نگرشی بر محاسبه ی آنتروپی ماکزیمم به کمک نرم‌افزار 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 توزیع‌ها متناظر با یک سری از قیود می‌باشد. در این راستا برخی از توزیع‌هایی را که دارای آنتروپی ماکزیمم تحت قیود معین می‌باشد را به دست می آوریم. ضمن معرفی برآوردگر آنتروپی ماکزیمم، ارتباط این برآوردگر را با برآوردگر درستنمایی ماکزیمم مورد بررسی قرار می‌دهیم. در این ارتباط برنامه‌ای که به کمک نرم‌افزار مطلب نوشته شده است ارائه می‌دهیم و به کمک چند مثال ارتباط این دو برآوردگر را روشن می‌سازیم.

واژه‌های کلیدی: آنتروپی شانون، برآورد آنتروپی ماکزیمم، برآورد پارامتر، برآورد درستنمایی ماکزیمم، توزیع آنتروپی ماکزیمم.

۱ مقدمه

روش برآورد درستنمایی ماکزیمم(MLE) یکی از قدیمی‌ترین و پراهمیت‌ترین روش ها در نظریه‌ی برآوردهاست. 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]، نرمال‌سازی مجانبی برآوردگر درستنمایی ماکزیمم را ارئه داد. روش MLE با وجود این که از لحاظ شهودی مورد توجه است، بهترین روش نیست و نمی‌تواند همیشه مورد استفاده قرار بگیرد. در حقیقت هیچ قضیه‌ای برای برآورد پارامترها وجود ندارد که ثابت کند که این روش، بجز در موارد مجانبی، دارای خواص بهینه است. شانون[۹]، برای اولین بار اندازه‌ای را معرفی کرد که میزان اطلاع منبع را اندازه‌گیری می‌کرد وی این اندازه را آنتروپی نامید. اصل ماکزیمم آنتروپی (ME)، ابتدا توسط جینز[٦]، 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 می‌کند که آنتروپی یک سیستم مکانیکی همیشه افزایشی است. کاپور[۸] و کواسان و ژانگ ، وانگ و چانگ [۱۰]اصل آنتروپی ماکزیمم را از دیدگاه‌های گوناگون مورد بررسی قرار داده‌اند.

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

فیزیک و مهندسی مطرح شده است، به همین دلیل بسیاری گمان می کنند 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، نظریه ی مخابرات، اقتصاد، تجارت، ژئوفیزیک و... کاربرد دارد.

افرادی مانند جعفری [۳]، گرندار[۲] و کاپور[۷] به بررسی ارتباط بین برآوردگر درستنمایی ماکزیمم و آنتروپی ماکزیمم پرداخته‌اند. دراین مقاله این مفاهیم را بررسی کرده و ارتباط بین دو برآورد معرفی شده را بیان خواهیم کرد. علاوه براین، براساس برنامه‌ی نرم‌افزاری که با ایده ازمقاله ی 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 $(X_۱, X_۲, ..., X_n)$، یک بردار تصادفی با تابع چگالی احتمال $\theta \in \Theta$ ،$f_\theta(x_۱, x_۲, ..., x_n)$، باشد. تابع$L(\theta; x_۱, x_۲, ..., x_n) = f_\theta(x_۱, x_۲, ..., x_n)$ 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 ،$\hat{\theta}$، مقداری است که تابع درستنمایی را ماکزیمم می‌کند. یعنی $L(\hat{\theta}\ ; x_۱, x_۲, ..., x_n) = \sup_{\theta\in\Theta}\ L(\theta; x_۱, x_۲, ..., x_n)$. 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 صورت وجود باید در معادله‌ی زیر صدق کند:

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = ۰,\ j = ۱, ..., k,\ \ \theta = (\theta_۱, \theta_۲, ..., \theta_k).$$

در راه یافتن MLE، ممکن است با مشکلاتی مواجه شویم. ممکن است عبارت $\frac{\partial L}{\partial \theta} = ۰$ چندین ریشه داشته باشد، ممکن است تابع درستنمایی به ازای هر مقدار در θ، مشتق پذیر نباشد و یا حل معادله $\frac{\partial L}{\partial \theta} = ۰$ مشکل باشد، در این شرایط از روش‌های عددی برای یافتن برآوردگردرستنمایی ماکزیمم استفاده می‌کنیم. در ادامه چند ویژگی این برآوردگر ارائه می‌شود:

- برآورد MLE ممکن است یکتا نباشد. حتی اگربرآورد MLE یکتا باشد، لزومی ندارد که نااریب باشد.
- اگر T آماره‌ی بسنده برای خانواده چگالی احتمال $\{f_\theta(.), \theta \in \Theta\}$ 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، معادله ی درستنمایی حل یکتایی برای $\hat{\theta}$ ارائه می‌دهد. در این شرایط برآورد MLE، کاراترین برآوردگر است.

