

پیش‌گویی مقادیر رکوردهای آینده در توزیع نمایی

ابراهیم امینی سرشت^۱، سعید بگرضایی^۲

چکیده:

در این مقاله می‌خواهیم بر اساس مشاهدات اولین n رکورد بالایی از توزیع نمایی، برآورد حداکثر درست‌نمایی پارامتر این توزیع را به دست آوریم. سپس روی مسئله پیش‌گویی نقطه‌ای مقادیر رکوردهای بالایی آینده در توزیع نمایی بر اساس نگرش‌های کلاسیک و بیز و تحت توابع زیان درجه دوم و لاینکس متمرکز می‌شویم. در پایان نیز از طریق شبیه‌سازی مونت کارلو به مقایسه عددی پیش‌گویی‌های نقطه‌ای به دست آمده خواهیم پرداخت.

واژه‌های کلیدی: بیز، پیش‌گویی، توزیع پیشین، توزیع گامای معکوس، توزیع نمایی، مقادیر رکورد.

۱ مقدمه

تابع چگالی $f(x; \theta)$ مشاهده کرده‌ایم. در این صورت چگالی توأم مقادیر اولین n -رکورد بالایی (آرنولد و همکاران ۱۹۹۸) به صورت:

$$f(x, \theta) = \prod_{i=1}^{n-1} h(x_i; \theta) f(x_n; \theta) \quad (1)$$

$$-\infty < x_1 < x_2 < \dots < x_n < \infty$$

محاسبه می‌گردد که در آن $h(x_i; \theta) = \frac{f(x_i; \theta)}{1 - F(x_i; \theta)}$ و $x = (x_1, x_2, \dots, x_n)$ ممکن است یک بردار باشد که در آن فضای پارامتر است. برای مطالعه بیشتر در مورد آماره‌های رکوردی می‌توانید به احسن‌الله^۳ (۱۹۹۵) یا آرنولد^۴، بالاکریشنان^۵ و ناگاراچا^۶ (۱۹۹۸) مراجعه کنید.

بسیاری از محققین به مطالعه فاصله پیش‌گویی برای مشاهدات آینده از توزیع نمایی پرداخته‌اند. احسن‌الله (۱۹۸۰) پیش‌گویی مقادیر رکوردهای آینده از این توزیع را مورد مطالعه قرار داد. دانسمور^۷ (۱۹۸۳) مسئله فاصله پیش‌گویی بیزی برای مقادیر رکوردهای آینده از توزیع نمایی را در نظر گرفت. باساک^۸ و بالاکریشنان (۲۰۰۳) پیش‌گویی‌های میانه شرطی (CMP)، میانه نارایی (MUP) و حداکثر درست‌نمایی (MLP) را در توزیع

فرض کنید $\{X_i, i \geq 1\}$ یک دنباله از متغیرهای تصادفی پیوسته، مستقل و هم‌توزیع با تابع توزیع تجمعی $F(x; \theta)$ و تابع چگالی $f(x; \theta)$ که در آن θ یک پارامتر (ممکن است یک بردار از پارامترها باشد) است، مشاهده X_j یک مقدار رکورد بالایی است اگر مقدارش از همه مشاهدات قبلی بزرگتر باشد. بنابراین X_j یک رکورد بالایی است اگر $X_j > X_i$ برای هر $j > i$. همچنین دنباله زمان‌های رکوردهای بالایی $\{U(n), n \geq 1\}$ به صورت زیر تعریف می‌شود:

$$U(1) = 1$$

و برای $n \geq 2$

$$U(n) = \min\{j : j > U(n-1), X_j > X_{U(n-1)}\}.$$

دنباله مقادیر رکوردهای بالایی به صورت $X_{U(n)}, n = 1, 2, \dots$ تعریف می‌شود.

فرض کنید مقادیر اولین n رکورد بالایی را به صورت $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(n)} = x_n$ از تابع توزیع تجمعی $F(x; \theta)$ و

^۱دانشجوی دکتری-گروه آمار دانشگاه رازی

^۲کارشناس ارشد-گروه آمار دانشگاه رازی

^۳Ahsanullah

^۴Arnold

^۵Balakrishnan

^۶Nagaraja

^۷Dunsmore

^۸Basak

نمایی به دست آوردند. از مطالعات انجام شده در سال‌های اخیر می‌توان به جاهین^۹ (۲۰۰۴) و احمدی^{۱۰} و همکاران (۲۰۰۷) اشاره کرد.

$$F(x; \theta) = 1 - \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0, \quad \theta > 0. \quad (۴)$$

