

نواحی اطمینان توأم

اکبر اصغرزاده^۱، حامد یحیایی^۲، موسی عبدی^۳

چکیده:

فاصله اطمینان یکی از مهم ترین مباحث در آمار ریاضی است که ارتباط تنگاتنگی با آزمون فرض آماری دارد. در مبحث فاصله اطمینان، هدف پیدا کردن فاصله ای تصادفی است که پارامتر مجهول را با احتمال زیاد در برداشته باشد. فاصله اطمینان و انواع آن به طور گسترده در بسیاری از کتاب های آمار ریاضی مورد بحث قرار گرفته است. از آنجایی که بیشتر توزیع های آماری، بیش از دو پارامتر دارند لذا پیدا کردن ناحیه اطمینان توأم یا فواصل اطمینان هم زمان برای جفت پارامترها حائز اهمیت است. در این مقاله ضمن معرفی ناحیه اطمینان توأم، چگونگی بدست آوردن آن در قالب چند مثال تشریح می شود.

واژه های کلیدی: کمیت محوری، ناحیه اطمینان دقیق، ناحیه اطمینان تقریبی.

۱ مقدمه

بدست می آید. چون کمیت های محوری

$$Q_1 = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1),$$

$$Q_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi_{(n-1)}^2,$$

مستقل می باشند، لذا با استفاده از جداول استاندارد می توانیم مقادیر a ، b و c را طوری پیدا کنیم که

$$P \left[-a < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < a \right] = 1 - \gamma_1,$$

$$P \left[b < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} < c \right] = 1 - \gamma_2,$$

که a صدک بالایی $\frac{\gamma_1}{2}$ توزیع نرمال استاندارد و b و c به ترتیب صدک های بالایی و پایینی $\frac{\gamma_2}{2}$ توزیع خی دو با $n - 1$ درجه آزادی می باشند. چون دو کمیت محوری مستقل هستند می توانیم بنویسیم:

$$P \left[-a < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < a, b < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} < c \right] \\ = (1 - \gamma_1)(1 - \gamma_2).$$

با حل گزاره احتمالی فوق بر حسب μ و σ^2 یک ناحیه اطمینان

توأم با ضریب اطمینان $(1 - \gamma_1)(1 - \gamma_2)$ برای (μ, σ^2) ناحیه $R(\underline{X})$

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از جامعه ای با تابع چگالی احتمال $f(x; \theta_1, \theta_2)$ باشد که θ_1 و θ_2 پارامترهای مجهول می باشند. یک ناحیه اطمینان توأم $100(1 - \gamma)\%$ برای (θ_1, θ_2) یک ناحیه تصادفی $R(\underline{X})$ می باشد به طوری که

$$P[(\theta_1, \theta_2) \in R(\underline{X})] = 1 - \gamma.$$

ناحیه اطمینان توأم تعمیمی دو بعدی از فاصله اطمینان است که مجموعه ای از نقاط در یک فضای دو بعدی می باشد که اغلب به عنوان یک بیضی حول نقطه برآورد شده $(\hat{\theta}_1, \hat{\theta}_2)$ نشان داده می شود، هر چند به شکل های دیگر نیز می تواند باشد. ناحیه اطمینان توأم به طور مفصل در کتاب های استاندارد آمار ریاضی مورد بحث قرار نگرفته است. تنها توصیف از ناحیه اطمینان توأم، در کتب آمار ریاضی، ناحیه اطمینان توأم برای میانگین و واریانس توزیع نرمال می باشد که در مرجع [۴] مورد بحث قرار گرفته است.

