

نگرشی بر محاسبه ی آنتروپی ماکزیمم به کمک نرم افزار مطلب و برآوردگر درستنمایی ماکزیمم

فاطمه عباس پور^۱، غلامرضا محتشمی برزادران^۲ و یحیی محتشمی^۳

چکیده:

یکی از مباحثی که نقش مهمی در استنباط آماری دارد، ماکزیمم سازی آنتروپی در یک کلاس از توزیع ها متناظر با یک سری از قیود می باشد. در این راستا برخی از توزیع هایی را که دارای آنتروپی ماکزیمم تحت قیود معین می باشد را به دست می آوریم. ضمن معرفی برآوردگر آنتروپی ماکزیمم، ارتباط این برآوردگر را با برآوردگر درستنمایی ماکزیمم مورد بررسی قرار می دهیم. در این ارتباط برنامه ای که به کمک نرم افزار مطلب نوشته شده است ارائه می دهیم و به کمک چند مثال ارتباط این دو برآوردگر را روشن می سازیم.

واژه های کلیدی: آنتروپی شانون، برآورد آنتروپی ماکزیمم، برآورد پارامتر، برآورد درستنمایی ماکزیمم، توزیع آنتروپی ماکزیمم.

۱ مقدمه

روش برآورد درستنمایی ماکزیمم (MLE) یکی از قدیمی ترین و پراهمیت ترین روش ها در نظریه ی برآوردها است. این روش توسط فیشر [۱] در انتقاد از روش برآورد گشتاوری مورد استفاده قرار گرفت. وی همچنین به معرفی سازگاری، کارایی و بسندگی پرداخت. همچنین هابر [۵]، نرمال سازی مجانبی برآوردگر درستنمایی ماکزیمم را ارائه داد. روش MLE با وجود این که از لحاظ شهودی مورد توجه است، بهترین روش نیست و نمی تواند همیشه مورد استفاده قرار بگیرد. در حقیقت هیچ قضیه ای برای برآورد پارامترها وجود ندارد که ثابت کند که این روش، بجز در موارد مجانبی، دارای خواص

بهینه است. شانون [۹]، برای اولین بار اندازه ای را معرفی کرد که میزان اطلاع منبع را اندازه گیری می کرد و این اندازه را آنتروپی نامید. اصل ماکزیمم آنتروپی (ME)، ابتدا توسط جینز [۶]، بیان شد که یک بحث در مکانیک آماری است و این مسئله مطرح شد که ماکزیمم آنتروپی با قانون دوم ترمودینامیک مرتبط است که بیان می کند که آنتروپی یک سیستم مکانیکی همیشه افزایشی است. کاپور [۸] و کواسان و ژانگ، وانگ و چانگ [۱۰] اصل آنتروپی ماکزیمم را از دیدگاه های گوناگون مورد بررسی قرار داده اند.

عبارت آنتروپی از بیش از یکصد و پنجاه سال پیش در

^۱ دانشجوی کارشناسی ارشد دانشگاه فردوسی مشهد
^۲ عضو هیئت علمی دانشگاه فردوسی مشهد
^۳ دانشجوی کارشناسی دانشگاه فردوسی مشهد

صورت وجود باید در معادله‌ی زیر صدق کند:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = 0, \quad j = 1, \dots, k, \quad \theta = (\theta_1, \theta_2, \dots, \theta_k).$$

در راه یافتن MLE ، ممکن است با مشکلاتی مواجه شویم. ممکن است عبارت $\frac{\partial L}{\partial \theta} = 0$ چندین ریشه داشته باشد، ممکن است تابع درستنمایی به ازای هر مقدار در θ ، مشتق پذیر نباشد و یا حل معادله $\frac{\partial L}{\partial \theta} = 0$ مشکل باشد، در این شرایط از روش‌های عددی برای یافتن برآوردگر درستنمایی ماکزیمم استفاده می‌کنیم. در ادامه چند ویژگی این برآوردگر ارائه می‌شود:

- برآورد MLE ممکن است یکتا نباشد. حتی اگر برآورد MLE یکتا باشد، لزومی ندارد که ناریب باشد.

- اگر T آماره‌ی بسنده برای خانواده چگالی احتمال $\{f_\theta(\cdot), \theta \in \Theta\}$ باشد و برآوردگر MLE برای پارامتر θ یکتا باشد، آنگاه برآوردگر MLE تابعی از T است و اگر برآوردگر ML پارامتر θ یکتا نباشد، می‌توانیم آن را به گونه‌ای بیابیم که تابعی از T باشد.