تابع زیان $L(\theta, \delta)$ بیانگر مقدار زبانی است که آماردان متحمل می‌شود وقتی که θ وضعیت درست طبیعت بوده و او عمل δ را اتخاذ می‌کند. رایج‌ترین تابع زیان در مسائل برآورد و پیش‌بینی $L(\theta, \delta) = (\delta - \theta)^2$ که به تابع زیان درجه دوم معروف است. این تابع زیان متقارن است و بیش برآوردی و کم برآوردی را به طور یکسان جریمه می‌کند. با افزایش اندازه بیش برآوردی و کم برآوردی مقدار این تابع به سرعت بزرگ می‌شود. تابع زیان دیگری که در این مقاله با آن سر و کار داریم به فرم:

$$L(\theta, \delta) = b \exp\{c(\delta - \theta)\} - c(\delta - \theta) - 1$$

یا

$$L(\theta, \delta) = b \exp\{c\Delta\} - c\Delta - 1$$

در این بخش برآورد پارامتر مجهول θ در توزیع نمایی بر اساس مقادیر مشاهده شده از اولین n رکورد بالایی از این توزیع را در نظر می‌گیریم. فرض کنید اولین n رکورد بالایی را به صورت $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(n)} = x_n$ مشاهده کرده‌ایم. از (1)، (3) و (4)، تابع چگالی توأم $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(n)} = x_n$ (مرجع [4] را ببینید) به صورت:

$$f(x; \theta) = \theta^{-n} \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0, \quad \theta > 0 \quad (۵)$$

خواهد بود که در آن $x = (x_1, x_2, \dots, x_n)$ است. برآوردگر حداکثر درست‌نمایی (MLE) برای θ به صورت زیر است:

$$\hat{\theta} = \frac{X_{U(n)}}{n}. \quad (۶)$$

از آن‌جا که چگالی حاشیه‌ای $X_{U(n)}$ (مرجع [4] را ببینید) به صورت:

$$\begin{aligned} f_n(x; \theta) &= f(x; \theta) \frac{[-\ln(1 - F(x; \theta))]^{n-1}}{(n-1)!} \\ &= \frac{1}{\theta^n \Gamma(n)} x^{n-1} \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0 \end{aligned} \quad (۷)$$

است، یعنی $X_{U(n)} \sim \text{gamma}(n, \theta)$ ، به راحتی می‌توان ثابت کرد که:

$$E(\hat{\theta}) = \theta,$$

تابع زیان $L(\theta, \delta)$ بیانگر مقدار زبانی است که آماردان متحمل می‌شود وقتی که θ وضعیت درست طبیعت بوده و او عمل δ را اتخاذ می‌کند. رایج‌ترین تابع زیان در مسائل برآورد و پیش‌بینی $L(\theta, \delta) = (\delta - \theta)^2$ که به تابع زیان درجه دوم معروف است. این تابع زیان متقارن است و بیش برآوردی و کم برآوردی را به طور یکسان جریمه می‌کند. با افزایش اندازه بیش برآوردی و کم برآوردی مقدار این تابع به سرعت بزرگ می‌شود. تابع زیان دیگری که در این مقاله با آن سر و کار داریم به فرم:

$$L(\theta, \delta) = b \exp\{c(\delta - \theta)\} - c(\delta - \theta) - 1$$

یا

$$L(\theta, \delta) = b \exp\{c\Delta\} - c\Delta - 1$$

است که به تابع زیان لاینکس یا خطی نمایی شهرت دارد که در آن $\Delta = \delta - \theta$.

این تابع نسبت به δ اکیداً محدب است. در این تابع $b > 0$ پارامتر مقیاس و $c \neq 0$ پارامتر شکل است. بزرگی c بیانگر درجه تقارن تابع مذکور است. در حالتی که $c < 0$ باشد، $L(\Delta)$ وقتی که $\Delta < 0$ است به صورت نمایی و وقتی که $\Delta > 0$ است به صورت خطی افزایش می‌یابد. در حالتی که $c > 0$ باشد، $L(\Delta)$ وقتی که $\Delta < 0$ است به صورت خطی و وقتی که $\Delta > 0$ است به صورت نمایی افزایش می‌یابد. زمانی که c نزدیک صفر باشد این تابع زیان تقریباً با تابع زیان درجه دو برابر است. در این مقاله حالت $b = 1$ را در نظر می‌گیریم و داریم:

$$L(\theta, \delta) = \exp\{c(\delta - \theta)\} - c(\delta - \theta) - 1. \quad (۲)$$

تعریف ۱.۱. گوییم متغیر تصادفی X دارای توزیع نمایی یک پارامتری با پارامتر مقیاس θ است و آنرا با $X \sim \text{Exp}(\theta)$ نمایش می‌دهیم، هرگاه تابع چگالی آن به صورت:

$$f(x; \theta) = \frac{1}{\theta} \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0, \quad \theta > 0 \quad (۳)$$

^۹JaheenLTR^{۱۰}Ahmadi

به این ترتیب $\hat{\theta}$ یک برآوردگر ناریب θ است. از (۵) می توان نشان و داد که $X_{U(n)}$ آماره بسنده کامل است و از این رو $\hat{\theta}$ برآوردگر ناریب به طور یکنواخت دارای کمترین واریانس ($UMVUE$) برای θ است.

خواهد بود و از این رو برآوردگر حداکثر درستنمایی پیشگوی θ ($PMLE$) و پیشگوی حداکثر درستنمایی Y به ترتیب به صورت زیر خواهند بود:

$$\begin{aligned}\hat{\theta}_{PMLE} &= \frac{1}{n+1} X_{U(n)}, \\ \hat{\theta}_{MLP} &= \frac{s}{n+1} X_{U(n)}.\end{aligned}\quad (10)$$

۲.۱.۳ پیش گویی میانه شرطی

چگالی شرطی Y به شرط $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}$ (بنابر خاصیت مارکف) به صورت:

$$f(y | x; \theta) = \frac{y - x_n}{\theta^{s-n} \Gamma(s-n)} \exp \left\{ \frac{y - x_n}{\theta} \right\}, x_n < y \quad (11)$$

است. میانه Y به شرط $X_{U(n)}$ ، پیشگوی میانه شرطی (CMP) نامیده می شود (مرجع [11] را ببینید).

