آزمون فرضیه و بازهٔ اطمینان دقیق برای میانگین توزیع نمایی تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول

حسین نادب[۱] و حمزه ترابی[۲]

تاریخ دریافت: ۱۳۹۴/۸/۷
تاریخ پذیرش: ۱۳۹۵/۸/۳۰

چکیده:
نمونه‌های سانسور شده در آزمایش‌های مربوط به آزمون‌های طول عمر مطرح می‌شوند؛ یعنی هنگامی که آزمایشگر، زمان‌های از کار افتادگی تمام واحدهای موجود در آزمون طول عمر را مشاهده نمی‌کند. در سال‌های اخیر، استنباط بر پایهٔ نمونه‌های سانسور شده بسیار مورد توجه قرار گرفته است، به‌طوری که در مورد پارامترهای توزیع‌های مختلفی مانند نرمال، نمایی، گاما، رایلی، وایبول، لگ‌نرمال، گوسی معکوس، لوژستیک، لاپلاس و پارتو بر اساس نمونه‌های سانسور شده استنباط صورت گرفته است.

در این مقاله، روشی برای انجام آزمون فرضیه و یافتن بازهٔ اطمینان دقیق برای میانگین توزیع نمایی تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول[۳] پیشنهاد می‌شود. سپس با استفاده از شبیه‌سازی، عملکرد بازهٔ اطمینان پیشنهادی مورد ارزیابی قرار می‌گیرد. سرانجام روش‌های پیشنهادی، روی یک مجموعه از داده‌های واقعی اجرا می‌شود.

واژه‌های کلیدی: توزیع نمایی، سانسور دورگهٔ پیش‌روندهٔ نوع اول، آزمون فرضیه، بازهٔ اطمینان، احتمال پوشش[۴].

۱ مقدمه

در بیشتر پژوهش‌های مربوط به داده‌های طول عمر، عوامل بازدارنده‌ای همچون زمان و هزینه باعث می‌شوند که آزمایش‌ها پیش از شکست همهٔ واحدها به پایان برسند. به داده‌های برآمده از این آزمایش‌ها، داده‌های سانسور شده گفته می‌شود. در مطالعهٔ داده‌های سانسور شده، توزیع نمایی از اهمیت بسزایی برخوردار است.

اپشتاین [۱۴] در سال ۱۹۵۴ سانسور دورگهٔ نوع اول را پیشنهاد کرد. تحت این نوع سانسور، توزیع نمایی بسیار مورد توجه قرار گرفت تا این‌که بالاکریشنان و باسو [۲] در سال ۱۹۹۵ آن را از لحاظ نظری و کاربردی گسترش دادند. به‌شرط مشاهدهٔ حداقل یک شکست، چن و بهاتاچاریا [۶] در سال ۱۹۸۸ از روش تابع مولد گشتاور شرطی برای به دست آوردن توزیع برآوردگر ماکسیمم درستنمایی برای پارامتر مقیاس در مدل مقیاسی استفاده کردند. چایلدز و همکاران [۹] در سال ۲۰۰۳ از همین روش برای ساده کردن تابع چگالی استفاده کردند. در مدل سانسور دورگهٔ پیش‌روندهٔ نوع اول، چایلدز و همکاران [۸] در سال ۲۰۰۸ و کوندو و ژواردر [۱۵] در سال ۲۰۰۶ نیز از همین روش استفاده کردند. چایلدز و همکاران [۷] در سال ۲۰۱۲ روشی را برای آماره‌های ترتیبی به کار بردند. آن‌ها نتایج خود را برای داده‌های ارائه‌شده در بارلو و همکاران [۳] به کار گرفتند و حدود بالایی و پایینی را برای پارامتر مقیاس به دست آوردند. سرانجام کرامر و بالاکریشنان [۱۰] در سال ۲۰۱۳ توزیع دقیق برآوردگرهای ماکسیمم درست‌نمایی را تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول در توزیع نمایی به دست آوردند. در این مقاله با به کار گیری توزیع دقیق برآوردگر ماکسیمم درست‌نمایی ارائه‌شده در کرامر و بالاکریشنان [۱۰]، روشی برای انجام آزمون فرضیه و به دست آوردن بازه اطمینان برای میانگین توزیع نمایی پیشنهاد می‌شود.

