

برآورد و پیشگویی بیزی پارامترهای توزیع نمایی یک و دو پارامتری بر اساس نمونه‌های سانسور شده مضاعف نوع دوم از آماره‌های ترتیبی دنباله‌ای

مهرانگیز فلاحتی نایینی^۱

چکیده:

در این مقاله، آماره‌های ترتیبی دنباله‌ای معرفی می‌شوند، سپس بر اساس نمونه سانسور شده مضاعف نوع دوم از آماره‌های ترتیبی دنباله‌ای، برآوردگر بیز برای پارامترهای توزیع نمایی یک و دو پارامتری تحت فرض توزیع پیشین گامای وارون و تابع زیان توان دوم خطا به دست آمده و همچنین پیشگویی بیزی زمان‌های شکست آتی بررسی شده است. در ادامه، در مثالی کاربردی برآوردگر بیز و برآوردگرهای غیربیزی هم‌چون بهترین برآوردگر ناریب خطی (BLUE) و برآوردگرهای ماکسیمم درست‌نمایی تقریبی (AMLE) به دست آورده می‌شوند.

واژه‌های کلیدی: آماره‌های ترتیبی دنباله‌ای، سانسور مضاعف نوع دوم، برآورد بیز، پیشگویی بیز، بهترین برآوردگر ناریب خطی، برآوردگر ماکسیمم درست‌نمایی تقریبی.

۱ مقدمه

که در آن، $\alpha_1, \dots, \alpha_n$ مقادیر حقیقی و مثبت هستند که پارامتر مدل نامیده می‌شوند و F تابع توزیع مطلقاً پیوسته متناظر با تابع چگالی f است. نرخ خطر عناصر باقی‌مانده بعد از j -امین شکست به صورت $\frac{\alpha_{j+1}f}{1-F}$ به دست می‌آید که در واقع α_{j+1} نشانگر اثر j -امین شکست بر روی $n-j$ عنصر باقی‌مانده در حال کار است.

در فرمول بالا اگر $\alpha_1 = \alpha_2 = \dots = \alpha_n = 1$ یا $F_1 = F_2 = \dots = F_n$ مدل به مدل آماره‌های ترتیبی معمولی تبدیل می‌شود.

تعریف ۱.۱. گمپس [۵]. فرض کنید $1 \leq i < n$ ، $Y_j^{(i)}$ فرض کنید $1 \leq j \leq n-i+1$ ، متغیرهای تصادفی مستقل با توزیع F_i باشند که $F_1, \dots, F_n$ اکیداً صعودی و توابعی پیوسته هستند و $F_1^{-1}(1) \leq \dots \leq F_n^{-1}(1)$ در این صورت اگر $X_j^{(1)} = Y_j^{(1)}$ ، $1 \leq j \leq n$ ،

$$X_*^{(1)} = \min\{X_1^{(1)}, \dots, X_n^{(1)}\},$$

و برای $2 \leq i \leq n$

$$X_j^{(i)} = F_i^{-1}(F_i(Y_j^{(i)})(1 - F_i(X_*^{(i-1)})) + F_i(X_*^{(i-1)})),$$

و

$$X_*^{(i)} = \min\{X_j^{(i)}, 1 \leq j \leq n-i+1\},$$

یک زیرمجموعه مهم از آماره‌های ترتیبی تعمیم‌یافته، آماره‌های ترتیبی دنباله‌ای است که برای مدل‌بندی ساختار نوعی سیستم k از n دنباله‌ای مورد استفاده قرار می‌گیرد. به صورت معمول سیستم k از n سیستمی است که با n عنصر شروع به کار می‌کند و این سیستم تا زمانی که حداقل k یا بیشتر از k عنصر آن کار می‌کند، سالم است و به کار کردن ادامه می‌دهد. در اغلب موارد فرض می‌شود که عناصر این سیستم به صورت مستقل کار می‌کنند اما این فرض در همه سیستم‌ها به صورت کامل صدق نمی‌کند و در برخی مواقع ممکن است شکست و خرابی یک عنصر فشار بیشتر و یا خطری را برای عناصر باقی‌مانده به وجود آورد (در یک کار گروهی با کم شدن یکی از اعضا برای اتمام کار فشار بیشتری بر افراد باقی‌مانده وارد می‌شود).

سیستم k از n دنباله‌ای اولین بار توسط گمپس [۶] در سال ۱۹۹۵ معرفی شد که در این سیستم بعد از هر شکست، توزیع طول عمر جدیدی بر عناصر باقی‌مانده تطبیق داده می‌شود. بنابراین در آماره‌های ترتیبی دنباله‌ای توزیع آماره‌ها با تعداد مؤلفه‌های در حال کار مرتبط است و این شرط توسط α_j اعمال می‌شود. در واقع توابع توزیع $F_1, \dots, F_n$ از رابطه‌ی زیر به دست می‌آیند:

$$F_j = 1 - (1 - F)^{\alpha_j},$$

به صورت معمول در آزمایش‌های مختلف تمامی زمان‌های شکست قابل دسترس و یا قابل استفاده نیستند. به عنوان مثال ممکن است در یک آزمون طول عمر برای ذخیره هزینه و زمان، آزمون بعد از تعداد مشخصی از شکست‌های مشاهده شده متوقف شود. در چنین مواردی در استنباط آماری نمی‌توان تمام زمان‌های شکست را مشاهده و بنابراین باید انتخاب این زمان‌ها را محدود کرد. در این صورت زمان لازم برای رسیدن به q -امین شکست، یک متغیر تصادفی است.

مدل آماره‌های ترتیبی دنباله‌ای نیز برای چنین داده‌هایی می‌تواند مورد استفاده قرار بگیرد. فرض کنید j_1 تا j_q -امین مشاهده رخ دهد ($x^{(j_1)} \leq \dots \leq x^{(j_q)}$) و باقی داده‌ها در دسترس نباشند، متغیرهای تصادفی متناظر با این مشاهدات را با نماد

$$X_*^{(j_1)} \leq \dots \leq X_*^{(j_q)}, \quad 1 < j_1 < \dots < j_q < n,$$

نشان داده و آماره‌های ترتیبی دنباله‌ای سانسور شده مضاعف نوع دوم نامیده می‌شود.

اگر تنها $1 + p$ -امین تا q -امین زمان شکست مشاهده شود و p تا از کوچکترین و $n - q$ تا از بزرگترین زمان‌های شکست سانسور شده باشند، متغیرهای تصادفی متناظر را با

$$X_*^{(p+1)} \leq X_*^{(p+2)} \leq \dots \leq X_*^{(q)},$$

نشان داده و آماره‌های ترتیبی دنباله‌ای سانسور شده دوگانه نوع دوم نامیده می‌شود. در حالت خاص اگر $p = 0$ ، متغیرهای تصادفی $X_*^{(1)} \leq X_*^{(2)} \leq \dots \leq X_*^{(q)}$ آماره‌های ترتیبی دنباله‌ای سانسور شده نوع دوم را تشکیل می‌دهند.

در این مقاله بر روی نمونه‌هایی از داده‌های سانسور شده مضاعف نوع دوم از آماره‌های ترتیبی دنباله‌ای با توزیع $Exp(\mu, \sigma_i)$ کار می‌شود که در واقع q_i مشاهده از i -امین نمونه به طوری که $1 \leq i \leq s$ ، گرفته شده است. در این روش همه نمونه‌ها دارای پارامتر مکان یکسان μ و پارامترهای مقیاس $\sigma_1, \dots, \sigma_s$ هستند. به صورت کلی مجموعه داده‌ها را می‌توان به صورت زیر نمایش داد:

$$\mu = x_i^{(j_{i0})} \leq x_i^{(j_{i1})} \leq \dots \leq x_i^{(j_{iq_i})},$$

$$j_{i0} = 0 < j_{i1} < \dots < j_{iq_i} \leq n_i, \quad 1 \leq i \leq s,$$

و آماره‌های ترتیبی دنباله‌ای متناظر نیز با نماد $X_{*i}^{(j_{ik})}$ که $1 \leq i \leq s$ ، $1 \leq k \leq q_i$ نشان داده می‌شود.

آنگاه متغیرهای تصادفی $X_*^{(1)}, \dots, X_*^{(n)}$ ، آماره‌های ترتیبی دنباله‌ای نامیده می‌شوند.

