = \frac{1}{s}\psi(s) + \log(1-\rho^2) - \frac{\log r}{s}$

۲) $E\left(\frac{X^2-2\rho XY+Y^2}{1-\rho^2}\right)^s = \frac{N}{sr}$

آنگاه توزیع ماکسیمم آنتروپی توزیع کاتز [12] با تابع چگالی احتمال زیر است:

$$f(x,y) = sr^{\frac{N}{s}} \frac{(x^2 - 2\rho xy + y^2)}{\pi\Gamma\left(\frac{N}{s}\right)(1-\rho^2)^{N-\frac{1}{2}}} \exp\left\{-r\left(\frac{x^2-2\rho xy+y^2}{1-\rho^2}\right)^s\right\}.$$

که در این توزیع $N,r,s > 0$ و $-1 < \rho < 1$ می باشد.با جایگذاری $N = 1$، $s = 1$ و $r = \frac{1}{2}$ در این توزیع به توزیع نرمال دومتغیره می رسیم.

-- اگر داشته باشیم:

$$E(X^2 - 2\rho XY + Y^2) = \frac{1-\rho^2}{a}, \quad E\left(e^{-a\left(\frac{X^2-2\rho XY+Y^2}{1-\rho^2}\right)}\right) = \frac{(b+1)e^{-b}-1}{b(1-e^{-b})}$$

انگاه توزیع ماکسیمم آنتروپی، توزیع گامبل [13] است که دارای تابع چگالی احتمال زیر می باشد:

$$f(x,y) = \frac{ab(1-\rho^2)^{\pi\{1-e\exp -b\}}}{-\frac{1}{2}} e^{-\frac{a(x^2-2\rho xy+y^2)}{1-\rho^2}} e^{-be^{-\frac{a(x^2-2\rho xy+y^2)}{1-\rho^2}}}$$

۳.۱.۲ توزیع های ماکزیمم آنتروپی روی بازه $(0,\infty)\times(0,\infty)$

-- اگر $E\left(\log\left(1+\frac{X}{\sigma_1}+\frac{Y}{\sigma_2}\right)\right) = \frac{2\alpha-1}{\alpha(\alpha-1)}$، آنگاه توزیع ماکسیمم آنتروپی توزیع پاره تو دومتغیره ماردیا [14] با تابع چگالی دومتغیره زیر است:

$$f(x,y) = \frac{\alpha(\alpha-1)}{\sigma_1\sigma_2}\left(1+\frac{X}{\sigma_1}+\frac{Y}{\sigma_2}\right)^{-(alpha+1)}, \quad x,y > 0.$$

-- هنگامی که $E(\log(b+X+Y)) == \frac{2a+3+(\log b)(a+1)(a+2)}{(a+1)(a+2)}$ معلوم باشد، توزیع پاره‌تودومتغیره با تابع چگالی احتمال زیر دارای ماکزیمم آنتروپی است.

$$f(x,y) = \frac{(a+1)(a+2)b^{a+1}}{(b+x+y)^{a+3}}, \quad x,y > 0, \quad a,b > 0.$$

---

[12] $Kotz - type\ Elliptical\ Distribution$
[13] $Gumbel - Type\ Elliptical\ distribution$
[14] $Mardias\ Bivariate\ Pareto\ Distribution$

-- با داشتن مقدار $E(\log(\lambda_1 X + \lambda_2 Y + b)) = 2a + 3 + (\log b)(a+1)(a+2)$، توزیع دومتغیره آمیخته نمایی [۱۵] با تابع چگالی زیر ماکزیمم آنتروپی را دارد.

$$f(x,y) = \frac{\lambda_1 \lambda_2 (a+1)(a+2) b^{a+1}}{(\lambda_1 x + \lambda_2 y + b)^{a+3}}, \quad x,y > 0, \quad \lambda_1, \lambda, a, b > 0.$$

-- با داشتن محدودیت های

$$E(X) = \frac{\theta_2}{\theta_1}, E(\log X) = \psi(\theta_2), E(XY) = \frac{1}{\theta_3}$$

توزیع نمایی گاما دومتغیره [۱۶] با تابع چگالی

$$f(x,y) = \frac{\theta_1^{\theta_2} \theta_3}{\Gamma(\theta_2)} x^{\theta_2} \exp\{-(\theta_1 x + \theta_3 xy)\}$$

توزیع ماکزیمم آنتروپی است.

نتیجه‌گیری

با استفاده از اصل ماکزیمم آنتروپی می توان هنگامی که یکی یا بیشتر از گشتاورها معلوم باشند، توزیع های احتمال را به دست آورد.هدف از تحقیق حاضر به دست آوردن توزیع ماکزیمم آنتروپی در توزیع های دو متغیره تحت محدودیت های مختلف بوده است.

مراجع

Kagan, A. M., Linnik, Y. V. and Rao, C, R. (1973), *Characterization Problems in Mathematical Statistics*, Wiley, New York.

Kapur, J. N. (1989), *Maximum Entropy Models in Sciences and Engineering*, New York, John Wiley.

Shannon, C. E. (1948), *A Mathematical theory of communication*, Bell System Technichal Journal, 27, 379-423.

[۱۵] Bivariate Exponential Mixture Distribution

[۱۶] *Bivariate Gamma Exponential Distribution*

Zografos, K. and Nadarajah, S. (2005), *Expressions for Renyi and Shannon entropies for bivariate distributions*, information science,**170**, 173-189.

Balakrishnan, N. and Lai, C.D .(2009), *Continous bivarite distribution.* Sprringer Dordrecht Heidelberg London New York.

Djafari, M. (1991), *A matlab program to calculate the maximum entropy distribution*, Proc. of the 11 the Int, Maxent Workshop, Seattle,USA.

Jaynes, E.T. (1957). *Information theory and statistical mechaanics*, Physical Review, **106**, 629-630.