**مجله علوم آماری، بهار و تابستان ۱۳۹۴**

**جلد ۹، شماره ۱، ص ۱۰۱–۱۱۸**

# تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی

**شهرام منصوری**

**دانشکده علوم ریاضی، دانشگاه شهید بهشتی**

**تاریخ دریافت: ۱۳۹۳/۴/۲۰ تاریخ آخرین بازنگری: ۱۳۹۴/۲/۲۴**

**چکیده :** بنا بر اصل ماکسیمم آنتروپی جینز، در میان تمام توابع توزیع احتمال که در قیود معین صدق می‌کنند توزیعی باید انتخاب شود که دارای ماکسیمم آنتروپی است. در این مقاله روشی برای به دست آوردن تابع چگالی احتمال توام دو متغیره با معلوم بودن توزیع‌های حاشیه‌ای و ضریب همبستگی در یک ناحیه معین با ماکسیمم کردن اندازه‌های آنتروپی تانیجا و بورگز ارائه و مثال‌هایی زده شده است. برای حالاتی که نتوان مسئله را به‌صورت تحلیلی حل کرد، روشی عددی نیز پیشنهاد و نحوه اجرای آن در یک مثال توضیح داده شده است.

**واژه‌های کلیدی** : اندازه آنتروپی تانیجا، اندازه آنتروپی بورگز، اصل ماکسیمم آنتروپی، اسپلاین‌ها.

---

آدرس الکترونیکی مسئول مقاله: شهرام منصوری، sh_mansouri@sbu.ac.ir

کد موضوع‌بندی ریاضی (۲۰۱۰): $62B10$

## ۱ مقدمه

اندازه آنتروپی تانیجا[۱] برای $r > ۱$ و اندازه آنتروپی بورگز[۲] یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x)$ به‌ترتیب به‌صورت

$$H_T(f) = -۲^{r-۱}\int_R (f(x))^r \ln f(x)dx \quad (۱)$$

$$H_B(f) = \int_R \ln f(x)dx \quad (۲)$$

تعریف شده‌است.

اصل ماکسیمم آنتروپی اولین بار توسط جینز (۱۹۵۷) بیان شد. تعیین توزیع‌های با ماکسیمم آنتروپی یک متغیره توسط تعدادی از محققین از جمله رضا (۱۹۶۱)، جینز (۱۹۶۸)، گسیو (۱۹۷۷)، شور (۱۹۷۸) مورد مطالعه قرار گرفته است. اگر تابع توزیع احتمال توام دو متغیره تصادفی معلوم باشد می‌توان توزیع‌های حاشیه‌ای را به دست آورد ولی اگر توزیع‌های حاشیه‌ای معلوم باشند توزیع‌های توام زیادی می‌توان تعیین کرد که دارای همان توزیع‌های حاشیه‌ای باشند. فرشه (۱۹۵۱) کرانهایی برای اینگونه توزیع‌ها به دست آورد و نلسون (۱۹۸۷،۱۹۹۱) با معلوم بودن توزیع‌های حاشیه‌ای و ضریب همبستگی بین متغیرها روشی برای به دست آوردن توزیع توام ارائه داد. در خصوص توزیع‌های توام گسسته ماکسیمم آنتروپی مطالعاتی توسط آماردانها از جمله پاتیل و جوشی (۱۹۶۸)، و کاپور (۱۹۸۱،۱۹۸۹) صورت گرفته است. همزمان در خصوص توزیع‌های پیوسته نیز مطالعاتی توسط پاتیل و جوشی (۱۹۶۸) کاپور (۱۹۸۱)، میورویزن و بدفورد (۱۹۹۷) و میلر و ویهان (۲۰۰۲) انجام شده است. منصوری و همکاران (۲۰۰۵) بر اساس آنتروپی بیزی یک فرم کلی برای تابع توزیع احتمال چند متغیره ماکسیمم آنتروپی با شرط معلوم بودن توابع چگالی حاشیه‌ای و ماتریس کوواریانس متغیرهای تصادفی ارائه کردند. همچنین روش ماکسیمم آنتروپی به‌عنوان یک ابزار سودمند در تعیین توزیع متغیرهای تصادفی توسط فیلیپس و همکاران (۲۰۰۶) و دیددیک و همکاران (۲۰۰۷) استفاده شده است. پاشا و همکاران (۲۰۰۸) توزیع احتمال توام چند

[۱] Taneja

[۲] Burgs

متغیره را تحت برخی قیود بر اساس اندازه آنتروپی شانون ارائه دادند. دانگ و همکاران (۲۰۱۳) چهار پارامتر توزیع‌احتمال ماکسیمم آنتروپی را بر اساس روش ماکسیمم درستنمایی برآورد نموده و مسئله را فرمول‌بندی کردند.

در این مقاله روش تعیین تابع چگالی احتمال توام دو متغیره ماکسیمم آنتروپی تانیجا با معلوم بودن یکی از توابع چگالی احتمال حاشیه‌ای و تابع چگالی احتمال توام دو متغیره ماکسیمم آنتروپی بورگز با معلوم بودن هر دو تابع چگالی احتمال حاشیه‌ای و ضریب همبستگی در بخش دوم مورد بررسی قرار می‌گیرد، سپس یک روش عددی برای تعیین تابع چگالی احتمال توام دو متغیره ماکسیمم آنتروپی بورگز در بخش سوم ارائه خواهد شد. بحث و نتیجه‌گیری نیز در بخش چهارم بیان شده است.