حالت اول. θ معلوم. در این حالت پیشگوی میانه شرطی به صورت $\hat{Y}_{CMP} = K(X_{U(n)}, \theta)$ خواهد بود که در آن $P(K(X_{U(n)}, \theta) | x) = \frac{1}{2}$ با توجه به عبارت (۱۱):

$$\int_{x_n}^{K(x_n, \theta)} \frac{(y - x_n)^{s-n-1}}{\theta^{s-n} \Gamma(s-n)} \exp \left\{ \frac{-(y - x_n)}{\theta} \right\} dy = \frac{1}{2}$$

خواهد بود. با تغییر متغیر $y - x_n = t$ خواهیم داشت:

$$\int_0^{K(x_n, \theta) - x_n} \frac{t^{s-n-1}}{\theta^{s-n} \Gamma(s-n)} \exp \left\{ \frac{-t}{\theta} \right\} dt = \frac{1}{2}.$$

بنابراین:

$$\hat{Y}_{CMP} = X_{U(n)} + \theta \text{Med}[Z(s-n)]$$

که در آن $Z(s-n) \sim \text{gamma}(s-n, 1)$ و $\text{Med}[V]$ ، میانه V متغیر تصادفی است. چون $\frac{1}{2} \chi_{2(s-n), 0.5}^2$ ، $\text{Med}[Z(s-n)] = \frac{1}{2} \chi_{2(s-n), 0.5}^2$ ، پیشگوی حداکثر درستنمایی Y به صورت:

$$\hat{Y}_{CMP} = X_{U(n)} + \frac{\theta}{2} \chi_{2(s-n), 0.5}^2$$

خواهد بود که در آن $\chi_{r,p}^2$ ، $100p$ -امین صدک از توزیع χ^2 با درجه آزادی r است.

حالت دوم. θ مجهول. هنگامی که θ مجهول باشد، $\hat{\theta}$ با θ (۳) جایگزین می گردد و پیشگوی حداکثر درستنمایی Y به ترتیب به صورت زیر خواهند بود:

$$\hat{Y}_{CMP} = X_{U(n)} \left(1 + \frac{1}{2n} \chi_{2(s-n), 0.5}^2 \right). \quad (12)$$

۳ پیش گویی مقادیر رکوردهای بالایی

پیش گویی نقطه‌ای مقادیر رکوردهای بالایی آینده با استفاده از مقادیر رکوردهای بالایی گذشته مورد توجه بسیاری از محققین قرار گرفته است. برای جزئیات بیشتر می توانید مراجع [2]، [7]، [5] و [10] را ببینید.

۱.۳ پیش گویی کلاسیک

۱.۱.۳ پیش گویی حداکثر درستنمایی

فرض کنید که مقادیر اولین n رکورد بالایی را از تابع چگالی $f(x; \theta)$ به صورت $x = (x_1, x_2, \dots, x_n)$ مشاهده کرده ایم. می خواهیم $Y = X_{U(s)}$ ، $s > n$ را پیش گویی کنیم.

تابع درستنمایی پیش گوی Y و θ (فرض کنید θ مجهول است) به وسیله باساک و بالا کریشن (۲۰۰۳) به صورت زیر تعریف می شود:

$$\begin{aligned}L(y, \theta; x) &= f(y; \theta) \frac{[H(y; \theta) - H(x_n; \theta)]^{s-n-1}}{\Gamma(s-n)} \\ &\times \prod_{j=1}^n h(x_j; \theta), \quad x_n < y\end{aligned}$$

که $H(t; \theta) = -\ln(1 - F(t; \theta))$ است. لگاریتم تابع درستنمایی پیشگوی Y و θ به شکل زیر:

$$\begin{aligned}l &= \sum_{j=1}^n \ln h(x_j; \theta) - \ln \Gamma(s-n) + \ln f(y; \theta) \\ &+ (s-n-1) \ln[H(y; \theta) - H(x_n; \theta)], \quad x_n < y \quad (\lambda)\end{aligned}$$

است. با استفاده از ۳، ۴ و ۸ خواهیم داشت:

$$\begin{aligned}l &= -s \ln \theta - \ln \Gamma(s-n) \\ &+ (s-n-1) \ln(y - x_n) - \frac{y}{\theta}.\end{aligned}\quad (9)$$

بنابراین معادلات لگاریتم درستنمایی به صورت:

$$\frac{\partial l}{\partial \theta} = \frac{\theta(s-n-1) - (y-x_n)}{\theta(y-x_n)} = 0$$

۲.۳ پیش‌گویی بیزی

به دست می‌آید که در آن $\beta(.,.)$ تابع بتای کامل است. با استفاده از (۱۶) پیش‌گویی بیز نقطه‌ای Y را به صورت زیر محاسبه می‌کنیم:

$$\begin{aligned} E_{h^*}(Y | x) &= \int_{x_n}^{\infty} y h^*(y | x) dy \\ &= \frac{\beta(n + \alpha - 1, s - n + 1)}{(x_n + \beta)^{n + \alpha - 1}} \\ &+ \frac{\beta(n + \alpha, s - n)}{(x_n + \beta)^{n + \alpha}} x_n. \end{aligned}$$