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع نرمال با پارامترهای μ و σ^2 باشند، به طوری که $\mu \in R$ و $\sigma^2 \in R^+$. یک ناحیه اطمینان توأم $100(1 - \gamma)\%$ برای (μ, σ^2) بصورت زیر

^۱عضو هیئت علمی گروه آمار دانشگاه مازندران، a.asgharzadeh@umz.ac.ir

^۲کارشناسی ارشد آمار دانشگاه مازندران، yahyaeihamed@yahoo.com

^۳عضو هیأت علمی گروه آمار مجتمع آموزش عالی بم، me.abdi@bam.ac.ir

است که

نظر اکیداً یکنوا (صعودی اکید یا نزولی اکید) باشد. روش محوری مبتنی بر دو قضیه می باشد.

قضیه ۱.۲. فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی به اندازه n از جامعه ای با تابع توزیع $F(x; \theta_1, \theta_2)$ باشد و $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ آماره های ترتیبی متناظر باشند. در این صورت

$$Y_{(1)} < Y_{(2)} < \dots < Y_{(n)},$$

که

$$Y_{(i)} = -\ln[1 - F(X_{(i)}; \theta_1, \theta_2)] \quad i = 1, 2, \dots, n,$$

آماره های ترتیبی متناظر یک نمونه تصادفی به اندازه n از توزیع نمایی استاندارد خواهند بود.

قضیه ۲.۲. اگر $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$ آماره های ترتیبی متناظر با یک نمونه تصادفی n تایی از توزیع نمایی استاندارد باشند و داشته باشیم

$$Z_1 = nY_{(1)}$$

$$Z_2 = (n-1)(Y_{(2)} - Y_{(1)})$$

⋮

$$Z_n = Y_{(n)} - Y_{(n-1)}.$$

در این صورت فواصل آماره های ترتیبی $Z_1, Z_2, \dots, Z_n$ متغیر های تصادفی مستقل و هم توزیع از توزیع نمایی استاندارد می باشند.

اثبات: برای اثبات به [۱] مراجعه شود.

براساس دو قضیه فوق می توان دو جفت کمیت محوری برای پیدا کردن ناحیه اطمینان توأم برای (θ_1, θ_2) به صورت زیر معرفی کرد. چون $Z_1, \dots, Z_n$ متغیرهای تصادفی مستقل و هم توزیع از توزیع نمایی استاندارد می باشند. بنابراین،

$$V = 2Z_1 = 2nY_{(1)},$$

دارای توزیع خی دو با ۲ درجه آزادی و

$$U = 2 \sum_{i=2}^n Z_i = 2 \left[\sum_{i=1}^n Y_{(i)} - nY_{(1)} \right],$$

$$R(\underline{X}) = \left\{ (\mu, \sigma^2) : \bar{X} - a \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + a \frac{\sigma}{\sqrt{n}}, \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{c} < \sigma^2 < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{b} \right\}.$$

توجه کنید که اگر بخواهیم ضریب اطمینان برابر $(1-\gamma)$ باشد کافی است مقادیر a, b و c به صورت زیر انتخاب شوند:

$$a = z_{\frac{1-\gamma}{2}}, \quad b = \chi^2_{(\frac{1+\gamma}{2}), (n-1)}, \quad c = \chi^2_{(\frac{1-\gamma}{2}), (n-1)}.$$

روش فوق چون مبتنی بر کمیت های محوری Q_1 و Q_2 می باشد روش محوری نامیده می شود. اما سوالی که برای بدست آوردن ناحیه اطمینان توأم بوجود می آید این است که آیا انتخاب کمیت های محوری Q_1 و Q_2 منحصر به فرد است؟ همچنین در هر مسئله آماری برای بدست آوردن ناحیه اطمینان توأم همیشه کمیت های محوری وجود دارند؟ در این مقاله این پرسش ها مورد بحث قرار خواهند گرفت و نشان داده می شود که کمیت های محوری برای محاسبه ناحیه اطمینان توأم لزوماً منحصر به فرد نیستند. همچنین نشان داده می شود که در حالت کلی نمی توان از روش محوری یک ناحیه اطمینان توأم پیدا کرد. اما در حالت های خاص در برخی از مدل های آماری با ارائه کمیت های محوری که توابعی از فواصل آماره های ترتیبی هستند می توان یک ناحیه اطمینان توأم پیدا کرد. با توجه به اینکه ناحیه اطمینان توأم کمتر در کتاب های آمار ریاضی مورد بحث قرار گرفته است. این مقاله می تواند برای آشنایی دانشجویان و محققان رشته آمار در این زمینه سودمند باشد. برای مطالعه بیشتر می توان به مراجع [۵-۱۱] و [۱۴-۱۷] مراجعه کرد.