با توجه به این تعریف، توزیع شرطی متغیر تصادفی $X_1^{(i)}$ برای $1 \leq i \leq n$ ، بنابر استقلال $Y_1^{(i)}$ و $X_*^{(i-1)}$ ، به صورت زیر به دست می‌آید:

$$\begin{aligned} P(X_1^{(i)} \leq t | X_*^{(i-1)} = s) &= P\left(F_i(Y_1^{(i)}) \leq \frac{F_i(t) - F_i(s)}{1 - F_i(s)}\right) \\ &= \frac{F_i(t) - F_i(s)}{1 - F_i(s)} = G_i(t|s). \end{aligned}$$

این توزیع شرطی، نشان دهنده این است که بعد از i -امین خرابی در سیستم، شکست بعدی از $X_1^{(i)}, \dots, X_{n-i+1}^{(i)}$ با توزیع $G_i(t|s)$ به دست می‌آید. بنابر تعریف

$$X_*^{(i)} = \min\{X_j^{(i)}, 1 \leq j \leq n - i + 1\},$$

و همچنین تعریف $G_i(t|s)$ ، خواهیم داشت:

$$\begin{aligned} P(X_*^{(r)} > t | X_*^{(r-1)} = s) &= (P(X_1^{(r)} > t | X_*^{(r-1)} = s))^{n-r+1} \\ &= \left(\frac{1 - F_r(t)}{1 - F_r(s)}\right)^{n-r+1}, \end{aligned}$$

$$P(X_*^{(r)} < t | X_*^{(r-1)} = s) = 1 - \left(\frac{1 - F_r(t)}{1 - F_r(s)}\right)^{n-r+1}.$$

با مشتق گیری نسبت به t داریم:

$$\begin{aligned} f_{X_*^{(r)} | X_*^{(r-1)}}(t|s) &= (n - r + 1) \frac{f_r(t)}{1 - F_r(s)} \\ &\times \left(\frac{1 - F_r(t)}{1 - F_r(s)}\right)^{n-r}. \end{aligned}$$

بنابراین تابع چگالی توأم $X_*^{(1)}, \dots, X_*^{(r)}$ با استفاده از خاصیت مارکف به صورت زیر به دست می‌آید:

$$\begin{aligned} f_{X_*^{(1)}, \dots, X_*^{(r)}}(x_1, \dots, x_r) &= \\ \frac{n!}{(n-r)!} \prod_{i=1}^r \left(\frac{1 - F_i(x_i)}{1 - F_i(x_{i-1})}\right)^{n-i} \frac{f_i(x_i)}{1 - F_i(x_{i-1})}. \end{aligned}$$

مثال ۲.۱. اگر $F(x_i) = 1 - e^{-x_i}$ ، آنگاه با استفاده از رابطه‌ی $F_i(x_i) = 1 - (1 - F(x_i))^{\alpha_i}$ تابع چگالی توأم $X_*^{(1)}, \dots, X_*^{(n)}$ به صورت زیر به دست می‌آید:

$$f_{X_*^{(1)}, \dots, X_*^{(n)}}(x_1, \dots, x_n) = n! \prod_{i=1}^n \alpha_i e^{-(x_i - x_{i-1})(n-i+1)\alpha_i}.$$

۲ تابع چگالی آماره‌های ترتیبی دنباله‌ای ۳ برآورد و پیشگویی بیزی

سانسور شده مضاعف نوع دوم

در سال‌های اخیر مطالعات زیادی بر روی روش‌های بیزی با نمونه‌های سانسور شده نوع دوم و همچنین نمونه‌های سانسور شده پیش‌رونده نوع دوم بر اساس آماره‌های ترتیبی دنباله‌ای و یا آماره‌های تعمیم‌یافته، توسط جاهین [۴]، احمدی و دوست پرست [۱] و سلطان [۸] انجام گرفته است. شنک و همکاران [۷] به محاسبه برآورد و پیشگویی‌های بیزی بر اساس نمونه‌های سانسور شده مضاعف نوع دوم از آماره‌های ترتیبی دنباله‌ای در توزیع نمایی یک و دو پارامتری پرداختند. در این قسمت با استفاده از یافته‌های بخش قبل، برآوردگر بیزی برای پارامترهای توزیع نمایی یک و دو پارامتری به دست می‌آید و همچنین پیشگویی بیزی برای زمان شکست آینده ارائه می‌گردد.

قضیه ۱.۲. فواصل متوالی استاندارد شده آماره‌های ترتیبی دنباله‌ای، تحت توزیع نمایی استاندارد، متغیرهای مستقل و هم‌توزیع با توزیع نمایی استاندارد هستند؛ در واقع

$$W_{*i}^{(1)} = \frac{\gamma_{i1}}{\sigma_i} (X_{*i}^{(1)} - \mu), \quad 1 \leq i \leq s,$$

$$W_{*i}^{(j)} = \frac{\gamma_{ij}}{\sigma_i} (X_{*i}^{(j)} - X_{*i}^{(j-1)}), \quad 2 \leq j \leq n_i,$$

با $\gamma_{ij} = \alpha_{ij}(n_i - j + 1)$ به صورت نمایی استاندارد توزیع شده‌اند. اثبات. به کمپس [۵] رجوع شود. □

با در نظر گرفتن تعداد نمونه‌ها برابر با ۱، قضیه زیر بیان می‌شود:

۱.۳ توزیع نمایی یک پارامتری

اگر $\sigma_1 = \dots = \sigma_s = \sigma$ به صورت گامای وارون توزیع شده باشد، آنگاه برآورد بیزی σ و پیشگویی بیزی زمان شکست آینده با فرض وجود نمونه‌هایی از آماره‌های ترتیبی سانسور شده مضاعف نوع دوم به صورت زیر به دست می‌آیند.

قضیه ۲.۲. الف) تابع چگالی آماره‌های ترتیبی دنباله‌ای سانسور شده مضاعف نوع دوم

$$X_{*}^{(j_1)} \leq \dots \leq X_{*}^{(j_q)}, \quad 1 \leq j_1 < j_2 < \dots < j_q \leq n,$$

تحت توزیع $Exp(\mu, \sigma)$ به صورت زیر به دست می‌آید:

برآورد پارامتر مقیاس

لم ۱.۳. فرض کنید

$$\mathbf{X}_{*i} = (X_{*i}^{(j_{i1})}, \dots, X_{*i}^{(j_{iq_i})}), \quad i = 1, \dots, s,$$

$$\mathbf{x}_{*i} = (x_{*i}^{(j_{i1})}, \dots, x_{*i}^{(j_{iq_i})}), \quad i = 1, \dots, s,$$

$$\mathbf{X} = (\mathbf{X}_{*1}, \dots, \mathbf{X}_{*s}),$$

$$\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_s),$$

و

$$\mathbf{K} = \{\mathbf{k} = (k_{11}, \dots, k_{sq_s}) | k_{ip} \in \{j_{i,p-1} + 1, \dots, j_{ip}\},$$

$$1 \leq p \leq q_i, 1 \leq i \leq s\},$$

$$\Psi(\mathbf{k}) = \prod_{i=1}^s \prod_{p=1}^{q_i} a_{k_{ip}}^{(j_{i,p-1})}(j_{ip}), \quad Q = \sum_{i=1}^s q_i,$$

$$V_{\mathbf{k}} = \sum_{i=1}^s \sum_{p=1}^{q_i} V_{ik_{ip}}^p, \quad V_{ik_{ip}}^p = \gamma_{ik}(x_i^{(j_{ip})} - x_i^{(j_{i,p-1})}),$$

و

$$\pi_{\Sigma}(\sigma) \propto \sigma^{-(b+1)} e^{-\frac{a}{\sigma}}, \quad \sigma > 0, a > 0, b > 0.$$

$$f_{\mathbf{X}_{*}}(\mathbf{x}) = \sigma^{-q} \left(\prod_{p=1}^q \prod_{k=j_{p-1}+1}^{j_p} \gamma_k \right) \times \prod_{p=1}^q \left(\sum_{k=j_{p-1}+1}^{j_p} a_k^{(j_{p-1})}(j_p) \exp\left(\frac{-\gamma_k}{\sigma}(x^{j_p} - x^{j_{p-1}})\right) \right),$$

که در آن، $\mathbf{x}_{*} = (x_{*}^{(j_1)}, \dots, x_{*}^{(j_q)})$ ، $\mathbf{X}_{*} = (X_{*}^{(j_1)}, \dots, X_{*}^{(j_q)})$

$$\text{و } j_0 = 0, \mu = X_{*}^{(j_0)} \leq X_{*}^{(j_1)} \leq \dots \leq X_{*}^{(j_q)}$$

$$a_k^{(j_{p-1})}(j_p) = \prod_{\substack{l=j_{p-1}+1 \\ l \neq k}}^{j_p} \frac{1}{\gamma_l - \gamma_k},$$

$$\text{که } j_{p-1} + 1 \leq k \leq j_p, 1 \leq p \leq q$$

ب) تابع چگالی شرطی $X_{*}^{j_r}$ به شرط $\mathbf{X}_{*} = \mathbf{x}_{*}$ برابر است

با:

$$f_{X_{*}^{(j_r)} | \mathbf{X}_{*}}(x^{(j_r)} | \mathbf{x}) = \frac{1}{\sigma} \left(\prod_{k=j_q+1}^{j_r} \gamma_k \right) \sum_{j=j_q+1}^{j_r} a_j^{j_q}(j_r) \exp\left(\frac{-\gamma_j}{\sigma}(x^{(j_r)} - x^{(j_q)})\right),$$

$$\text{که } x^{(j_q)} < x^{(j_r)} \text{ و } j_q < j_r$$

در این صورت چگالی پسین Σ به شرط x با فرض $Q + b > 0$,

برابر است با:

$$\pi_{\Sigma|X}(\sigma|x) = c^{-1} \sigma^{-(Q+b+1)} \sum_{k \in K} \Psi(k) e^{-\frac{V_k + a}{\sigma}},$$