• فرض کنید که $f_\theta,\ \theta \in \Theta$ خانواده‌ای از تابع چگالی f_θ و تابع درستنمایی $L(\theta)$ که $\Theta \subseteq R_k$ و $k \geq ۱$ باشد. $h : \Theta \to \Lambda$ نگاشتی از Θ به Λ باشد که Λ بازه‌ای در R_p $(۱ \leq p \leq k)$ 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 $P = (p_۱, p_۲, ..., p_n)$ باشد که برای هر i، $p_i \geq ۰$ و$\sum_{i=۱}^n p_i = ۱$ باشد، اندازه‌ی آنتروپی شانون به صورت $H(X) = -\sum_{i=۱}^n p_i \ln p_i$ تعریف می‌شود.

با توجه به تعریف آنتروپی می‌توان 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 $p_n, ..., p_۲, p_۱$ است، یعنی ماکزیمم نسبی در صورت وجود، ماکزیمم مطلق است. همچنین همواره $H(X) \leq \ln n$ است و تساوی هنگامی برقرار است که $p_i = \frac{۱}{n}$، $i = ۱, ۲, ..., n$.

اصل آنتروپی ماکزیمم :

فرض کنید متغیر تصادفی X، مقادیر $x_n, ..., x_۲, x_۱$ را با احتمالات متناظر $p_n, ..., p_۲, p_۱$ اتخاذ کند که در آن به ازای هر i، $p_i \geq ۰$ باشد و $\sum_{i=۱}^n p_i = ۱$ و همچنین

$$\sum_{i=۱}^n p_i \varphi_r(x_i) = \mu_r, \quad r = ۱, ۲, ..., M, \qquad (۱)$$

که توابع $\varphi_r(x)$ ، $r = ۱, ۲, ..., M$، 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 می‌کنیم در این صورت با توجه به شرط $\sum_{i=۱}^n p_i = ۱$ داریم:

$$p(x_i) = \frac{\exp(-\sum_{r=۰}^M \lambda_r \varphi_r(x_i))}{\sum_{i=۱}^n \exp(-\sum_{r=۰}^M \lambda_r \varphi_r(x_i))},$$

که $M + ۱$ پارامتر $\lambda_M, ..., \lambda_۱, \lambda_۰$، در صورت قابل حل بودن دستگاه، از حل $M + ۱$ معادله‌ی غیرخطی زیر به دست می‌آید:

$$\sum_{i=۱}^n \varphi_r(x_i) \exp\{-\sum_{r=۰}^M \lambda_r \varphi_r(x_i)\} = \mu_r, \quad r = ۰, ۱, ..., M.$$

نکته ۱ توابع $\varphi_M, ..., \varphi_۲, \varphi_۱$ مقادیری معلوم هستند، ولی $\lambda_M, ..., \lambda_۲, \lambda_۱$ نامعلوم هستند و مقادیر $\mu_M, ..., \mu_۲, \mu_۱$ ممکن است معین نشده باشند در این صورت لازم است که آن ها را برآورد کنیم. اگر برآورد آنتروپی ماکزیمم پارامترهای $\lambda_M, ..., \lambda_۲, \lambda_۱$ را که یافته‌ایم در در تابع آنتروپی جایگزین کنیم، ماکزیمم تابع آنتروپی به دست می آید که آن را با H_{max} نشان می‌دهیم. بنابراین $H_{max} = \lambda_۰ + \lambda_۱\mu_۱ + ... + \lambda_M\mu_M$ است. همچنین H_{max} تابعی مقعر از $\mu_M, ..., \mu_۲, \mu_۱$ است.

اگر $Q=(q_۱,q_۲,...,q_n)$ تابع چگالی احتمال دلخواهی باشد که در شرایط اولیه صدق می‌کند، در این صورت:

$$H_{max}-H=\sum_{i=۱}^{n}q_i\ln p_i\geq ۰,$$

که در آن $H=-\sum_{i=۱}^{n}q_i\ln q_i$ می‌باشد. به نامساوی $H_{max}-H\geq ۰$ نامساوی شانون گفته می‌شود و $H=H_{max}$ است اگر و فقط اگر تابع p برابر تابع q باشد.

