یک مسئله بهینه‌سازی مقید برای تعیین کوچک‌ترین ناحیه اطمینان توزیع پارتو تحت داده‌های سانسور پیش‌رونده نوع دوم

مرجان زارع، اکبر اصغر زاده[۱]

گروه آمار، دانشگاه مازندران

تاریخ دریافت: ۱۳۹۶/۷/۴ تاریخ پذیرش: ۱۳۹۷/۸/۸

چکیده: در این مقاله، یک مسئله بهینه‌سازی مقید برای تعیین کوچک‌ترین ناحیه اطمینان توأم برای پارامترهای توزیع پارتو بر اساس نمونه‌های سانسور شده پیش‌رونده نوع دوم فرمول‌بندی و حل می‌شود. تابع هدف، مساحت ناحیه اطمینان و قید مسئله، سطح اطمینان مشخص است. کوچک‌ترین ناحیه اطمینان توأم پیشنهادی برای نمونه‌های کامل و نمونه‌های سانسور شده نوع دوم از راست نیز معتبر است. مساحت کوچک‌ترین ناحیه اطمینان پیشنهادی و مساحت ناحیه اطمینان متعادل مورد مقایسه قرار می‌گیرند. در انتها، دو مثال عددی برای تشریح روش بهینه‌سازی پیشنهادی ارائه می‌شود.

واژه‌های کلیدی: نواحی اطمینان توأم، سانسور پیش‌رونده نوع دوم، روش لاگرانژ، توزیع پارتو

رده‌بندی موضوعی (۲۰۱۰): ۶۲F۲۵، ۶۲F۳۰، ۶۲N۰۱.

۱- مقدمه

تابع چگالی احتمال و تابع توزیع تجمعی توزیع پارتو با پارامتر شکل $\alpha > ۰$ و پارامتر دقت $\tau > ۰$ به‌صورت زیر می‌باشند.

$$f(x;\alpha,\tau) = \tau\alpha(\tau x)^{-\alpha-۱}$$

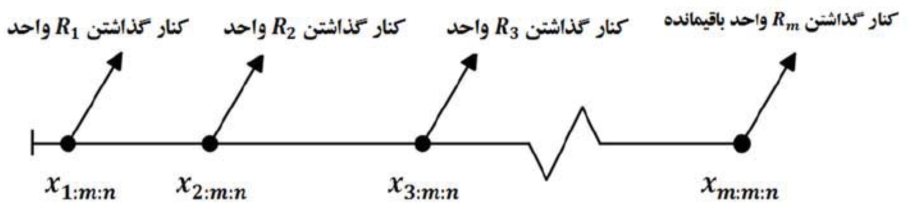
$$F(x;\alpha,\tau) = ۱-(\tau x)^{-\alpha} \qquad x \geq ۱/\tau.$$

۱- آدرس الکترونیکی نویسنده مسئول مقاله: a.asgharzadeh@umz.ac.ir

توزیع پارتو اولین بار توسط پارتو (۱۹۸۷) به‌عنوان یک مدل برای توزیع درآمد معرفی شد اما امروزه به‌عنوان یک مدل آماری در زمینه‌های مختلفی چون بیمه، اقتصاد، مهندسی، آب‌شناسی و قابلیت اعتماد به کار می‌رود. برای مطالعه بیشتر روی توزیع پارتو می‌توان به کتاب جانسون و همکاران [۱] مراجعه کرد.

به دست آوردن نواحی اطمینان توأم برای پارامترهای مجهول در بسیاری از مطالعات حائز اهمیت است. به‌طور خاص تعیین ناحیه اطمینان توأم برای پارامترهای توزیع پارتو اغلب موردنیاز است. در مقالات، مطالعات زیادی در زمینه‌ی ناحیه اطمینان توأم انجام‌شده است با این وجود کوچک‌ترین ناحیه اطمینان توأم خیلی موردمطالعه قرار نگرفته است. کوچک‌ترین ناحیه اطمینان توأم پارتو بر اساس نمونه‌های سانسور شده نوع دوم دوطرفه توسط فرناندز [۳-۲] موردمطالعه قرار گرفت. فرناندز [۴]، روش استنباطی اتکایی را برای پیدا کردن کوچک‌ترین ناحیه اطمینان توأم پارتو پیشنهاد نمود. اصغر زاده و همکاران [۵] ناحیه اطمینان توأم متعادل را برای پارامترهای توزیع پارتو بر اساس داده‌های رکوردی ارائه کرده‌اند. اخیراً اصغر زاده و همکاران [۶]، کوچک‌ترین ناحیه اطمینان توأم را برای پارامترهای توزیع رایلی دو پارامتری موردمطالعه قرار داده‌اند. در این مقاله، بر اساس نمونه‌های سانسور شده پیش‌رونده نوع دوم، یک روش بهینه‌سازی مقید برای پیدا کردن کوچک‌ترین ناحیه اطمینان توأم پارتو ارائه می‌شود.