[۱] دانشجوی دکترای آمار، دانشگاه یزد
[۲] عضو هیات علمی گروه آمار، دانشگاه یزد

[۳] Type-I progressive hybrid censoring

[۴] coverage probability

۲ سانسور دورگهٔ پیش‌روندهٔ نوع اول

فرض کنید n واحد آزمایشی در یک آزمون طول عمر قرار دارند. در زمان رخداد نخستین شکست، $r_۱$ واحد از $n-۱$ واحد باقی‌مانده به‌طور تصادفی از آزمایش حذف می‌شوند. سپس در زمان رخداد دومین شکست، $r_۲$ واحد از $n-r_۱-۲$ واحد باقی‌مانده به‌طور تصادفی از آزمایش حذف می‌شوند. به همین ترتیب آزمایش تا لحظهٔ $T^*=\min(t_۰, X_{m:m:n})$ ادامه می‌یابد، که $t_۰$ (زمان آستانه‌ای) مقداری ثابت، معلوم و از پیش تعیین‌شده و $X_{m:m:n}$ زمان رخداد mامین شکست است، که m نیز مقداری ثابت، معلوم و از پیش تعیین‌شده است. در این طرح، سانسور تعداد شکست‌های مشاهده‌شده با D نشان داده می‌شود و یک متغیر تصادفی است. واضح است که اگر $t_۰ \to \infty$، سانسور دورگهٔ پیش‌روندهٔ نوع اول به سانسور پیش‌روندهٔ نوع دوم تبدیل می‌شود.

فرض کنید $X_{۱:m:n},\dots,X_{m:m:n}$ یک نمونهٔ سانسور شدهٔ پیش‌روندهٔ نوع دوم با بردار حذفیات $\boldsymbol{r}=(r_۱,\dots,r_m)$ از جامعه‌ای با تابع توزیع F باشد. متغیرهای تصادفی $X_۱^{(I)},\dots,X_m^{(I)}$ و D را به‌صورت زیر تعریف می‌کنیم:

$$X_j^{(I)}=\min(X_{j:m:n},t_۰),\ ۱\le j\le m,$$
$$D=\sum_{j=۱}^{m}\mathrm{I}_{(-\infty,t_۰]}(X_{j:m:n}),$$

که D همان تعداد شکست‌های مشاهده‌شده است. واضح است که D دارای تکیه‌گاه $\{۰,\dots,m\}$ است.

۳ تابع درست‌نمایی و برآورد

فرض کنید جامعهٔ مورد نظر دارای توزیع نمایی با میانگین θ ($\mathrm{Exp}(\theta)$) باشد. کرامر و بالاکریشنان [۱۰] تابع درست‌نمایی پارامتر θ، برآوردگر ماکسیمم درست‌نمایی پارامتر θ و تابع چگالی $\hat\theta$ به‌شرط $D\ge ۱$ را به دست آوردند که در این بخش به مرور آن‌ها می‌پردازیم.

فرض کنید $Z_۱^{(I)},\dots,Z_D^{(I)}$ آماره‌های ترتیبی تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول با زمان آستانه‌ای $t_۰$ از جامعه‌ای با توزیع $\mathrm{Exp}(\theta)$ باشند. تابع درست‌نمایی پارامتر θ به‌صورت زیر است:

$$L(\theta)=\frac{\prod_{j=۱}^{D}\gamma_j}{\theta^D}\exp\left\{-\frac{۱}{\theta}\left[\sum_{j=۱}^{D}(۱+r_j)Z_j^{(I)}+\gamma_{D+۱}t_۰\right]\right\},$$
$$۰\le Z_۱^{(I)}\le\dots\le Z_D^{(I)}\le t_۰. \qquad (۱)$$