- اگر شرایط نظم برای برقراری نامساوی کرامر-رائو، برقرار باشد و θ متعلق به مجموعه‌ی اعداد حقیقی باشد و اگر برآوردگر $\hat{\theta}$ از θ ، کران پایین نامساوی کرامر-رائو را نتیجه دهد، معادله‌ی درستنمایی حل یکتایی برای $\hat{\theta}$ ارائه می‌دهد. در این شرایط برآورد MLE ، کاراترین برآوردگر است.

فیزیک و مهندسی مطرح شده است، به همین دلیل بسیاری گمان می‌کنند که بحث در مورد ماکزیمم سازی آنتروپی تنها منحصر به فیزیک و علوم مهندسی است، در حالی که این بحث در بسیاری از علوم دیگر مانند ریاضیات، آمار، نظریه‌ی مخابرات، اقتصاد، تجارت، ژئوفیزیک و... کاربرد دارد.

افرادی مانند جعفری [۳]، گزندار [۲] و کاپور [۷] به بررسی ارتباط بین برآوردگر درستنمایی ماکزیمم و آنتروپی ماکزیمم پرداخته‌اند. در این مقاله این مفاهیم را بررسی کرده و ارتباط بین دو برآورد معرفی شده را بیان خواهیم کرد. علاوه بر این، براساس برنامه‌ی نرم‌افزاری که با ایده از مقاله‌ی جعفری [۳]، برآورد آنتروپی ماکزیمم پارامترها را به ازای هر تعداد قید دلخواه، نوشته و مفاهیم فوق را به کمک چند مثال ارائه خواهیم کرد.

۲ برآوردگر درستنمایی ماکزیمم

فرض کنید $(X_1, X_2, \dots, X_n)$ یک بردار تصادفی با تابع چگالی احتمال $f_\theta(x_1, x_2, \dots, x_n)$ ، $\theta \in \Theta$ باشد. تابع $L(\theta; x_1, x_2, \dots, x_n) = f_\theta(x_1, x_2, \dots, x_n)$ تابع درستنمایی نامیده می‌شود و تابعی از θ است. معمولاً فرض می‌کنیم که مشاهدات نمونه‌ای تصادفی هستند. اصل درستنمایی ماکزیمم بیان می‌کند که برآورد ماکزیمم درستنمایی پارامتر θ ، مقدار $\hat{\theta}$ است که تابع درستنمایی را ماکزیمم می‌کند. یعنی $L(\hat{\theta}; x_1, x_2, \dots, x_n) = \sup_{\theta \in \Theta} L(\theta; x_1, x_2, \dots, x_n)$ با توجه به یکنوا بودن تابع لگاریتم، می‌توان به جای تابع درستنمایی، لگاریتم تابع درستنمایی را ماکزیمم کرد. بنابراین برآوردگر درستنمایی ماکسیمم پارامتر θ ، $\hat{\theta}$ در

ازای هر i ، $p_i \geq 0$ باشد و $\sum_{i=1}^n p_i = 1$ و همچنین

$$\sum_{i=1}^n p_i \varphi_r(x_i) = \mu_r, \quad r = 1, 2, \dots, M, \quad (1)$$

که توابع $\varphi_r(x)$ ، $r = 1, 2, \dots, M$ ، تابع معلوم هستند و μ_r امیدریاضی این مقادیر است. برای ماکزیمم کردن آنتروپی با توجه به شرایط در نظر گرفته شده از روش ضرایب لاگرانژ استفاده می کنیم در این صورت با توجه به شرط $\sum_{i=1}^n p_i = 1$ داریم:

$$p(x_i) = \frac{\exp(-\sum_{r=0}^M \lambda_r \varphi_r(x_i))}{\sum_{i=1}^n \exp(-\sum_{r=0}^M \lambda_r \varphi_r(x_i))},$$

که $M+1$ پارامتر $\lambda_0, \lambda_1, \dots, \lambda_M$ ، در صورت قابل حل بودن دستگاه، از حل $M+1$ معادله ی غیرخطی زیر به دست می آید:

$$\sum_{i=1}^n \varphi_r(x_i) \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x_i)\} = \mu_r, \quad r = 0, 1, \dots, M.$$

نکته ۱ توابع $\varphi_1, \varphi_2, \dots, \varphi_M$ مقادیری معلوم هستند، ولی $\lambda_1, \lambda_2, \dots, \lambda_M$ نامعلوم هستند و مقادیر $\mu_1, \mu_2, \dots, \mu_M$ ممکن است معین نشده باشند در این صورت لازم است که آن ها را برآورد کنیم. اگر برآورد آنتروپی ماکزیمم پارامترهای $\lambda_1, \lambda_2, \dots, \lambda_M$ را که یافته ایم در تابع آنتروپی جایگزین کنیم، ماکزیمم تابع آنتروپی به دست می آید که آن را با H_{max} نشان می دهیم. بنابراین $H_{max} = \lambda_0 + \lambda_1 \mu_1 + \dots + \lambda_M \mu_M$ است. همچنین H_{max} تابعی مقعر از $\mu_1, \mu_2, \dots, \mu_M$ است.