که در آن، مقدار ثابت c از رابطه‌ی زیر به دست می‌آید:

$$c = \Gamma(Q + b) \sum_{k \in K} \Psi(k) (V_k + a)^{-(Q+b)}.$$

بنابراین با استفاده از فرضیات لم پیشین، برآوردگر بیز تحت

تابع زیان توان دوم خطا و k -امین گشتاور پسین σ برای

$Q + b - 1 > 0$ ، به ترتیب برابرند با:

$$\begin{aligned} \hat{\sigma} &= E(\Sigma|X = x) \\ &= c^{-1} \sum_{k \in K} \Psi(k) \frac{\Gamma(Q + b - 1)}{(V_k + a)^{Q+b-1}}, \end{aligned} \quad (۱)$$

$$\begin{aligned} \widehat{\sigma^k} &= E(\Sigma^k|X = x) \\ &= c^{-1} \sum_{k \in K} \Psi(k) \frac{\Gamma(Q + b - k)}{(V_k + a)^{Q+b-k}}. \end{aligned} \quad (۲)$$

و واریانس پسین برآورد بیز σ تحت تابع زیان توان دوم خطا به

صورت زیر محاسبه می‌شود:

$$\begin{aligned} Var(\hat{\sigma}|X = x) &= E(\Sigma^2|X = x) - E^2(\Sigma|X = x) \\ &= (\widehat{\sigma^2} - \hat{\sigma}^2). \end{aligned}$$

پیشگویی زمان شکست آینده

در نظر بگیرید

$$x_i^{(j_{i1})} \leq x_i^{(j_{i2})} \leq \dots \leq x_i^{(j_{iq_i})}, \quad i = 1, \dots, s,$$

$\sum_{i=1}^s q_i$ زمان‌های شکست مشاهده شده و $x_i^{(n_i)} \leq \dots \leq x_i^{(j_{iq_i}+1)}$ نمونه‌ی مرتب شده‌ای از مشاهدات آینده i -امین نمونه باشد.

لم ۲.۳. فرض کنید $r \in \mathbb{N}$ و $r \geq 2$ و برای $l, j \in \{1, \dots, r\}$

$\gamma_l > 0$ و اگر $l \neq j$ ، آنگاه $\gamma_l \neq \gamma_j$ و

$$a_j(r) = a_j^{(0)}(r) = \prod_{\substack{l=1 \\ l \neq j}}^r \frac{1}{\gamma_l - \gamma_j}, \quad r \geq 2, \quad a_j(1) = 1.$$

در این صورت برای $k = l_0 \in \mathbb{N}$ و $l_r = 0$ داریم:

$$\left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r \frac{1}{\gamma_{k+1}} a_j(r) = \sum_{l_1=0}^k \sum_{l_2=0}^{l_1} \dots \sum_{l_{r-1}=0}^{l_{r-2}} \left(\gamma_1^{l_{r-1}} \dots \gamma_{r-1}^{l_1-l_2} \gamma_r^{k-l_1} \right)^{-1},$$

و برای حالت خاص $k = 0, k = 1$ و $k = 2$ به ترتیب خواهیم داشت:

$$\left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r \frac{1}{\gamma_j} a_j(r) = 1, \quad (۳)$$

$$\left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r \frac{1}{\gamma_j^2} a_j(r) = \sum_{j=1}^r \frac{1}{\gamma_j}, \quad (۴)$$

$$\left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r \frac{1}{\gamma_j^3} a_j(r) = \sum_{j=1}^r \frac{1}{\gamma_j^2} + \sum_{\substack{l,k=1 \\ l < k}}^r \frac{1}{\gamma_l \gamma_k}. \quad (۵)$$

قضیه ۳.۳. اگر فرضیات لم ۱.۳ برقرار باشد، پیشگویی کننده زمان

شکست r_i امین آماره ترتیبی دنباله‌ای برابر است با:

$$\hat{X}_{*i}^{(r_i)} = x_i^{(j_{iq_i})} + \hat{\sigma} \sum_{k=j_{iq_i}+1}^{r_i} \frac{1}{\gamma_{ik}}, \quad j_{iq_i} < r_i, \quad (۶)$$

که $\hat{\sigma}$ برآوردگر بیز به دست آمده است.

فرض کنید نمونه‌ها دارای پارامترهای مقیاس متفاوت

$\sigma_1, \dots, \sigma_s$ هستند و چگالی پیشین Σ_i به صورت زیر است:

$$\pi_{\Sigma_i}(\sigma_i) \propto \sigma_i^{-(b_i+1)} \exp\left(-\frac{a_i}{\sigma_i}\right),$$

$$\pi_{\Sigma}(\sigma) = \prod_{i=1}^s \pi_{\Sigma_i}(\sigma_i), \quad \sigma = (\sigma_1, \dots, \sigma_s),$$

و علاوه بر تعاریف موجود در لم ۱.۳ داشته باشیم:

$$K_i = \{k = (k_{i1}, \dots, k_{iq_i}) | k_{ip} \in \{j_{i,p-1} + 1, \dots, j_{ip}\},$$

$$1 \leq p \leq q_i\},$$

$$V_{ik} = \sum_{p=1}^{q_i} V_{ik_{ip}}^p,$$

$$\Psi_i(k) = \prod_{p=1}^{q_i} a_{k_{ip}}^{(j_{i,p-1})}(j_{ip}),$$

$$c_i = \sum_{k \in K_i} \Psi_i(k) \frac{\Gamma(q_i + b_i)}{(V_{ik} + a_i)^{q_i+b_i}},$$

در این صورت برآوردگر بیز و k -امین گشتاور پسین σ_i برای

$q_i + b_i - k > 0$ به ترتیب برابر است با:

$$\hat{\sigma}_i = c_i^{-1} \sum_{k \in K_i} \Psi_i(k) \frac{\Gamma(q_i + b_i - 1)}{(V_{ik} + a_i)^{q_i+b_i-1}},$$

و بنابراین برآوردگر بیز σ و μ تحت تابع زیان توان دوم خطا برابر با مقادیر زیر است:

$$\hat{\sigma} = E(\Sigma | \mathbf{X} = \mathbf{x}) = \frac{1}{(Q+b-2)} \left\{ \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \left((H_{\mathbf{k}}(M_0))^{-(Q+b-2)} - (H_{\mathbf{k}}(0))^{-(Q+b-2)} \right) \right. \\ \left. / \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \left((H_{\mathbf{k}}(M_0))^{-(Q+b-1)} - (H_{\mathbf{k}}(0))^{-(Q+b-1)} \right) \right\}$$

$$\hat{\mu} = E(\Delta | \mathbf{X} = \mathbf{x}) = \left\{ M_0 \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) (H_{\mathbf{k}}(M_0))^{-(Q+b-1)} / \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \left((H_{\mathbf{k}}(M_0))^{-(Q+b-1)} - (H_{\mathbf{k}}(0))^{-(Q+b-1)} \right) \right\} \\ - \hat{\sigma} \left(c + \sum_{i=1}^s \gamma_{i1} \right)^{-1}.$$

اثبات. با توجه به روابط قبلی و این نکته که

$$I_{[\mu, \infty)}(z) I_{[0, M]}(\mu) = I_{\{\mu \leq z\}} I_{\{\mu \leq M\}} = I_{\{\mu \leq M_0\}} = I_{[0, M_0]}(\mu),$$

چگالی پسین و برآوردگرهای بیز به راحتی محاسبه می شود. $\square$

پیشگویی زمان شکست آینده

فرض کنید

$$V_{\mathbf{k}}(t, \mu) = \gamma_{ij}(t - x_i^{(j_{iq_i})}) + H_{\mathbf{k}}(\mu), \quad t \in [x_i^{(r_i)}, \infty).$$

در این صورت چگالی پسین پیشگویی کننده $X_{*i}^{(r_i)}$ به شرط $\mathbf{X}$ برابر:

$$f_{X_{*i}^{(r_i)} | \mathbf{X}}(x_i^{(r_i)} | \mathbf{x}) = c^{-1} \left(\prod_{k=j_{iq_i}+1}^{r_i} \gamma_{ik} \right) \sum_{k=j_{iq_i}+1}^{r_i} a_l^{j_{iq_i}}(r_i) \\ \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \frac{\Gamma(Q+b)}{c + \sum_{i=1}^s \gamma_{i1}} \left((V_{\mathbf{k}}(x_i^{(r_i)}, M_0))^{-(Q+b)} \right. \\ \left. - (V_{\mathbf{k}}(x_i^{(r_i)}, 0))^{-(Q+b)} \right),$$

است و از آنجایی که پیشگویی کننده $X_{*i}^{(r_i)}$ برابر با میانگین $X_{*i}^{(r_i)}$ به شرط $\mathbf{X} = \mathbf{x}$ است، با انتگرال گیری جزء به جزء مقدار زیر به دست می آید:

$$\hat{X}_{*i}^{(r_i)} = x_i^{(j_{iq_i})} + \hat{\sigma} \sum_{k=j_{iq_i}+1}^{r_i} \frac{1}{\gamma_{ik}}, \quad j_{iq_i} < r_i.$$

اگر نمونه ها دارای پارامترهای مقیاس متفاوت $\sigma_1, \dots, \sigma_s$ و چگالی پیشین توأم σ و μ برای $\sigma_1, \dots, \sigma_s > 0$ به صورت زیر باشد:

$$\pi_{(\Sigma, \Delta)}(\sigma, \mu) \propto \prod_{i=1}^s \left(\sigma_i^{-(b_i+1)} \exp \left(-\frac{a_i - c_i \mu}{\sigma_i} \right) \right),$$

$$\hat{\sigma}_i^k = c_i^{-1} \sum_{\mathbf{k} \in K_i} \Psi_i(\mathbf{k}) \frac{\Gamma(q_i + b_i - k)}{(V_{i\mathbf{k}} + a_i)^{q_i + b_i - k}}.$$