نکته ۲ اگر فضای نمونه متناهی باشد و هیچ شرطی بجز $\sum_{i=۱}^{n}p_i=۱$ 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 $\sum_{i=۱}^{n}p_i=۱$ 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 شرط های زیر

$$\begin{aligned}\int_a^b f(x)dx &= ۱\\ \int_a^b f(x)\varphi_r(x)dx &= \mu_r,\quad r=۱,...,M,\end{aligned}$$

ماکزیمم کنیم. در این صورت فرم کلی تابع چگالی‌ای که از ماکزیمم کردن آنتروپی به دست می آید به صورت زیراست:

$$f(x)=\frac{\exp\{-\sum_{r=۰}^{M}\lambda_r\varphi_r(x)\}}{\int_a^b\exp\{-\sum_{r=۰}^{M}\lambda_r\varphi_r(x)\}dx},$$

که در آن $\lambda_۱,\lambda_۲,...,\lambda_M$ از حل معادلات غیرخطی زیر به دست می آیند:

$$\int_a^b\varphi_r(x)\exp\{-\sum_{r=۰}^{M}\lambda_r\varphi_r(x)\}dx=\mu_r,\quad r=۱,...,M.$$

اگر تابع $f(x)$—ای که از حل معادلات قبل حاصل می‌شود در آنتروپی قرار دهیم، مقدار ماکزیمم آنتروپی به دست می آید که آن را با H_{max} نشان می‌دهیم. این تابع دارای خاصیت تقعر اکید نسبت به $\mu_۱,\mu_۲,...,\mu_M$ است و اگر X یک متغیر تصادفی تعریف شده روی $(-\infty,\infty)$ باشد و مقادیر $E(X)$ و $E(X^۲)$ موجود و برابر m و $\alpha^۲$ باشد آن گاه $MEPD$، توزیع نرمال با میانگین m و واریانس $\alpha^۲-m^۲$ است و $H_{max}=\frac{۱}{۲}\ln ۲\pi e(\alpha^۲-m^۲)$.

۴ ارتباط بین روش برآورد MLE و ME

دو روش 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

$P=\{p(x)|E(\varphi_r(x))\}=\mu_r$ ، $r=۰,۱,...,M$ باشد

و $p(x)$، تابع چگالی متغیر تصادفی X تعریف شده باشد. که $\mu_\circ = ۱$ و $\varphi_\circ(x) = ۱$ است و $\varphi_r(x)$ ، $r = ۱, ۲, ..., M$، M تابع معلوم هستند، دراین 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، ماکزیمم کردن $H(X)$ با شرط

$$E(\varphi_r(X)) = \int_D \varphi_r(x)f(x)dx = \mu_r, \quad r = ۱, ۲, ..., M, \quad (۲)$$

است که با حل کلاسیک این مسئله، $p(x) = \exp\{-\sum_{r=\circ}^M \lambda_r\varphi_r(x)\}$، تابع چگالی‌ای است که آنتروپی را ماکزیمم می‌کند. در این راستا برای برآورد $M + ۱$ پارامتر $\lambda_\circ, \lambda_۱, ..., \lambda_M$ باید معادلات غیرخطی $G_r(\lambda) = \int \varphi_r(x)\exp\{-\sum_{r=\circ}^M \lambda_r\varphi_r(x)\}dx = \mu_r$، $r = \circ, ۱, ..., M$ را حل کنیم. که در ادامه روش عددی برای حل این معادلات ارائه می‌شود. سری تیلور $G_r(\lambda)$ حول λ° را به صورت زیر در نظر می‌گیریم :

$$G_r(\lambda) = G_r(\lambda^\circ) + (\lambda - \lambda^\circ)^t[\partial G_r(\lambda)]_(\lambda = \lambda^\circ) = \mu_r, \quad (۳)$$

که در آن r مقادیر ۰ تا M رااخذ می‌کند و نیز قرار می‌دهیم

$$\delta = \lambda - \lambda^\circ, \ v = [\mu_\circ - G_\circ(\lambda^\circ), ..., \mu_M - G_M(\lambda^\circ)]^t$$