• فرض کنید که $\theta \in \Theta$ ، f_θ خانواده ای از تابع چگالی f_θ و تابع درستنمایی $L(\theta)$ که $\Theta \subseteq R_k$ و $k \geq 1$ باشد. $h: \Theta \rightarrow \Lambda$ نگاشتی از Θ به Λ باشد که Λ بازه ای در R_p ($1 \leq p \leq k$) است. در این صورت اگر $\hat{\theta}$ ، برآورد MLE پارامتر θ باشد، آن گاه $h(\hat{\theta})$ ، برآورد MLE تابع $h(\theta)$ است.

۳ برآوردگر آنتروپی ماکزیمم (ME)

اگر X یک متغیر تصادفی گسسته با تابع احتمال $P = (p_1, p_2, \dots, p_n)$ باشد که برای هر i ، $p_i \geq 0$ و $\sum_{i=1}^n p_i = 1$ باشد، اندازه ی آنتروپی شانون به صورت $H(X) = -\sum_{i=1}^n p_i \ln p_i$ تعریف می شود.

با توجه به تعریف آنتروپی می توان گفت که آنتروپی، نسبت به ترتیب قرار گرفتن احتمالات متقارن است. همچنین آنتروپی یک پیشامد غیرممکن، همان آنتروپی یک پیشامد با احتمال صفر است. علاوه بر آن تابع آنتروپی همواره نامنفی است. آنتروپی تابعی مقعر نسبت به $p_1, p_2, \dots, p_n$ است، یعنی ماکزیمم نسبی در صورت وجود، ماکزیمم مطلق است. همچنین همواره $H(X) \leq \ln n$ است و تساوی هنگامی برقرار است که $p_i = \frac{1}{n}$ ، $i = 1, 2, \dots, n$.

اصل آنتروپی ماکزیمم:

فرض کنید متغیر تصادفی X ، مقادیر $x_1, x_2, \dots, x_n$ را با احتمالات متناظر $p_1, p_2, \dots, p_n$ اتخاذ کند که در آن به

اگر $Q = (q_1, q_2, \dots, q_n)$ تابع چگالی احتمال دلخواهی

باشد که در شرایط اولیه صدق می‌کند، در این صورت:

$$H_{max} - H = \sum_{i=1}^n q_i \ln p_i \geq 0,$$

که در آن $H = -\sum_{i=1}^n q_i \ln q_i$ می‌باشد. به نامساوی

$H_{max} - H \geq 0$ نامساوی شانن گفته می‌شود و

$H_{max} = H$ است اگر و فقط اگر تابع p برابر تابع q باشد.

نکته ۲ اگر فضای نمونه متناهی باشد و هیچ شرطی

بجز $\sum_{i=1}^n p_i = 1$ وجود نداشته باشد، توزیع احتمال

ماکزیمم آنتروپی ($MEPD$)، توزیع یکنواخت است. در

حالی که اگر فضای نمونه نامتناهی باشد، $MEPD$ وجود

ندارد و نیز اگر علاوه بر شرط $\sum_{i=1}^n p_i = 1$ میانگین

حسابی نیز داده شده باشد یعنی $E(X) = m$ باشد،

$MEPD$ توزیع هندسی خواهد بود.

۱.۳ اصل آنتروپی ماکزیمم در حالت

پیوسته

در حالت پیوسته برای بازه‌ی بسته‌ی $[a, b]$ و تابع

چگالی $f(x)$ ، تابع آنتروپی در صورت وجود انتگرال، به

فرم $H(X) = -\int_a^b f(x) \ln f(x) dx$ تعریف می‌شود. در

این صورت اصل آنتروپی ماکزیمم بیان می‌کند که باید

آنتروپی را با توجه به شرط‌های زیر

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

$$\int_a^b f(x) \varphi_r(x) dx = \mu_r, \quad r = 1, \dots, M,$$

ماکزیمم کنیم. در این صورت فرم کلی تابع چگالی‌ای

که از ماکزیمم کردن آنتروپی به دست می‌آید به صورت

زیراست:

$$f(x) = \frac{\exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\}}{\int_a^b \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx},$$