سانسور پیش‌رونده نوع دوم، تعمیمی از سانسور نوع دوم است. در طرح سانسور پیش‌رونده نوع دوم از راست، n واحد در زمان $t=0$ در آزمون طول عمر قرار می‌گیرند. بلافاصله بعد از مشاهده اولین شکست، r_1 واحد سالم از بین $n-1$ واحد دیگر به‌تصادف از آزمایش خارج می‌شوند. با مشاهده شکست دوم، r_2 واحد سالم از بین $(n-r_1-2)$ واحد موجود در آزمایش سانسور می‌شوند. این روند ادامه می‌یابد تا اینکه در زمان شکست m-ام تمام $r_m=n-r_1-r_2-\ldots-r_{m-1}-m$ واحد باقیمانده از آزمایش کنار گذاشته می‌شوند. در این طرح سانسور مقادیر r_1، r_2،...، r_m و در نتیجه m از قبل مشخص و ثابت در نظر گرفته می‌شوند. لذا زمان‌های سانسور همگی تصادفی می‌باشند. شکل۱ نمایشی از این وضعیت را ارائه می‌دهد. به زمان‌های شکست $(X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n})$ نمونه سانسور شده پیش‌رونده نوع دوم گویند. در حالت خاص هنگامی‌که $r_1=r_2=\ldots=r_{m-1}=0$ و $r_m=n-m$، این طرح به طرح سانسور نوع دوم متداول تبدیل می‌شود که در آن تنها اولین m آماره ترتیبی مشاهده می‌شوند. همچنین اگر $m=n$ و $r_1=r_2=\ldots=r_m=0$، آنگاه n آماره ترتیبی بدون سانسور (نمونه کامل) حاصل خواهد شد.



شکل (۱): فرآیند تولید آماره‌های ترتیبی سانسور پیش‌رونده نوع دوم از راست

در سال‌های اخیر مطالعات زیادی روی روش‌های استنباطی با داده‌های سانسور پیش‌رونده نوع دوم انجام‌شده است، برای مطالعه بیشتر روی کارهای انجام‌شده می‌توان به بالاکریشنان و آگاروالا [۷] و بالاکریشنان و کرامر [۸] مراجعه کرد. این مقاله کلاسی از نواحی اطمینان توأم را برای (α,τ) توزیع پارتو بر پایه‌ی نمونه سانسور پیش‌رونده نوع دوم ارائه می‌دهد که شامل ناحیه پیشنهادشده توسط کوش و کایا [۹] به‌عنوان ناحیه‌ی متعادل است. ازآنجایی‌که در انتخاب یک ناحیه اطمینان معمولاً کوچک کردن اندازه آن سودمند است، یک مسئله‌ی بهینه‌سازی مقید به‌منظور تعیین کوچک‌ترین ناحیه اطمینان در کلاس پیشنهادی با سطح اطمینان خواسته‌شده فرمول‌بندی و حل می‌شود.

ساختار مقاله به‌صورت زیر است: در بخش دوم، بر اساس نمونه‌ی سانسور پیش‌رونده نوع دوم از توزیع پارتو یک ناحیه اطمینان دقیق برای (α,τ) به‌عنوان ناحیه اطمینان متعادل ارائه می‌شود. در بخش سوم، عبارت صریحی برای مساحت ناحیه اطمینان توأم محاسبه و یک مسئله‌ی برنامه‌ریزی غیرخطی برای پیدا کردن کوچک‌ترین ناحیه اطمینان فرمول‌بندی و حل می‌شود. جواب بهینه با استفاده از روش لاگرانژی به دست می‌آید. در بخش چهارم، دو مثال عددی برای تشریح روش پیشنهادی ارائه می‌شوند. نتیجه‌گیری و کاربردهای نواحی اطمینان پیشنهادی در بخش پنجم ارائه می‌شود.

۲- ناحیه اطمینان توأم متعادل

فرض کنید $\underline{\mathrm{X}}=(\mathrm{X}_1,\ldots,\mathrm{X}_m)$ یک نمونه سانسور پیش‌رونده نوع دوم با طرح سانسور پیش‌رونده $r=(\mathrm{r}_1,\ldots,\mathrm{r}_m)$ باشد. در ادامه برای راحتی در نوشتن به‌جای $\mathrm{X}_{i:\mathrm{m:n}}$، X_i را می‌نویسیم. تعریف کنید

$$Y_i=-\log(1-F(\mathrm{X}_i;\alpha,\tau))=\alpha\log(\tau\mathrm{X}_i),\qquad i=1,2,\ldots,m,$$

آنگاه $Y_1 < Y_2 < \ldots < Y_m$ آماره‌های ترتیبی سانسور پیش‌رونده نوع دوم از توزیع نمایی استاندارد می‌باشند(بالاکریشنان و آگاروالا [[۷]). چون فواصل

$$Z_1 = nY_1,$$

$$Z_i = (n - \sum_{j=1}^{i-1} R_j - i + 1)\left(Y_i - Y_{i-1}\right), \qquad i = 2,3,\ldots,m, \tag{1}$$