## ۲ توزیع‌های احتمال دو متغیره ماکسیمم آنتروپی تانیجا و بورگز

در این بخش ابتدا روابط مورد نیاز از حساب تغییرات و سپس لم و قضایای لازم برای تعیین تابع توزیع احتمال ماکسیمم آنتروپی بیان گردیده و از آن برای به‌دست آوردن تابع چگالی احتمال ماکسیمم آنتروپی دو متغیره با معیارهای اندازه آنتروپی تانیجا و بورگز تحت برخی قیود استفاده می‌شود.

فرض کنید، باید $y = y(x)$ را چنان تعیین شود که جواب بهینه $\int_{x_1}^{x_2} F(x, y, y')dx$ باشد، که در آن $y' = \frac{dy}{dx}$، $y(x_1) = y_1,\ y(x_2) = y_2$ و $F$ یک تابع معلوم با مشتقات جزئی موجود باشد. در این صورت اویلر(۱۷۴۴) ثابت کرده $y = y(x)$ در معادله اویلر به‌صورت

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0, \tag{۳}$$

صدق می‌کند. حال اگر علاوه بر شرایط فوق $y$ در قیود $\int_{x_1}^{x_2} G_i(x, y, y')dx = A_i$ برای $i = 1, \ldots, n,$ نیز صدق کند، که در آن $G_i$ ها توابعی معلوم هستند، برای حل مسئله به روش لاگرانژ، داریم

$$L \;=\; \int_{x_1}^{x_2} F(x, y, y')dx - \sum_{i=1}^{n} \lambda_i\left(\int_{x_1}^{x_2} G_i(x, y, y')dx - A_i\right)$$

$$= \int_{x_1}^{x_2} \{F(x,y,y') - \sum_{i=1}^{n} \lambda_i G(x,y,y')\} dx + \sum_{i=1}^{n} \lambda_i A_i$$

اکنون با قرار دادن $F(x,y,y') - \sum_{i=1}^{n} \lambda_i G_i(x,y,y')$ به‌جای $F$ در معادله (۳) جواب بهینه به‌دست می‌آید که این تکنیک، روش اویلر- لاگرانژ نامیده می‌شود (روساک، ۲۰۰۲).

**لم ۱** (منصوری و همکاران، ۲۰۰۵): اگر $L^2$ مجموعه توابع اندازه‌پذیر باشد که مربع آنها انتگرال پذیر است، آنگاه برای هر دو تابع $h_1$ و $h_2$ در $L^2$، $h_1 = h_2$ است اگر و تنها اگر برای هر تابع دلخواه $K$ تساوی زیر برقرار باشد

$$\int_R K(x) h_1(x) dx = \int_R K(x) h_2(x) dx \qquad (۴)$$

**قضیه ۱** : اگر $g(x) \in L^2$ تابع چگالی احتمال حاشیه‌ای یک خانواده از توزیع‌های احتمال دو متغیره با محمل

$$E = \{(x,y) \in R^2 | a < x < b, C_1(x) < y < C_2(x)\}$$

باشد، آنگاه ضابطه تابع چگالی احتمال ماکسیمم آنتروپی دو متغیره تانیجا عبارت است از

$$f_M(x,y) = \frac{g(x)}{C_2(x) - C_1(x)} \qquad (۵)$$

**برهان** با توجه به لم ۱ معلوم بودن $g$ معادل معلوم بودن $E(M(X))$ برای هر تابع دلخواه $M$ است. فرض کنید

$$\mu = E(M(X)) \qquad (۶)$$

از طرفی در حالت خاص $\int\int_E M(x) f dxdy = \mu$ به‌ازای $M(x) = 1$، $\mu = 1$ است، لذا قید تابع چگالی احتمال بودن $f$ نیز در شرط (۶) اعمال گردیده است. حال برای

به‌دست آوردن تابع چگالی احتمال ماکسیمم آنتروپی تحت شرط (۶) باید (۱) را ماکسیمم کرد. اکنون با روش اویلر- لاگرانژ، لاگرانژین را می‌توان به‌صورت

$$L = -2^{r-1}\int\int_E f^r \ln f dxdy + \lambda_1(\int\int_E M(x) f dxdy - \mu)$$
$$= \int\int_E \{-2^{r-1} f^r \ln f + \lambda_1 M(x) f\} dxdy - \lambda_1 \mu$$