بنابراین پیش‌گویی بیز نقطه‌ای تحت تابع زیان درجه دوم برابر مقدار زیر است:

$$\hat{Y}_{BS} = \frac{s + \alpha - 1}{n + \alpha - 1} X_{U(n)} + \frac{\beta(s - n)}{n + \alpha - 1}. \quad (17)$$

همچنین پیش‌گویی بیز نقطه‌ای تحت تابع زیان لاینکس (تابع زیان (۲)) به صورت زیر خواهد بود:

$$\hat{Y}_{BL} = \frac{-1}{c} \ln(E_{h^*}(\exp\{-cY\})). \quad (18)$$

۱.۲.۳ پیش‌گویی بیز تجربی

فرض کنید α و β یعنی پارامترهای توزیع پیشین θ هر دو مجهول باشند. فرض کنید در شرایطی که مشاهدات $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}$ را در اختیار داریم از قبل همه مشاهدات m نمونه رکوردی و هر یک به حجم n را به صورت:

$$X_{j,U(1)}, X_{j,U(2)}, \dots, X_{j,U(n)}, \quad j = 1, 2, \dots, m$$

در اختیار داشته باشیم. فرض می‌شود نمونه j -ام از توزیع $Exp(\theta_j)$ به دست آمده باشد. با توجه به (۶) برآوردگر حداکثر درست‌نمایی θ_j به صورت $Z_j = \frac{X_{j,U(n)}}{n}$ است. چگالی شرطی Z_j به شرط θ_j به صورت:

$$f(z_j | \theta_j) = \frac{n^n}{\Gamma(n)\theta_j^n} z_j^{n-1} \exp\left\{-\frac{nz_j}{\theta_j}\right\}, \quad z_j > 0 \quad (19)$$

و توزیع پیشین برای θ_j به شکل زیر:

$$\pi(\theta_j) = \frac{\beta^\alpha \exp\left\{-\frac{\beta}{\theta_j}\right\}}{\Gamma(\alpha)\theta_j^{\alpha+1}}, \quad \alpha > 2, \theta_j > 0 \quad (20)$$

در این بخش مسئله پیش‌گویی نقطه‌ای مقادیر رکوردهای بالایی آینده را بر اساس نگرش بیز در نظر می‌گیریم. با فرض آن که θ مجهول باشد، توزیع پیشین برای θ را گامای معکوس با پارامترهای $\alpha > 2$ و $\beta > 0$ در نظر می‌گیریم ($\theta \sim IG(\alpha, \beta)$) بنابراین داریم:

$$\pi(\theta) = \frac{\beta^\alpha \exp\left\{-\frac{\beta}{\theta}\right\}}{\Gamma(\alpha)\theta^{\alpha+1}}, \quad \theta > 0 \quad (13)$$

در این صورت تابع چگالی پیشین θ در رابطه زیر صدق می‌کند

$$\pi(\theta | x) \propto f(x | \theta)\pi(\theta). \quad (14)$$

که در آن تابع چگالی توام $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}$ در x و در (۵) و $\pi(\theta)$ تابع چگالی پیشین داده شده در (۱۳) است. با جایگذاری (۵) و (۱۳) در (۱۴) و پس از نرمال‌سازی تابع چگالی پیشین θ به صورت

$$\pi(\theta | x) = \frac{(\beta + x_n)^{n+\alpha} \exp\left\{-\frac{(\beta+x_n)}{\theta}\right\}}{\Gamma(n+\alpha)\theta^{n+\alpha+1}}, \quad \theta > 0 \quad (15)$$

به دست خواهد آمد.

فرض کنید مقادیر اولین n رکورد بالایی را به صورت $Exp(\theta)$ از $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(n)} = x_n$ مشاهده کرده‌ایم. می‌خواهیم بر اساس چنین نمونه‌ای مقدار $Y = X_{U(s)}$, $s > n$ را پیش‌بینی کنیم. تابع چگالی پیش‌گویی بیز Y به شرط x (مرجع [8] را ببینید) به صورت زیر تعریف می‌شود:

$$h^*(y | x) = \int_{\Theta} f_{Y|x}(y | x)\pi(\theta | x)d\theta.$$

با ترکیب چگالی پسین داده شده در (15) و چگالی شرطی داده شده در (11) و انتگرال‌گیری نسبت به θ ، چگالی پیش‌گویی بیز Y به شرط n رکورد بالایی گذشته به شکل زیر:

$$h^*(y | x) = \frac{(y - x_n)^{s-n-1}(\beta + x_n)^{n+\alpha}}{\beta(n + \alpha, s - n)(y + \beta)^{s+\alpha}}, \quad y > x_n \quad (16)$$

^{۱۱}Schafer

^{۱۲}Feduccia

پیشگوهای بیز و بیز تجربی، تحت تابع زیان لاینکس را برای مقادیر داده شده c حساب می‌کنیم.

(۶) توان دوم انحرافات $(y^* - y)^2$ را محاسبه می‌کنیم که در آن y^* هر کدام از پیشگوهای محاسبه شده در مراحل چهار و پنج است.