متغیرهای تصادفی مستقل و هم توزیع از توزیع نمایی استاندارد می‌باشند بنابراین

$$V = 2Z_1 = 2nY_1$$

دارای توزیع خی دو با ۲ درجه آزادی و

$$U = 2\sum_{i=2}^{m} Z_i = 2\sum_{i=1}^{m} (r_i + 1)\, Y_i - 2nY_1$$

دارای توزیع خی دو با $2(m-1)$ درجه آزادی است و U و V متغیرهای مستقل می‌باشند. حال فرض کنید

$$T_1 = \frac{U/2(m-1)}{V/2} = \frac{U}{(m-1)V} = \frac{\sum_{i=1}^{m} (r_i + 1)\, Y_i - nY_1}{n(m-1)\, Y_1}$$

و

$$T_2 = U + V = 2\sum_{i=1}^{m} (r_i + 1)\, Y_i \,.$$

به‌راحتی دیده می‌شود که T_1 دارای توزیع فیشر با درجات آزادی $2(m-1)$ و ۲ و همچنین T_2 دارای توزیع خی دو با درجه آزادی $2m$ است. علاوه بر این می‌توان اثبات کرد که T_1 و T_2 از هم مستقل هستند. از کمیت‌های محوری T_1 و T_2 برای پیدا کردن نواحی اطمینان استفاده می‌شود.

فرض کنید برای $\gamma \in (0,1)$، $F_{\nu_1,\nu_2,\gamma}$ چندک γام توزیع F با درجات آزادی ν_1 و ν_2 و $\chi^2_{\nu,\gamma}$ چندک γام توزیع χ^2 با درجه آزادی ν را نشان دهد و همچنین فرض کنید ناحیه اطمینان $100(1-\gamma)\%$ برای (α,τ) مدنظر باشد. کمیت‌های T_1 و T_2 را می‌توان به‌صورت زیر نوشت

$$T_1(\tau) \equiv T_1(\tau; \underline{X}, r) = \frac{\sum_{i=1}^{m} (r_i + 1)\log X_i - n\log X_1}{n(m-1)\log(\tau X_1)} \tag{2}$$

$$T_{۲}(\alpha,\tau) \equiv T_{۲}(\alpha,\tau;\underline{X},\mathrm{r}) = ۲\alpha\sum_{i=۱}^{m}(r_i+۱)\log \mathrm{X}_i + n\log\tau. \quad (۳)$$

داریم

$$\Pr\left\{F_{۲m-۲,۲;۱-\sqrt{۱-\gamma}} < \mathrm{T}_{۱}(\tau) < \infty\right\} = \sqrt{۱-\gamma}\,,$$

$$\Pr\left\{\chi^{۲}_{۲m;\frac{۱-\sqrt{۱-\gamma}}{۲}} < T_{۲}(\alpha,\tau) < \chi^{۲}_{۲m;\frac{۱+\sqrt{۱-\gamma}}{۲}}\right\} = \sqrt{۱-\gamma}$$

لذا

$$\Pr\left\{F_{۲m-۲,۲;۱-\sqrt{۱-\gamma}} < \mathrm{T}_{۱}(\tau) < \infty, \chi^{۲}_{۲m;\frac{۱-\sqrt{۱-\gamma}}{۲}} < T_{۲}(\alpha,\tau) < \chi^{۲}_{۲m;\frac{۱+\sqrt{۱-\gamma}}{۲}}\right\} = ۱-\gamma,$$

یا به‌طور معادل،

$$\Pr\left\{\frac{۱}{X_{۱}} < \tau < \frac{۱}{X_{۱}}\exp\left\{\frac{\sum_{i=۱}^{m}(r_i+۱)\log X_i - n\log X_{۱}}{n(m-۱)\mathrm{F}_{۲m-۲,۲;۱-\sqrt{۱-\gamma}}}\right\}, \frac{\chi^{۲}_{۲m;\frac{۱-\sqrt{۱-\gamma}}{۲}}}{U_{۲}(\tau)} < \alpha < \frac{\chi^{۲}_{۲m;\frac{۱+\sqrt{۱-\gamma}}{۲}}}{U_{۲}(\tau)}\right\} = ۱-\gamma$$

که

$$U_{۲}(\tau) = \sum_{i=۱}^{m}(\mathrm{r}_i+۱)\log(\tau X_i).$$

بنابراین ناحیه اطمینان توأم $(۱-\gamma)۱۰۰\%$ برای (α,τ) به‌صورت زیر به دست می‌آید.

$$\left(\frac{۱}{X_{۱}} < \tau < \frac{۱}{X_{۱}}\exp\left\{\frac{\sum_{i=۱}^{m}(r_i+۱)\log X_i - n\log X_{۱}}{n(m-۱)\mathrm{F}_{۲m-۲,۲;۱-\sqrt{۱-\gamma}}}\right\}, \frac{\chi^{۲}_{۲m;\frac{۱-\sqrt{۱-\gamma}}{۲}}}{U_{۲}(\tau)} < \alpha < \frac{\chi^{۲}_{۲m;\frac{۱+\sqrt{۱-\gamma}}{۲}}}{U_{۲}(\tau)}\right).$$