به این ترتیب با همان روش قضیه قبل ثابت می شود که پیشگویی کننده زمان شکست r_i امین آماره ترتیبی دنباله ای برابر است با:

$$\hat{X}_{*i}^{(r_i)} = x_i^{(j_{iq_i})} + \hat{\sigma}_i \sum_{k=j_{iq_i}+1}^{r_i} \frac{1}{\gamma_{ik}}, \quad j_{iq_i} < r_i.$$

۲.۳ توزیع نمایی دو پارامتری

اگر $\sigma_1 = \dots = \sigma_s = \sigma$ و مقیاس μ و σ دارای چگالی پیشین توأم زیر باشند:

$$\pi_{(\Sigma, \Delta)}(\sigma, \mu) \propto \sigma^{-(b+1)} \exp \left(-\frac{a - c\mu}{\sigma} \right) I_{[0, M]}(\mu),$$

و علاوه بر شرایط لم ۱.۳ داشته باشیم:

$$z = \min\{x_1^{(j_{11})}, \dots, x_s^{(j_{s1})}\},$$

$$M_0 = \min\{z, M\},$$

$$x_i^{(j_{i0})} = 0, \quad 1 \leq i \leq s,$$

$$V_{ik}^p = \gamma_{ik}(x_i^{(j_{ip})} - x_i^{(j_{i,p-1})}),$$

$$V_{\mathbf{k}} = \sum_{i=1}^s \sum_{p=1}^{q_i} V_{ik_{ip}}^p,$$

$$c \neq - \sum_{i=1}^s \gamma_{i1},$$

$$H_{\mathbf{k}}(t) = V_{\mathbf{k}} + a - t(c + \sum_{i=1}^s \gamma_{i1}),$$

$$H_{\mathbf{k}}(M_0) > 0, \quad \mathbf{k} \in K,$$

آنگاه چگالی پسین توأم (Σ, Δ) به شرط $\mathbf{X} = (X_{*1}, \dots, X_{*s})$ برای $Q+b-1 > 0$ به صورت زیر به دست می آید:

$$\pi_{(\Sigma, \Delta)} | \mathbf{X}((\sigma, \mu) | \mathbf{x}) = c^{-1} \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \\ \sigma^{-(Q+b-1)} \exp \left(-\frac{H_{\mathbf{k}}(\mu)}{\sigma} \right) I_{[0, M_0]}(\mu),$$

که $\mathbf{x} = (x_1, \dots, x_s)$ و مقدار ثابت c نیز برابر است با:

$$c = \frac{\Gamma(Q+b-1)}{c + \sum_{i=1}^s \gamma_{i1}} \sum_{\mathbf{k} \in K} \Psi(\mathbf{k})$$

$$\left((H_{\mathbf{k}}(M_0))^{-(Q+b-1)} - (H_{\mathbf{k}}(0))^{-(Q+b-1)} \right).$$

آنگاه با فرض این که $H_{i,k}(t) = V_{i,k} + a_i - t(c_i + \gamma_{i1})$ و $V_{i,k} =$ توزیع $Exp(\mu, \sigma_i)$ باشد، آنگاه

$$E(X_{*i}^{(j_{ik})}) = \mu + \sigma_i \sum_{j=1}^{j_{ik}} \frac{1}{\gamma_{ij}} =: \mu + \sigma_i S_1(i, k),$$

$$Var(X_{*i}^{(j_{ik})}) = \sigma_i^2 \sum_{j=1}^{j_{ik}} \frac{1}{\gamma_{ij}^2} =: \sigma_i^2 S_2(i, k),$$

$$Cov(X_{*i}^{(j_{ik})}, X_{*i}^{(j_{il})}) = Var(X_{*i}^{(j_{ik})}) =: \sigma_i^2 S_2(i, k).$$

تعاریف زیر را در نظر بگیرید.

$$\mathbf{X}_i = (X_{*i}^{(j_{i1})}, \dots, X_{*i}^{(j_{ir_i})})^t, \quad 1 \leq i \leq s,$$

$$\mathbf{X} = (\mathbf{X}_1^t, \dots, \mathbf{X}_s^t)^t,$$

$$\mathbf{1}_{r_i} = (1, \dots, 1)^t \in \mathbb{R}^{r_i}, \quad i = 1, \dots, s,$$

$$\mathbf{1} = (1, \dots, 1)^t \in \mathbb{R}^{\sum_{i=1}^s r_i},$$

$$\mathbf{X}_{iv} = \mathbf{X}_i - v \mathbf{1}_{r_i}, \quad v \in \mathbb{R},$$

$$\mathbf{X}_v = \mathbf{X} - v \mathbf{1}, \quad v \in \mathbb{R},$$

$$\boldsymbol{\alpha}_i = (S_1(i, 1), \dots, S_1(i, r_i))^t, \quad 1 \leq i \leq s,$$

$$\boldsymbol{\alpha} = (\alpha_1^t, \dots, \alpha_s^t)^t \in \mathbb{R}^{\sum_{i=1}^s r_i},$$

$$B_i = (S_2(i, k, l))_{1 \leq k, l \leq r_i}, \quad 1 \leq i \leq s,$$

$$B = \text{diag}(B_1, \dots, B_s) \in \mathbb{R}^{\sum_{i=1}^s r_i \times \sum_{i=1}^s r_i}.$$

قضیه ۲.۴. با فرض وجود نمونه سانسور شده مضاعف نوع دوم از

توزیع نمایی یک پارامتری، اگر $\sigma_1 = \dots = \sigma_s = \sigma$ ، آنگاه بهترین

برآوردگر ناریب خطی σ برابر:

$$\sigma^* = \frac{\boldsymbol{\alpha}^t B^{-1} \mathbf{X} \boldsymbol{\mu}}{\boldsymbol{\alpha}^t B^{-1} \boldsymbol{\alpha}},$$

است و بنابراین واریانس این برآوردگر نیز برابر با مقدار زیر خواهد

$$Var(\sigma^*) = \frac{\sigma^2}{\boldsymbol{\alpha}^t B^{-1} \boldsymbol{\alpha}}.$$

به طور ساده تر می توان گفت بهترین برآوردگر ناریب خطی σ

برابر است با:

$$\sigma^* = \left(\sum_{i=1}^s \sum_{k=1}^{r_i} \frac{s_1^2(i, k)}{s_2(i, k)} \right)^{-1} \sum_{i=1}^s \sum_{k=1}^{r_i} \frac{s_1(i, k)}{s_2(i, k)} (X_{*i}^{(j_{ik})} - X_{*i}^{(j_{i, k-1})}), \quad X_{*i}^{(j_{i0})} = \mu$$

که در آن،

$$s_1(i, k) = S_1(i, k) - S_1(i, k-1)$$

$$= \sum_{j=1}^{j_{ik}} \frac{1}{\gamma_{ij}} - \sum_{j=1}^{j_{i, k-1}} \frac{1}{\gamma_{ij}},$$

چگالی پسین (Σ, Δ) به شرط $\mathbf{X} = (\mathbf{X}_{*1}, \dots, \mathbf{X}_{*s})$ ، $\sum_{p=1}^{q_i} V_{ik_{ip}}^p$ برابر:

$$\pi_{(\Sigma, \Delta)|\mathbf{X}}((\boldsymbol{\sigma}, \mu)|\mathbf{x}) \propto$$

$$\sum_{\mathbf{k} \in \mathbf{K}} \Psi(\mathbf{k}) \left(\prod_{i=1}^s \sigma_i^{-(q_i + b_i + 1)} \exp\left(-\frac{H_{i, \mathbf{k}}(\mu)}{\sigma_i}\right) \right) I_{[0, M_0]}(\mu).$$

است. یا می توان گفت

$$\pi_{(\Sigma, \Delta)|\mathbf{X}}((\boldsymbol{\sigma}, \mu)|\mathbf{x}) =$$

$$c^{-1} \sum_{\mathbf{k} \in \mathbf{K}} \Psi(\mathbf{k}) \left(\prod_{i=1}^s \sigma_i^{-(q_i + b_i + 1)} \exp\left(-\frac{H_{i, \mathbf{k}}(\mu)}{\sigma_i}\right) \right) I_{[0, M_0]}(\mu),$$

که c برابر است با:

$$c = \sum_{\mathbf{k} \in \mathbf{K}} \Psi(\mathbf{k}) \int_0^{M_0} \prod_{i=1}^s \frac{\Gamma(q_i + b_i)}{(H_{i, \mathbf{k}}(\mu))^{q_i + b_i}} d\mu.$$

انتگرال بالا مقدار روشن و صریحی ندارد و بایستی به روش های

عددی محاسبه گردد و بنابراین برآوردگر بیز σ_i و μ برابر با میانگین

توزیع پسین به دست آمده هستند.

۴ بهترین برآوردگر ناریب خطی (BLUE)

در این بخش، بهترین برآورد ناریب خطی برای پارامترهای توزیع

نمایی یک و دو پارامتری با فرض پارامترهای مقیاس یکسان

به دست آمده است.