به‌طوری که

$$\gamma_k=\sum_{j=k}^{m}(۱+r_j).$$

وقتی $D\ge ۱$، با ماکسیمم کردن تابع $L(\theta)$ نسبت به θ داریم:

$$\hat\theta=\frac{۱}{D}\left[\sum_{j=۱}^{D}(۱+r_j)Z_j^{(I)}+\gamma_{D+۱}t_۰\right].$$

کرامر و بالاکریشنان [۱۰] تابع چگالی $\hat\theta$ به‌شرط $D\ge ۱$ را بر حسب توابع بی‌اسپلاین[۵] به دست آوردند. قبل از آوردن تابع چگالی شرطی، نخست تفاضلات تقسیم شده[۶] و توابع بی‌اسپلاین را تعریف می‌کنیم. با استفاده از باردن و فیرز [۴]، تفاضلات تقسیم‌شده به‌صورت زیر تعریف می‌شود:

تعریف ۱.۳. (تفاضلات تقسیم‌شده) فرض کنید g تابعی باشد که مقادیرش در یک مجموعه از نقاط (گره‌های) $x_۰,\dots,x_n$ معلوم یا قابل محاسبه باشد. در این‌جا فرض می‌شود که این نقاط دوبه‌دو متمایز هستند اما نیازی به مرتب شدن بر روی خط حقیقی ندارند. تفاضلات تقسیم‌شدهٔ اول بین x_i و $x_{i+۱}$ به‌صورت زیر تعریف می‌شوند:

$$g[x_i,x_{i+۱}]=\frac{g(x_{i+۱})-g(x_i)}{x_{i+۱}-x_i}.$$

تفاضلات تقسیم‌شدهٔ دوم بین x_i، $x_{i+۱}$ و $x_{i+۲}$ چنین تعریف می‌شوند:

$$g[x_i,x_{i+۱},x_{i+۲}]=\frac{g[x_{i+۱},x_{i+۲}]-g[x_i,x_{i+۱}]}{x_{i+۲}-x_i},$$

و سرانجام تفاضل تقسیم‌شدهٔ nام بین $x_۰$، $x_۱$، ... و x_n به‌کمک از

$$g[x_۰,\dots,x_n]=\frac{g[x_۱,\dots,x_n]-g[x_۰,\dots,x_{n-۱}]}{x_n-x_۰}.$$

در این مقاله تفاضل تقسیم‌شدهٔ nام تابع g بین نقاط $x_۰$، $x_۱$، ... و x_n مورد نظر است که با نماد $\triangle\{g(.)\ :\ x_۰,\dots,x_n\}$

[۵] B-Spline

[۶] divided differences

نمایش داده می‌شود و با توجه به علی [۱] به‌صورت زیر به دست می‌آید:

$$\triangle\{g(\cdot) : x_0, \ldots, x_n\} = \sum_{i=0}^{n} \frac{g(x_i)}{\prod_{j=0, j\neq i}^{n}(x_i - x_j)}.$$

تعریف ۲.۳. (بی‌اسپلاین) فرض کنید $\boldsymbol{t} = \{t_j\}$ دنباله‌ای غیر نزولی (متناهی یا نامتناهی) باشد. jامین بی‌اسپلاین (نرمال‌شدهٔ) مرتبهٔ k برای گره‌های دنبالهٔ $\boldsymbol{t}$ با نماد $\mathrm{B}_{j,k,\boldsymbol{t}}$ نشان داده می‌شود و به‌صورت زیر تعریف می‌شود:

$$\mathrm{B}_{j,k,\boldsymbol{t}}(x) = (t_{j+k} - t_j)\ \triangle\{[\cdot - x]_+^{k-1} : t_j, \ldots, t_{j+k}\}, \quad \forall x \in \mathbb{R},$$

به‌طوری که $[x]_+ = \max(0, x)$.

تعریف اولیهٔ (غیر نرمال) بی‌اسپلاین توسط کاری و شوئنبرگ [۱۱] در سال ۱۹۶۶ به‌صورت زیر ارائه‌شده است:

$$\mathrm{M}_{j,k,\boldsymbol{t}}(x) = \frac{k}{t_{j+k} - t_j}\mathrm{B}_{j,k,\boldsymbol{t}}(x).$$

برای اطلاع بیشتر در مورد توابع بی‌اسپلاین به [۱۱، ۱۲، ۱۳] مراجعه شود.