نوشت، چون $-2^{r-1} f^r \ln f + \lambda_1 M(x) f$ تابعی مقعر و پیوسته بر حسب $f$ است جواب بهینه آنتروپی را ماکسیمم می‌کند، لذا تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی بنا به رابطه (۳) جواب معادله

$$\frac{\partial}{\partial f}[-2^{r-1} f^r \ln f + \lambda_1 M(x) f] = 0$$

است، که از حل آن، $f$ تابعی بر حسب $x$ به دست می‌آید، لذا ضابطه تابع چگالی احتمال ماکسیمم آنتروپی تانیجا به‌صورت

$$f_M(x,y) = f_M(x);\quad (x,y) \in E$$

است. برای به‌دست آوردن $f_M(x)$ با توجه به اینکه $g(x)$ تابع چگالی احتمال حاشیه‌ای $f_M(x,y)$ است، داریم

$$g(x) = \int_{C_1(x)}^{C_2(x)} f_M(x) dy = f_M(x)(C_2(x) - C_1(x))$$

پس تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی تانیجا عبارت است از

$$f_M(x,y) = f_M(x) = \frac{g(x)}{C_2(x) - C_1(x)};\quad (x,y) \in E$$

**مثال ۱** : اگر محمل $(X,Y)$ به‌صورت

$$E = \{(x,y) \in R^2 | 0 < x < 1, 0 < y < \sqrt{1-x^2}\}$$

بوده و $X$ متغیر تصادفی با توزیع $Beta(3,2)$ باشد، یعنی تابع چگالی احتمال آن به‌صورت

$$g(x) = \frac{\Gamma(5)}{\Gamma(2)\Gamma(3)} x^2(1-x) = 6x^2(1-x);\quad 0 < x < 1$$

باشد، در این صورت تابع چگالی احتمال ماکسیمم آنتروپی تانیجا با استفاده از رابطه (۵) برابر است با

$$f_M(x,y) = 6x^2\sqrt{\frac{1-x}{1+x}};\quad 0<x<1,\ 0<y<\sqrt{1-x^2}$$

**تذکر ۱** : قضیه ۱ قابل تعمیم برای توزیع‌های چند متغیره است. هرگاه $g(x_1)$ تابع چگالی احتمال حاشیه‌ای متغیر تصادفی $X_1$ از بردار تصادفی $(X_1,\dots,X_n)$ با محمل

$$\begin{aligned} E \quad = \quad & \{(x_1,\dots,x_n)\in R^n | a<x_1<b, A_1(x_1)<x_2<A_2(x_1), \\ & B_1(x_1,x_2)<x_3<B_2(x_1,x_2),\dots, \\ & L_1(x_1,\dots,x_{n-1})<x_n<L_2(x_1,\dots,x_{n-1})\} \end{aligned}$$

باشد، آنگاه تابع چگالی احتمال ماکسیمم آنتروپی تانیجا تابعی از $x_1$ به‌صورت زیر است

$$f_M(x_1,\dots,x_n) = \frac{g(x_1)}{\int_{A_1(x_1)}^{A_2(x_1)}\int_{B_1(x_1,x_2)}^{B_2(x_1,x_2)}\cdots\int_{L_1(x_1,\dots,x_{n-1})}^{L_2(x_1,\dots,x_{n-1})} dx_n dx_{n-1}\dots dx_2} \qquad (x_1,\dots,x_n)\in E$$

**قضیه ۲** : اگر $g,h\in L^2$ توابع چگالی حاشیه‌ای یک خانواده از توزیع‌های احتمال توام دو متغیره با محمل $E\subseteq R^2$ باشند، آنگاه ضابطه تابع چگالی احتمال ماکسیمم آنتروپی بورگز به‌صورت

$$f_M(x,y) = \frac{1}{f_1(x)+f_2(y)} \quad (x,y)\in E \tag{7}$$

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

$$\begin{cases} \int_R \frac{1}{f_1(x)+f_2(y)}dy = g(x) \\ \int_R \frac{1}{f_1(x)+f_2(y)}dx = h(y) \end{cases} \tag{8}$$

به‌دست می‌آیند.

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

$$\mu_1 = E(M(X)), \ \ \mu_2 = E(N(Y)) \tag{۹}$$

نیز معلوم خواهند بود. با توجه به لم ۱ معلوم بودن $g$ و $h$ معادل معلوم بودن $E(M(X))$ و $E(N(Y))$ است از طرفی در حالت خاص $\int\int_E M(x)fdxdy = \mu_1$ به‌ازای $M(x) = 1$، $\mu_1 = 1$ است، لذا قید تابع چگالی احتمال بودن $f$ نیز در شرط (۹) اعمال گردیده است. بنابراین برای به دست آوردن تابع چگالی احتمال ماکسیمم آنتروپی تحت شرایط (۹) باید (۲) را ماکسیمم نمود. حال با روش اویلر- لاگرانژ، لاگرانژین را می‌توان به‌صورت

$$\begin{aligned} L &= \int\int_E \ln f dxdy \\ &+ \lambda_1(\int\int_E M(x)fdxdy - \mu_1) + \lambda_2(\int\int_E N(x)fdxdy - \mu_2) \\ &= \int\int_E \{\ln f + (\lambda_1 M(x) + \lambda_2 N(y))f\}dxdy - (\lambda_1\mu_1 + \lambda_2\mu_2) \end{aligned}$$

نوشت، چون $\ln f + (\lambda_1 M(x) + \lambda_2 N(y))f$ تابعی مقعر و پیوسته بر حسب $f$ است جواب بهینه آنتروپی را ماکسیمم می‌کند، لذا تابع چگالی احتمال ماکسیمم آنتروپی بورگز بنا به رابطه (۳) جواب معادله

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

است، بنابراین

$$f_M(x,y) = \frac{-1}{\lambda_1 M(x) + \lambda_2 N(y)}$$

واضح است که $f_M$ تابعی از حاصل جمع دو تابع یکی بر حسب $x$ و دیگری بر حسب $y$ است، لذا

$$f_M(x,y) = \frac{1}{f_1(x) + f_2(y)}; \quad (x,y) \in E$$

از آنجا که $g$ و $h$ توابع چگالی حاشیه‌ای هستند، بنابراین $f_1$ و $f_2$ از (۸) قابل حصول خواهند بود.