۳- ناحیه اطمینان توأم بهین

حال برای ثابت‌های $a>۰$ و $۰<b_{۱}<b_{۲}$ فرض کنید داشته باشیم

$$\Pr\left\{a < \mathrm{T}_{۱} < \infty, b_{۱} < T_{۲} < b_{۲}\right\} = \left\{۱-F_{۱}(\mathrm{a})\right\}\left\{F_{۲}(\mathrm{b}_{۲}) - F_{۲}(\mathrm{b}_{۱})\right\} = ۱-\gamma,$$

که در آن $F_{۱}$ و $F_{۲}$ به ترتیب توابع توزیع $T_{۱}$ و $T_{۲}$ می‌باشند. بر اساس ثابت‌های a، $b_{۱}$ و $b_{۲}$ ناحیه اطمینان $(۱-\gamma)۱۰۰\%$ و مساحت ناحیه متناظر برای (α,τ) به‌صورت زیر می‌باشند.

$$R_{1-\gamma} = \left\{ \frac{1}{X_1} < \tau < U_1(\mathrm{a}), \frac{b_1}{U_2(\tau)} < \alpha < \frac{b_2}{U_2(\tau)} \right\},$$

$$\mathrm{Area}(R_{1-\gamma}) = \int_{1/X_1}^{U_1(\mathrm{a})} \frac{b_2 - b_1}{U_2(\tau)} d\tau,$$

که

$$U_1(\mathrm{a}) = \frac{1}{X_1} \exp\left\{ \frac{\sum_{i=1}^{m} (r_i + 1)\log X_i - n \log X}{n(m-1)a} \right\}.$$

اکنون مسئله بهینه‌سازی غیرخطی برای تعیین ناحیه اطمینان $(1-\gamma)100\%$ با کمترین مساحت به‌صورت زیر فرمول‌بندی می‌شود

$$\text{Minimize} \qquad \int_{1/X_1}^{U_1(\mathrm{a})} \frac{b_2 - b_1}{U_2(\tau)} d\tau$$

$$\text{Subject to} \qquad \{1 - F_1(\mathrm{a})\}\{F_2(\mathrm{b}_2) - F_2(\mathrm{b}_1)\} = 1-\gamma,$$

$$\alpha > 0, \; 0 < b_1 < b_2 .$$

تابع لاگرانژ عبارت است از:

$$H(\mathrm{z},\beta) = \int_{1/X_1}^{U_1(\mathrm{a})} \frac{b_2 - b_1}{U_2(\tau)} d\tau + \beta\left[\{1 - F_1(\mathrm{a})\}\{F_2(\mathrm{b}_2) - F_2(\mathrm{b}_1)\} - (1-\gamma)\right],$$

که $z = (a, \mathrm{b}_1, \mathrm{b}_2)$ و β ضریب لاگرانژ است. مشتق‌های تابع لاگرانژ نسبت به a، b_1، b_2 و β به ترتیب عبارتند از:

$$\frac{\partial H(\mathrm{z},\beta)}{\partial a} = U_1'(\mathrm{a}) \frac{b_2 - b_1}{U_2(\mathrm{U}_1(\mathrm{a}))} - \beta f_1(a)\left(F_2(b_2) - F_2(b_1)\right),$$

$$\frac{\partial H(\mathrm{z},\beta)}{\partial b_1} = \int_{1/X_1}^{U_1(\mathrm{a})} \frac{-1}{U_2(\tau)} d\tau - \beta\{1 - F_1(\mathrm{a})\} f_2(\mathrm{b}_1),$$

$$\frac{\partial H(\mathrm{z},\beta)}{\partial b_2} = \int_{1/X_1}^{U_1(\mathrm{a})} \frac{-1}{U_2(\tau)} d\tau + \beta\{1 - F_1(\mathrm{a})\} f_2(\mathrm{b}_2),$$

$$\frac{\partial H(\mathrm{z},\beta)}{\partial \beta} = \{1 - F_1(\mathrm{a})\}\{F_2(\mathrm{b}_2) - F_2(\mathrm{b}_1)\} - (1-\gamma),$$

که در آن‌ها f_1 و f_2 به ترتیب مشتق‌های F_1 و F_2 می‌باشند. با قرار دادن مشتقات جزئی تابع $H(z,\beta)$ برابر صفر می‌توانیم جواب بهینه $z^* = (a^*, b_1^*, b_2^*)$ را به دست آوریم. جواب بهینه از حل سه معادله زیر به‌طور هم‌زمان به دست می‌آید.

$$\int_{1/X_1}^{U_1(\mathrm{a})} \frac{1}{U_2(\tau)} d\tau = \frac{\{1-F_1(\mathrm{a})\} f_2(\mathrm{b}_2) U_1'(\mathrm{a})(\mathrm{b}_1-\mathrm{b}_2)}{U_2(\mathrm{U}_1(\mathrm{a})) f_1(a) \{F_2(\mathrm{b}_2) - F_2(\mathrm{b}_1)\}},$$

$$\{1-F_1(\mathrm{a})\}\{F_2(\mathrm{b}_2) - F_2(\mathrm{b}_1)\} = 1-\gamma,$$

$$f_2(\mathrm{b}_1) = f_2(\mathrm{b}_2)\,.$$

۴- مثال‌های عددی

اکنون دو مثال عددی برای تشریح روش پیشنهادشده در این مقاله ارائه می‌شوند.