پس از تعریف بی‌اسپلاین، با استفاده از کرامر و بالاکریشنان [۱۰]، تابع چگالی $\hat{\theta}$ به‌شرط $D \geq 1$ به‌صورت زیر بیان می‌شود:

$$\begin{aligned} f^{\hat{\theta}|D\geq 1}(t) = &\frac{1}{1 - \mathrm{e}^{-nt_0\theta}} \sum_{d=1}^{m} \left(\prod_{j=1}^{d} \gamma_j\right) \\ &\times \frac{t_0^d}{(d-1)!\theta^d} \times \mathrm{B}_{d-1}\left(dt|\gamma_{d+1}t_0,\right. \\ &\left.\ldots, \gamma_1 t_0\right)\mathrm{e}^{-dt\theta}, \quad t \geq 0, \end{aligned} \tag{2}$$

که در آن $\mathrm{B}_{d-1}(x|x_0, \ldots, x_d)$ نشان‌دهندهٔ صفر امین بی‌اسپلاین غیر نرمال مرتبهٔ d (درجهٔ $d-1$) با گره‌های $x_0, \ldots, x_d$ است که به‌صورت زیر ساده می‌شود:

$$\mathrm{B}_{d-1}(x|x_0, \ldots, x_d) = \sum_{i=0}^{d} \frac{d[x_i - x]_+^{d-1}}{\prod_{j=0, j\neq i}^{d}(x_i - x_j)}, \quad \forall x \in \mathbb{R}.$$

۴ آزمون فرضیه

در این بخش به انجام آزمون فرضیه در مورد پارامتر θ پرداخته می‌شود. آزمون‌هایی که در این مقاله مورد بحث قرار گرفته‌اند عبارت‌اند از:

$$\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} \tag{3}$$

و

$$\begin{cases} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} \tag{4}$$

و

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0. \end{cases} \tag{5}$$

با استفاده از تابع درست‌نمایی (۱) داریم:

$$\frac{L(\theta_2)}{L(\theta_1)} = (\frac{\theta_1}{\theta_2})^D e^{D(\frac{1}{\theta_1} - \frac{1}{\theta_2})\hat{\theta}}.$$

روشن است که عبارت بالا برای هر $0 < \theta_1 < \theta_2$، تابعی صعودی از $\hat{\theta}$ است. پس این خانواده در $\hat{\theta}$ دارای ویژگی MLR است.

اکنون آزمون (۳) را در نظر بگیرید. بنا بر قضیهٔ کارلین-روبین در کسلا و برگر [۵]، آزمونی با ناحیهٔ رد $\hat{\theta} > k_1$ یک آزمون UMP سطح α است که مقدار k_1 با حل معادلهٔ زیر به دست می‌آید:

$$\begin{aligned} \alpha &= P_{\theta_0}(\tau > k_1 | D \geq 1) \\ &= \frac{1}{1 - \mathrm{e}^{-nt_0\theta_0}} \sum_{d=1}^{m} \left(\prod_{j=1}^{d} \gamma_j\right) \frac{t_0^d}{(d-1)!\theta_0^d} \\ &\times \int_{k_1}^{\infty} \mathrm{B}_{d-1}\left(dt|\gamma_{d+1}t_0, \ldots, \gamma_1 t_0\right) \mathrm{e}^{-dt\theta_0}\, dt. \end{aligned} \tag{6}$$

با استدلالی مشابه، آزمونی با ناحیه رد $\hat{\theta} < k_2$ یک آزمون UMP سطح α برای انجام آزمون (۴) است که مقدار k_2 با حل معادلهٔ زیر به دست می‌آید:

$$\begin{aligned} \alpha = &\frac{1}{1 - \mathrm{e}^{-nt_0\theta_0}} \sum_{d=1}^{m} \left(\prod_{j=1}^{d} \gamma_j\right) \frac{t_0^d}{(d-1)!\theta_0^d} \\ &\times \int_{0}^{k_2} \mathrm{B}_{d-1}\left(dt|\gamma_{d+1}t_0, \ldots, \gamma_1 t_0\right) \mathrm{e}^{-dt\theta_0}\, dt. \end{aligned}$$

توجه کنید که معادلات به دست‌آمده جواب صریح ندارند، اما با معلوم بودن تمام پارامترها، می‌توان آن‌ها را با نرم‌افزارهای قدرتمند ریاضی مانند Mathematica حل کرد.