که در آن $\lambda_M, \dots, \lambda_2, \lambda_1$ از حل معادلات غیرخطی زیر

به دست می‌آیند:

$$\int_a^b \varphi_r(x) \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx = \mu_r, \quad r = 1, \dots, M.$$

اگر تابع $f(x)$ -ای که از حل معادلات قبل حاصل

می‌شود در آنتروپی قرار دهیم، مقدار ماکزیمم آنتروپی به

دست می‌آید که آن را با H_{max} نشان می‌دهیم. این تابع

دارای خاصیت تقعر اکید نسبت به $\mu_M, \dots, \mu_2, \mu_1$ است

و اگر X یک متغیر تصادفی تعریف شده روی $(-\infty, \infty)$

باشد و مقادیر $E(X)$ و $E(X^2)$ موجود و برابر m و α^2

باشد آن گاه $MEPD$ ، توزیع نرمال با میانگین m و

واریانس $\alpha^2 - m^2$ است و $H_{max} = \frac{1}{2} \ln 2\pi e(\alpha^2 - m^2)$.

۴ ارتباط بین روش برآورد MLE و

ME

دوروش برآورد آنتروپی ماکزیمم (ME) و درست‌نمایی

ماکزیمم (MLE)، دو شیوه‌ی استنتاج متفاوت از توزیع

احتمال بردار X ارائه می‌دهند. موضوع هر دوی آنها،

انتخاب یک تابع توزیع احتمال برای X است به طوری

که بهترین ارائه از مشاهدات باشد. هنگامی که مقدار

امید ریاضی توابعی داده شده، معلوم باشد از روش ME

استفاده می‌کنیم و هنگامی که نمونه‌ای تصادفی با تابعی

معین داشته باشیم از روش MLE استفاده می‌کنیم.

فرض کنید که

$P = \{p(x) | E(\varphi_r(x))\} = \mu_r, \quad r = 0, 1, \dots, M$ باشد

و $p(x)$ تابع چگالی متغیر تصادفی X تعریف

شده باشد. که $\mu_0 = 1$ و $\varphi_0(x) = 1$ است و (۴)
 $\varphi_r(x)$, $r = 1, 2, \dots, M$ تابع معلوم هستند،

در این صورت اگر تنها داده‌های معلوم روی X , μ_r ها

باشند داریم $\hat{P}(x) = \max_{p \in P} \{-\int_D p(x) \ln p(x) dx\}$

در واقع هدف، ماکزیم کردن $H(X)$ با شرط

$$E(\varphi_r(X)) = \int_D \varphi_r(x) f(x) dx = \mu_r, \quad r = 1, 2, \dots, M, \quad (2)$$

است که با حل کلاسیک این مسئله،

$p(x) = \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\}$ تابع چگالی ای است

که آنتروپی را ماکزیم می‌کند. در این راستا برای برآورد

$M+1$ پارامتر $\lambda_0, \lambda_1, \dots, \lambda_M$ باید معادلات غیرخطی

$$G_r(\lambda) = \int \varphi_r(x) \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx = \mu_r$$

$r = 0, 1, \dots, M$ را حل کنیم. که در ادامه روش عددی

برای حل این معادلات ارائه می‌شود. سری تیلور $G_r(\lambda)$

حول λ^0 را به صورت زیر در نظر می‌گیریم:

$$G_r(\lambda) = G_r(\lambda^0) + (\lambda - \lambda^0)^t [\partial G_r(\lambda)]_{(\lambda = \lambda^0)} = \mu_r, \quad (3)$$

که در آن r مقادیر 0 تا M را اخذ می‌کند و نیز قرار

می‌دهیم

$$\delta = \lambda - \lambda^0, \quad v = [\mu_0 - G_0(\lambda^0), \dots, \mu_M - G_M(\lambda^0)]^t$$

و ماتریس G را به صورت زیر تعریف می‌کنیم:

$$G = (g_{rk}) = \left(\frac{\partial G_r(\lambda)}{\partial \lambda_k} \right)_{(\lambda = \lambda^0)} r, k = 0, 1, \dots, M,$$

در این صورت (۳) به صورت $G\delta = v$ تبدیل می‌شود.