**تذکر ۲** : حکم قضیه ۱ به‌طور مشابه برای توزیع‌های ماکسیمم آنتروپی بورگز نیز قابل اثبات است، به عبارت دیگر هرگاه $g(x)$ تابع چگالی احتمال حاشیه‌ای یک خانواده از توزیع‌های احتمال دو متغیره با محمل

$$E = \{(x, y) \in R^2 | a < x < b, C_1(x) < y < C_2(x)\}$$

باشد، آنگاه ضابطه تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی بورگز به‌صورت زیر است.

$$f_M(x, y) = \frac{g(x)}{C_2(x) - C_1(x)} \tag{10}$$

**مثال ۲** : اگر محمل $(X, Y)$ به‌صورت

$$E = \{(x, y) \in R^2 | 0 < x < 1, x^2 < y < x\}$$

باشد و $X$ دارای توزیع $Beta\,(2, 2)$ با تابع چگالی احتمال

$$g(x) = \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} x(1 - x) = 6x(1 - x); \quad 0 < x < 1$$

باشد آنگاه تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی بورگز با استفاده از رابطه (۱۰) به‌صورت

$$f_M(x, y) = \frac{6x(1 - x)}{x - x^2} = 6; \quad 0 < x < 1,\ x^2 < y < x$$

است، یعنی $(X, Y)$ دارای توزیع یکنواخت روی ناحیه $E$ است.

**تذکر ۳** : تذکر ۱ قابل تعمیم برای توزیع‌های چند متغیره است، هرگاه $g(x_1)$ تابع چگالی احتمال حاشیه‌ای متغیر تصادفی $X_1$ از بردار تصادفی $(X_1, \dots, X_n)$ با محمل

$$\begin{aligned} E \;=\; & \{(x_1, \dots, x_n) \in R^n | a < x < b, A_1(x_1) < x_2 < A_2(x_1), \\ & B_1(x_1, x_2) < x_3 < B_2(x_1, x_2), \dots, \\ & L_1(x_1, \dots, x_{n-1}) < x_n < L_2(x_1, \dots, x_{n-1})\} \end{aligned}$$

باشد آنگاه تابع چگالی احتمال ماکسیمم آنتروپی بورگز تابعی از $x_1$ به‌صورت زیر است.

$$f_M(x_1,\ldots,x_n)=\frac{g(x_1)}{\int_{A_1(x_1)}^{A_2(x_1)}\int_{B_1(x_1,x_2)}^{B_2(x_1,x_2)}\int_{L_1(x_1,\ldots,x_{n-1})}^{L_2(x_1,\ldots,x_{n-1})}dx_ndx_{n-1}\ldots dx_2}$$
$$(x_1,\ldots,x_n)\in E$$

**قضیه ۳** : اگر $g,h\in L^2$ توابع چگالی حاشیه‌ای یک خانواده از توزیع‌های احتمال توام دو متغیره با محمل $E\subseteq R^2$ و ضریب همبستگی $\rho$ باشند، آنگاه ضابطه تابع چگالی احتمال ماکسیمم آنتروپی بورگز به‌صورت

$$f_M(x,y)=\frac{1}{f_1(x)+f_2(y)+\lambda xy};\quad (x,y)\in E \tag{۱۱}$$

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

$$\begin{cases}\int_R f_M(x,y)dy=g(x)\\ \int_R f_M(x,y)dx=h(y)\\ \int\int_E xyf_M(x,y)dxdy=\rho\sigma_X\sigma_Y+\mu_X\mu_Y\end{cases} \tag{۱۲}$$

به‌دست می‌آیند.

**برهان** چون توابع چگالی حاشیه‌ای $f(x,y)$ معلومند، پس برای توابع دلخواه $M(X)$ و $N(Y)$

$$\mu_1=E(M(X)),\quad \mu_2=E(N(Y)) \tag{۱۳}$$

معلوم هستند. با توجه به لم ۱ معلوم بودن $g$ و $h$ معادل معلوم بودن $E(M(X))$ و $E(N(Y))$ است و چون ضریب همبستگی مقدار معلوم $\rho$ است، پس عبارت

$$E(XY)=\rho\sigma_X\sigma_Y+\mu_X\mu_Y=C \tag{۱۴}$$

نیز معلوم است. حال برای تعیین تابع چگالی احتمال ماکسیمم آنتروپی تحت شرایط (۱۳) و (۱۴) باید (۲) را ماکسیمم نمود. از روش اویلر- لاگرانژ، لاگرانژین را می‌توان به‌صورت

$$L=\int\int_E \ln f dxdy+\lambda'(\int\int_E xyfdxdy-C)$$

$$
\begin{aligned}
&+ \quad \lambda_1(\int\int_E M(x)fdxdy - \mu_1) + \lambda_2(\int\int_E N(x)fdxdy - \mu_2) \\
&= \quad \int\int_E \{\ln f + (\lambda' xy + \lambda_1 M(x) + \lambda_2 N(y))f\}dxdy \\
&- \quad (\lambda' C + \lambda_1\mu_1 + \lambda_2\mu_2).
\end{aligned}
$$