۲ روش محوری برای پیدا کردن ناحیه اطمینان توأم

برای بدست آوردن ناحیه اطمینان توأم ابتدا باید به دنبال پیدا کردن دو کمیت محوری باشیم که یکی از آنها تنها وابسته به یک پارامتر و دیگری به هر دو پارامتر بستگی داشته باشند. همچنین کمیت محوری که به تک پارامتر وابسته است باید نسبت به پارامتر مورد

دارای توزیع خبی دو با $2(n-1)$ درجه آزادی می باشد. افزون بر این می توان دید که U و V متغیرهای تصادفی مستقل می باشند.

همچنین فرض کنید

$$T_1 = \frac{U/2(n-1)}{V/2} = \frac{\sum_{i=1}^n Y_{(i)} - nY_{(1)}}{n(n-1)Y_{(1)}},$$

و

$$T_2 = U + V = 2 \sum_{i=1}^n Y_{(i)}.$$

به راحتی دیده می شود که T_1 دارای توزیع فیشر با درجات آزادی $2(n-1)$ و 2 درجه آزادی و T_2 دارای توزیع خبی دو با $2n$ درجه آزادی می باشد. علاوه بر این به راحتی می توان ثابت کرد که T_1 و T_2 از هم مستقل اند [۴].

جفت های (U, V) و (T_1, T_2) کمیت های محوری هستند که می توانند برای پیدا کردن ناحیه اطمینان توأم مورد استفاده قرار گیرند. مثلاً براساس کمیت های محوری T_1 و T_2 داریم:

۱.۴ توزیع نمایی

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع نمایی

دو پارامتری با تابع چگالی

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

و تابع توزیع تجمعی

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

باشد. چون دامنه توزیع به پارامتر وابسته است لذا شرایط نظم در اینجا برقرار نیست، بنابراین نمی توان ناحیه اطمینان توأم تقریبی را در این توزیع پیدا کرد. اما به روش محوری می توان ناحیه اطمینان توأم را برای (μ, θ) بدست آورد. فرض کنید تصادفی فوق باشند. بنابراین طبق تعریف $Y_{(i)}$ ها در قضیه ۱ داریم:

$$Y_{(i)} = -\ln[1 - F(X_{(i)}; \alpha, \beta)] = \frac{X_{(i)} - \mu}{\theta}, \quad i = 1, 2, \dots, n$$

که آماره های ترتیبی از توزیع نمایی استاندارد هستند. لذا جفت کمیت های محوری (U, V) و (T_1, T_2) بصورت زیر بدست می آیند.

$$V = 2nY_{(1)} = \frac{2n(X_{(1)} - \mu)}{\theta},$$

$$P\left[F_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n-2, 2)} < T_1 < F_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n-2, 2)}, \chi_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n)}^2 < T_2 < \chi_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n)}^2\right] = 1 - \gamma.$$

حال اگر T_1 یا T_2 تنها به یکی از پارامترهای مجهول وابسته و نسبت به آن اکیداً یکنوا باشد و دیگری به هر دو پارامتر وابسته باشد، می توان با حل گزاره احتمالی فوق براساس θ_1 و θ_2 ناحیه اطمینان توأم را پیدا کرد.