سیستم فوق بر حسب δ حل می‌شود. حال قرار می‌دهیم

$\lambda = \lambda^0 + \delta$ و تکرار عددی را ادامه می‌دهیم تا اینکه δ

مقداری بسیار کوچک شود. ماتریس G متقارن است

و عناصر آن به صورت زیر هستند:

$$g_{rk} = g_{kr} = - \int \varphi_r(x) \varphi_k(x) \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx, \quad (4)$$

که در آن r و k مقادیر 0 تا M را می‌گیرند. پس لازم

است $\frac{M(M-1)}{2}$ انتگرال به فرم (۴) محاسبه شود که

بوسیله ی برنامه‌ی نرم‌افزار مطلب که در پیوست

ارائه شده قابل محاسبه است. حال اگر به جای

داشتن μ_r ها، n نمونه $X_1, X_2, \dots, X_n$ از متغیرهای

تصادفی X را داشته باشیم و فرض کنیم که متغیر

تصادفی X دارای تابع چگالی احتمال $p(x)$ به فرم

$$p(x) = \frac{1}{z(\lambda)} \exp \sum_{r=0}^M \lambda_r \varphi_r(x)$$

باشد که $\lambda_0, \dots, \lambda_M$ پارامترهای آن هستند و نیز $z(\lambda)$ ثابت نرمال‌سازی است.

حال می‌خواهیم با استفاده از روش MLE ، برآورد پارامترها

را بیابیم. در این صورت تابع درستمایی به صورت

$$L(\lambda) = \prod_{i=1}^n p(x_i, \lambda) = \prod_{i=1}^n \frac{1}{z(\lambda)} \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x_i)\}$$

و لگاریتم تابع درستمنمایی برابر

$$\ell(\lambda) = -n \ln z(\lambda) - \sum_{i=1}^n \sum_{r=0}^M \lambda_r \varphi_r(x_i)$$

بنابراین λ ، برآورد MLE پارامتر λ ، از حل معادلات

$$r = 0, 1, \dots, M, \quad \frac{\partial \ln z(\lambda)}{\partial \lambda_r} = \frac{1}{n} \sum_{i=1}^n \varphi_r(x_i)$$

می‌شود که در آن $z(\lambda) = \int \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx$

است. به طور معادل داریم:

$$\int \varphi_r(x) \exp\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\} dx = \frac{1}{n} \sum_{i=1}^n \varphi_r(x_i), \quad r = 0, 1, 2, \dots, M, \quad (5)$$

با مقایسه ی معادله ی (۲) با (۵) به این نتیجه می‌رسیم

که μ_r ها با مقدار برآورد تجربی خود جایگزین شده‌اند.

یعنی $\mu_r = \frac{1}{n} \sum_{i=1}^n \varphi_r(x_i)$; $r = 0, 1, \dots, M$ است.

نامحدود از شرایط را به دست آوریم. در ادامه به بررسی چند نمونه از این مسائل می‌پردازیم.

توزیع نرمال:

توزیع نرمال را که دارای چگالی‌ای به فرم زیر است را در نظر می‌گیریم:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

این توزیع می‌تواند به عنوان یک توزیع دارای آنتروپی ماکزیمم با شرایط $\phi_1(x) = x$, $\phi_0(x) = 1$ و $\phi_2(x) = x^2$ در نظر گرفته شود. در نتیجه تابع چگالی دارای آنتروپی ماکزیمم به صورت زیر خواهد بود:

$$f(x; \mu, \sigma^2) = e^{-\lambda_0 - \lambda_1 x - \lambda_2 x^2}. \quad (7)$$

بامعلوم بودن مقادیر μ_1 و μ_2 می‌توانیم λ_0 , λ_1 و λ_2 را معین کنیم. برای نشان دادن روابط بین برآورد ML و ME به صورت زیر عمل می‌کنیم:

ابتدا فرض می‌کنیم که μ_1 و μ_2 مقادیری معلوم‌اند. به این ترتیب یافتن برآورد ME پارامترها به سادگی انجام می‌گیرد. برای چند مقدار μ_1 و μ_2 ، در جدول زیر برآورد ME پارامترهای λ_0 , λ_1 و λ_2 را به دست آورده‌ایم:

که با استفاده از روش‌های عددی معرفی شده می‌توان برآورد پارامترهای λ را یافت. در واقع ارتباط بین این دو روش برآورد را می‌توان به صورت زیر خلاصه کرد:

فرض می‌کنیم که $\varphi_r(x)$ و μ_r , $r = 0, 1, \dots, M$ داده شده باشد می‌خواهیم توزیع با ماکزیمم مقدار آنتروپی را بیابیم. در این صورت $p(x; \lambda)$ به صورت زیر خواهد بود:

$$p(x; \lambda) = \exp\left\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\right\} \quad (6)$$

ضرایب لاگرانژ λ در این رابطه از حل معادلات $G_r(\lambda) = \int \varphi_r(x) \exp\left\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\right\} dx = \mu_r$ حاصل می‌شود.