اکنون آزمون (۵) را در نظر بگیرید. برای انجام دادن این آزمون، آزمون‌های زیر را در نظر بگیرید:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} \tag{7}$$

و

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} \tag{8}$$

فرض کنید ناحیهٔ رد آزمون (۵) برابر با R و نواحی رد آزمون‌های (۷) و (۸) به‌ترتیب برابر با $R_1 = \{\hat{\theta} < c_1\}$ و $R_2 = \{\hat{\theta} > c_2\}$ باشد. در آزمون (۵) واضح است که زمانی فرضیهٔ $\theta = \theta_0$ در مقابل فرضیهٔ $\theta \neq \theta_0$ رد می‌شود که فرضیه $\theta = \theta_0$ در مقابل لااقل یکی از فرضیه‌های $\theta < \theta_0$ یا $\theta > \theta_0$ رد شود. بنابراین $R = R_1 \cup R_2$. اکنون فرض کنید اندازهٔ آزمون (۵) برابر با α و اندازهٔ آزمون‌های (۷) و (۸) به‌ترتیب برابر با α_1 و α_2 باشد، که $0 < \alpha_1 + \alpha_2 < 1$. در نتیجه:

$$\begin{aligned}\alpha &= P_{\theta_0}(R_1 \cup R_2)\\ &= P_{\theta_0}(< c_1 \text{ یا } > c_2)\\ &= P_{\theta_0}(< c_1) + P_{\theta_0}(> c_2)\\ &= \alpha_1 + \alpha_2.\end{aligned}$$

بنا بر این می‌توان با در نظر گرفتن $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ به‌عنوان یک حالت خاص به نتیجهٔ دلخواه رسید.

۵ بازهٔ اطمینان

در این بخش در مورد بازهٔ اطمینان برای پارامتر θ بحث می‌شود. برای این منظور در ابتدا دو قضیه از کسلا و برگر [۵] و لهمن و رومئو [۱۶] بیان می‌شود.

قضیهٔ ۱.۵. فرض کنید $f_\theta(x)$ خانواده‌ای از توابع چگالی دارای خاصیت MLR در T و F_θ نشان‌دهندهٔ تابع توزیع آمارهٔ T متناظر با مقدار θ باشد. در این صورت اگر $\theta_2 > \theta_1$ باشد، خواهیم داشت:

$$F_{\theta_2}(t) \leq F_{\theta_1}(t), \qquad \forall t;$$

یعنی $F_\theta(t)$ به‌ازای هر t نسبت به θ نزولی است.

قضیهٔ ۲.۵. فرض کنید T یک آماره با تابع توزیع پیوستهٔ F_θ متناظر با مقدار θ باشد، که $F_\theta(t)$ به‌ازای هر t نسبت به θ نزولی است. همچنین فرض کنید $\alpha = \alpha_1 + \alpha_2$، که $0 < \alpha < 1$ است. اگر $\theta_L(t)$ و $\theta_U(t)$ در روابط

$$F_{\theta_U(t)}(t) = \alpha_1, \qquad F_{\theta_L(t)}(t) = 1 - \alpha_2$$

صدق کنند، $[\theta_L(T), \theta_U(T)]$ یک بازهٔ اطمینان برای θ با ضریب اطمینان $1 - \alpha$ خواهد بود.

توجه کنید که اگر α_1 و α_2 هر دو مخالف صفر باشند، بازهٔ اطمینان دوطرفه و اگر $\alpha_1 = 0$ یا $\alpha_2 = 0$ باشد، بازه‌های اطمینان یک‌طرفه به دست می‌آیند.

