برآورد پارامترهای توزیع لوماکس تحت داده‌های سانسور با استفاده از الگوریتم EM و تقریب لیندلی

روشنک زمان، پرویز نصیری*؛
گروه آمار، دانشگاه پیام نور، ص.پ. ۴۶۹۷-۱۹۳۹۵، تهران، ایران

دریافت ۹۶/۰۴/۲۳ پذیرش ۹۶/۱۱/۳۰

چکیده

برآورد پارامتر توزیع های آماری از بحث‌های مهم استنباط آماری است. با توجه به کاربردهای توزیع لوماکس در تجارت، اقتصاد، علوم آماری، نظریه صف، مدل‌بندی ترافیکی اینترنت و غیره در این مقاله پارامترهای توزیع لوماکس تحت داده‌های سانسور نوع دوم با استفاده از الگوریتم EM و تقریب لیندلی برآورد می‌شوند. از آن‌جا که انتخاب توزیع‌های پیشین و توابع زیان، نقش مهمی در برآورد بیزی ایفا می‌کند. از این رو، برآورد بیزی با انتخاب توزیع پیشین مناسب تحت توابع زیان میانگین مربع خطا، لاینکس و آنتروپی ارائه می‌شود. از آن‌جاکه معادلات نرمال به‌دست آمده از روش‌های برآورد، توابع صریح از پارامترها نیستند، اقدام به برآورد پارامترها با استفاده از روش‌های خطا، از جمله الگوریتم EM و تقریب لیندلی خواهد شد. و در پایان با استفاده از ملاک میانگین توان دوم خطا، براوردگرها مقایسه و نتایج حاصل از شبیه‌سازی نشان می‌دهد که برآوردگر بیزی بهتر از برآورد بیشینۀ درست‌نمایی عمل می‌کند و با افزایش حجم نمونه در حالی که تعداد شکست ثابت است، دقت برآوردگر بیش‌تر می‌شود.

واژه‌های کلیدی: الگوریتم EM ، تقریب لیندلی، توزیع لوماکس، داده‌های سانسور شده، روش‌های برآورد، میانگین مربعات خطا

مقدمه

در تحلیل داده‌های مربوط به طول عمر، که بسیاری از شاخه‌های آمار کاربردی و بررسی‌های پزشکی مورد توجه است، اغلب به‌دلیل حذف واحدهایی از آزمایش که سانسور نامیده می‌شود، با داده‌های سانسور شده روبرو هستیم. که در این مقاله از سانسور نوع دوم[1] استفاده شده است. در این نوع داده‌ها بررسی تا رخداد پیشامد مورد نظر، زمان شکست r مؤلفۀ اول، $X_{(1)} < \cdots < X_{(r)}$ ادامه می‌یابد. که r یک مقدار ثابت از پیش تعیین شده است. در حالت کلی تابع درست‌نمایی تحت داده‌های سانسورشده نوع دوم برای تابع چگالی $f(.)$ و تابع توزیع $F(.)$ بدین‌صورت است [4]:

$$L(\theta; x) = \frac{n!}{(n-r)!} \prod_{i=1}^{r} f(x_i)[1 - F(x_r)]^{n-r} \qquad (۱)$$

که در آن منظور از x_r مقدار مشاهده r امین زمان شکست است.

متغیر تصادفی X دارای توزیع لوماکس[2] با پارامترهای شکل α و پارامتر مقیاس β است، که تابع چگالی و تابع توزیع آن به ترتیب برابر است با:

*نویسنده مسئول pnasiri45@yahoo.com

1. Type II censoring
2. Lomax

$$f_X(x) = \alpha\beta(1+\beta x)^{-(\alpha+1)}; x > 0 \quad (۲)$$

$$F(x) = 1 - (1+\beta x)^{-\alpha} \quad (۳)$$

توزیع لوماکس که به توزیع پارتو نوع دوم نیز مشهور است، یکی از توزیع‌های دم‌کلفت است که در تجارت، اقتصاد، علوم آماری، نظریهٔ صف و مدل‌بندی ترافیکی اینترنت استفاده می‌شود. لوماکس در [6] این توزیع را برای تجزیه و تحلیل داده‌های عدم موفقیت تجاری به‌کار برده است. بررسی‌های زیادی در مورد برآورد پارامترهای توزیع لوماکس در دو رهیافت کلاسیک و بیزی انجام شده است، که خلاصهٔ مفیدی از این تحقیقات را می‌توان در مقالهٔ جانسون و همکاران [3] ملاحظه کرد. مارشال و الکین [7] نشان دادند، که توزیع لوماکس به‌عنوان توزیع طول عمر می‌تواند استفاده شود. برآورد پارامتر مقیاس توزیع لوماکس تحت داده‌های سانسور فزایندهٔ نوع دوم را اصغرزاده و ولی‌اللهی بررسی کرده‌اند [1]. و اکاشا با استفاده از روش بیز تجربی برآورد پارامترهای توزیع لوماکس را تحت داده‌های سانسور شده نوع دوم فقط با در نظر گرفتن تابع زیان مربع خطا ارائه کرده است [8]. با توجه اهمیت این توزیع و روش‌های برآورد در ادامه مقاله، در بخش ۲، برآورد بیشینهٔ درست‌نمایی پارامترهای شکل و مقیاس توزیع لوماکس با استفاده از الگوریتم EM و در بخش ۳، برآورد بیزی پارامترها تحت توابع زیان توان دوم خطا، لاینکس و آنتروپی با فرض توزیع پیشین مزدوج گاما محاسبه شده است و در بخش ۴، تقریب لیندلی برآوردهای بیزی ارائه می‌شود. در بخش ۵ با استفاده از شبیه‌سازی براوردگرها مقایسه شده است.