۳ ناحیه اطمینان تقریبی

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از جامعه ای با تابع چگالی $f(x; \alpha, \beta)$ و تابع توزیع $F(x; \alpha, \beta)$ ، به طوری که پارامترهای α و β مجهول می باشند. می دانیم که برآوردگر درستنمایی ماکزیمم $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ تحت شرایط نظم دارای توزیع

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \simeq N_2 \left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \Sigma_{\alpha, \beta}^{-1} \right).$$

به طوری $\Sigma_{\alpha, \beta}^{-1}$ وارون ماتریس اطلاع فیشر می باشد و

$$\Sigma_{\alpha, \beta} = - \begin{pmatrix} E\left(\frac{\partial^2 \ln L}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \ln L}{\partial \alpha \partial \beta}\right) \\ E\left(\frac{\partial^2 \ln L}{\partial \alpha \partial \beta}\right) & E\left(\frac{\partial^2 \ln L}{\partial \beta^2}\right) \end{pmatrix}.$$

که معادل است با

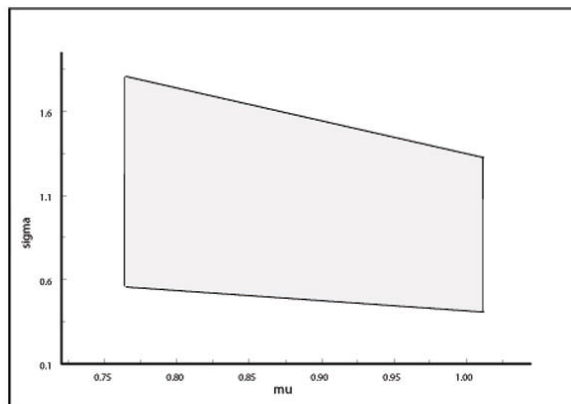
$$P \left[X_{(1)} - \frac{\theta}{2n} \chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2)} < \mu < X_{(1)} - \frac{\theta}{2n} \chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2)}, \frac{2 \sum_{i=1}^n (X_{(i)} - X_{(1)})}{\chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n-2)}} < \theta < \frac{2 \sum_{i=1}^n (X_{(i)} - X_{(1)})}{\chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n-2)}} \right] = 1 - \gamma.$$

به صورت مشابهی مشاهده می شود که T_1 تنها وابسته به پارامتر μ است که بر حسب آن اکیدا تابعی صعودی است و T_2 به هر دو پارامتر وابسته است. لذا ناحیه اطمینان توأم $(1-\gamma)100\%$ برای (μ, θ) بر حسب کمیت های محوری (T_1, T_2) بصورت زیر بدست می آید:

$$P \left[F_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n-2, 2)} < T_1 < F_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n-2, 2)}, \chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n)} < T_2 < \chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n)} \right] = 1 - \gamma,$$

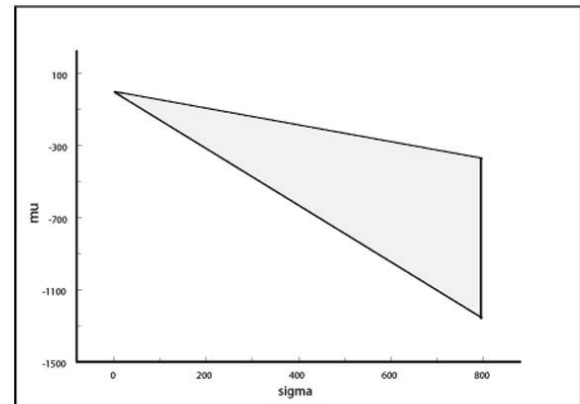
که معادل است با

$$P \left[X_{(1)} - \frac{\sum_{i=1}^n (X_{(i)} - X_{(1)})}{n(n-1)F_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n-2, 2)}} < \mu < X_{(1)} - \left(\frac{\sum_{i=1}^n (X_{(i)} - X_{(1)})}{n(n-1)F_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n-2, 2)}} \right), \frac{2 \sum_{i=1}^n (X_{(i)} - \mu)}{\chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n)}} < \theta < \frac{2 \sum_{i=1}^n (X_{(i)} - \mu)}{\chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n)}} \right] = 1 - \gamma.$$