روش MLE :

فرض می‌کنیم که $\varphi_r(x)$, $r = 0, 1, \dots, M$ و نمونه‌ی $x_1, \dots, x_n$ با تابع چگالی به فرم (۶) داده شده باشد. می‌خواهیم پارامترهای $\lambda_0, \dots, \lambda_M$ را برآورد کنیم که با توجه به برآورد MLE متغیرهای μ_r داریم:

$$\begin{aligned} G_r(\lambda) &= \int \varphi_r(x) \exp\left\{-\sum_{r=0}^M \lambda_r \varphi_r(x)\right\} dx \\ &= \frac{1}{n} \sum_{i=1}^n \varphi_r(x_i), \end{aligned}$$

جعفری [۳]، به کمک یک برنامه نرم‌افزار مطلب، برآورد ME پارامترها در توزیع گاما و چهارتایی^۴ را به دست آورده است. در این مقاله با بازنویسی برنامه‌ی مورد ذکر، توانستیم برآورد ME پارامترها، برای تعداد دلخواه و

جدول ۱. برآورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی فرم (۷).

آنتروپی	λ_2	λ_1	λ_0	μ_2	μ_1
۱.۴۰۸	-۰.۰۰۶	۰	۱.۵	۱.۵	۰
۱.۲۱۷	۰.۴۵۶	-۰.۶۲۳	۱.۰۷	۱	۰.۵
۱.۲۹۳	-۰.۳۴۳	-۰.۱	۲	۲	۰.۲

حال فرض می‌کنیم که مقادیر μ_1 و μ_2 نامعلوم باشند. توزیع نرمال $\mu = ۰.۲$ و $\sigma^2 = ۱$ است. برای سه حجم نمونه به این ترتیب با استفاده از شبیه‌سازی، آن‌ها را برآورد می‌کنیم. فرض می‌کنیم که مقدار نظری پارامترهای λ_0 ، λ_1 و λ_2 ، در جدول زیر آمده‌است:

جدول ۲. برآورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی فرم (۷) به ازای برآورد ML پارامترهای μ_1 و μ_2 .

آنتروپی	λ_2	λ_1	λ_0	μ_2	μ_1	حجم نمونه
۱.۳۶۸	۰.۱۸۳	-۰.۲۱۲	۱.۲۰۵	۱.۱۶۵	۰.۲۳۸	۵۰۰
۱.۳۵۹	۰.۲۱۱	-۰.۲۲۵	۱.۱۷۵	۱.۱۳۱	۰.۲۴۵	۷۰۰
۱.۳۹۹	۰.۵۲	-۰.۲۲۸۸	۰.۹۲۴۵	۱.۰۱۱	۰.۲۲	۱۵۰۰

توزیع لگ نرمال: اگر $\mu_1 = E(\log(X))$ و نتیجه:

$$\mu_2 = E(\log(X))^2$$

معلم باشد، آن گاه توزیع لگ نرمال، چگالی با آنتروپی ماکزیمم خواهد بود. در

در جدول (۳) برای سه مقدار معلوم μ_1 و μ_2 ، برآورد پارامترهای λ_0 ، λ_1 و λ_2 ارائه شده‌است:

جدول ۳. برآورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی فرم (۸).

آنتروپی	λ_2	λ_1	λ_0	μ_2	μ_1
۰.۶۹۳	۰.۰۰۷	۰.۰۲۸	۰.۶۹۴	۱	-۰.۳
۰.۶۷۱	-۰.۰۰۲	۰.۱۹۲	۰.۷۷۱	۱.۵	-۰.۵
۰.۶۳۲	-۰.۲۰۵	-۰.۶۰۵	۰.۷۷۸	۱.۳	-۰.۲

حال فرض می‌کنیم که μ_1 و μ_2 ، مقادیری نامعلوم باشند. پس با در نظر گرفتن مقدار نظری پارامترهای توزیع لگ نرمال به صورت $m = \text{mean}(\log(X)) = -۰.۷$ و $\sigma^2 = \text{var}(\log(X)) = ۱$ و به کمک شبیه‌سازی توزیع

فوق، μ_1 و μ_2 را برآورد می‌کنیم و برای سه حجم نمونه مختلف، برآورد λ_0 ، λ_1 و λ_2 را به صورت زیر محاسبه می‌کنیم:

جدول ۴. برآورد ME پارامترهای λ_0 ، λ_1 و λ_2 در چگالی فرم (۸) به ازای برآورد ML پارامترهای μ_1 و μ_2 .