اکنون فرض کنید $Z_1^{(I)}, \ldots, Z_D^{(I)}$ آماره‌های ترتیبی تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول با زمان آستانه‌ای t_0 از جامعه‌ای با توزیع $\mathrm{Exp}(\theta)$ و هدف به دست آوردن بازهٔ اطمینان با ضریب $1 - \alpha$ برای θ باشد. می‌دانیم که خانوادهٔ چگالی‌های مربوط، در $\hat{\theta}$ دارای خاصیت MLR است. بنابراین با به‌کارگیری قضایای ۱.۵ و ۲.۵ می‌توان بازهٔ اطمینان دلخواه را به دست آورد.

توجه کنید که در این حالت، بازه‌های اطمینان، شکل صریحی ندارند و باید با نرم‌افزارهای قدرتمندی مانند Mathematica محاسبه شوند.

۶ مطالعات شبیه‌سازی

در این بخش، با استفاده از شبیه‌سازی، عملکرد بازهٔ اطمینان ارائه‌شده برای پارامتر θ که در بخش پیشین ارائه شد، ارزیابی می‌شود. برای ارزیابی عملکرد بازهٔ اطمینان، اغلب مفهوم میانگین طول بازه و احتمال پوشش مطرح می‌شود. برای این منظور، به‌ازای $n = 10$، $m = 5$ و مقادیر مختلف θ و t_0، چهار نوع طرح سانسور به کار گرفته شده است. پس از شبیه‌سازی ۱۰۰ نمونهٔ سانسور شدهٔ دورگهٔ پیش‌روندهٔ نوع اول از توزیع نمایی با میانگین θ و تعیین بازه‌های اطمینان ۹۵ درصدی برای هر نمونه، میانگین طول این بازه‌ها و نسبت بازه‌هایی که پارامتر θ را شامل می‌شوند (احتمال پوشش) محاسبه شده و در جدول ۱ آورده شده است. همان‌گونه که در جدول ۱ مشاهده می‌شود، برای بردار ثابت $\boldsymbol{r}$ و مقدار ثابت θ، با افزایش مقدار t_0، میانگین طول بازهٔ اطمینان کاهش می‌یابد که نتیجه‌ای قابل انتظار است؛ چون با افزایش t_0، میانگین تعداد شکست‌های مشاهده‌شده (D) افزایش می‌یابد و مشاهدات بیشتری در دسترس قرار می‌گیرد. با استدلالی مشابه، برای بردار ثابت $\boldsymbol{r}$ و مقدار ثابت t_0، با افزایش θ، میانگین طول بازهٔ اطمینان افزایش می‌یابد. همچنین مطابق انتظار، احتمال‌های پوشش نیز به ۰/۹۵ نزدیک هستند.

جدول ۱. میانگین طول بازه‌های اطمینان و احتمال پوشش آن‌ها برای طرح‌های مختلف سانسور و مقادیر مختلف θ و t_0.

احتمال پوشش	میانگین طول	$t.$	θ	r
۰/۹۶	۲/۸۲	۳	۱	(۵, ۰, ۰, ۰, ۰)
۰/۹۵	۲/۷۴	۴		
۰/۹۴	۲/۶۸	۵		
۰/۹۷	۴/۷۱	۳	۱/۵	
۰/۹۵	۴/۲۳	۴		
۰/۹۶	۴/۱۹	۵		
۰/۹۶	۲/۷۰	۳	۱	(۱, ۱, ۱, ۱, ۱)
۰/۹۸	۲/۵۷	۴		
۰/۹۵	۲/۴۴	۵		
۰/۹۶	۴/۲۴	۳	۱/۵	
۰/۹۵	۳/۷۶	۴		
۰/۹۲	۳/۷۰	۵		
۰/۹۲	۲/۶۵	۳	۱	(۱, ۱, ۲, ۰, ۱)
۰/۹۶	۲/۵۶	۴		
۰/۹۷	۲/۵۱	۵		
۰/۹۰	۴/۷۵	۳	۱/۵	
۰/۹۴	۳/۹۹	۴		
۰/۹۵	۳/۷۶	۵		
۰/۹۵	۲/۷۵	۳	۱	(۰, ۰, ۰, ۰, ۵)
۰/۹۶	۲/۴۶	۴		
۰/۹۷	۲/۴۴	۵		
۰/۹۴	۳/۹۳	۳	۱/۵	
۰/۹۲	۳/۷۴	۴		
۰/۹۶	۳/۷۰	۵		