(۷) مراحل دو تا شش را هزار مرتبه تکرار می‌کنیم و برآورد ریسک را به صورت میانگین مربع انحرافات در هزار تکرار در نظر می‌گیریم.

محاسبات مربوط به شبیه‌سازی فوق در جدول ۱ آمده‌اند. توجه شود که $E(R_m(Y^*))$ به معنای برآورد ریسک پیش‌گوی بیز تجربی (Y^*) در حالتی که m نمونه رکوردی و هر یک به حجم n در اختیار داشته باشیم، است.

۵ نتایج شبیه‌سازی

(۱) اگر n بزرگ باشد ($n \geq 5$) برآورد ریسک پیش‌گوی بیز تحت تابع زیان درجه دوم کمتر از برآورد ریسک پیش‌گوی میانه شرطی است و برای n های کوچک ($n = 2, 3, 4$) برآورد ریسک پیش‌گوی بیز تحت تابع زیان درجه دوم در پیش‌گویی رکوردهای دورتر کمتر از برآورد ریسک پیش‌گوی میانه شرطی است.

(۲) تقریباً همه جا برآورد ریسک پیش‌گوی میانه شرطی کمتر از برآورد ریسک پیش‌گوی حداکثر درستی است.

(۳) همواره برآورد ریسک پیش‌گوی بیز تحت تابع زیان درجه دوم کمتر از برآورد ریسک پیش‌گوی حداکثر درستی است.

(۴) با دور شدن c از صفر برآورد ریسک پیش‌گوی بیز و بیز تجربی تحت تابع زیان لاینکس بزرگ می‌شود و رشد این افزایش در حالتی که c منفی باشد بسیار بیشتر از حالتی است که c مثبت است.

(۵) هنگامی که c به صفر نزدیک می‌شود برآورد ریسک پیش‌گوی بیز و بیز تجربی تحت تابع زیان لاینکس به ترتیب به برآورد ریسک پیش‌گوی بیز و بیز تجربی تحت تابع زیان درجه دوم میل خواهند کرد.

(۶) هنگامی که m و n افزایش می‌یابند برآورد ریسک پیش‌گوهای معرفی شده کاهش پیدا می‌کنند.

(۷) برای هر مقدار دیگری از $\alpha > 2$ و $\beta > 0$ نتایج شبیه‌سازی به صورت بالا است.

است. پیرو چیفر^{۱۱} و فیداکشا^{۱۲} و با استفاده از (۱۹) و (۲۰)، چگالی حاشیه‌ای Z_j برای $j = 1, 2, \dots, m$ برابر مقدار زیر است:

$$f(z_j) = \int_0^\infty f(z_j | \theta_j) \pi(\theta_j) d\theta_j \\ = \frac{n^n \beta^\alpha z_j^{n-1}}{\beta(\alpha, n)(nz_j + \beta)^{n+\alpha}}, \quad z_j > 0. \quad (21)$$

با استفاده از (۲۱) برآورد گشتاوری α و β را به صورت:

$$\tilde{\alpha} = \frac{2nS_2 - (n+1)S_1^2}{nS_2 - (n+1)S_1^2}, \quad \tilde{\beta} = (\tilde{\alpha} - 1)S_1 \quad (22)$$

به دست می‌آوریم که در آن $S_2 = \sum_{i=1}^m \frac{Z_i^2}{m}$ و $S_1 = \sum_{i=1}^m \frac{Z_i}{m}$ است. با جایگذاری $\tilde{\alpha}$ و $\tilde{\beta}$ در (۱۷) و (۱۸) به ترتیب پیش‌گوهای بیز تجربی تحت تابع زیان درجه دوم خطا و لاینکس (تابع زیان (۲)) به دست خواهند آمد.

۴ محاسبات عددی

پیش‌گوهای حداکثر درستی، میانه شرطی، بیز و بیز تجربی (تحت توابع زیان درجه دوم و لاینکس) برای n و s داده شده، بر اساس شبیه‌سازی مونت کارلو طی مراحل زیر با هم مقایسه می‌شوند:

(۱) به ازای n در نظر گرفته شده، برای مقادیر داده شده α و β ($\alpha = 3.5, \beta = 1.1$)، θ_j ، $j = 1, 2, \dots, m$ را از توزیع پیشین (۲۰) تولید می‌کنیم و سپس به ازای هر j یک نمونه تصادفی از توزیع $gamma(n, \theta_j/n)$ تولید کرده و آنرا Z_j می‌نامیم. و سپس بر اساس این نمونه‌ها $\tilde{\alpha}$ و $\tilde{\beta}$ را از (۲۲) حساب می‌کنیم.

(۲) برای مقادیر داده شده α و β ، θ را از توزیع پیشین (۱۵) تولید می‌کنیم.

(۳) نمونه تصادفی $v_1, v_2, \dots, v_n, \dots, v_s$ را از $Exp(\theta)$ تولید می‌کنیم و قرار می‌دهیم:

$$x_n = \sum_{j=1}^n v_j, \quad y = \sum_{j=1}^s v_j.$$

(۴) پیشگوهای حداکثر درستی و میانه شرطی (در حالت θ مجهول)، بیز و بیز تجربی، تحت تابع زیان درجه دو را محاسبه می‌کنیم.