قضیه ۱.۴. فرض کنید $B \in \mathbb{R}^{n \times n}$ یک ماتریس معین مثبت

متقارن، $A \in \mathbb{R}^{n \times (k+1)}$ یک ماتریس طرح با رتبه $k+1$ و $\mathbf{X}$

و $\Delta \in \mathbb{R}^n$ دو بردار تصادفی باشند. در این صورت اگر

$$\mathbf{X} = A\boldsymbol{\beta} + \boldsymbol{\sigma},$$

یک مدل خطی با شرایط $E(\mathbf{X}) = A\boldsymbol{\beta}$ ، $Cov(\mathbf{X}) = Cov(\Delta)$

و $E(\Delta) = 0$ باشد، آنگاه بهترین برآوردگر ناریب خطی $\boldsymbol{\beta}$

برابر:

$$\hat{\boldsymbol{\beta}} = (A^t B^{-1} A)^{-1} A^t B^{-1} \mathbf{X}$$

است.

اگر $X_{*i}^{(j_{ik})}$ ، j_k -امین آماره ترتیبی دنباله ای از i -امین نمونه با

۵ برآوردگر ماکسیم درستنمایی (MLE)

هدف این قسمت به دست آوردن برآورد ماکسیم درستنمایی برای نمونه سانسور شده مضاعف نوع دوم است، به این ترتیب طبق محاسبه این برآوردگر را برای حالت ساده آماره های ترتیبی دنباله ای، یعنی $\alpha = 1$ (آماره های ترتیبی معمولی) بررسی می کنیم. فرض کنید نمونه ی تصادفی سانسور شده مضاعف نوع دوم به صورت زیر باشد:

$$Y_1 = X^{(j_1)}, Y_2 = X^{(j_2)}, \dots, Y_q = X^{(j_q)}.$$

در این صورت، با در نظر گرفتن توزیع نمایی یک پارامتری با تابع چگالی و تابع توزیع

$$f(x) = \frac{1}{\sigma} e^{-x/\sigma},$$

$$F(x) = 1 - e^{-x/\sigma},$$

تابع درستنمایی بر اساس نمونه سانسور شده مضاعف، به صورت زیر است:

$$L = \text{const} \times \sigma^{-q} \left[1 - e^{-\frac{y_1}{\sigma}} \right]^{j_1-1} e^{\frac{\sum_{i=1}^q y_i + (n-j_q)y_q}{\sigma}} \times \prod_{i=1}^{q-1} \left[e^{-\frac{y_i}{\sigma}} - e^{-\frac{y_{i+1}}{\sigma}} \right]^{j_{i+1}-j_i-1}.$$

با برابر صفر قرار دادن مشتق لگاریتم این تابع براساس σ داریم:

$$q\sigma + (j_1 - 1)y_1 \frac{e^{-\frac{y_1}{\sigma}}}{[1 - e^{-\frac{y_1}{\sigma}}]} - \sum_{i=1}^q y_i - (n - j_q)y_q - \sum_{i=1}^{q-1} [j_{i+1} - j_i - 1] \left\{ \frac{-y_{i+1}e^{-\frac{y_{i+1}}{\sigma}} + y_i e^{-\frac{y_i}{\sigma}}}{-e^{-\frac{y_{i+1}}{\sigma}} + e^{-\frac{y_i}{\sigma}}} \right\} = 0.$$

با توجه به این که این معادله، معادله یکنوا بر حسب σ است، برآورد ماکسیم درستنمایی یکتایی را نتیجه می دهد اما به دلیل این که جواب صریحی از این معادله به دست نمی آید با تقریبی بر روی این تابع برآورد ماکسیم درستنمایی تقریبی را محاسبه می کنیم.

برای تقریب زدن از روش بالاکریشن که در مرجع [۲] معرفی شده است، استفاده می کنیم:

$$\frac{-y_{i+1}e^{-\frac{y_{i+1}}{\sigma}} + y_i e^{-\frac{y_i}{\sigma}}}{-e^{-\frac{y_{i+1}}{\sigma}} + e^{-\frac{y_i}{\sigma}}} \approx \gamma_i + \delta_i \frac{y_i}{\sigma} + (1 - \delta_i) \frac{y_{i+1}}{\sigma},$$

که در آن،

$$\begin{aligned} p_i^* &= \frac{j_i}{n+1}, & q_i^* &= 1 - p_i^*, \\ \delta_i &= \frac{q_i^*}{q_i^* - q_{i+1}^*} - \frac{q_i^* q_{i+1}^*}{(q_i^* - q_{i+1}^*)^2} \log\left(\frac{q_i^*}{q_{i+1}^*}\right), \\ \gamma_i &= \frac{q_{i+1}^* \log q_{i+1}^* - q_i^* \log q_i^*}{q_i^* - q_{i+1}^*} + \delta_i \log q_i^* + (1 - \delta_i) \log q_{i+1}^*, \end{aligned}$$

$$\begin{aligned} s_2(i, k) &= S_2(i, k) - S_2(i, k-1) \\ &= \sum_{j=j_{i,k-1}+1}^{j_{i,k}} \frac{1}{\gamma_{ij}^2}, \\ j_{i0} &= 0, i = 1, \dots, s \text{ و } S_1(i, 0) = S_2(i, 0) = 0 \end{aligned}$$

قضیه ۳.۴. با فرض وجود نمونه سانسور شده مضاعف نوع دوم از توزیع نمایی دو پارامتری، اگر $\sigma_1 = \dots = \sigma_s = \sigma$ ، آنگاه بهترین برآوردگر نارایب خطی μ و σ به ترتیب برابر

$$\mu^* = \left(\frac{\alpha^t B^{-1} \alpha 1^t B^{-1} - \alpha^t B^{-1} 1 \alpha^t B^{-1}}{(\alpha^t B^{-1} \alpha)(1^t B^{-1} 1) - (\alpha^t B^{-1} 1)^2} \right) X,$$

$$\sigma^* = \left(\frac{1^t B^{-1} 1 \alpha^t B^{-1} - 1^t B^{-1} \alpha 1^t B^{-1}}{(\alpha^t B^{-1} \alpha)(1^t B^{-1} 1) - (\alpha^t B^{-1} 1)^2} \right) X.$$

هستند که واریانس و کواریانس این دو برآوردگر برابر با مقادیر زیر است:

$$\text{Var}(\mu^*) = \frac{\sigma^2(\alpha^t B^{-1} \alpha)}{(\alpha^t B^{-1} \alpha)(1^t B^{-1} 1) - (\alpha^t B^{-1} 1)^2},$$

$$\text{Var}(\sigma^*) = \frac{\sigma^2(1^t B^{-1} 1)}{(\alpha^t B^{-1} \alpha)(1^t B^{-1} 1) - (\alpha^t B^{-1} 1)^2},$$

$$\text{Cov}(\mu^*, \sigma^*) = \frac{-\sigma^2(\alpha^t B^{-1} 1)}{(\alpha^t B^{-1} \alpha)(1^t B^{-1} 1) - (\alpha^t B^{-1} 1)^2}.$$

بهترین برآوردگر نارایب خطی σ و μ به صورت ساده تر برابر:

$$\begin{aligned} \sigma^* &= \frac{1}{K} \left[\left(\sum_{i=1}^s \frac{1}{s_2(i, 1)} \right) \left(\sum_{i=1}^s \sum_{k=1}^{r_i} \frac{s_1(i, k)}{s_2(i, k)} \right. \right. \\ &\quad \left. \left. (X_{*i}^{(j_i, k)} - X_{*i}^{(j_i, k-1)}) \right) - \left(\sum_{i=1}^s \frac{s_1(i, 1)}{s_2(i, 1)} \right) \left(\sum_{i=1}^s \frac{1}{s_2(i, 1)} X_{*i}^{(j_1)} \right) \right], \\ \mu^* &= \frac{1}{\sum_{i=1}^s \frac{1}{s_2(i, 1)}} \left(\sum_{i=1}^s \frac{1}{s_2(i, 1)} X_{*i}^{(j_1)} - \sum_{i=1}^s \frac{s_1(i, 1)}{s_2(i, 1)} \sigma^* \right), \end{aligned}$$

است که در آن،

$$\begin{aligned} K &= \left(\sum_{i=1}^s \sum_{k=1}^{r_i} \frac{s_1^2(i, k)}{s_2(i, k)} \right) \left(\sum_{i=1}^s \frac{1}{s_2(i, 1)} \right) \\ &\quad - \left(\sum_{i=1}^s \frac{s_1(i, 1)}{s_2(i, 1)} \right)^2 > 0. \end{aligned}$$

بنابراین برای واریانس و کواریانس این دو برآوردگر نیز مقادیر

زیر به دست می آید:

$$\begin{aligned} \text{Var}(\sigma^*) &= \frac{\sigma^2 \sum_{i=1}^s \frac{1}{s_2(i, 1)}}{K}, \\ \text{Var}(\mu^*) &= \frac{\sigma^2 \sum_{i=1}^s \sum_{k=1}^{r_i} \frac{s_1^2(i, k)}{s_2(i, k)}}{K}, \\ \text{Cov}(\mu^*, \sigma^*) &= -\sigma^2 \frac{\sum_{i=1}^s \frac{s_1(i, 1)}{s_2(i, 1)}}{K}. \end{aligned}$$

و بنابراین برآورد ماکسیمم درستنمایی تقریبی σ برابر خواهد شد خواهد بود.