برآورد بیشینهٔ درست‌نمایی

تابع درست‌نمایی توزیع لوماکس تحت داده‌های سانسوریده نوع دوم، برای نمونه تصادفی $X_n, \dots X_2, X_1$ ، با توجه به رابطهٔ (۱) برابر است با:

$$\mathrm{L}(\theta; \mathrm{x}) = \alpha^r\beta^r \prod_{i=1}^{r} B_i^{-(\alpha+1)}\, B_r^{-\alpha(n-r)} \quad (۴)$$

که در آن $B_i = 1 + \beta x_i$ و $B_r = 1 + \beta x_r$. و در سراسر مقاله برای خلاصه‌نویسی فرمول‌ها از این نمادها استفاده شده است. پس از لگاریتم‌گیری از رابطهٔ (۴) می‌توان نوشت:

$$l(\theta; x) = rln\alpha + rln\beta - (\alpha+1)\sum_{i=1}^{r}\log B_i - \alpha(n-r)\log B_r \quad (۵)$$

با مشتق‌گیری از لگاریتم درست‌نمایی نسبت به پارامترهای مدل معادلات نرمال بدین‌صورت به‌دست می‌آید:

$$\frac{\partial l(\theta,x)}{\partial\alpha} = \frac{r}{\alpha} - \sum_{i=1}^{r}\log B_i - (n-r)\log B_r \quad (۶)$$

$$\frac{\partial l(\theta,x)}{\partial\beta} = \frac{r}{\beta} - (\alpha+1)\sum_{i=1}^{r}\frac{x_i}{B_i} - \alpha(n-r)\frac{x_r}{B_r} \quad (۷)$$

از آن‌جاکه معادلات نرمال حاصل شده، تابع صریح از پارامترها نیستند، برای براورد پارامترها از الگوریتم EM که دمپستر و همکاران معرفی کردند [2]، استفاده می‌شود. الگوریتم EM شامل دو مرحله، گام E یا مرحلهٔ محاسبه امید ریاضی و گام M یا مرحلهٔ بیشینه‌سازی است که در ادامه آورده می‌شود.

لگاریتم تابع درست‌نمایی بر اساس مشاهدات کامل برابر است با:

$$l(\theta; x) = nln\alpha + nln\beta - (\alpha+1)\left[\sum_{i=1}^{r}\log B_i + \sum_{i=r+1}^{n}\log B_i\right] \quad (۸)$$

که در گام E، عبارت $E(logL(\theta;x))$ بدین‌صورت محاسبه می‌شود. که در آن فضای پارامتر $\theta = (\alpha, \beta)$ است.

$$E\big(logL(\theta;x)\big) = nln\alpha + nln\beta - (\alpha+1)\left[\sum_{i=1}^{r}\log B_i + \sum_{i=r+1}^{n}\mathrm{E}(\log B_i|X_i > x_r)\right] \quad (۹)$$

اما در گام M برای بیشینه‌سازی عبارت $E\big(logL(\theta;x)\big)$ معادلهٔ

$$\frac{\partial E\big(logL(\theta;x)\big)}{\partial\theta} = E\left(\frac{\partial l(\theta|x)}{\partial\theta}\right) = 0 \quad (۱۰)$$

در نظر گرفته می‌شود. بنابراین برای براورد پارامترهای α و β توزیع لوماکس داریم:

$$E\left(\frac{\partial l(\theta|x)}{\partial\alpha}\right)=\frac{n}{\alpha}-\left[\sum_{i=1}^{r}\log B_i+\sum_{i=r+1}^{n}\mathrm{E}(\log B_i|X_i>x_r)\right]$$
$$=\frac{n}{\alpha}-\left[\sum_{i=1}^{r}\log B_i+\sum_{i=r+1}^{n}k_1(\alpha,\beta)\right] \qquad (۱۱)$$

که در آن مقدار k_1 برابر است با:

$$k_1(\alpha,\beta)=\mathrm{E}(\log B_i|X_i>x_r)=\int_{x_r}^{\infty}\log B_i\,\frac{\alpha\beta B_i^{-(\alpha+1)}}{B_r^{-\alpha}}\,dx_i \qquad (۱۲)$$

با تغییر متغیر $x_i=\frac{x_r}{y}$ و رابطهٔ $dx_i=\frac{-x_r}{y^2}dy$، نتیجه می‌شود:

$$k_1(\alpha,\beta)=\alpha\beta x_r B_r^{\alpha}\int_0^1\log\left(\frac{y+\beta x_r}{y}\right)\frac{(y+\beta x_r)^{-\alpha-1}}{y^{-\alpha+1}}dy \qquad (۱۳)$$

و هم‌چنین

$$E\left(\frac{\partial l(\theta|x)}{\partial\beta}\right)=\frac{\mathrm{n}}{\beta}-(\alpha+1)\left[\sum_{\mathrm{i}=1}^{\mathrm{r}}\frac{\mathrm{x_i}}{B_i}+\sum_{i=r+1}^{n}E\left(\frac{\mathrm{x_i}}{B_i}\middle|X_r>x_i\right)\right]$$
$$=\frac{\mathrm{n}}{\beta}-(\alpha+1)\left[\sum_{\mathrm{i}=1}^{\mathrm{r}}\frac{\mathrm{x_i}}{B_i}+\sum_{i=r+1}^{n}k_2(\alpha,\beta)\right] \qquad (۱۴)$$