نوشت، حال چون $\ln f + (\lambda' xy + \lambda_1 M(x) + \lambda_2 N(y))f$ تابعی مقعر و پیوسته بر حسب $f$ است جواب بهینه آنتروپی را ماکسیمم می‌کند، لذا تابع چگالی احتمال ماکسیمم آنتروپی بورگز بنا به رابطه (۳) جواب معادله

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

است. بنابراین

$$f_M(x,y) = \frac{-1}{\lambda_1 M_1(x) + \lambda_2 N(y) + \lambda' xy}$$

لذا ضابطه تابع چگالی احتمال ماکسیمم آنتروپی بورگز به‌صورت

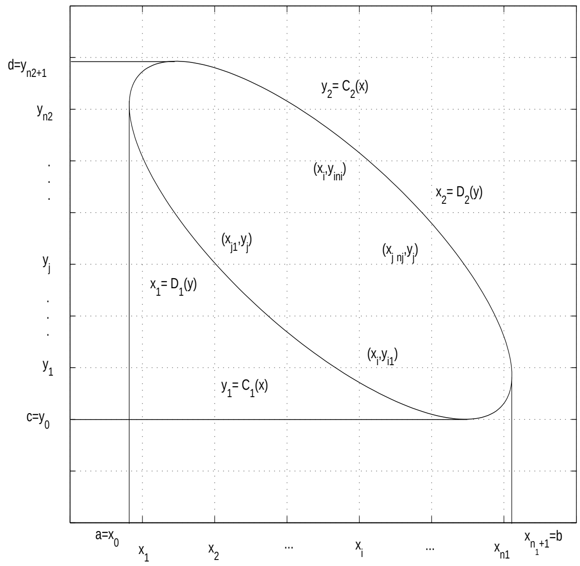
$$f_M(x,y) = \frac{1}{f_1(x) + f_2(y) + \lambda xy} \quad (x,y) \in E$$

است. از آنجا که توابع چگالی حاشیه‌ای و ضریب همبستگی معلوم هستند، بدیهی است که $f_1$، $f_2$ و $\lambda$ از (۱۲) قابل حصول خواهند بود.

## ۳ روش عددی

با توجه به اینکه معمولا دستگاه معادلات (۸) یا (۱۲) به‌صورت تحلیلی قابل حل نیستند، برای به‌دست آوردن تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی از روش‌های عددی استفاده می‌شود. فرض کنید قرار است تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی را روی ناحیه

$$
\begin{aligned}
E &= \{(x,y) \in R^2 | a < x < b, C_1(x) < y < C_2(x)\} \\
&= \{(x,y) \in R^2 | c < y < d, D_1(y) < x < D_2(y)\}
\end{aligned}
$$




شکل ۱: شبکه‌بندی $E$ در $R^2$

که در شکل ۱ نشان داده شده است برای توابع چگالی $g$ و $h$ با ضریب همبستگی معلوم $\rho$ تعیین شود. با استفاده از قضیه ضابطه تابع چگالی احتمال ماکسیمم آنتروپی به‌صورت

$$f_M(x,y) = \frac{1}{f_1(x) + f_2(y) + \lambda xy}; \quad (x,y) \in E \qquad (۱۵)$$

است. بنابراین توزیع‌های حاشیه‌ای به‌صورت

$$g(x) = \int_{C_1(x)}^{C_2(x)} \frac{1}{f_1(x) + f_2(y) + \lambda xy} dy; \quad a < x < b, \qquad (۱۶)$$

$$h(y) = \int_{D_1(y)}^{D_2(y)} \frac{1}{f_1(x) + f_2(y) + \lambda xy} dx; \quad c < y < d. \qquad (۱۷)$$

هستند. ناحیه $E$ در شکل ۱ را می‌توان به‌صورت

$$a = x_0 < x_1 < \cdots < x_{n_1+1} = b,$$

$$c = y_0 < y_1 < \cdots < y_{n_2+1} = d,$$

$$\Delta x_i = x_i - x_{i-1} = \Delta, \quad i = 2, \ldots, n_1,$$
$$\Delta y_j = y_j - y_{j-1} = \Delta, \quad j = 2, \ldots, n_2,$$

شبکه‌بندی کرد. از (۱۶) و (۱۷) برای هر $1 \leq i \leq n_1$ و $1 \leq j \leq n_2$ با به کارگیری روش ذوزنقه داریم

$$\begin{aligned}
g(x_i) &= \frac{1}{2}\Big[\frac{1}{f_1(x_i) + f_2(C_1(x_i)) + \lambda x_i C_1(x_i)} + \frac{1}{f_1(x_i) + f_2(y_{i1}) + \lambda x_i y_{i1}}\Big] \\
&\times (y_{i1} - C_1(x_i)) + \frac{1}{2}\sum_{k=2}^{n_i}\Big[\frac{1}{f_1(x_i) + f_2(y_{ik}) + \lambda x_i y_{ik}} \\
&+ \frac{1}{f_1(x_i) + f_2(y_{i,k-1}) + \lambda x_i y_{i,k-1}}\Big](y_{ik} - y_{i,k-1}) \\
&+ \frac{1}{2}\Big[\frac{1}{f_1(x_i) + f_2(y_{in_i}) + \lambda x_i y_{in_i}} \\
&+ \frac{1}{f_1(x_i) + f_2(C_2(x_i)) + \lambda x_i C_2(x_i)}\Big](C_2(x_i) - y_{in_i}),
\end{aligned}$$