۶ مثال‌های کاربردی

مثال اول

نمونه شبیه‌سازی شده در این مثال که در مرجع [۳] آمده است، مربوط به زمان‌های شکست مشاهده شده از سیستمی است که طول عمر مؤلفه‌های آن دارای توزیع نمایی یک پارامتری با $\sigma = 20$ هستند. در این نمونه 10 زمان شکست به صورت مضاعف سانسور شده‌اند. همچنین بایستی یادآور شد که در این سیستم $\alpha = 1$ در نظر گرفته شده است؛ در واقع می‌توان گفت بعد از اولین شکست در این سیستم، توزیع طول عمرهای باقی‌مانده تغییر نمی‌کند و بنابراین مدلی که پیش‌رو داریم مدلی از آماره‌های ترتیبی معمولی است. نمونه مورد نظر به‌قرار زیر است:

$$x_1 = 0.961, x_2 = 0.990, x_3 = 1.565, x_4 = 2.031$$

$$x_5 = 2.204, x_6 = 2.340, x_7 = 3.642, x_8 = 6.008$$

$$x_9 = 6.538, x_{10} = 7.145, x_{11} = -, x_{12} = -,$$

$$x_{13} = -, x_{14} = 11.937, x_{15} = 15.433, x_{16} = 18.234,$$

$$x_{17} = 18.307, x_{18} = 22.096, x_{19} = -, x_{20} = -,$$

$$x_{21} = -, x_{22} = 28.799, x_{23} = 30.692, x_{24} = 30.737,$$

$$x_{25} = 33.702, x_{26} = 34.245, x_{27} = -, x_{28} = -,$$

$$x_{29} = -, x_{30} = -,$$

بهترین برآوردگر نارایب خطی σ برابر با $\sigma^* = 19.9426$ و برآوردگر ماکسیمم درستنمایی تقریبی σ نیز برابر با مقدار $\bar{\sigma} = 19.9712$ به دست می‌آید. نتایج محاسبه برآوردگر بیز و همچنین خطای استاندارد پسین برای a و b های متفاوت در جدول (۱) آمده است که $\sqrt{(\hat{\sigma}^2 - \sigma^2)}$ $se(\hat{\sigma}) = \sqrt{Var(\hat{\sigma} | \mathbf{X} = \mathbf{x})} = \sqrt{(\hat{\sigma}^2 - \sigma^2)}$ همان‌گونه که مشاهده می‌شود برای حالت $a = 1$ و $b = 1$ نزدیکترین مقدار را به مقدار واقعی دارد اما خطای استاندارد پسین در حالت $a = 0$ و $b = 2$ کمترین مقدار را می‌گیرد.

$$\bar{\sigma} = \left\{ \sum_{i=1}^q y_i + \sum_{i=0}^{q-1} [j_{i+1} - j_i - 1] \delta_i y_i + (1 - \delta_i) y_{i+1} + (n - j_q) y_q \right\} / \left\{ q - \sum_{i=0}^{q-1} [j_{i+1} - j_i - 1] \gamma_i \right\}.$$

با در نظر گرفتن توزیع نمایی دو پارامتری با تابع چگالی و تابع توزیع

$$f(x) = \frac{1}{\sigma} e^{-\frac{(x-\mu)}{\sigma}}, \quad x \geq \mu, \sigma > 0,$$

$$F(x) = 1 - e^{-\frac{(x-\mu)}{\sigma}}, \quad x \geq \mu, \sigma > 0,$$

تابع درستنمایی بر اساس نمونه سانسور شده مضاعف، به صورت زیر است:

$$L = const \times \sigma^{-q} [1 - e^{-\frac{y_1 - \mu}{\sigma}}]^{j_1 - 1} e^{\frac{\mu}{\sigma} (n - j_1 + 1)} e^{-\frac{[\sum_{i=1}^q y_i + (n - j_q) y_q]}{\sigma}} \times \prod_{i=1}^{q-1} \left[e^{-\frac{y_i}{\sigma}} - e^{-\frac{y_{i+1}}{\sigma}} \right]^{j_{i+1} - j_i - 1}.$$

مشخص است که اگر $j_1 = 1$ ، تابع بالا تابعی یکنوا در μ است و بنابراین برآورد ماکسیمم درستنمایی μ برابر:

$$\bar{\mu} = y_1 = x^{(j_1)} = x^{(1)},$$

است و اگر $j_1 > 1$ ، طبق معادله درستنمایی زیر

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma} \left[(n - j_1 + 1) - (j_1 - 1) \frac{e^{-\frac{(y_1 - \mu)}{\sigma}}}{1 - e^{-\frac{(y_1 - \mu)}{\sigma}}} \right] = 0,$$

برآورد ماکسیمم درستنمایی μ برابر:

$$\bar{\mu} = y_1 + \bar{\sigma} \log \left[\frac{n - j_1 + 1}{n} \right].$$

خواهد بود. معادله درستنمایی برای σ نیز به صورت زیر است:

$$\frac{\partial \log L}{\partial \sigma} = \frac{-1}{\sigma^2} \left[q\sigma + (n - j_1 + 1)y_1 - \left\{ \sum_{i=1}^q y_i + (n - j_q)y_q \right\} - \sum_{i=1}^{q-1} (j_{i+1} - j_i - 1) \left\{ \frac{y_i e^{-\frac{y_i}{\sigma}} - y_{i+1} e^{-\frac{y_{i+1}}{\sigma}}}{e^{-\frac{y_i}{\sigma}} - e^{-\frac{y_{i+1}}{\sigma}}} \right\} \right] = 0.$$

مانند بخش قبل با توجه به این که این معادله، معادله‌ای یکنوا بر اساس σ است، برآورد ماکسیمم درستنمایی یکتایی را نتیجه می‌دهد اما به دلیل این که جواب صریحی از این معادله به دست نمی‌آید، با به کارگیری تقریب بالاکریشن بر روی این معادله، برآورد ماکسیمم درستنمایی تقریبی را محاسبه می‌کنیم و به این ترتیب برآورد ماکسیمم درستنمایی σ برابر:

$$\bar{\sigma} = \left\{ \sum_{i=1}^q y_i + (n - j_q)y_q + \sum_{i=1}^{q-1} (j_{i+1} - j_i - 1)(\delta_i y_i + (1 - \delta_i)y_{i+1}) - (n - j_1 + 1)y_1 \right\} / \left\{ q - \sum_{i=1}^{q-1} (j_{i+1} - j_i - 1)\gamma_i \right\}.$$

مثال دوم

۷ شبیه سازی

این مثال شامل پنج نمونه شبیه سازی شده است که این نمونه ها مربوط به زمان های شکست مشاهده شده از یک سیستم 2 از 5 دنباله ای هستند و طول عمر مؤلفه های این سیستم دارای توزیع نمایی یک پارامتری با $\sigma = 100$ است.

هر نمونه شامل پنج مشاهده است که برخی از این مشاهدات سانسور شده اند و به صورت کلی در هر نمونه زمان شکست پنجم مشاهده نمی شود زیرا سیستم بعد از شکست چهارم خود به خود از کار می افتد.

همچنین در این سیستم i -امین شکست بر توزیع طول عمر $n - i$ مؤلفه باقی مانده تأثیر می گذارد، به این ترتیب که مجموعه پارامترهای α برابر مقادیر زیر است:

$$\alpha_1 = 1, \alpha_2 = 1.2, \alpha_3 = 1.4, \alpha_4 = 1.6, \alpha_5 = 1.8.$$

پیوست

اثبات قضیه ها

قضیه ۲.۲

در این مثال ابتدا مقادیر $\hat{\sigma}$ و $se(\hat{\sigma})$ را محاسبه و سپس با توجه به این که در نمونه آخر تنها سه زمان شکست مشاهده شده است، زمان شکست چهارم پیشگویی می شود. نمونه مورد نظر به قرار زیر است:

اثبات. با توجه به قضیه ۱.۲، برای $1 \leq l \leq r$ داریم:

$$X_*^{(l)} = \mu + \sum_{k=1}^l \frac{\sigma}{\gamma_k} W_*^{(k)} = \mu + \sum_{k=1}^l U_*^{(k)}.$$

و با در نظر گرفتن $q = 1$ داریم:

$$\begin{aligned} f_{X_*^{(j_1)}}(x^{j_1}) &= f_{\mu + \sum_{k=1}^{j_1} U_*^{(k)}}(x^{(j_1)}) \\ &= f_{\sum_{k=1}^{j_1} U_*^{(k)}}(x^{(j_1)} - \mu), \end{aligned}$$

بنابراین تابع چگالی توأم آماره های ترتیبی دنباله ای سانسور شده مضاعف نوع دوم تحت توزیع نمایی به صورت زیر به دست می آید:

$$f_{\mathbf{X}_*}(\mathbf{x}) = \prod_{p=1}^q f_{\sum_{k=j_{p-1}}^{j_p} U_*^{(k)}}(x^{(j_p)} - x^{(j_{p-1})}).$$