که در آن $k_2(\alpha,\beta)$ برابر است با:

$$k_2(\alpha,\beta)=\alpha\beta {x_r}^2 B_r^{\alpha}\int_0^1\frac{(y+\beta x_r)^{-\alpha-2}}{y^{-\alpha+1}}dy \qquad (۱۵)$$

بنابراین برآوردگر پارامترها بدین‌صورت به‌دست می‌آیند:

$$\hat{\alpha}_{EM}=\frac{n}{\sum_{i=1}^{r}\log B_i+(n-r)k_1(\alpha,\beta)} \qquad (۱۶)$$

$$\hat{\beta}_{EM}=\frac{n}{(\alpha+1)\left[\sum_{i=1}^{r}\frac{\mathrm{x_i}}{B_i}+(n-r)k_2(\alpha,\beta)\right]} \qquad (۱۷)$$

برآورد بیزی

از آن‌جاکه در براورد بیزی، تابع زیان نقش مهم دارد. از این‌رو، ابتدا توابع زیان مربع خطا[1]، لاینکس[2] و آنتروپی[3] برای برآورد بیزی پارامترهای αو β به‌ترتیب بدین‌صورت آورده می‌شوند.

$$L_s\left(d(\theta),\hat{d}(\theta)\right)=(\hat{d}(\theta)-d(\theta))^2 \qquad (۱۸)$$

$$L_L\left(d(\theta),\hat{d}(\theta)\right)=e^{c(\hat{d}(\theta)-d(\theta))}-c\left(\hat{d}(\theta)-d(\theta)\right)-1\ \ ,c\neq 0 \qquad (۱۹)$$

$$L_E\left(d(\theta),\hat{d}(\theta)\right)=\left(\frac{\hat{d}(\theta)}{d(\theta)}\right)^{\nu}-\nu\log\left(\frac{\hat{d}(\theta)}{d(\theta)}\right)-1\ \ ,\nu\neq 0 \qquad (۲۰)$$

در آن $\hat{d}(\theta)$، برآوردگر پارامتر $d(\theta)$ است. ضمناً مقادیر c وν ثابت هستند. به‌عنوان مثال مقدار c در تابع زیان لاینکس به‌منظور بیش براورد و کم براورد در ارائهٔ براوردگر در نظر گرفته می‌شود. و در حالتی که مقدار آن به‌سمت صفر میل می‌کند، تابع زیان حاصل، تابع زیان کم‌ترین مربعات می‌شود.

از آن‌جاکه به‌راحتی نمی‌توان توزیع پیشین مزدوج ارائه داد، از این‌رو، برای توزیع‌های پیشین به‌ترتیب دارای توزیع $\alpha\sim\Gamma(b_1,b_2)$ و$\beta\sim\Gamma(b_3,b_4)$، در نظر گرفته می‌شوند. **بنابراین** تابع چگالی احتمال پیشین توأم $\theta=(\alpha,\beta)$ بدین‌صورت در نظر گرفته می‌شود.

1. Squared error loss function
2. Linex loss function
3. Entropy loss function

$$\pi(\theta) = \left(\alpha^{b_1-1} e^{-b_2\alpha}\right)\left(\beta^{b_3-1} e^{-b_4\beta}\right) \qquad \alpha, \beta > 0, b_1, b_2 b_3, b_4 > 0 \tag{۲۱}$$

با توجه توزیع پیشین، تابع چگالی پسین توأم α و β برابر است با:

$$\pi(\theta|x) \propto L(x|\theta)\pi(\theta) = \prod_{i=1}^{r} \alpha\beta B_i^{-(\alpha+1)} B_r^{-\alpha(n-r)} \left(\alpha^{b_1-1} e^{-b_2\alpha}\right)\left(\beta^{b_3-1} e^{-b_4\beta}\right)$$
$$= B_r^{-\alpha(n-r)} \alpha^{r+b_1-1} \beta^{r+b_3-1} e^{-b_2\alpha - b_4\beta} \prod_{i=1}^{r} B_i^{-(\alpha+1)} \tag{۴}$$

برای توزیع پسین به‌دست آمده، در ادامه برآورد بیزی پارامترها تحت توابع زیان مربع خطا، لاینکس و آنتروپی ارائه می‌شود.

برآورد بیزی تحت تابع زیان مربع خطا

برآورد بیزی α و β تحت تابع زیان مربع خطا (L_S) یا میانگین پسین، به‌ترتیب برابرند با:

$$\hat{\alpha}_S = E(\alpha|x) = \int \alpha\pi(\theta|x) d\theta$$
$$= \int_0^\infty \int_0^\infty B_r^{-\alpha(n-r)} \alpha^{r+b_1} \beta^{r+b_3-1} e^{-b_2\alpha - b_4\beta} \prod_{i=1}^{r} B_i^{-(\alpha+1)} \, d\alpha d\beta \tag{۲۳}$$

$$\hat{\beta}_S = E(\beta|x) = \int \beta\pi(\theta|x) d\theta$$
$$= \int_0^\infty \int_0^\infty B_r^{-\alpha(n-r)} \alpha^{r+b_1-1} \beta^{r+b_3} e^{-b_2\alpha - b_4\beta} \prod_{i=1}^{r} B_i^{-(\alpha+1)} \, d\alpha d\beta \tag{۲۴}$$

برآورد بیزی تحت تابع زیان لاینکس

تحت تابع زیان لاینکس (L_L)، برآورد بیزی به صورت

$$\hat{d}_L(\theta) = \frac{-1}{c} \ln\left\{E_\theta\left(e^{-cd(\theta)}|x\right)\right\} \tag{۲۵}$$