و

$$\begin{aligned}
h(y_j) &= \frac{1}{2}\Big[\frac{1}{f_1(D_1(y_j)) + f_2(y_j) + \lambda D_1(y_j) y_j} + \frac{1}{f_1(x_{j1}) + f_2(y_j) + \lambda x_{j1} y_j}\Big] \\
&\times (x_{j1} - D_1(y_j)) + \frac{1}{2}\sum_{k=2}^{n_j}\Big[\frac{1}{f_1(x_{jk}) + f_2(y_j) + \lambda x_{jk} y_j} \\
&+ \frac{1}{f_1(x_{j,k-1}) + f_2(y_j) + \lambda x_{j,k-1} y_j}\Big](x_{jk} - x_{j,k-1}) \\
&+ \frac{1}{2}\Big[\frac{1}{f_1(x_{jn_j}) + f_2(y_j) + \lambda x_{jn_j} y_j} \\
&+ \frac{1}{f_1(D_2(y_j)) + f_2(y_j) + \lambda D_2(y_j) y_j}\Big](D_2(y_j) - x_{jn_j})
\end{aligned}$$

با در نظر گرفتن تقریب‌های

$$f_2(C_1(x_i)) = f_2(y_{i1}), \quad f_2(C_2(x_i)) = f_2(y_{in_i}) \quad 1 \leq i \leq n_1,$$
$$f_1(D_1(y_j)) = f_1(x_{j1}), \; f_1(D_2(y_j)) = f_1(x_{jn_j}) \quad 1 \leq j \leq n_2.$$

به‌ازای هر $i = 1, \ldots, n_1$ داریم

$$g(x_i) = \frac{1}{f_1(x_i) + f_2(y_{i1}) + \lambda x_i y_{i1}}(y_{i1} - C_1(x_i))$$

$$\begin{aligned}
&+ \quad \frac{\Delta}{۲}[\frac{۱}{f_۱(x_i)+f_۲(y_{i۱})+\lambda x_i y_{i۱}} + ۲(\frac{۱}{f_۱(x_i)+f_۲(y_{i۲})+\lambda x_i y_{i۲}} \\
&+ \quad \cdots + \frac{۱}{f_۱(x_i)+f_۲(y_{i,(n_i-۱)})+\lambda x_i y_{i,(n_i-۱)}}) \\
&+ \quad \frac{۱}{f_۱(x_i)+f_۲(y_{in_i})+\lambda x_i y_{in_i}}] \\
&+ \quad \frac{۱}{f_۱(x_i)+f_۲(y_{in_i})+\lambda x_i y_{in_i}}(C_۲(x_i)-y_{in_i}), \qquad (۱۸)
\end{aligned}$$

و به‌ازای هر $j=۱,\ldots,n_۲$، داریم

$$\begin{aligned}
h(y_j) \quad &= \quad \frac{۱}{f_۱(x_{j۱})+f_۲(y_j)+\lambda x_{j۱} y_j}(x_{j۱}-D_۱(y_j)) \\
&+ \quad \frac{\Delta}{۲}[\frac{۱}{f_۱(x_{j۱})+f_۲(y_j)+\lambda x_{j۱} y_j} + ۲(\frac{۱}{f_۱(x_{j۲})+f_۲(y_j)+\lambda x_{j۲} y_j} \\
&+ \quad \cdots + \frac{۱}{f_۱(x_{j,(n_j-۱)})+f_۲(y_j)+\lambda x_{j,(n_j-۱)} y_j}) \\
&+ \quad \frac{۱}{f_۱(x_{jn_i})+f_۲(y_j)+\lambda x_{jn_i} y_j}] \\
&+ \quad \frac{۱}{f_۱(x_{jn_j})+f_۲(y_j)+\lambda x_{jn_j} y_j}(D_۲(y_j)-x_{jn_j}) \qquad (۱۹)
\end{aligned}$$

حال با استفاده از قید $EXY=\int\int_{R^۲} xyf_M(x,y)dxdy=\rho\sigma_X\sigma_Y+\mu_X\mu_Y$ و با فرض اینکه $EXY=C$ رابطه دیگر به‌صورت

$$\int\int_E xy\frac{۱}{f_۱(x)+f_۲(y)+\lambda xy}dxdy=C \qquad (۲۰)$$

به‌دست می‌آید، برای محاسبه انتگرال (۲۰) ناحیه $E$ به المان‌های مربعی به‌طول $\Delta$ که مرکز هر مربع $(x_i,y_j)$ است افراز شود، مجموع

$$\Delta^۲\sum_{(x_i,y_j)\in E} x_i y_j\frac{۱}{f_۱(x_i)+f_۲(y_j)+\lambda x_i y_j}=C \qquad (۲۱)$$