آنتروپی	λ_2	λ_1	λ_0	μ_2	μ_1	حجم نمونه
۰.۵۹۸	۰.۱۹۸	۰.۷۸۳	۰.۸۱۸	۱.۳۵۴	-۰.۶۳۵	۵۰۰
۰.۵۵۸	۰.۲۲۴	۰.۹۲۲	۰.۸۷۵	۱.۴۹۴	-۰.۷۰۷	۷۰۰
۰.۷۳۱	۰.۴۴۱	۱.۵۶۶	۱.۱۴۵	۱.۴۹۸	-۰.۶۸۷	۱۵۰۰

توزیع کوشی: اگر $\mu = E(\log(1 + X^2))$ ، مقداری معلوم باشد، آن گاه توزیع کوشی چگالی با آنتروپی ماکزیمم خواهد بود و بنابراین:

آنتروپی	λ_1	λ_0	μ
۰.۵۹۷	۰.۹	۰.۱۴۷	۰.۵
۰.۶۹۲	۰.۰۷۸	۰.۶۳۸	۰.۷
۰.۵۴۴	-۱.۰۹۵	۱.۶۳۹	۱

$$f(x) = \exp\{\lambda_0 + \lambda_1 \log(1 + x^2)\} \quad (9)$$

و اگر μ ، نامعلوم باشد و اگر پارامتر مکان آن را صفر و پارامتر مقیاس آن را یک در نظر بگیریم، به کمک شبیه سازی می توانیم μ را برآورد کنیم و به این ترتیب، برآورد ME پارامترهای λ_0 و λ_1 را بیابیم:

در جدول (۵) برای سه مقدار معلوم μ ، برآورد پارامترهای λ_0 و λ_1 آمده است:

جدول ۶. برآورد ME پارامترهای λ_0 و λ_1 در چگالی فرم (۹) به ازای برآورد ML پارامتر μ .

آنتروپی	λ_1	λ_0	μ
۰.۵۹۷	۰.۹	۰.۱۴۷	۰.۵
۰.۶۹۲	۰.۰۷۸	۰.۶۳۸	۰.۷
۰.۵۴۴	-۱.۰۹۵	۱.۶۳۹	۱

توزیع گاوسی معکوس: توزیع گاوسی معکوس را که دارای تابع چگالی به فرم:

$$f(x) = \exp\{\lambda_0 + \lambda_1 \log(x) + \lambda_2 x + \frac{\lambda_3}{x}\}, \quad (10)$$

باتوجه به مقادیر μ_1 ، μ_2 و μ_3 می توانیم λ_0 ، λ_1 و λ_2 و λ_3 را معین کنیم. در جدول (۷) برای چندین مقدار معلوم μ_1 ، μ_2 و μ_3 ، برآورد پارامترهای λ_0 و λ_1 و λ_2 و λ_3 را به دست آورده ایم:

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right\}, \quad x > 0, \lambda > 0, \mu > 0,$$

است را در نظر می گیریم. این توزیع می تواند به عنوان توزیع دارای آنتروپی ماکزیمم هنگامی که میانگین حسابی و هندسی و هارمونیک معلوم باشند در نظر گرفته شود. در این صورت فرم چگالی آنتروپی ماکزیمم به

جدول ۷. برآورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی فرم (۱۰).

آنتروپی	λ_3	λ_2	λ_1	λ_0	μ_3	μ_2	μ_1
۰.۱۲۳	۱.۱۳	-۸.۶۹	۸.۸۶	۰.۰۲۶	۳	۱	-۰.۵
۰.۶۴	-۰.۰۵۴	۳.۳۲	-۱.۴۹	-۱.۳۰۶	۲.۵	۱	-۰.۲
۰.۰۳۳	۰.۶۵۹	-۸.۳۷	۷.۳۵	۱۰.۲۹۹	۳	۱.۲	-۰.۳

۵ نتیجه گیری

در این مقاله ضمن معرفی ارتباط بین ME و MLE ، به دنبال یافتن توزیعی بودیم که دارای آنتروپی ماکزیمم تحت شرایط و قیود خاص است و با حل معادلات غیرخطی‌ای که از قیود حاصل می‌شود و با استفاده از برنامه‌ی نرم‌افزار مطلب که در پیوست آمده، مقادیر آنتروپی ماکزیمم و فرم تابع چگالی دارای آنتروپی ماکزیمم را به دست آوردیم. در واقع ایده‌ی این برنامه‌ی نرم‌افزار از مقاله‌ی جعفری گرفته شده بود.