است. برآورد بیزی α و β به‌صورت (۲۶) به‌دست می‌آیند:

$$\hat{\alpha}_L = \frac{-1}{c} \ln\{E(e^{-c\alpha}|x)\} \tag{۲۶}$$

که در آن

$$E(e^{-c\alpha}|x) = \int_0^\infty \int_0^\infty B_r^{-\alpha(n-r)} \alpha^{r+b_1-1} \beta^{r+b_3-1} e^{-(b_2+c)\alpha - b_4\beta} \prod_{i=1}^{r} B_i^{-(\alpha+1)} d\alpha d\beta \tag{۲۷}$$

و هم‌چنین

$$\hat{\beta}_L = \frac{-1}{c} \ln\left\{E\left(e^{-c\beta}|x\right)\right\} \tag{۲۸}$$

که در آن

$$E\left(e^{-c\beta}\middle|x\right) = \int_0^\infty \int_0^\infty B_r^{-\alpha(n-r)} \alpha^{r+b_1-1} \beta^{r+b_3-1} e^{-b_2\alpha - (b_4+c)\beta} \prod_{i=1}^{r} B_i^{-(\alpha+1)} d\alpha d\beta \tag{۲۹}$$

برآورد بیزی تحت تابع زیان آنتروپی

برآورد بیزی تحت تابع زیان آنتروپی(L_E)، بدین‌صورت است.

$$\hat{d}_L(\theta) = \left\{E_\theta\left(\left(d(\theta)\right)^{-\nu}|X\right)\right\}^{\frac{-1}{\nu}} \tag{۳۰}$$

بنابراین برآورد بیزی α و β، تحت تابع زیان آنتروپی، برابر است با:

$$\hat{\alpha}_E = \{E(\alpha^{-\nu}|X)\}^{\frac{-1}{\nu}}, \hat{\beta}_E = \{E(\alpha^{-\nu}|X)\}^{\frac{-1}{\nu}} \tag{۳۱}$$

که در آن

$$E(\alpha^{-\nu}|X) = \int_0^\infty \int_0^\infty B_r{}^{-\alpha(n-r)} \alpha^{r+b_1-\nu-1} \beta^{r+b_3-1} e^{-b_2\alpha-b_4\beta} \prod_{i=1}^r B_i{}^{-(\alpha+1)} d\alpha d\beta \quad (۳۲)$$

$$E(\beta^{-\nu}|X) = \int_0^\infty \int_0^\infty B_r{}^{-\alpha(n-r)} \alpha^{r+b_1-1} \beta^{r+b_3-\nu-1} e^{-b_2\alpha-b_4\beta} \prod_{i=1}^r B_i{}^{-(\alpha+1)} d\alpha d\beta \quad (۳۳)$$

با توجه به برآوردهای بیزی به‌دست آمده برای پارامترهای مدل بر اساس داده‌های سانسور نوع دوم، مشاهده می‌شود که برآوردگرها به‌صورت تحلیلی قابل محاسبه نیستند. بنابراین نیازمند به روش‌های عددی هستیم که در بخش بعدی به‌روش تقریب لیندلی[۱] ارائه می‌شود.

تقریب لیندلی

در این بخش، برآورد بیزی تقریبی پارامترهای α و β را با استفاده از روش لیندلی به‌دست می‌آوریم. لیندلی روش تقریبی برای ارزیابی امید پسین $U(\theta)$ را بدین‌صورت توسعه داد [5].

$$E(U(\theta)|X) = \frac{\int U(\theta) e^{l(\theta)+\rho(\theta)} d\theta}{\int e^{l(\theta)+\rho(\theta)}\, d\theta} \quad (۳۴)$$

که برآورد بیز $U(\theta)$ است، در آن $\rho(\theta) = \log \pi(\theta)$ ، لگاریتم توزیع پیشین و $l(\theta)$ لگاریتم تابع درست‌نمایی است. برای جزئیات بیش‌تر لیندلی (۱۹۸۰) را ببینید. تقریب لیندلی برای $\theta = (\theta_1, \theta_2)$، به‌صورت (۳۵) کاهش می‌یابد:

$$E(U(\theta)|X) = U(\theta) + \frac{A}{2} + \rho_1 A_{12} + \rho_2 A_{21} + \frac{1}{2}[l_{30}B_{12} + l_{21}c_{12} + l_{12}c_{21} + l_{03}B_{21}] \quad (۳۵)$$

که

$$A = \sum_{i=1}^{2} (U_{i1}\delta_{i1} + U_{i2}\delta_{i2}) = U_{11}\delta_{11} + U_{12}\delta_{12} + U_{21}\delta_{21} + U_{22}\delta_{22}$$

$$A = \sum_{i=1}^{2}\sum_{j=1}^{2} U_{ij}\delta_{ij}\, , l_{ks} = \frac{\partial^{k+s} L}{\partial\theta_1^k\, \partial\theta_2^s} \ \ , k,s = 0,1,2,3 \ \ , k+s = 3, i,j = 1,2$$

$$\rho_i = \frac{\partial\rho}{\partial\theta_i} \ \ , \qquad U_i = \frac{\partial U}{\partial\theta_i} \ \ , \qquad U_{ij} = \frac{\partial^2 U}{\partial\theta_i\, \partial\theta_j}$$

برای $i \neq j$

$$A_{ij} = U_i\delta_{ii} + U_j\delta_{ji}, B_{ij} = \left(U_i\delta_{ii} + U_j\delta_{ij}\right)\delta_{ii}, C_{ij} = 3U_i\delta_{ii}\delta_{ij} + U_j\left(\delta_{ii}\delta_{jj} + 2\delta_{ij}^2\right)$$

که δ_{ij} عنصر (i,j) ام از ماتریس معکوس $\left\{-l_{ij} = -l_{ij} = \frac{\partial^2 l}{\partial\theta_i\, \partial\theta_j}\right\}$ $(i,j = 1,2)$ است.