(۵) یک نمونه‌ی هزارتایی از (۱۶) به روش نمونه‌گیری برشی^{۱۳} (برای جزئیات بیشتر به مرجع [6] مراجعه کنید) تولید می‌کنیم و

جدول ۱: ریسک برآورد شده (ER) در پیش‌گویی Y برای مقادیر مختلف m, n, c در هزار تکرار (α = ۳/۵, β = ۱/۱)

n	s	ER(\hat{Y}_{MLP})	ER(\hat{Y}_{CMP})	ER(\hat{Y}_{BS})	ER _{m=1} (\hat{Y}_{EBS})	ER _{m=2} (\hat{Y}_{EBS})	c	ER(\hat{Y}_{BL})	ER _{m=1} (\hat{Y}_{EHL})	ER _{m=2} (\hat{Y}_{EHL})
۳	۳	۳۶,۶۰۰	۱۶,۷۹۳	۲۰,۳۶۲	۱۷,۶۹۳	۱۷,۷۱۵	-1	۲۸۳,۸۱۸	۷۲۹,۲۲۸	۷۵۲,۱۰۶
							-1	۴۴,۱۷۵	۶۵,۰۵۶	۵۲,۳۴۸
							-1	۴۰,۳۶۵	۱۷,۵۵۳	۱۷,۷۲۶
							1	۳۰,۳۷۷	۱۷,۵۵۶	۱۷,۷۱۸
							1	۲۱,۰۷۰	۱۸,۱۴۷	۱۸,۰۵۷
1	۲۴,۱۸۹	۲۲,۲۹۷	۲۲,۱۰۸							
۳	۴	۵۲,۹۲۹	۳۵,۷۲۳	۳۶,۷۵۰	۳۷,۱۱۶	۳۸,۶۹۶	-1	۴۰۵,۹۱۱	۱۳۶۷,۸۷۰	۴۰۴,۷۲۲
							-1	۱۸۸,۲۸۹	۱۷۷,۳۲۱	۴۲,۰۵۳
							-1	۴۶,۲۸۱	۳۷,۷۸۲	۳۹,۰۶۷
							1	۴۶,۳۶۲	۳۷,۶۸۱	۳۹,۰۶۹
							1	۴۹,۴۹۷	۳۷,۵۴۶	۴۰,۰۶۶
1	۶۰,۹۰۸	۵۲,۰۲۰	۵۲,۳۷۴							
۳	۵	۸۳,۲۹۱	۸۳,۵۰۳	۷۶,۰۹۹	۶۹,۸۳۳	۵۵,۲۷۹	-1	۶۶۰,۵۲۱	۲۵۸۹,۶۸۲	۱۵۵۱,۳۷۷
							-1	۲۱۱,۵۲۰	۵۰۰,۰۵۲	۲۵۲,۴۷۹
							-1	۷۷,۵۰۶	۷۲,۳۲۲	۵۸,۸۶۰
							1	۷۷,۵۵۱	۷۱,۵۴۲	۵۸,۶۶۷
							1	۸۶,۳۰۵	۵۹,۱۲۷	۶۰,۱۵۰
1	۱۱۲,۶۴۱	۸۸,۷۷۸	۹۵,۰۱۷							

n	s	ER(\hat{Y}_{MLP})	ER(\hat{Y}_{CMP})	ER(\hat{Y}_{BS})	ER _{m=1} (\hat{Y}_{EBS})	ER _{m=2} (\hat{Y}_{EBS})	c	ER(\hat{Y}_{BL})	ER _{m=1} (\hat{Y}_{EHL})	ER _{m=2} (\hat{Y}_{EHL})
۳	۶	۱۲۱,۶۰۲	۱۳۸,۸۱۷	۱۱۵,۰۰۰	۸۴,۵۵۹	۸۱,۸۷۷	-1	۸۶۵,۶۲۱	۹۱۷,۲۳۷	۲۸۲,۹۸۲
							-1	۲۳۵,۳۷۲	۲۱۲,۰۲۶	۸۹,۰۸۶
							-1	۱۱۷,۰۴۱	۸۷,۸۰۵	۸۳,۶۰۸
							1	۱۱۶,۹۵۳	۸۷,۸۹۷	۸۳,۶۷۷
							1	۱۲۶,۹۹۳	۹۹,۴۲۴	۸۹,۴۵۶
1	۱۷۳,۲۹۵	۱۵۶,۲۱۴	۱۲۱,۹۸۳							
۳	۷	۱۷۲,۷۵۳	۱۹۵,۶۹۳	۱۶۹,۱۸۶	۱۲۴,۶۵۲	۱۱۱,۲۷۹	-1	۹۰۶,۸۲۵	۳۵۵,۳۰۵	۱۳۸۷,۰۶۲
							-1	۳۶۹,۱۲۲	۳۵۵,۳۰۵	۲۸۵,۷۴۸
							-1	۱۷۷,۴۱۷	۳۵۵,۳۰۵	۱۲۰,۶۶۳
							1	۱۷۷,۶۸۱	۳۵۵,۳۰۵	۱۲۰,۷۲۰
							1	۱۹۹,۵۲۸	۳۵۵,۳۰۵	۱۳۷,۹۵۷
1	۲۷۸,۰۰۹	۳۵۵,۳۰۶	۲۴۲,۱۸۵							
۳	۸	۲۱,۷۵۳	۲۰,۵۹۶	۲۱,۰۵۷	۲۱,۱۷۶	۱۹,۰۹۹	-1	۳۱۵,۸۰۷	۲۰۸,۰۲۴	۶۶۶,۲۲۰
							-1	۳۵,۷۱۲	۲۵۷,۷۷۸	۵۳,۶۳۰
							-1	۲۱,۱۲۵	۲۱,۳۵۰	۱۹,۲۵۹
							1	۲۱,۱۴۶	۲۱,۳۱۵	۱۹,۲۶۲
							1	۲۲,۰۸۷	۲۰,۵۲۴	۱۹,۹۱۰
1	۲۵,۷۱۵	۲۳,۳۹۵	۲۳,۳۰۰							