مثال ۱: داده‌های زیر زمان‌های خرابی نوعی عایق الکترونیکی که تحت‌فشار (با ولتاژ ثابت) قرارگرفته است را نشان می‌دهد. این داده‌ها از نلسون [۱۰] اقتباس شده‌اند و قبلاً نیز توسط اصغر زاده و همکاران [۱۱] به کار رفته‌اند. داده‌ها عبارتند از:

۰/۳۵ ۰/۵۹ ۰/۹۶ ۰/۹۹ ۱/۶۹ ۱/۹۷ ۲/۰۷ ۲/۵۸ ۲/۷۱ ۲/۹۰
۳/۶۷ ۳/۹۹ ۵/۳۵ ۱۳/۷۷ ۲۵/۵۰

به کمک آزمون کولموگروف- اسمیرنوف می‌توان نشان داد که توزیع پارتو با پارامترهای $\alpha = 0/51$ و $\tau = 2/857$ به داده‌های فوق برازش می‌شود. برای محاسبه ناحیه اطمینان متعادل و کوچک‌ترین ناحیه اطمینان، ۳ طرح سانسور پیش‌رونده را به‌صورت زیر در نظر می‌گیریم:

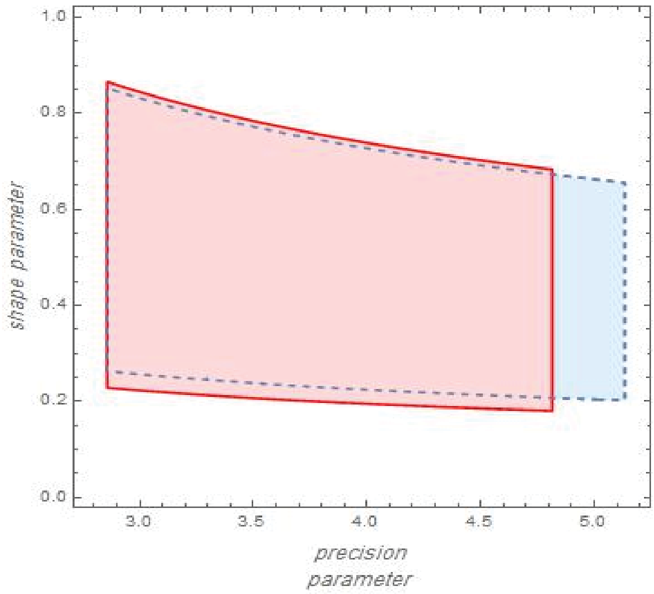
طرح سانسور ۱: $\mathrm{r}_1 = \mathrm{r}_2 = \ldots = \mathrm{r}_{15} = 0,\ r = (\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_{15}),\ n = m = 15,\ 1-\gamma = 0/95$.

این حالت متناظر با نمونه کامل $\underline{\mathrm{X}} = (\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_{15})$ است. در این حالت، نواحی اطمینان متعادل و بهین به ترتیب به‌صورت زیر می‌باشند.

$$R^b_{0/95} = \left\{(\alpha,\tau): 2/85714 < \tau < 5/13456, \frac{15/3903}{U_2(\tau)} < \alpha < \frac{49/9138}{U_2(\tau)}\right\},$$

$$R^*_{0/95} = \left\{(\alpha,\tau): 2/85714 < \tau < 4/81512, \frac{13/3622}{U_2(\tau)} < \alpha < \frac{50/7002}{U_2(\tau)}\right\}.$$

مساحت‌های $R^{b}_{۰/۹۵}$ و $R^{*}_{۰/۹۵}$ به ترتیب عبارتند از ۱/۱۵۹ و ۱/۰۹۵؛ بنابراین روش بهینه، کاهش $۱-\text{Area}[R^{*}_{۰/۹۵}]/\text{Area}[R^{b}_{۰/۹۵}] \approx ۵/۵۳۲$ در مساحت را نشان می‌دهد. شکل ۲ نواحی اطمینان ۹۵٪ متعادل و بهینه برای (α,τ) را نشان می‌دهد.



شکل (۲): نواحی اطمینان ۹۵٪ متعادل (خط تیره) و بهینه (خطوط یکپارچه) در طرح ۱، مثال ۱

طرح سانسور ۲: $r=(r_۱,r_۲,\ldots,r_{۱۰})=(۰,۰,\ldots,۰,۵)$ ، $m=۱۰$ ، $n=۱۵$ ، $۱-\gamma=۰/۹۵$

در این حالت نمونه $\underline{X}=(X_۱,X_۲,\ldots,X_{۱۰})$ عبارت است از:

۲/۹۰ ۲/۷۱ ۲/۵۸ ۲/۰۷ ۱/۹۷ ۱/۶۹ ۰/۹۹ ۰/۹۶ ۰/۵۹ ۰/۳۵

و این حالت متناظر باحالت نمونه سانسور نوع دوم از راست است. در این حالت ناحیه اطمینان متعادل و کوچک‌ترین ناحیه اطمینان به ترتیب عبارتند از:

$$R^{b}_{۰/۹۵}=\left\{(\alpha,\tau):۲/۸۵۷۱۴<\tau<۶/۴۸۷۵۱,\frac{۸/۵۷۳۶۹}{U_۲(\tau)}<\alpha<\frac{۳۶/۷۱۴۱}{U_۲(\tau)}\right\}$$

$$R^{*}_{۰/۹۵}=\left\{(\alpha,\tau):۲/۸۵۷۱۴<\tau<۵/۸۴۹۷,\frac{۶/۸۳۴۸۹}{U_۲(\tau)}<\alpha<\frac{۳۷/۴۵۴۴}{U_۲(\tau)}\right\}$$

که مساحت آن‌ها به ترتیب ۱/۶۴۹۹۸ و ۱/۵۲۳۴۵ می‌باشند و روش بهینه، مساحت ناحیه را ۷/۶۶۸٪ کاهش می‌دهد.

طرح سانسور ۳: $n=۱۵,\ \ m=۷,\ \ r=(r_۱,r_۲,...,r_۷)=(۰,۰,۳,۰,۳,۰,۲),\ \ ۱-\gamma=۰/۹۵.$ در این حالت داده‌ها عبارتند از:

۰/۳۵ ۰/۵۹ ۰/۹۶ ۱/۶۹ ۱/۹۷ ۲/۵۸ ۲/۹۰

و نواحی اطمینان ۹۵٪ متعادل و بهین عبارتند از:

$$R^{b}_{۰/۹۵}=\left\{(\alpha,\tau):۲/۸۵۷۱۴<\tau<۷/۹۴۳۵۷,\frac{۵/۶۶۰۲۷}{U_۲(\tau)}<\alpha<\frac{۲۶/۰۳۰۳}{U_۲(\tau)}\right\},$$

$$R^{*}_{۰/۹۵}=\left\{(\alpha,\tau):۲/۸۵۷۱۴<\tau<۹/۵۳۶۱,\frac{۴/۸۸۵۷۷}{U_۲(\tau)}<\alpha<\frac{۲۸/۳۸۰۶}{U_۲(\tau)}\right\},$$

با مساحت‌های متناظر ۲/۵۰۳۹ و ۲/۲۱۷۶ . لذا مقدار کاهش در مساحت $R^{b}_{۰/۹۵}$ نسبت به مساحت $R^{*}_{۰/۹۵}$، برابر ۱۱/۴۳۶٪ است.

مثال ۲: داده‌های جدول ۱، یک نمونه سانسور پیش‌رونده نوع دوم را نشان می‌دهد که توسط کوش و کایا [۹] از توزیع پارتو با پارامترهای $\alpha=۲$ و $\tau=۰/۳۳$ شبیه‌سازی شده‌اند. این داده‌ها قبلاً توسط پارسی و همکاران [۱۲] نیز استفاده‌شده است که فواصل اطمینان هم‌زمان را برای پارامترهای توزیع پارتو موردمطالعه قرار دادند.

جدول (۱): نمونه شبیه‌سازی از توزیع پارتو $(m=۸, n=۱۹)$

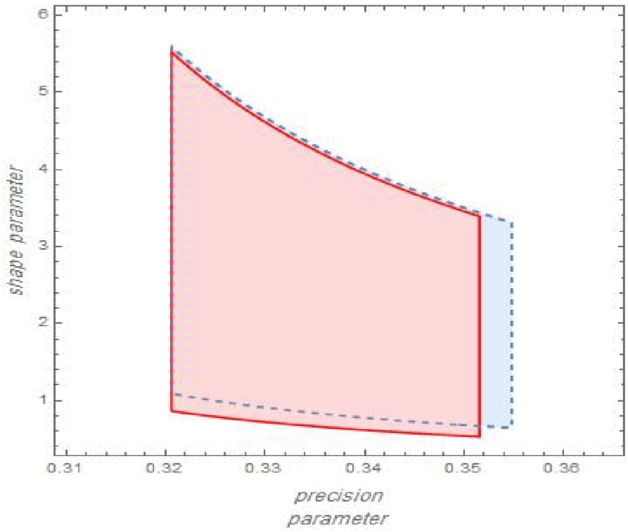
i	۱	۲	۳	۴	۵	۶	۷	۸
r	۰	۰	۳	۰	۳	۰	۰	۵
$x_{i;m;n}$	۳/۱۱۹۲	۳/۱۲۰۸	۳/۱۷۲۷	۳/۲۰۴۶	۳/۳۵۴۸	۳/۵۱۳۸	۴/۱۶۹۷	۴/۳۴۵۶

بر اساس داده‌های جدول ۱، نواحی اطمینان ۹۵٪ متعادل و بهینه به ترتیب عبارتند از:

$$R^{b}_{۰/۹۵}=\left\{(\alpha,\tau):۰/۳۲۰۶<\tau<۰/۳۵۴۸,\frac{۶/۰۶۸۳۹}{U_۲(\tau)}<\alpha<\frac{۳۱/۲۰۷}{U_۲(\tau)}\right\},$$

$$R^*_{۰/۹۵}=\left\{(\alpha,\tau):۰/۳۲۰۵۹۵<\tau<۰/۳۵۱۵۴۵,\frac{۴/۸۱۳۲۲}{U_۲(\tau)}<\alpha<\frac{۳۰/۷۹۸۷}{U_۲(\tau)}\right\}.$$