با توجه به تابع چگالی پیشین $\alpha \sim \Gamma(b_1, b_2)$ و $\beta \sim \Gamma(b_3, b_4)$ داریم:

$$\rho = (b_1 - 1)Ln\, \alpha + (b_3 - 1)Ln\, \beta - b_2\alpha - b_4\beta \quad (۳۶)$$

$$\rho_1 = \frac{\partial\rho}{\partial\alpha} = \frac{b_1 - 1}{\alpha} - b_2, \rho_2 = \frac{\partial\rho}{\partial\beta} = \frac{b_3 - 1}{\beta} - b_4$$

$$l_{30} = \frac{2r}{\alpha^3},$$

$$l_{21} = 0$$

$$l_{12} = -\sum \frac{x_i^2}{B_i{}^2} + (n-r)\frac{x_r^2}{B_r{}^2}$$

$$l_{03} = \frac{2r}{\beta^3} - (\alpha+1)\sum_{i=1}^{r} \frac{2x_i^3}{B_i{}^3} - \alpha(n-r)\frac{2x_r^3}{B_r{}^3}$$

1. Lindley approximation

تقریب لیندلی برآوردهای بیزی پارامترها تحت تابع زیان مربع خطا

با جاگذاری $u(\theta)=\alpha$ در رابطهٔ (۳۵) تقریب لیندلی $\hat{\alpha}_s$ به‌دست می‌آید.

$$\left.\begin{array}{lll} U_1=\dfrac{\partial u}{\partial\alpha}=1\,, & U_{11}=\dfrac{\partial^2 U}{\partial\alpha\partial\alpha}=0\,, & U_{12}=\dfrac{\partial^2 U}{\partial\alpha\partial\beta}=0 \\ U_2=\dfrac{\partial u}{\partial\beta}=0\,, & U_{21}=\dfrac{\partial^2 u}{\partial\beta\partial\alpha}=0\,, & U_{22}=\dfrac{\partial^2 U}{\partial\beta\partial\beta}=0 \end{array}\right\}\Rightarrow A=0$$

$$l_{30}=\frac{2r}{\alpha^3},A_{12}=\delta_{11},A_{21}=\delta_{12},C_{12}=3\delta_{11}\delta_{12},C_{21}=\delta_{11}\delta_{22}+2\delta_{21}^2$$

$$B_{21}=\delta_{21}\delta_{22},B_{12}=\delta_{11}^2$$

$$\delta_{11}=\frac{\frac{r}{\beta^2}-\frac{(\alpha+1)x_i^2}{B_i{}^2}-\frac{\alpha(n-r)x_r^2}{(1+\beta x_r)^2}}{D_1D_4-D_2D_3},\delta_{21}=\delta_{12}=\frac{-\sum\frac{x_i}{B_i}-\frac{(n-r)x_r}{1+\beta x_r}}{D_1D_4-D_2D_3}$$

$$\delta_{22}=\frac{\frac{r}{\alpha^2}}{D_1D_4-D_2D_3}$$

که در آن منظور از D_1 و D_2 و D_3 و D_4 بدین‌صورت است:

$$D_1=\frac{r}{\alpha^2}\,,D_2=D_3=\sum\frac{x_i}{B_i}+\frac{(n-r)x_r}{1+\beta x_r},D_4=\frac{r}{\beta^2}-\frac{(\alpha+1)x_i^2}{B_i{}^2}-\frac{\alpha(n-r)x_r^2}{(1+\beta x_r)^2},$$

و در نهایت داریم:

$$\hat{\alpha}_{BS}^{Lindley}=\alpha+R_1 \tag{۳۷}$$

که در آن R_1 برابر است با:

$$R_1=\rho_1\delta_{11}+\rho_2\delta_{12}+\frac{1}{2}[l_{30}\delta_{11}^2+3l_{21}\delta_{11}\delta_{12}+l_{12}(\delta_{11}\delta_{22}+2\delta_{12}^2)+l_{03}\delta_{21}\delta_{22}] \tag{۳۸}$$

همچنین با جای‌گذاری $U(\theta)=\beta$ در رابطهٔ (۳۵) تقریب لیندلی $\hat{\beta}_S$ بدین‌صورت به‌دست می‌آید:

$$U_1=\frac{\partial u}{\partial\alpha}=0\quad,\quad U_2=\frac{\partial u}{\partial\beta}=1,U_{11}=U_{12}=U_{21}=U_{22}=0\quad\Rightarrow\quad A=0$$

$$A_{12}=\delta_{21},A_{21}=\delta_{22},C_{12}=\delta_{11}\delta_{22}+2\delta_{12}^2,C_{21}=3\delta_{22}\delta_{21},B_{12}=\delta_{12}\delta_{11},B_{21}=\delta_{22}^2$$

$$\hat{\beta}_{BS}^{Lindley}=\beta+R_2 \tag{۳۹}$$

که منظور از R_2 رابطهٔ (۴۰) است.