n	s	ER(\hat{Y}_{MLP})	ER(\hat{Y}_{CMP})	ER(\hat{Y}_{BS})	ER _{m=1} (\hat{Y}_{EBS})	ER _{m=2} (\hat{Y}_{EBS})	c	ER(\hat{Y}_{BL})	ER _{m=1} (\hat{Y}_{EHL})	ER _{m=2} (\hat{Y}_{EHL})
۳	۵	۵۶,۷۶۶	۴۲,۹۹۱	۴۴,۵۸۲	۸۵,۰۶۰	۳۸,۰۲۸	-1	۵۴۱,۰۵۰	۷۵,۱۴۲	۷۶۹,۴۷۳
							-1	۱۱۸,۵۷۲	۷۵,۲۱۲	۱۰۰,۳۵۱
							-1	۴۵,۲۴۰	۷۵,۲۱۶	۳۷,۶۵۷
							1	۴۵,۳۰۶	۷۵,۲۱۶	۳۷,۶۶۶
							1	۴۸,۸۰۴	۷۵,۲۱۸	۳۹,۰۲۲
1	۶۳,۴۲۲	۷۵,۲۲۸	۵۸,۳۸۷							
۳	۶	۸۰,۵۷۳	۷۷,۳۸۵	۷۳,۶۶۸	۶۰,۲۲۳	۶۱,۶۰۴	-1	۹۰۸,۰۹۱	۱۵۹۵,۸۰۲	۱۲۸۸,۲۸۹
							-1	۲۶۷,۲۹۴	۳۳۹,۰۷۱	۱۷۷,۳۰۰
							-1	۷۵,۱۸۸	۶۳,۶۲۹	۶۳,۹۱۰
							1	۷۵,۲۰۹	۶۳,۴۹۹	۶۳,۷۶۵
							1	۸۲,۶۸۲	۶۵,۵۶۵	۶۵,۴۳۰
1	۱۱۷,۷۵۹	۱۰۵,۰۸۸	۱۰۳,۲۸۹							
۳	۷	۱۱۳,۰۸۳	۱۱۳,۸۱۳	۱۰۶,۲۳۷	۷۸,۷۱۷	۸۱,۰۶۰	-1	۱۲۹۵,۲۰۵	۱۶۴۳,۲۹۵	۶۴۴,۰۲۵
							-1	۶۳۹,۹۰۱	۳۶۸,۷۷۶	۱۰۹,۸۱۹
							-1	۱۱۳,۶۸۲	۸۱,۳۹۶	۸۵,۰۲۴
							1	۱۱۱,۹۸۴	۸۱,۰۲۶	۸۵,۱۲۳
							1	۱۱۸,۳۵۷	۸۵,۰۷۶	۹۶,۱۱۲
1	۱۷۲,۱۶۷	۱۵۱,۱۹۰	۱۵۱,۶۳۴							

n	s	$ER(\hat{Y}_{MLP})$	$ER(\hat{Y}_{CMP})$	$ER(\hat{Y}_{BS})$	$ER_{m=1}(\hat{Y}_{EBS})$	$ER_{m=2}(\hat{Y}_{EBS})$	c	$ER(\hat{Y}_{BL})$	$ER_{m=1}(\hat{Y}_{EBL})$	$ER_{m=2}(\hat{Y}_{EBL})$
۴	۵	۱۳,۱۷۱	۱۳,۶۷۱	۱۳,۶۱۴	۱۳,۹۱۳	۱۳,۹۱۹	-۱	۲۸۶,۶۵۶	۱۸,۰۷۰	۲۲۱,۱۵۸
							-۱	۲۵,۶۹۹	۱۸,۰۷۰	۱۳,۶۱۶
							-۱	۱۳,۷۸۸	۱۸,۰۷۰	۱۳,۰۳۴
							۱	۱۳,۷۸۲	۱۸,۰۷۰	۱۳,۰۳۷
							۱	۱۶,۱۵۷	۱۸,۰۷۰	۱۳,۲۷۲
۴	۶	۵۵,۴۳۶	۴۴,۱۹۲	۴۴,۹۷۸	۴۸,۶۸۹	۴۸,۲۷۱	-۱	۵۱۳,۹۷۶	۱۸۹,۶۰۰	۷۰۱,۶۶۱
							-۱	۱۱۳,۱۱۴	۱۷۶,۳۱۳	۸۲,۵۴۳
							-۱	۴۶,۱۸۸	۳۸,۷۲۶	۳۸,۷۸۲
							۱	۴۶,۱۵۷	۳۸,۶۱۴	۳۸,۷۸۷
							۱	۴۸,۴۵۴	۳۹,۴۳۹	۴۱,۰۲۰
۴	۷	۶۸,۰۳۲	۵۹,۹۳۱	۵۹,۵۸۳	۶۹,۴۹۰	۶۹,۱۹۶	-۱	۶۶۴,۹۶۹	۱۳۷,۵۲۳	۹۱۷,۸۶۱
							-۱	۱۴۸,۱۰۷	۲۹۲,۶۳۷	۱۱۰,۴۲۰
							-۱	۶۱,۴۹۷	۵۱,۲۸۷	۵۰,۲۳۸
							۱	۶۱,۴۲۲	۵۱,۰۶۷	۵۰,۱۹۰
							۱	۶۶,۹۴۴	۵۲,۰۶۰	۵۲,۷۴۷
۴	۸	۹۷,۱۴۲	۸۷,۱۹۲	۸۷,۱۹۲	۹۷,۱۴۲	۹۷,۱۴۲	-۱	۹۷,۱۴۲	۸۷,۱۹۲	۸۵,۶۸۰