مساحت‌های $R^b_{۰/۹۵}$ و $R^*_{۰/۹۵}$ به ترتیب ۰/۱۱۶۶۹ و ۰/۱۱۱۵۸ می‌باشند؛ بنابراین روش بهینه، کاهش ۴/۳۸۳۶٪ در مساحت را نشان می‌دهد. شکل ۳ نواحی اطمینان ۹۵٪ متعادل و بهین را برای (α,τ) در مثال فوق نشان می‌دهد.



شکل (۳): نواحی اطمینان ۹۵٪ متعادل (خط تیره) و بهینه (خطوط یکپارچه) در مثال ۲

۵- کاربردها و نتیجه‌گیری

در این مقاله، نواحی اطمینان دقیق برای پارامترهای توزیع پارتو تحت سانسور پیش‌رونده نوع دوم مطالعه شده است. به کمک بهینه‌سازی مقید و با مینیمم کردن تابع لاگرانژ مربوطه، کوچک‌ترین نواحی اطمینان توأم پیشنهاد شد.

نتایج نشان می‌دهد که رویکرد پیشنهادی یک روش مؤثر برای تعیین بهترین مجموعه‌های اطمینان است. کاربردهای نواحی اطمینان پیشنهادی در انجام آزمون فرض‌های آماری و برآورد فاصله‌ای تابعی از پارامترها است. به‌عنوان مثال برای انجام آزمون آماری دوطرفه $H_۰:(\alpha,\tau)=(\alpha_۰,\tau_۰)$ در مقابل $H_۱:(\alpha,\tau)\neq(\alpha_۰,\tau_۰)$ در سطح معنی‌داری γ، اگر ناحیه اطمینان توأم $۱۰۰(۱-\gamma)\%$ به دست آمده شامل $(\alpha_۰,\tau_۰)$ باشد فرض $H_۰$ در سطح γ رد نمی‌شود.

از طرفی به کمک ناحیه اطمینان فوق می‌توان فواصل اطمینان را برای توابعی از پارامترهای توزیع پارتو و همچنین باندهای اطمینان هم‌زمان و نقطه‌ای را برای تابع توزیع تجمعی پارتو به دست آورد. برای مثال برای نقطه‌ی $t \geq X_{1}$، روشن است که مجموعه مقادیر $F(\mathrm{t};\alpha,\tau)$ با $(\alpha,\tau) \in \mathrm{R}^{*}_{1-\gamma}$ یک‌فاصله اطمینان را برای تابع توزیع تجمعی پارتو در t تشکیل می‌دهد. به همین ترتیب، باند اطمینان هم‌زمان را برای تابع توزیع تجمعی پارتو به کمک کوچک‌ترین ناحیه اطمینان $\mathrm{R}^{*}_{1-\gamma}$ می‌توان به‌صورت زیر تعریف کرد.

$$B^{*}_{1-\gamma} = \left\{(\mathrm{x}, \mathrm{F}(\mathrm{x};\alpha,\tau)) : (\alpha,\tau) \in \mathrm{R}^{*}_{1-\gamma}, x > 1/\tau\right\},$$

که معادل با مجموعه زیر است:

$$B^{*}_{1-\gamma} = \left\{(\mathrm{x}, [1-(\tau x)^{-\alpha}]\mathrm{I}(x > 1/\tau)) : \mathrm{c}^{*}_{1} < \tau < \mathrm{c}^{*}_{2}, \frac{b^{*}_{1}}{U_{2}(\tau)} < \alpha < \frac{b^{*}_{2}}{U_{2}(\tau)}, x > 1/\tau\right\}$$

که $I(.)$ بیانگر تابع نشانگر است.

اینک برای انجام آزمون آماری برای داده‌های مربوط به زمان‌های خرابی در مثال ۱، با توجه به اینکه بر اساس طرح ۱ (نمونه‌ی کامل) برآوردهای ماکسیمم درست نمایی α و τ عبارتند از $\hat{\alpha} = ۰/۵۱$ و $\hat{\tau} = ۲/۸۵$، فرض کنید بخواهیم آزمون آماری $H_{0} : \alpha = ۰/۵, \tau = ۳$ را در مقابل $H_{1} : \alpha \neq ۰/۵, \tau \neq ۳$ در سطح ۵٪ انجام دهیم. با توجه به اینکه ناحیه اطمینان ۹۵٪ برای (α,τ) بر اساس داده‌های فوق عبارت است از:

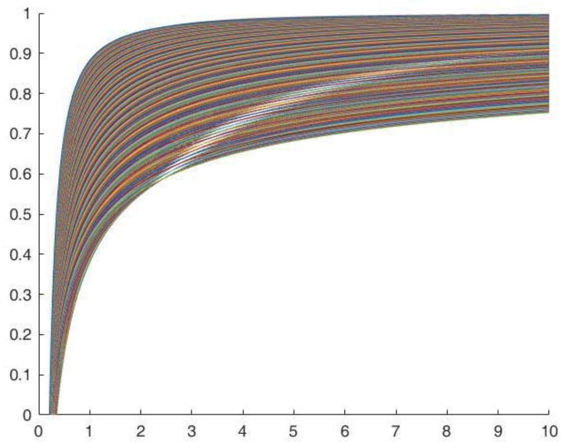
$$R^{*}_{۰/۹۵} = \left\{(\alpha,\tau) : ۲/۸۵۷۱۴ < \tau < ۴/۸۱۵۱۲, \frac{۱۳/۳۶۲۲}{U_{2}(\tau)} < \alpha < \frac{۵۰/۷۰۰۲}{U_{2}(\tau)}\right\}$$

و این ناحیه اطمینان نقطه‌ی $(۰/۵,۳)$ را در بردارد، فرض H_{0} در سطح ۵٪ رد نمی‌شود.

از طرفی باند اطمینان ۹۵٪ برای تابع توزیع پارتو عبارت است از

$$B^{*}_{۰/۹۵} = \left\{(\mathrm{x}, [1-(\tau x)^{-\alpha}]\mathrm{I}(x > 1/\tau)) : ۲/۸۵۷۱۴ < \tau < ۴/۸۱۵۱۲, \frac{۱۳/۳۶۲۲}{U_{2}(\tau)} < \alpha < \frac{۵۰/۷۰۰۲}{U_{2}(\tau)}\right\}.$$

شکل ۴ باند اطمینان مربوطه را نشان می‌دهد.



شکل (۴): باند اطمینان ۹۵٪ برای تابع توزیع پارتو بر اساس داده‌های مثال ۱، سانسور طرح ۱

سپاس‌گزاری

نویسندگان مقاله از داوران محترم برای ارائه پیشنهادهای سازنده‌شان در بهبود این مقاله سپاس‌گزاری می‌نمایند.

منابع

[1] Johnson, N.L.and Kotz, Balakrishnan, N. (1994). *Continuous Uinvariare Distributions*, Vol. 1, 2nd ed. John Wiley & Sons, New York.

[2] Fernandez, A.J. (2012). Minimizing the area of a Pareto confidence region, *European Journal of Operational research*, **221**, 205–212.

[3] Fernandez, A.J. (2013). Smallest Pareto confidence regions and applications, *Computational Statistics and Data Analysis*, **62**, 11–25.

[4] Fernandez, A.J. (2014). Computing optimal confidence sets for Pareto models under progressive censoring, *Journal of Computational and Applied Mathematics*, **258**, 168–180

[5] Asgharzadeh, A., Abdi, M. and Kus, C. (2011). Interval estimation for the two-parameter Pareto distribution based on record values, *Selçuk Journal of Applied Mathematics*, 149–161. Special Issue.

[6] Asgharzadeh, A., Fernandez, A.J. and Abdi, M. (2017). Confidence sets for two-parameter Rayleigh distribution under progressive censoring, *Applied Mathematical modelling*, **47**, 656–667.

[7] Balakrishnan, N., and Aggarwala, R. (2000). *Progressive censoring: Theory, Methods and Applications*, Birkhauser Publishers, Boston.

[8] Balakrishnan, N. and Cramer, E. (2014). *The Art of Progressive Censoring*, Springer, New York.

[9] Kus, C. and Kaya, M.F. (2007). Estimation for the parameters of the Pareto distribution under progressive censoring, *Communications in Statistics - Theory and Methods*, **36**, 1359–1365.

[10] Nelson, W. B. (1970). Statistical methods for accelerated life test datathe inverse power law model, *General Electric Co. Tech. Rep.71-C011, New York: Schenectady*.

[11] Asgharzadeh, A., Mohammadpour, M. and Ganji, Z. M. (2014). Estimation and reconstruction Based on Left Censored Data from Pareto Model, *Journal of Iranian Statistical Society*, **13**, 151–175.

[12] Parsi, S., Ganjali, M. and Sanjari Farsipour, N. (2010). Simultaneous confidence intervals for the parameters of Pareto distribution under progressive censoring, *Communications in Statistics-Theory and Methods*, **39**, 94–106.

A Constrained Optimization Problem to Determine the Smallest Pareto Confidence Region Under Progressive Type-II Censoring

Marjan Zare and Akbar Asgharzadeh

Department of Statistics, University of Mazandaran, Babolsar, Iran.

Abstract

In this paper, a constrained optimization problem is formulated and solved to determine the smallest joint confidence region for Pareto parameters based on the progressively Type-II censored samples. The objective function is the area of the confidence region and the problem constraint is the specified confidence level. The proposed joint confidence region is also valid for the complete samples and right censored samples. The area of the smallest proposed confidence region and the area of the balanced confidence region are compared. Finally, two numerical examples are presented to describe the proposed optimization method.

***Keywords*:** Joint confidence regions, progressive censoring, Lagrangian method, Pareto distribution.

Subject Classification (2010): 62N01, 62F30, 62F25.