$$R_2=\rho_1\delta_{21}+\rho_2\delta_{22}+\frac{1}{2}[l_{30}\delta_{12}\delta_{11}+l_{21}(\delta_{11}\delta_{22}+2\delta_{12}^2)+3l_{12}\delta_{22}\delta_{21}+l_{03}\delta_{22}^2] \tag{۴۰}$$

تقریب لیندلی برآوردهای بیزی پارامترها تحت تابع زیان لاینکس

برای به‌دست آوردن تقریب لیندلی $\hat{\alpha}_L$ با جای‌گذاری $U(\theta)=e^{-C\alpha}$ در رابطهٔ (۳۵) داریم:

$$U_1=\frac{\partial u}{\partial\alpha}=-Ce^{-C\alpha}\;,u_2=0\,,\;u_{12}=0\,,u_{21}=0\;,u_{11}=C^2e^{-C\alpha}\;,u_{22}=0$$

$$A=C^2e^{-C\alpha}\delta_{11},A_{12}=-Ce^{-C\alpha}\delta_{11},A_{21}=-Ce^{-C\alpha}\delta_{12},C_{12}=-3Ce^{-C\alpha}\delta_{11}\delta_{12}$$

$$C_{21}=-Ce^{-C\alpha}(\delta_{11}\delta_{22}+2\delta_{12}^2),B_{21}=-Ce^{-C\alpha}\delta_{21}\delta_{22},B_{12}=-Ce^{-C\alpha}\delta_{11}^2$$

$$\hat{\alpha}_{BL}^{Lindley}=-\frac{1}{C}Ln\{E(e^{-C\alpha}|x)\}=\alpha-\frac{1}{C}\,Ln\,\left(1+\frac{1}{2}C^2\delta_{11}-CR_1\right) \tag{۴۱}$$

به‌طور مشابه برای تقریب لیندلی $\hat{\beta}_L$ با جای‌گذاری $U(\theta)=e^{-C\beta}$ داریم:

$$\hat{\beta}_{BL}^{Lindley} = \beta - \frac{1}{C}\, Ln\, \left(1 + \frac{1}{2}C^2\delta_{22} - CR_2\right) \qquad (۴۲)$$

تقریب لیندلی برآوردهای بیزی پارامترها تحت تابع زیان آنتروپی

به‌منظور به‌دست آوردن تقریب لیندلی $\hat{\alpha}_E$ با جای‌گذاری $U(\theta) = \alpha^{-v}$ در رابطهٔ (۳۵) داریم:

$$U_1 = \frac{\partial u}{\partial \alpha} = -v\alpha^{-v-1}\,, U_2 = 0\,, u_{11} = v(v+1)\alpha^{-v-2}\,, u_{12} = 0\,,$$

$$u_{21} = 0\,, u_{22} = 0, A = v(v+1)\alpha^{-v-2}\delta_{11}, A_{12} = -v\alpha^{-v-1}\delta_{11},$$

$$A_{21} = -v\alpha^{-v-1}\delta_{12}, C_{12} = -3v\alpha^{-v-1}\delta_{11}\delta_{12}, C_{21} = -v\alpha^{-v-1}(\delta_{11}\delta_{22} + 2\delta_{12}^2)$$

$$B_{21} = -v\alpha^{-v-1}\delta_{21}\delta_{22}, B_{12} = -v\alpha^{-v-1}\delta_{11}^2$$

$$\hat{\alpha}_{BE}^{Lindley} = \alpha\left\{1 + \frac{1}{2}v(v+1)\alpha^{-2}\delta_{11} - v\alpha^{-1}R_1\right\}^{-1/v} \qquad (۴۳)$$

به‌طور مشابه می‌توان نشان داد که برای $U(\theta) = \beta^{-v}$ داریم:

$$\hat{\beta}_{BE}^{Lindley} = \beta\left\{1 + \frac{1}{2}v(v+1)\alpha^{-2}\delta_{11} - v\alpha^{-1}R_2\right\}^{-1/v} \qquad (۴۴)$$

لازم به ذکر است که با جای‌گذاری $\hat{\alpha}_{EM}$ و $\hat{\beta}_{EM}$ در روابط به‌دست آمده، برآورد ها ارزیابی می‌شوند.

نتایج شبیه‌سازی

در مقایسهٔ برآوردگرهای بررسی شده، نمونه‌هایی به حجم $n = 20,40,60,80,100$ برای تعداد شکست‌های $r = 5,10,15$ از توزیع لوماکس در حالی که $\alpha = 2.5$ و $\beta = 0.49$ است، تولید شده است. در برآورد بیشینهٔ درست‌نمایی برای حل معالات نرمال از الگوریتم EM و در برآورد بیزی از تقریب لیندلی استفاده شد. مقادیر ثابت برای توزیع‌های پیشین به‌صورت $b_1 = 2, b_2 = 1, b_3 = 2, b_4 = 1, c = v = 1$ فرض شده‌اند. برای گرفتن نتایج بهتر تعداد تکرارها ۱۰۰۰۰ بار در نظر گرفته شده است. برای مقایسه براوردگرها میانگین توان دوم خطا (MSE) در حالت‌های مختلف محاسبه شده است. که نتایج شبیه‌سازی در جدول‌های ۱ و ۲ آمده است. منظور از $MSE(\hat{\alpha}_{BS})$، $MSE(\hat{\alpha}_{BL})$ و $MSE(\hat{\alpha}_{BE})$ به‌ترتیب میانگین توان دوم خطای برآوردگر بیز تحت تابع زیان مربع خطا، لاینکس و آنتروپی برای پارامتر $\boldsymbol{\alpha}$ و هم‌چنین $MSE(\hat{\beta}_{BS})$، $MSE(\hat{\beta}_{BL})$ و $MSE(\hat{\beta}_{BE})$ به‌ترتیب میانگین توان دوم خطای برآوردگر بیز تحت تابع زیان مربع خطا، لاینکس و آنتروپی برای پارامتر β است. چنان‌که در هر دو جدول چشم‌گیر است با افزایش n وقتی r ثابت است توان دوم خطا کاهش یافته است. هم‌چنین نتایج نشان می‌دهد که در تمام حالات برآوردگر بیزی بهتر از برآوردگر بیشینهٔ درست‌نمایی عمل می‌کند. قابل ذکر است از مقایسهٔ برآوردگرها تحت توابع زیان نتیجه می‌شود که به‌ترتیب برآوردگرهای تحت تابع زیان لاینکس، آنتروپی و مربع خطا بهتر عمل می‌کنند.