را می‌توان به‌عنوان تقریبی برای انتگرال (۲۰) به‌کار برد. بدیهی است هر چه تعداد خانه‌های شبکه بیشتر شود تعداد جملات مجموع بیشتر شده و قسمت بیشتری از ناحیه $E$ پوشانده و تقریب بهتری به‌دست می‌آید. چون در روابط (۱۸)، (۱۹) و (۲۱) $n_۱$ مجهول به ازای $f_۱(x_i);\ i=۱,\ldots,n_۱$، $n_۲$ مجهول به‌ازای $f_۲(y_j);\ j=۱,\ldots,n_۲$ و با احتساب مجهول $\lambda$ جمعا $n_۱+n_۲+۱$ مجهول داریم

از طرفی $n_۱$ معادله از رابطه (۱۸)، $n_۲$ معادله از رابطه (۱۹) و یک معادله از رابطه (۲۱) به‌دست می‌آید که جمعا $n_۱ + n_۲ + ۱$ معادله خواهد شد. بنابراین مجهولات از نظر تئوری قابل محاسبه هستند حال از حل دستگاه معادلات (۱۸)، (۱۹) و (۲۱) و $\hat{f}_۱$ و $\hat{f}_۲$ را برای هر $x_i$ و $y_j$ به‌دست آورده و $\hat{\lambda}$ را مشخص می‌شود. برای داده‌های فوق با برازش اسپلاین درجه ۳ و $\hat{f}_۱$، $\hat{f}$ و $\hat{\lambda}$ را به‌دست آورده، سپس برآورد تابع چگالی توام با ماکسیمم آنتروپی با استفاده از رابطه (۱۱) تعیین می‌شود.

**مثال ۳** : فرض کنید $E = \{(x,y) \in R^۲ | ۰ < x + y < ۱, ۰ < x < ۱\}$ و توزیع‌های حاشیه‌ای به‌صورت

$$g(x) = \frac{۱}{\ln ۴ - ۱} \frac{۱ - x}{۱ + x}, ۰ < x < ۱, \quad h(y) = \frac{۱}{\ln ۴ - ۱} \frac{۱ - y}{۱ + y}, ۰ < y < ۱$$

بوده و $\rho = \frac{۱}{۲۰}$ باشد برای به‌دست آوردن تابع چگالی توام ماکسیمم آنتروپی، به‌دلیل تقارن $E$ و یکسان بودن توابع چگالی حاشیه‌ای قرار می‌دهیم $f_۲ = f_۱$. با استفاده از (۱۸)، (۱۹) و (۲۱) داریم

$$f(۰) = ۰/۲۸, f(۰/۱) = ۰/۳۳, f(۰/۲) = ۰/۳۷ f(۰/۳) = ۰/۴۲,$$
$$f(۰/۴) = ۰/۴۷, f(۰/۵) = ۰/۵۳ f(۰/٦) = ۰/۵۹, f(۰/۷) = ۰/٦۵,$$
$$f(۰/۸) = ۰/۷۲ f(۰/۹) = ۰/۷۹, f(۱) = ۰/۸٦, \lambda = ۰/۳۸$$

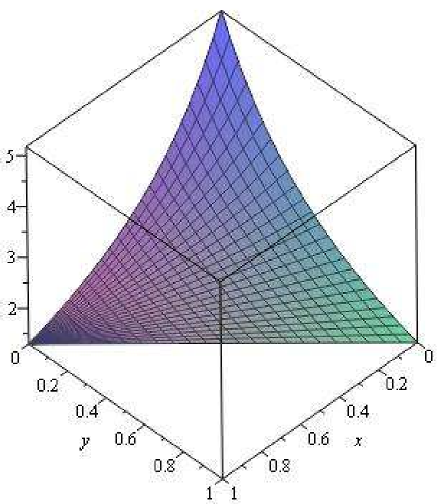
حال با برازش اسپلاین درجه ۳ به داده‌های تابع

$$\hat{f}(x) = \begin{cases} ۰/۲۸ + ۰/۳۹x + ۰/۸۱x^۳ & ۰ < x < ۰/۱ \\ ۰/۲۹ + ۰/۳٦x + ۰/۳۱x^۲ - ۰/۲۲x^۳ & ۰/۱ < x < ۰/۲ \\ ۰/۲۸ + ۰/۳۹x + ۰/۱۴x^۲ + ۰/۰٦x^۳ & ۰/۲ < x < ۰/۳ \\ ۰/۲۹ + ۰/۳۸x + ۰/۲۱x^۲ - ۰/۰۲x^۳ & ۰/۳ < x < ۰/۴ \\ ۰/۲۹ + ۰/۳۸x + ۰/۱۸x^۲ + ۰/۰۰۵x^۳ & ۰/۴ < x < ۰/۵ \\ ۰/۲۹ + ۰/۳۸x + ۰/۲۰x^۲ - ۰/۰۰۵x^۳ & ۰/۵ < x < ۰/٦ \\ ۰/۲۸ + ۰/۴۰x + ۰/۱٦x^۲ + ۰/۰۱٦x^۳ & ۰/٦ < x < ۰/۷ \\ ۰/۳۱ + ۰/۲۹x + ۰/۳۲x^۲ - ۰/۰۵۸x^۳ & ۰/۷ < x < ۰/۸ \\ ۰/۱٦ + ۰/۸۲x - ۰۳۴x^۲ + ۰/۲۱x^۳ & ۰/۸ < x < ۰/۹ \\ ۰/۹۲ - ۱.٦۹x + ۲/۴۴x^۲ - ۰/۲x^۳ & ۰/۹ < x < ۱ \end{cases}$$

به‌دست می‌آید که با استفاده از آن تابع چگالی احتمال ماکسیمم آنتروپی بورگز به‌صورت

$$f_M(x,y)=\frac{۱}{\hat{f}_۱(x)+\hat{f}_۲(y)+۰/۳۸xy}\quad ۰<x+y<۱,\ ۰<x<۱$$

حاصل خواهد شد که نمودار آن در شکل ۲ نشان داده شده است.