۷ مثال‌های کاربردی

در این بخش، دو مثال کاربردی برای اجرای آزمون فرضیه و تعیین فاصلهٔ اطمینان بیان می‌شود.

مثال ۱.۷. داده‌های موجود در بارلو و همکاران [۳] را در نظر بگیرید. داده‌ها شامل ۶ مشاهدهٔ ۴, ۹, ۱۱, ۱۸, ۲۷, ۳۸، با بردار حذفیات $r = (۰, ۰, ۰, ۰, ۰, ۴)$ و زمان آستانه‌ای $t. = ۵۰$ هستند. فرض کنید هدف، انجام دادن آزمون $H. : \theta = ۳۰$ در مقابل $H_۱ : \theta \neq ۳۰$ با فرض $\alpha_۱ = \alpha_۲ = ۰/۰۲۵$ باشد. در این حالت $\hat{\theta} = ۴۳/۱۷$ و

$$R = \hat{\theta} < ۱۱/۰۰۹۵ \text{ یا } \hat{\theta} > ۶۱/۷۲۳۷.$$

چون مقدار $\hat{\theta}$ در ناحیهٔ رد آزمون قرار ندارد، دلیلی برای رد فرضیهٔ صفر وجود ندارد. مقادیر بحرانی با نرم‌افزار *Mathematica* محاسبه شده‌اند. برنامهٔ محاسبهٔ مقادیر بحرانی در پیوست آمده است.

مثال ۲.۷. شرایطی مشابه با مثال ۱.۷ را در نظر بگیرید. فرض کنید هدف، تعیین یک بازهٔ اطمینان برای θ با ضریب اطمینان ۰/۹۵ باشد. در این حالت با قرار دادن $\alpha_۱ = \alpha_۲ = ۰/۰۲۵$ در روابط قضیهٔ ۲.۵، بازهٔ اطمینان به‌صورت [۲۲/۱۹۸۵, ۱۱۸/۰۴۴۰]

به دست می‌آید. برنامهٔ محاسبهٔ بازهٔ اطمینان در پیوست آمده است.

نتیجه‌گیری

در این مقاله روشی برای اجرای آزمون فرضیه و به دست آوردن فاصلهٔ اطمینان برای میانگین توزیع نمایی تحت سانسور دورگهٔ پیش‌روندهٔ نوع اول پیشنهاد شد. سپس با استفاده از شبیه‌سازی، عملکرد بازهٔ اطمینان پیشنهادی مورد ارزیابی قرار گرفت و سرانجام روش‌های پیشنهادی برای یک مجموعه‌داده به کار گرفته شد.

پیوست

برنامهٔ محاسبهٔ مقادیر بحرانی در مثال ۱.۷

```
r = {0, 0, 0, 0, 0, 4};
theta = 30;
t0 = 50;
m = Length[r];
gamma = Range[m + 1];
For[i = 1, i < m + 1, i++, gamma[[i]] = m - i + 1 + Sum[r[[j]]
, {j, i, m}]];
gamma[[m + 1]] = 0;
n = m + Total[r];
g = gamma t0;
f[t_, theta_, t0_] := 1.(1 - Exp[-n t0.theta])
(Sum[Product[gamma[[j]], {j, 1, d}]
(d t0^(d - 1)).(Factorial[d - 1]
(gamma[[1]] - gamma[[d + 1]]) theta^d)*
BSplineBasis[{d - 1, Sort[g[[1 ;; d + 1]]]}, 0, d t]
Exp[-d t.theta], {d, 1, m}] );
F[k_] := Integrate[f[t, theta, t0], {t, 0, k}];
FindRoot[F[k] == .025, {k, 10}]
FindRoot[F[k] == .975, {k, 50}]
```