جدول ۱. برآورد α برای حالت های مختلف n و r به همراه MSE

n	r	MLE		$Lindley$					
		$\hat{\alpha}$	$MSE(\hat{\alpha})$	$\hat{\alpha}_{BS}$	$MSE(\hat{\alpha}_{BS})$	$\hat{\alpha}_{BL}$	$MSE(\hat{\alpha}_{BL})$	$\hat{\alpha}_{BE}$	$MSE(\hat{\alpha}_{BE})$
20	5	$25.9783e^{-01}$	$6.2119e^{-02}$	$26.7359e^{-01}$	$3.0133e^{-02}$	$26.7351e^{-01}$	$3.0108e^{-02}$	$26.7353e^{-01}$	$3.0114e^{-02}$
	10	$28.2005e^{-01}$	$2.0464e^{-01}$	$27.3749e^{-01}$	$5.6402e^{-02}$	$27.3741e^{-01}$	$5.6367e^{-02}$	$27.3743e^{-01}$	$5.6377e^{-02}$
	15	$32.7928e^{-01}$	$7.9203e^{-01}$	$31.3202e^{-01}$	$3.9945e^{-01}$	$31.3195e^{-01}$	$3.9936e^{-01}$	$31.3197e^{-01}$	$3.9939e^{-01}$
40	5	$25.2426e^{-01}$	$1.5466e^{-02}$	$27.2093e^{-01}$	$4.8813e^{-02}$	$27.2088e^{-01}$	$4.8791e^{-02}$	$27.2089e^{-01}$	$4.8797e^{-02}$
	10	$25.7891e^{-01}$	$3.3748e^{-02}$	$23.1445e^{-01}$	$3.4426e^{-02}$	$23.1439e^{-01}$	$3.4451e^{-02}$	$23.1440e^{-01}$	$3.4447e^{-02}$
	15	$26.6611e^{-01}$	$6.6492e^{-02}$	$29.5466e^{-01}$	$2.0672e^{-02}$	$29.5459e^{-01}$	$2.0665e^{-02}$	$29.5461e^{-01}$	$2.0667e^{-02}$
60	5	$25.1199e^{-01}$	$7.2675e^{-03}$	$24.6388e^{-01}$	$1.3074e^{-03}$	$24.6389e^{-01}$	$1.3073e^{-03}$	$24.6389e^{-01}$	$1.3074e^{-03}$
	10	$25.3632e^{-01}$	$1.4650e^{-02}$	$24.2711e^{-01}$	$5.3124e^{-03}$	$24.2706e^{-01}$	$5.3197e^{-03}$	$24.2707e^{-01}$	$5.3184e^{-03}$
	15	$25.7155e^{-01}$	$2.3154e^{-02}$	$26.7827e^{-01}$	$3.1782e^{-02}$	$26.7821e^{-01}$	$3.1758e^{-02}$	$26.7823e^{-01}$	$3.1764e^{-02}$
80	5	$25.0739e^{-01}$	$4.3431e^{-03}$	$25.0585e^{-01}$	$4.1686e^{-05}$	$25.0594e^{-01}$	$4.5315e^{-05}$	$25.0592e^{-01}$	$4.4526e^{-05}$
	10	$25.2020e^{-01}$	$8.2875e^{-03}$	$26.0907e^{-01}$	$1.1897e^{-02}$	$26.0904e^{-01}$	$1.1890e^{-02}$	$26.0904e^{-01}$	$1.1892e^{-02}$
	15	$25.3966e^{-01}$	$1.2515e^{-02}$	$24.3496e^{-01}$	$4.2300e^{-03}$	$24.3491e^{-01}$	$4.2362e^{-03}$	$24.4349e^{-01}$	$4.2351e^{-03}$
100	5	$25.0436e^{-01}$	$2.7690e^{-03}$	$24.8976e^{-01}$	$1.1578e^{-04}$	$24.8995e^{-01}$	$1.1558e^{-04}$	$24.8991e^{-01}$	$1.1555e^{-04}$
	10	$25.1394e^{-01}$	$5.4134e^{-03}$	$25.6825e^{-01}$	$4.6598e^{-03}$	$25.6826e^{-01}$	$4.6606e^{-03}$	$25.6826e^{-01}$	$4.6604e^{-03}$
	15	$25.2816e^{-01}$	$8.0337e^{-03}$	$24.9777e^{-01}$	$5.1764e^{-03}$	$24.9775e^{-01}$	$5.3439e^{-03}$	$24.9775e^{-01}$	$5.31018e^{-03}$