از طرف دیگر می دانیم مجموع وزنی از متغیرهای مستقل و هم توزیع نمایی استاندارد به صورت ارلنگ تعمیم یافته توزیع شده اند، بنابراین تابع توزیع و تابع چگالی نتیجه محاسبه ها در جدول (۲) آمده است که با توجه به نتایج این جدول مشاهده می شود که در حالت $a = 0$ و $b = 2$ ، مقدار $\hat{\sigma}$ از دقت بیشتری برخوردار است.

$$x_1^{(1)} = -, x_1^{(2)} = -, x_1^{(3)} = 15.085,$$

$$x_1^{(4)} = 65.409,$$

$$x_2^{(1)} = 36.982, x_2^{(2)} = 48.055, x_2^{(3)} = -,$$

$$x_2^{(4)} = 197.405,$$

$$x_3^{(1)} = 29.657, x_3^{(2)} = 56.983,$$

$$x_3^{(3)} = 59.664, x_3^{(4)} = 75.441,$$

$$x_4^{(1)} = 4.738, x_4^{(2)} = 33.625,$$

$$x_4^{(3)} = 48.066, x_4^{(4)} = 89.756$$

$$x_5^{(1)} = 19.128, x_5^{(2)} = 58.242, x_5^{(3)} = 73.406.$$

بنابراین داریم:

$$\begin{aligned} c &= \sum_{\mathbf{k} \in K} \left(\prod_{i=1}^s \prod_{p=1}^{q_i} a_{k_{ip}}^{(j_{i,p-1})} (j_{ip}) \right) \\ &\times \int_0^\infty \sigma^{-(Q+b+1)} e^{-\frac{a}{\sigma}} e^{-\frac{\sum_{i=1}^s \sum_{p=1}^{q_i} V_{ik_{ip}}^p}{\sigma}} d\sigma \\ &= \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \int_0^\infty \sigma^{-(Q+b+1)} e^{-\frac{V_{\mathbf{k}}+a}{\sigma}} d\sigma \\ &= \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) \frac{\Gamma(Q+b)}{(V_{\mathbf{k}}+a)^{Q+b}}. \end{aligned}$$

سرانجام چگالی پسین Σ به شرط $\mathbf{X}$ برابر:

$$\pi_{\Sigma|\mathbf{X}}(\sigma|\mathbf{x}) = c^{-1} \sigma^{-(Q+b+1)} \sum_{\mathbf{k} \in K} \Psi(\mathbf{k}) e^{-\frac{V_{\mathbf{k}}+a}{\sigma}}.$$

□

است.

لم ۳.۲

اثبات. فرض کنید $X_1, \dots, X_r \stackrel{iid}{\sim} \text{Exp}(0, 1)$ در این صورت $Y = \sum_{j=1}^r Y_j = \sum_{j=1}^r \frac{1}{\gamma_j} X_j$ دارای توزیع ارلنگ تعمیم یافته با چگالی زیر است:

$$f_Y(t) = \left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r a_j(r) e^{-\gamma_j t}, \quad t \geq 0.$$

با توجه به چگالی بالا، k -مین گشتاور Y برابر:

$$E(Y^k) = \left(\prod_{j=1}^r \gamma_j \right) \sum_{j=1}^r a_j(r) \frac{\Gamma(k+1)}{\gamma_j^{k+1}} \quad (V)$$

است. اکنون تنها کافی است نشان دهیم:

$$E(Y^k) = k! \sum_{l_1=0}^k \sum_{l_2=0}^{l_1} \dots \sum_{l_{r-1}=0}^{l_{r-2}} \left(\gamma_1^{l_{r-1}} \dots \gamma_{r-1}^{l_1-l_2} \gamma_r^{k-l_1} \right)^{-1}. \quad (\Lambda)$$

برای اثبات از استقراء استفاده می کنیم. با توجه به این نکته که $E(Y_i^k) = k!/\gamma_i^k$ و این که Y_1 و Y_2 مستقل هستند، برای $r=2$ داریم:

$$E((Y_1 + Y_2)^k) = k! \sum_{l_1=0}^k (\gamma_1^{l_1} \gamma_2^{k-l_1})^{-1}.$$

اکنون بنابر اصل استقراء با فرض برقراری رابطه ی (۸) برای نمونه ی r -تایی، اثبات می شود که نمونه ی $r+1$ -تایی نیز در این رابطه صدق می کند:

$$\begin{aligned} E(Y^k) &= E\left(\sum_{j=1}^{r+1} Y_j\right)^k = E\left(\sum_{j=1}^r Y_j + Y_{r+1}\right)^k \\ &= k! \sum_{l_0=0}^k \sum_{l_1=0}^{l_0} \dots \sum_{l_{r-1}=0}^{l_{r-2}} \left(\gamma_1^{l_{r-1}} \dots \gamma_{r-1}^{l_1-l_2} \gamma_r^{l_0-l_1} \gamma_{r+1}^{k-l_0} \right)^{-1}. \end{aligned}$$

$$\sum_{k=j_{p-1}+1}^{j_p} U_*^{(k)} = \sum_{k=j_{p-1}+1}^{j_p} \frac{\sigma}{\gamma_k} W_*^{(k)}$$

$$\begin{aligned} F \sum_{k=j_{p-1}+1}^{j_p} \frac{\sigma}{\gamma_k} W_*^{(k)}(t) &= 1 - \left(\prod_{k=j_{p-1}+1}^{j_p} \gamma_k \right) \\ &\sum_{k=j_{p-1}+1}^{j_p} \frac{1}{\gamma_k} a_k^{(j_{p-1})}(j_p) e^{-\frac{\gamma_k}{\sigma} t}, \quad t \geq 0 \end{aligned}$$

۹

$$\begin{aligned} f \sum_{k=j_{p-1}+1}^{j_p} \frac{\sigma}{\gamma_k} W_*^{(k)}(t) &= \left(\prod_{k=j_{p-1}+1}^{j_p} \gamma_k \right) \sigma^{-1} \\ &\sum_{k=j_{p-1}+1}^{j_p} a_k^{(j_{p-1})}(j_p) e^{-\frac{\gamma_k}{\sigma} t}, \quad t \geq 0, \end{aligned}$$

هستند. بنابراین با قرار دادن این رابطه در معادله ی قبلی داریم:

$$\begin{aligned} f_{\mathbf{X}_*}(\mathbf{x}) \prod_{p=1}^q f \sum_{k=j_{p-1}+1}^{j_p} \frac{\sigma}{\gamma_k} W_*^{(k)}(x^{(j_p)} - x^{(j_{p-1})}) &= \\ \sigma^{-q} \left(\prod_{p=1}^q \prod_{k=j_{p-1}+1}^{j_p} \gamma_k \right) \prod_{p=1}^q \left(\sum_{k=j_{p-1}+1}^{j_p} a_k^{(j_{p-1})}(j_p) \right. & \\ \left. \exp\left(-\frac{\gamma_k}{\sigma}(x^{(j_p)} - x^{(j_{p-1})})\right) \right). \end{aligned}$$

لم ۳.۱

اثبات. می دانیم که

$$\begin{aligned} \pi_{\Sigma|\mathbf{X}}(\sigma|\mathbf{x}) &= \frac{\prod_{i=1}^s f_{\mathbf{X}_{*i}|\Sigma}(\mathbf{x}_i|\sigma) \times \pi_{\Sigma}(\sigma)}{\int_0^\infty \prod_{i=1}^s f_{\mathbf{X}_{*i}|\Sigma}(\mathbf{x}_i|\sigma) \times \pi_{\Sigma}(\sigma) d\sigma} \\ &= c^{-1} \sigma^{-(Q+b+1)} e^{-\frac{a}{\sigma}} \prod_{i=1}^s \prod_{p=1}^{q_i} \\ &\times \sum_{k=j_{i,p-1}+1}^{j_{ip}} a_k^{(j_{i,p-1})}(j_{ip}) e^{-\frac{V_{ik}^p}{\sigma}}, \end{aligned}$$

که در آن،

$$\begin{aligned} c &= \int_0^\infty \sigma^{-(Q+b+1)} e^{-\frac{a}{\sigma}} \\ &\prod_{i=1}^s \prod_{p=1}^{q_i} \sum_{k=j_{i,p-1}+1}^{j_{ip}} a_k^{(j_{i,p-1})}(j_{ip}) e^{-\frac{V_{ik}^p}{\sigma}} d\sigma. \end{aligned}$$

اما از آنجایی که

$$\prod_{i=1}^s \prod_{p=1}^{q_i} \sum_{k=1}^{z_{ip}} x_{ipk} = \sum_{\mathbf{k}=(k_{11}, \dots, k_{sq_s}), \substack{k_{\nu\varphi} \in (1, \dots, z_{\nu\varphi}), \\ 1 \leq \varphi \leq q_\nu, 1 \leq \nu \leq s}} \prod_{i=1}^s \prod_{p=1}^{q_i} x_{ipk_{ip}},$$

اکنون با انتقال اندیس‌های $l_0 \rightarrow l_1$ و $l_{r-1} \rightarrow l_r$ استقراء کامل و رابطه‌ی (۸) اثبات می‌شود. اینک با برابر قرار دادن (۷) و (۸) و آن بر طبق دو معادله (۳) و (۴)، پیش‌گویی‌کننده $\hat{X}_{*i}^{(r_i)}$ برابر تقسیم دو طرف بر $k!$ اثبات کامل می‌شود. □

$$\hat{X}_{*i}^{(r_i)} = x_i^{(j_{iq_i})} + \hat{\sigma} \sum_{k=j_{iq_i}+1}^{r_i} \frac{1}{\gamma_{ik}}, \quad j_{iq_i} < r_i,$$

□

است.