برنامهٔ محاسبهٔ بازهٔ اطمینان در مثال ۲.۷

```
r = {0, 0, 0, 0, 0, 4};
theta = 30;
t0 = 50;
m = Length[r];
gamma = Range[m + 1];
For[i = 1, i < m + 1, i++, gamma[[i]] = m - i + 1 + Sum[r[[j]]
, {j, i, m}]];
gamma[[m + 1]] = 0;
n = m + Total[r];
g = gamma t0;
f[t_, theta_, t0_] := 1.(1 - Exp[-n t0.theta])
(Sum[Product[gamma[[j]], {j, 1, d}] (d t0^(d - 1)).
(Factorial[d - 1] (gamma[[1]] - gamma[[d + 1]]) theta^d)*
BSplineBasis[{d - 1, Sort[g[[1 ;; d + 1]]]}, 0, d t]
Exp[-d t.theta], {d, 1, m}] );
th = 43.17;
F[theta_] := Integrate[f[x, theta], t0], {x, 0, th}];
FindRoot[F[theta] == .975, {theta, 20}]
FindRoot[F[theta] == .025, {theta, 120}]
```

مراجع

[1] Ali, M.M. (1973). Content of the frustum of simplex. *Pacific Journal of Mathematics*, **48**, 313–322.

[2] Balakrishnan, N. and Basu, A.P. (eds.) (1995). *The Exponential Distribution: Theory, Methods, and Applications*. Gordon and Breach Science, Newark, N.J.

[3] Barlow, R. E., Madansky, A., Proschan, F., and Scheuer, E.M. (1968). Statistical estimation procedures for the 'burn-in' process. *Technometrics*, **10**, 51–62.

[4] Burden, R.L. and Faires, J.D. (2011). *Numerical Analysis*, Ninth Edition. Brooks.Cole, Cencag Learning, Boston.

[5] Casella, G. and Berger, R.L. (2002) *Statistical Inference*, Second Edition. Duxbury, California.

[6] Chen, S.M. and Bhattacharyya, G.K. (1988). Exact confidence bounds for an exponential parameter under hybrid censoring. *Communications in Statistics: Theory and Methods*, **16**, 2429–2442.

[7] Childs, A., Balakrishnan, N. and Chandrasekar, B. (2012). Exact distribution of the MLEs of the parameters and of the quantiles of two-parameter exponential distribution under hybrid censoring. *Statistics*, **46**, 441–458.

[8] Childs, A., Chandrasekar, B., and Balakrishnan, N. (2008). Exact likelihood inference for an exponential parameter under progressive hybrid censoring schemes. *In Statistical models and methods for biomedical and technical systems*, 319-330. Birkhäuser Boston.

[9] Childs, A., Chandrasekar, B., Balakrishnan, N. and Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution. *Annals of the Institute of Statistical Mathematics*, **55**, 319–330.

[10] Cramer, E. and Balakrishnan, N. (2013). On some exact distributional results based on Type-I progressively hybrid censored data from exponential distributions. *Statistical Methodology*, **10**, 128-150.

[11] Curry, H. and Schoenberg, I. (1966). On Pólya frequency functions IV: the fundamental spline functions and their limits. *Journal d Analyse Mathematique*, **17**, 71-107.

[12] De Boor, C. (2001). *A Practical Guide to Splines*, Revised Edition. Springer, New York.

[13] De Boor, C. (1976). Splines as linear combinations of B-splines. A survey. Lorentz, G.G., Chui, C.K. and Schumaker, L.L. (eds.), Approximation Theory II. Academic Press, New York, pp. 1–47.

[14] Epstein, B. (1954). Truncated life tests in the exponential case. *The Annals of Mathematical Statistics*, **25**, 555–564.

[15] Kundu, D. and Joarder, A. (2006). Analysis of Type-II progressively hybrid censored data. *Computational Statistics and Data Analysis*, **50**, 2509–2528.

[16] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, Third Edition. Springer, New York.