جدول ۲. برآورد β برای حالت‌های مختلف n و r به‌همراه MSE

n	r	MLE		$Lindley$					
		$\hat{\beta}$	$MSE(\hat{\beta})$	$\hat{\beta}_{BS}$	$MSE(\hat{\beta}_{BS})$	$\hat{\beta}_{BL}$	$MSE(\hat{\beta}_{BL})$	$\hat{\beta}_{BE}$	$MSE(\hat{\beta}_{BE})$
20	5	$5.0765e^{-01}$	$2.1987e^{-03}$	$5.1943e^{-01}$	$8.6649e^{-04}$	$5.1943e^{-01}$	$8.6652e^{-04}$	$5.1984e^{-01}$	$8.9048e^{-04}$
	10	$5.4325e^{-01}$	$6.2888e^{-03}$	$5.2538e^{-01}$	$1.2524e^{-03}$	$5.2538e^{-01}$	$1.2524e^{-03}$	$5.2556e^{-01}$	$1.2648e^{-03}$
	15	$6.0606e^{-01}$	$1.9130e^{-02}$	$5.8445e^{-01}$	$8.9219e^{-03}$	$5.8445e^{-01}$	$8.9220e^{-03}$	$5.8456e^{-01}$	$8.9429e^{-03}$
40	5	$4.9447e^{-01}$	$5.9446e^{-04}$	$4.9213e^{-01}$	$5.0340e^{-06}$	$4.921e^{-01}$	$5.0374e^{-06}$	$4.929e^{-01}$	$8.7433e^{-06}$
	10	$5.0428e^{-01}$	$1.1848e^{-03}$	$4.5467e^{-01}$	$1.2481e^{-03}$	$4.5467e^{-01}$	$1.2480e^{-03}$	$4.5502e^{-01}$	$1.2238e^{-03}$
	15	$5.1899e^{-01}$	$2.1839e^{-03}$	$5.7194e^{-01}$	$6.7150e^{-03}$	$5.7194e^{-01}$	$6.7151e^{-03}$	$5.7219e^{-01}$	$6.7567e^{-03}$
60	5	$4.9228e^{-01}$	$2.7022e^{-04}$	$4.8106e^{-01}$	$8.0630e^{-05}$	$4.8107e^{-01}$	$8.0521e^{-05}$	$4.8228e^{-01}$	$5.9577e^{-05}$
	10	$4.9675e^{-01}$	$5.2997e^{-04}$	$4.7596e^{-01}$	$1.9711e^{-04}$	$4.7596e^{-01}$	$1.9706e^{-04}$	$4.7651e^{-01}$	$1.8171e^{-04}$
	15	$5.0297e^{-01}$	$8.0936e^{-04}$	$5.2174e^{-01}$	$1.0076e^{-03}$	$5.2174e^{-01}$	$1.0077e^{-03}$	$5.2212e^{-01}$	$1.1031e^{-03}$
80	5	$4.9141e^{-01}$	$1.6275e^{-04}$	$4.8823e^{-01}$	$5.0157e^{-06}$	$4.8824e^{-01}$	$4.9265e^{-06}$	$4.8992e^{-01}$	$5.4155e^{-08}$
	10	$4.9381e^{-01}$	$3.0393e^{-04}$	$5.0965e^{-01}$	$3.8664e^{-04}$	$5.0966e^{-01}$	$3.8675e^{-04}$	$5.1044e^{-01}$	$4.1818e^{-04}$
	15	$4.9731e^{-01}$	$4.4829e^{-04}$	$4.7646e^{-01}$	$1.8332e^{-04}$	$4.7646e^{-01}$	$1.8328e^{-04}$	$4.7694e^{-01}$	$1.7050e^{-04}$
100	5	$4.9083e^{-01}$	$1.0418e^{-04}$	$4.8448e^{-01}$	$3.3113e^{-05}$	$4.8449e^{-01}$	$3.2869e^{-05}$	$4.8656e^{-01}$	$1.1863e^{-05}$
	10	$4.9227e^{-01}$	$1.9901e^{-04}$	$5.0152e^{-01}$	$1.3309e^{-04}$	$5.0153e^{-01}$	$1.3318e^{-04}$	$5.0251e^{-01}$	$1.5648e^{-04}$
	15	$4.9527e^{-01}$	$2.9218e^{-04}$	$4.8882e^{-01}$	$2.3017e^{-04}$	$4.8883e^{-01}$	$2.3009e^{-04}$	$4.8945e^{-01}$	$2.1197e^{-04}$

منابع

1. Asgharzadeh A., Valiollahi R., "Estimation of the scale parameter of the Lomax distribution under progressive censoring", Int. J. Stat. Econ., 6 (2011) 37-48.
2. Dempster A. P., Laird N. M.,Rubin D. B., "Maximum likelihood from incomplete data via the EM algorithm", J. R. Stat. Soc. Ser. B, (1977) 1-38.
3. Johnson N. L., Kotz S., Balakrishnan N., "Continuous Univariate Distributions". 2nd edition. New York: Wiley (1994).

4. Lawless J. F., "Statistical models and methods for lifetime data", John Wiley & Sons, (2011).
5. Lindley D. V., "Approximate bayesian methods", Trab, estadística e Investig, Oper., 31 (1980) 223-245.
6. Lomax K. S., "Business failures: Another example of the analysis of failure data", J. Am. Stat. Assoc., 49 (1954) 847-852.
7. Marshall A. W., Olkin I., "Life distributions", 13. Springer, (2007).
8. Okasha H. M., "E-Bayesian Estimation for the Lomax distribution based on type-II censored data", J. Egypt. Math. Soc., 22 (2014) 489-495.