n	s	$ER(\hat{Y}_{MLP})$	$ER(\hat{Y}_{CMP})$	$ER(\hat{Y}_{BS})$	$ER_{m=1}(\hat{Y}_{EBS})$	$ER_{m=2}(\hat{Y}_{EBS})$	c	$ER(\hat{Y}_{BL})$	$ER_{m=1}(\hat{Y}_{EBL})$	$ER_{m=2}(\hat{Y}_{EBL})$
۵	۶	۳۵,۱۹۰	۱۵,۵۹۳	۱۵,۳۹۵	۱۶,۳۳۱	۱۶,۳۷۸	-۱	۳۳,۷۱۷	۲۶۵,۴۱۰	۴۳۲,۵۲۷
							-۱	۳۱,۰۰۷	۲۷,۷۲۱	۲۸,۸۸۳
							-۱	۱۵,۳۷۶	۱۶,۲۷۱	۱۶,۳۰۵
							۱	۱۵,۳۶۹	۱۶,۲۵۵	۱۶,۲۹۴
							۱	۱۵,۶۸۲	۱۶,۳۲۶	۱۶,۴۶۸
۵	۷	۴۵,۹۸۵	۲۹,۴۷۲	۲۸,۴۹۰	۲۶,۰۲۷	۲۶,۵۴۸	-۱	۴۶۹,۰۰۶	۷۶۳,۶۷۹	۸۷۳,۲۳۰
							-۱	۶۶,۰۳۶	۱۲۹,۷۹۰	۸۱,۵۱۷
							-۱	۲۸,۸۰۲	۲۶,۰۶۸	۲۶,۲۲۱
							۱	۲۸,۷۷۲	۲۵,۹۷۳	۲۶,۲۷۵
							۱	۲۹,۷۰۵	۲۶,۱۷۸	۲۶,۴۱۰
۵	۸	۲۶,۴۲۹	۱۶,۸۸۰	۱۶,۵۶۹	۱۳,۹۸۰	۱۳,۵۶۸	-۱	۲۸۲,۱۸۰	۲۶۹,۹۳۵	۳۵۷,۸۱۶
							-۱	۲۷,۳۳۲	۱۸,۵۱۵	۲۰,۶۴۷
							-۱	۱۶,۶۰۱	۱۶,۰۳۳	۱۶,۶۹۷
							۱	۱۶,۶۰۳	۱۶,۰۳۵	۱۶,۶۹۲
							۱	۱۵,۰۱۴	۱۶,۲۷۱	۱۶,۹۳۵

مراجع

[1] Ahmadi, J. and Doostparast, M.(2007). Statistical inference based on record data from pareto model, *Appl. Math. Comput*, **41**, 105-118 .

[2] Ahsanullah, M. (1980). Linear prediction of record values for the two parametr exponential distribution, *Ann. Inst. Statist. Math* **32** (A), 363-368.

[3] Ahsanullah, M.(1995). *Record Statistics*, Nova science, Hontington, NY.

[4] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*, John Wiley, New York.

[5] Basak, P. and Balakrishnan, N.(2003). Maximum likelihood prediction of future record statistic, In *Mathematical and Statistical Method in Reliability, Series on Quality, Reliability and Engineering, Volume 7*.

- [6] Dagpunar, J. S.(2007). *Simulation and Monte Carlo*, WILEY, England.
- [7] Dunsmore,IR.(1983). The future occurrence of records, *Ann. Inst. Statist. Math* **35**, 267-277.
- [8] Dunsmore,IR.(1974). The Bayesian predictive distribution in life testing mod els, *Technometrics* **16**, 455-460.
- [9] Jaheen, Z.F.(2004). Empirical Bayes analysis record statistics based on LINEX and Quadratic loss functions.*Computers and Mathematical with Applications*,A7, 947-954.
- [10] Raqab, M.Z. and Nagaraja, H.N.(1995). On some predictor of future order statistics, *Metron*, **53**, nos. 1-2, 185-204.
- [11] Schafer, R.E. and Feduccia, A.J. .(1972). Prior distributions fitted to observed reliability data, *IEEE Transactions on Reliability* **R-21**, 148-154.