و ماتریس G را به صورت زیر تعریف می‌کنیم:

$$G = (g_{rk}) = (\frac{\partial G_r(\lambda)}{\partial \lambda_k})_(\lambda = \lambda^\circ) r, k = \circ, ۱, ..., M,$$

در این صورت (۳) به صورت $G\delta = v$ تبدیل می‌شود. سیستم فوق بر حسب δ حل می‌شود. حال قرار می‌دهیم $\lambda = \lambda^\circ + \delta$ و تکرار عددی را ادامه می‌دهیم تا اینکه δ مقداری بسیار کوچک شود. ماتریس G متقارن است وعناصرآن به صورت زیر هستند:

$$g_{rk} = g_{kr} = -\int \varphi_r(x)\varphi_k(x)\exp\{-\sum_{r=\circ}^M \lambda_r\varphi_r(x)\}dx, \quad (۴)$$

که در آن r و k مقادیر ۰ تا M را می‌گیرند. پس لازم است $\frac{M(M-۱)}{۲}$ 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 $X_۱, X_۲, ..., X_n$ از متغیرهای تصادفی X را داشته باشیم و فرض کنیم که متغیر تصادفی X دارای تابع چگالی احتمال $p(x)$ به فرم $p(x) = \frac{۱}{z(\lambda)}\exp\sum_{r=\circ}^M \lambda_r\varphi_r(x)$ باشد که $\lambda_۱, ..., \lambda_M$، پارامترهای آن هستند و نیز $z(\lambda)$، ثابت نرمال‌سازی است. حال می‌خواهیم با استفاده از روش ML، برآورد پارامترها را بیابیم. در این صورت تابع درستنمایی به صورت

$$L(\lambda) = \prod_{i=۱}^n p(x_i, \lambda) = \prod_{i=۱}^n \frac{۱}{z(\lambda)}\exp\{-\sum_{r=\circ}^M \lambda_r\varphi_r(x_i)\}$$

و لگاریتم تابع درستنمایی برابر است با $\ell(\lambda) = -n\ln z(\lambda) - \sum_{i=۱}^n\sum_{r=۱}^M \lambda_r\varphi(x_i)$. بنابراین$\hat{\lambda}$، برآورد MLE پارامتر λ، از حل معادلات $\frac{\partial \ln z(\lambda)}{\partial \lambda_r} = \frac{۱}{n}\sum_{i=۱}^n \varphi_r(x_i)$، $r = ۱, ۲, ..., M$ حاصل می‌شود که در آن $z(\lambda) = \int \exp\{-\sum_{r=۱}^M \lambda_r\varphi_r(x)\}dx$ است. به طور معادل داریم:

$$\int \varphi_r(x)\exp\{-\sum_{r=۱}^M \lambda_r\varphi_r(x)\}dx = \frac{۱}{n}\sum_{i=۱}^n \varphi_r(x_i), \ r = ۱, ۲, ..., M, \quad (۵)$$

با مقایسه ی معادله ی (۲) با (۵) به این نتیجه می‌رسیم که μ_rها با مقدار برآورد تجربی خود جایگزین شده‌اند. یعنی$\mu_r = \frac{۱}{n}\sum_{i=۱}^n \varphi_r(x_i)$; $r = \circ, ۱, ..., M$ است.

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

روش ME:

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

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

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

روش MLE :

فرض می کنیم که $\varphi_r(x)$، $r=0,1,\ldots,M$ و نمونه ی $x_1,\ldots,x_n$، با تابع چگالی به فرم (۶) داده شده باشد. می‌خواهیم پارامترهای $\lambda_1,\ldots,\lambda_M$ را برآورد کنیم که با توجه به برآورد MLE متغیرهای μ_r داریم:

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

جعفری [۳] ، به کمک یک برنامه نرم‌افزار مطلب، برآورد ME–ی 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 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 $\phi_0(x)=1$، $\phi_1(x)=x$ و $\phi_2(x)=x^2$ در نظر گرفته شود. در نتیجه تابع چگالی دارای آنتروپی ماکزیمم به صورت زیر خواهد بود:

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

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:

[۴] Quartic

جدول ۱. برآوردME 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	−۰.۰۶	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. به این ترتیب با استفاده از شبیه‌سازی، آن ها را برآورد می‌کنیم. فرض می‌کنیم که مقدار نظری پارامترهای توزیع نرمال ۰.۲ $= \mu$ و ۱ $= \sigma^2$ است. برای سه حجم نمونه‌ی مختلف برآورد μ_1 و μ_2 و برآورد پارامترهای λ_0، λ_1 و λ_2 ، در جدول زیر آمده‌است:

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

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 $\mu_2 = E(\log(X))^2$ معلوم باشند، آن گاه توزیع لگ نرمال، چگالی با آنتروپی ماکزیمم خواهد بود. در نتیجه :

$$f(x) = \exp\{\lambda_0 + \lambda_1 \log(x) + \lambda_2 (\log(x))^2\}. \quad (8)$$

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. برآوردME 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	−.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 $m = mean(\log(X)) = -0.7$ و $\sigma^2 = var(\log(X)) = 1$ 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. برآوردME پارامترهای λ_0 ، λ_1 و λ_2 در چگالی فرم (۸) به ازای برآورد ML پارامترهای μ_2 و μ_1 .

حجم نمونه	μ_1	μ_2	λ_0	λ_1	λ_2	آنتروپی
۵۰۰	−۰.۶۲۵	۱.۳۵۴	۰.۸۱۸	۰.۷۸۲	۰.۱۹۸	۰.۵۹۸
۷۰۰	−۰.۷۰۷	۱.۴۹۴	۰.۸۷۵	۰.۹۲۲	۰.۲۲۴	۰.۵۵۸
۱۵۰۰	−۰.۶۸۷	۱.۴۹۸	۱.۱۴۵	۱.۵۶۶	۰.۴۴۱	۰.۷۳۱

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

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

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

جدول ۵. برآوردME پارامترهای λ_0 و λ_1 در چگالی فرم(۹).

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

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

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

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

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

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

است را در نظر می‌گیریم. این توزیع می‌تواند 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 صورت زیر خواهدبود:

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

باتوجه به مقادیر μ_1، μ_2 و μ_3، می‌توانیم λ_0، λ_1 وλ_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 وλ_2 و λ_3 را به دست آورده‌ایم:

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

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

۵ نتیجه گیری

در این مقاله ضمن معرفی ارتباط بین MLE و ME، به دنبال یافتن توزیعی بودیم که دارای آنتروپی ماکزیمم تحت شرایط و قیود خاص است و با حل معادلات غیرخطی‌ای که از قیود حاصل می شود و با استفاده از برنامه ی نرم‌افزار مطلب که در پیوست آمده، مقادیرآنتروپی ماکزیمم و فرم تابع چگالی دارای آنتروپی ماکزیمم را به دست آوردیم. در واقع ایده‌ی این برنامه‌ی نرم‌افزار از مقاله ی 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 مقاله ی جعفری این است که جعفری در مقاله‌ی خود برنامه ی نرم‌افزاری ارئه کرده‌است که تنها برای چند توزیع خاص و چند قید اولیه برآورد آنتروپی ماکزیمم پارامترها را یافته است، در این صورت برای یافتن برآوردME پارامترهامی‌بایست 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 داده های ترتیبی و فضایی دانشگاه فردوسی مشهد تشکر می نمایند.

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:

$function[lambda, p, entr] = me(mu, x, lambda0)$

$eps = 1e-6;$

$mu = [1, mu];$

$mu = mu'$

$xmin = x(1);$

$xmax = x(length(x));$

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

$if(nargin == 2)$

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

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

$else$

$lambda = lambda0(:);$

end

$n = length(lambda);$

$fin = fin - ۱(x);$

$iter = ۰;$

$while ۱$

$iter = iter + ۱;$

$p = exp(-(fin * lambda));$

$g = zeros(n, ۱);$

$for\ i = ۱ : n$

$g(i) = dx * sum(fin(:, i). * p);$

end

$entr(iter) = lambda' * g$

$gnk = zeros(n, n);$

$gnk(۱, :) = -g';$

$gnk(:, ۱) = -g;$

$for\ i = ۲ : n$

$for\ j = ۲ : i$

$gnk(i, j) = -dx * sum(fin(:, j). * fin(:, i). * p);$

end

end

$for\ i = ۲ : n$

$for\ j = i + ۱ : n$

$gnk(i, j) = gnk(j, i);$

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

$functionfin = fin - ۱(x)$

$m = input('enternumberoffinsyouwant')$

$fin = zeros(length(x), m);$

$for\ i = ۱ : m$

$fin(:, i) = input('enterfin');$

end

end