شکل ۲. ناحیه اطمینان در توزیع نمایی با استفاده از کمیت های

محوری T_1 و T_2



شکل ۱. ناحیه اطمینان در توزیع نمایی با استفاده از کمیت های محوری U و V

$$U = 2 \left[\sum_{i=1}^n Y_{(i)} - nY_{(1)} \right] = \frac{2}{\theta} \left[\sum_{i=1}^n (X_{(i)} - X_{(1)}) \right],$$

که V دارای توزیع خبی دو با ۲ درجه آزادی و U دارای توزیع خبی دو با $2(n-1)$ درجه آزادی و U و V مستقل می باشند. همچنین

$$T_1 = \frac{\sum_{i=1}^n Y_{(i)} - nY_{(1)}}{n(n-1)Y_{(1)}} = \frac{\sum_{i=1}^n (X_{(i)} - X_{(1)})}{n(n-1)(X_{(1)} - \mu)},$$

$$T_2 = 2 \sum_{i=1}^n Y_{(i)} = 2 \sum_{i=1}^n \frac{(X_{(i)} - \mu)}{\theta},$$

که T_1 دارای توزیع فیشر با درجات آزادی $2(n-1)$ و ۲ درجه آزادی و T_2 دارای توزیع خبی دو با $2n$ درجه آزادی و T_1 و T_2 مستقل می باشند. همانطور که مشاهده می شود کمیت محوری U تنها وابسته به پارامتر θ است که بر حسب آن اکیدا نزولی است و کمیت محوری V به هر دو پارامتر وابسته است و لذا براساس جفت کمیت محوری (U, V) می توان ناحیه اطمینان توأم $(1-\gamma)100\%$ برای (μ, θ) را بصورت زیر پیدا کرد:

$$P \left[\chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2)} < \frac{2n(X_{(1)} - \mu)}{\theta} < \chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2)}, \chi^2_{(\frac{1+\sqrt{1-\gamma}}{2}), (2n-2)} < \frac{2}{\theta} \sum_{i=1}^n (X_{(i)} - X_{(1)}) < \chi^2_{(\frac{1-\sqrt{1-\gamma}}{2}), (2n-2)} \right] = 1 - \gamma,$$

حال با استفاده از یک مجموعه داده واقعی روشهای بیان شده در این بخش تشریح می شود. داده های زیر زمان های بهبودی ۱۵ نفر مبتلا به بیماری سرطان خون را نشان می دهد که مرتب شده اند. [۱۳]

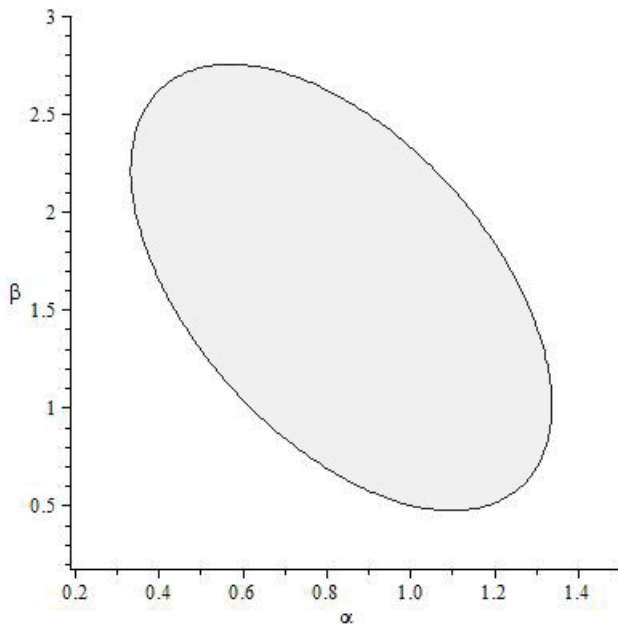
$$\Sigma_{\alpha, \beta} = \begin{pmatrix} n\psi'(\alpha) & \frac{n}{\beta} \\ \frac{n}{\beta} & \frac{n\alpha}{\beta^2} \end{pmatrix},$$