شکل ۲: نمودار تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی بورگز

## ۴ بحث و نتیجه‌گیری

برای تعیین تابع توزیع ماکسیمم آنتروپی می‌توان از اندازه‌های دیگر آنتروپی نظیر کاپور، رنی و ... نیز استفاده نمود. همچنین به‌دست آوردن ضابطه تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی تانیجا منجر به حل معادله لامبرت می‌شود که حل آن به‌عنوان یک مسئله جالب، قابل توجه خواهد بود. برای تعیین تابع چگالی احتمال دو متغیره ماکسیمم آنتروپی به روش عددی ممکن است از سایر روش‌ها نظیر بسط به توابع متعامد مانند هرمیت، بسل، همچنین سری فوریه، موجک‌ها و ...

نیز استفاده نمود که نیاز به مطالعه بیشتر دارد.

## تقدیر و تشکر

نویسنده لازم می‌داند که از داوران محترم مجله علوم آماری و هیئت تحریریه به دلیل دقت در بازبینی مقاله و پیشنهادات ارزنده‌شان که منجر به بهبود نسخه اولیه مقاله شده است، تشکر و قدردانی نماید.

## مراجع

Dudik, M., Phillips, S. J. and Schapire R. E. (2007), Maximum Entropy Density Estimation with Generalized Regularization and an Application to Species Distribution Modeling, *Journal of Machine Learning Research*, **8**, 1217-1260.

Dong, S., Tao, S., Lei, S. and Soares, C. G. (2013), Parameter Estimation of the Maximum Entropy Distribution of Significant Wave Height, Journal of Coastal Research, **29**, 597-604.

Euler, L. (1744), *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes Sive Solutio Problematis Isoperimetrici Latissimo Sensu Accepti*, Lausanne and Geneva, E65A, Opera Omnia.

Frechet M. (1951), Surles Tableaux de Correlation Dont les Marges Sont Donness, *Annales University Lyon*, **14**, 53-77.

Guiasu, S. (1971), Weighted Entropy, *Reports on Mathematics Physics*, **2**, 165-179.

Jaynes, E. T. (1957), Information Theory and Statistical Mechanics, *Physical Reviews*, **106**, 620-630.

Jaynes, E. T. (1968), Prior Probabilities, *IEEE Transaction*, **4**, 227-248.

Kapur J. N. (1981), Multivariate Multirectangular Distributions, *University Waterloo*

Kapur, J. N. (1989), *Maximum Entropy Models in Science and Engineering*, New Delhi, India, Wiley Eastern Limited, *Journal of Machine Learning Research*, **8**, 1217-1260.

Mansoury, S., Pasha, E. and Mohammadzadeh, M. (2005), Determination of Maximum Bayesian Entropy Probability Distribution, *Journal of Sciences Islamic Republic of Iran*, **16**, 339-345.

Meeuwissen, A. M. H. and Bedford, T. (1997), Minimaly Informative Distributions with Given Rank Correlation for Use in Uncertainty Analysis, *Journal of Statistics Computation and Simulation*, **57**, 143-17.

Miller, D. J. and Wei-han, L. (2002), On the Recovery of Joint Distributions from Limited Information, *Journal of Econometrics*, **107**, 259-274.

Nelsen R. B. (1987), Discrete Bivariate Distribution with Given Marginals and Correlation, *Comm. Simulation Computation*, **16**, 199-208.

Nelsen R. B. (1991), Copulas and Assocation in Advances in Probability Distribution with Given Marginals, *Kluwer Academic Publication*, Dordrecht/Boston/London.

Pasha, E. and Mansoury, S.(2008), Determination of Maximum Entropy Multivariate Probability Distribution under Some Constraints, *Applied Mathematical Sciences*, **2**, **57**, 2843-2849.

Phillips, S. J., Anderson, R. P. and Schapire, R. E. (2006), Maximum Entropy Modeling of Species Geographic Distributions, *Ecological Modelling*, **190**, 231-259.

Patil G. P. and Joshi S. W. (1968), A Dictionary and Bibliography of Discrete Distribution, *Oliver and Boyd, London.*

Reza, F. M. (1961), *Introduction to Information Theory*, McGraw-Hill, New York.

Russak, I. B. (2002), *Calculus of Variations MA 4311 Lecture Notes*, Navel Postgraduate School, California.

Shore, J. E. (1978), Derivation of Equilibrium and Time-Dependent Solution to M/M/$\infty$/N and M/M/$\infty$ Queuing Systems Using Entropy Maximation, *ASIAS Confrance Proceeding*, 483-487.