جعفری [۳]، به کمک یک برنامه نرم‌افزار مطلب، برآورد ME —ی پارامترها در توزیع گاما و توزیع چهارتایی را به دست آورده است. با بازنویسی برنامه‌ی فوق، توانستیم برآورد ME پارامترها، برای تعداد دلخواه و نامحدود از

شرایط را به دست آوریم. تفاوت برنامه‌ای که بازنویسی شده با برنامه‌ی مقاله‌ی جعفری این است که جعفری در مقاله‌ی خود برنامه‌ی نرم‌افزاری ارائه کرده‌است که تنها برای چند توزیع خاص و چند قید اولیه برآورد آنتروپی ماکزیمم پارامترها را یافته است، در این صورت برای یافتن برآورد ME پارامترهای بایست یک برنامه‌ی مخصوص به آن را می‌نوشتیم که کاری مشکل و وقت گیر بود، بنابراین برنامه‌ی فوق را بازنویسی کردیم و برنامه‌ای یافتیم که بدون هیچ محدودیتی و برای هر توزیعی برآورد آنتروپی ماکزیمم پارامترها را ارائه می‌کرد.

سپاسگزاری

نویسندگان این مقاله از حمایت قطب داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد تشکر می‌نمایند.

مراجع

- [1] Fisher, R.A. (1925). *Theory of Statistical Estimation*, proc.camb.phil.soc.22, 700-725.
- [2] Grendar, M.Jr. and Grendar, M. (2003). *Maximum probability and maximum entropy methods: bayesian interpretation*, arxiv:physics/0308005, v1, 1.
- [3] Djafari, Mohammad (1991). *A matlab program to calculate the maximum entropy distrubtion*, Proc. of the 11the Int, Maxent workshop, seattle, USA.
- [4] Rohatgi, K.V. and Ehsanes Saleh, A.K.MD . (2000). *An Introduction to Probability and Statistics*, 2nd ed. 409-421.
- [5] Huber, P.J. (1967). *The behaviour of maximum likelihood under nonstandard conditions*, Fifth Berkeley Symp.Math. Statist. Prob., 1, 175-194.

- [6] Jaynes, E.T. (1957). *Infotmation Theory and Statistical Mechanics*, Physical Reviews, 106:629-630.
- [7] Kapur, J.N. (1989). *Maximum Entropy Models in Science and Engineering*, Wiley Eastern Limited.
- [8] Kapur, J.N. and Kevasan, H.K. (1992). *Entropy Optimization principles with Applications*, Academic Press, Inc.
- [9] Shannon, C.E. (1948). *A Mathematical Theory of Communication*, Bell System Tech.J., 27, 379-423, 623-659.
- [10] Zhang, C.T., Wang, W. and Chang, C.Y. (2010). *Voltage stability analysis based on probabilistic power flow and maximum entropy*, Wiley IET Gener. Transm. Distrib., Vol. 4, Lss. 4, 530-537.

پیوست

برنامه مطلب تهیه شده برای به دست آوردن ماکزیمم آنتروپی تحت تعداد دلخواه قید داده شده، به صورت زیر است:

```
function[lambda,p,entr] = me(mu,x,lambdax)
```

```
eps = 1e-6;
```

```
mu = [1,mu];
```

```
mu = mu'
```

```
xmin = x(1);
```

```
xmax = x(length(x));
```

```
dx = x(2) - x(1);
```

```
if(nargin == 2)
```

```
lambda = zeros(size(mu));
```

```
lambda(1) = log(xmax - xmin);
```

```
else
```

```
lambda = lambda(x);
```

```
end
```

```

n = length(lambda);
fin = fin - \ (x);
iter = 0;
while \
iter = iter + \;
p = exp(-(fin * lambda));
g = zeros(n, \);
for i = \ : n
g(i) = dx * sum(fin(:, i) .* p);
end
entr(iter) = lambda' * g
gnk = zeros(n, n);
gnk(\, :) = -g';
gnk(:, \) = -g;
for i = \ : n
for j = \ : i
gnk(i, j) = -dx * sum(fin(:, j) .* fin(:, i) .* p);
end
end
for i = \ : n
for j = i + \ : n
gnk(i, j) = gnk(j, i);
end
end
v = mu - g;
delta = gnk v;
lambda = lambda + delta;
if(abs(delta./lambda) < eps)break, end

```

```

if(iter > ۲)
if(abs((entr(iter) - entr(iter - ۱))/entr(iter)) < eps),break,end
end
end
p = exp(-(fin * lambda));
plot(x,p);
entr = entr(:);
disp('Endoftheprogram');
end
function fin = fin - ۱(x)
m = input('enternumberof finsyouwant')
fin = zeros(length(x),m);
for i = ۱ : m
fin(:,i) = input('enter fin');
end
end
end

```