قضیه ۳.۳

اثبات. به دلیل استقلال شرطی $X_{*1}, \dots, X_{*s}$ به شرط $\Sigma = \sigma$ داریم:

$$f_{X_{*i}^{(r_i)} | \mathbf{X}, \Sigma}(t | (\mathbf{x}, \sigma)) = f_{X_{*i}^{(r_i)} | \mathbf{X}_{*i}, \Sigma}(t | (\mathbf{x}_i, \sigma)),$$

که $\sigma > 0$ و $t > x_i^{(j_{iq_i})}$

بنابراین تابع چگالی پیش‌گویی‌کننده $X_{*i}^{(r_i)}$ به شرط $\mathbf{X} = \mathbf{x}$ به صورت زیر به دست می‌آید:

$$\begin{aligned} & f_{X_{*i}^{(r_i)} | \mathbf{X}}(x_i^{(r_i)} | \mathbf{x}) \\ &= \int_0^\infty f_{(X_{*i}^{(r_i)}, \Sigma) | \mathbf{X}}((x_i^{(r_i)}, \sigma) | \mathbf{x}) d\sigma \\ &= \int_0^\infty f_{X_{*i}^{(r_i)} | (\mathbf{X}_{*i}, \Sigma)}(x_i^{(r_i)} | \mathbf{x}_i, \sigma) \\ & \quad \times \pi_{\Sigma | \mathbf{X}}(\sigma | \mathbf{x}) d\sigma \\ &= c^{-1} \left(\prod_{k=j_{iq_i}+1}^{r_i} \gamma_{ik} \right) \left(\sum_{k=j_{iq_i}+1}^{r_i} a_k^{(j_{iq_i})}(r_i) \right. \\ & \quad \left. \sum_{\mathbf{k} \in \mathbf{K}} \Psi(\mathbf{k}) \frac{\Gamma(Q+b+1)}{(V_{ik}^{(k)}(x_i^{(r_i)}))^{Q+b+1}} \right), \end{aligned}$$

که در آن،

$V_{ik}^{(k)}(x_i^{(r_i)}) = \gamma_{ik}(x_i^{(r_i)} - x_i^{(j_{iq_i})}) + V_k + a$ پیش‌گویی‌کننده $\hat{X}_{*i}^{(r_i)}$ از $X_{*i}^{(r_i)}$ برابر با میانگین $X_{*i}^{(r_i)}$ به شرط $\mathbf{X} = \mathbf{x}$ است و بنابراین

$$\begin{aligned} \hat{X}_{*i}^{(r_i)} &= E(X_{*i}^{(r_i)} | \mathbf{X} = \mathbf{x}) = \int_{x_i^{(j_{iq_i})}}^\infty t f_{X_{*i}^{(r_i)} | \mathbf{X}}(t | \mathbf{x}) dt = \\ & c^{-1} \left(\prod_{k=j_{iq_i}+1}^{r_i} \gamma_{ik} \right) \left(\sum_{k=j_{iq_i}+1}^{r_i} a_k^{(j_{iq_i})}(r_i) \right. \\ & \quad \left. \sum_{\mathbf{k} \in \mathbf{K}} \Psi(\mathbf{k}) \int_{x_i^{(j_{iq_i})}}^\infty t \frac{\Gamma(Q+b+1)}{(V_{ik}^{(k)}(t))^{Q+b+1}} dt \right), \end{aligned}$$

و با استفاده از انتگرال‌گیری جزء‌به‌جزء داریم:

$$\begin{aligned} & \int_{x_i^{(j_{iq_i})}}^\infty t (V_{ik}^{(k)}(t))^{-(Q+b+1)} dt = \\ & \frac{(V_k + a)^{-(Q+b)}}{(Q+b)} \left[\frac{x_i^{(j_{iq_i})}}{\gamma_{ik}} + \frac{V_k + a}{\gamma_{ik}^2(Q+b-1)} \right]. \end{aligned} \quad (9)$$

قضیه ۴.۲

اثبات. با در نظر گرفتن مدل خطی

$$\mathbf{X}_\mu = \alpha \sigma + \Delta,$$

با شرایط زیر

$$E(\mathbf{X}_\mu) = \sigma \alpha,$$

$$\text{Cov}(\mathbf{X}_\mu) = \text{Cov}(\Delta) = \sigma^2 B,$$

$$k=0, n = \sum_{i=1}^s r_i,$$

که ماتریس طرح، σ پارامتر مورد نظر و Δ بردار تصادفی است

و جایگذاری این مقادیر در قضیه ۱.۴ اثبات کامل می‌شود. □

قضیه ۴.۳

اثبات. اگر مدل خطی

$$\mathbf{X} = (\alpha, 1) \begin{pmatrix} \sigma \\ \mu \end{pmatrix} + \Delta,$$

با شرایط

$$E(\mathbf{X}) = (\alpha, 1) \begin{pmatrix} \sigma \\ \mu \end{pmatrix},$$

$$\text{Cov}(\mathbf{X}) = \text{Cov}() = \sigma^2 B,$$

$$k=1, n = \sum_{i=1}^s r_i,$$

که $(\alpha, 1)$ ماتریس طرح، بردار پارامتر مورد نظر و بردار تصادفی است را در نظر بگیریم آنگاه با جایگذاری این مقادیر در قضیه ۱.۴ نتیجه دلخواه به دست می‌آید. □

جدول ۱: نتایج برآورد بیزی پارامتر مدل با استفاده از چگالی پیشین گامای وارون برای a و b -های متفاوت

a	b	$\hat{\sigma}$	$se(\hat{\sigma})$
0	-1	21.6139	4.5067
1	0	20.7881	4.2466
0	0	20.7485	4.2368
1	1	19.9873	4.0007
0	1	19.9488	3.9933
1	2	19.2458	3.7776
0	2	19.2088	3.7717

جدول ۲: نتایج برآورد بیزی پارامتر مدل و پیش‌گویی زمان شکست چهارم از نمونه پنجم با استفاده از چگالی پیشین گامای وارون برای a و b -های متفاوت

a	b	$\hat{\sigma}$	$se(\hat{\sigma})$	$\hat{X}_{*5}^{(4)}$
0	-1	120.6417	30.2652	111.1065
1	0	113.9492	27.7374	109.0151
0	0	113.8932	27.7239	108.9976
1	1	107.9084	25.5312	107.1274
0	1	107.8554	25.5188	107.1108
1	2	102.4721	23.6022	105.4285
0	2	102.4217	23.5908	105.4128

جدول ۳: نتایج برآورد بیزی پارامتر مدل با استفاده از چگالی پیشین گامای وارون برای a و b های متفاوت

$k = 150$	a	b	$\bar{\sigma}$	$\bar{se}(\hat{\sigma})$
	1	0	31.5958	6.4565
	0	0	31.5557	6.4484
	1	1	30.3779	6.0825
	0	1	30.3393	6.0748
	1	2	29.2502	5.74326
	0	2	28.9091	5.7359
$k = 1000$	a	b	$\bar{\sigma}$	$\bar{se}(\hat{\sigma})$
	1	0	31.2674	6.3894
	0	0	31.2273	6.3812
	1	1	30.0622	6.0193
	0	1	30.0236	6.0116
	1	2	28.9462	5.6835
	0	2	28.9091	5.6762
$k = 1500$	a	b	$\bar{\sigma}$	$\bar{se}(\hat{\sigma})$
	1	0	31.3598	6.4081
	0	0	31.3197	6.3999
	1	1	30.1511	6.0369
	0	1	30.1125	6.0292
	1	2	29.0319	5.7001
	0	2	28.9948	5.6929

مراجع

- [1] Ahmadi, J. and Doostparast, M. (2008). Statistical inference based on k -records. *Mashhad Research Journal of Mathematical Sciences*. **1(1)**, 67–82.
- [2] Balakrishnan, N. (1990). On the maximum likelihood estimation of the location and scale parameters of exponential distribution based on multiply Type-II censored samples. *Applied Statistics*. **17**, 55–61.
- [3] Balasubramanian, K. and Balakrishnan, N. (1992). Estimation for one-and two-parameter exponential distributions under multiple Type-II censoring. *Statistische Hefte*. **33**, 203-216.
- [4] Jaheen, Z.F. (2005). Estimation based on generalized order statistics from the Burr model. *Communications in Statistics - Theory and Methods*. **34**, 785–794.
- [5] Kamps, U. (1995a). A concept of generalized order statistics. *Statistical Planning and Inference*, **48**, 1-23.
- [6] Kamps, U. (1995b). *A Concept of Generalized Order Statistics*. Teubner. Stuttgart.
- [7] Schenk, N., Burkschat, M., Cramer, E. and Kamps, U. (2011). Bayesian estimation and prediction with multiply Type-II censored samples of sequential order statistics from one- and two-parameter exponential distributions. *Statistical Planning and Inference*. **141**, 1575–1587.
- [8] Sultan, K.S. (2008). Bayesian estimates based on record values from the inverse Weibull life time model. *Quality Technology and Quantitative Management*. **5**, 363–374.