که در آن $\psi(t) = \frac{d}{dt}\Gamma(t)$ و $\psi'(t) = \frac{d}{dt}\psi(t)$ به ترتیب توابع دای گاما و تری گاما می باشند. حال فرم درجه دوم مرتبط یعنی

$$B(\alpha, \beta) = (\hat{\alpha} - \alpha, \hat{\beta} - \beta) \Sigma_{\hat{\alpha}, \hat{\beta}} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix},$$

برای n های بزرگ تقریباً دارای توزیع خبی دو با ۲ درجه آزادی است. در نتیجه یک ناحیه اطمینان تقریبی $\gamma\%(1 - \gamma)$ برای پارامترهای (α, β) عبارت است از:

$$\{(\alpha, \beta) : B(\alpha, \beta) < \chi^2_{(2)}(\gamma)\}.$$



شکل ۳. ناحیه اطمینان تقریبی در توزیع گاما

با توجه به توضیحات این بخش، یک ناحیه اطمینان تقریبی ۹۵٪ برای (α, β) با داده های زیر که یک نمونه شبیه سازی شده مرتب شده ۲۰ تایی از توزیع گاما با پارامترهای $\alpha = 1$ و $\beta = 1$ است محاسبه می شود.

0.022 , 0.036 , 0.052 , 0.170 , 0.241 , 0.271

0.459 , 0.898 , 0.948 , 1.027 , 1.146 , 1.244

1.315 , 1.779 , 1.795 , 2.101 , 2.626 , 2.787

2.831 , 2.921

1.013 , 1.034 , 1.109 , 1.169 , 1.266 , 1.509

1.533 , 1.563 , 1.716 , 1.929 , 1.965 , 2.061

2.344 , 2.546 , 2.626

در [۱۳] نشان داده شده است که توزیع نمایی دو پارامتری به این داده ها به خوبی برازش می شود. با توجه به فرمولهای بدست آمده در این بخش و همچنین با استفاده از داده های فوق یک ناحیه اطمینان توأم ۹۵٪ برای (μ, θ) بر اساس کمیت های محوری U و V بصورت زیر بدست می آید:

$$2.3317 < \theta < 797.0999,$$

$$1.013 - \frac{\theta}{30} \times 8.7362 < \mu$$

$$< 1.013 - \frac{\theta}{30} \times 0.0256$$

شکل (۱) ناحیه اطمینان توأم فوق را نشان می دهد.

همچنین براساس کمیت محوری های T_1 و T_2 ناحیه اطمینان

توأم ۹۵٪ برای (μ, θ) بصورت زیر بدست می آید

$$0.7644 < \mu < 1.0124,$$

$$\frac{50.766 - 30\mu}{49.9007} < \theta < \frac{50.766 - 30\mu}{15.3963}.$$

شکل (۲) تصویری از ناحیه اطمینان توأم فوق می باشد.

۲.۴ توزیع گاما

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه مستقل از توزیع گاما با

پارامترهای $\alpha > 0$ و $\beta > 0$ و تابع چگالی احتمال

$$f(x; \mu, \theta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \quad x > 0,$$

باشد.

برآوردگر درستنمایی ماکزیمم $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ دارای توزیع مجانبی

نرمال می باشد یعنی

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \simeq N_2 \left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \Sigma_{\alpha, \beta}^{-1} \right).$$

ابتدا با استفاده از روش های عددی، برآوردهای درستنمایی

ماکزیمم پارامترهای α و β را بدست می آوریم که عبارتند از $\hat{\alpha} = 0.8343$ و $\hat{\beta} = 1.6163$. در نتیجه ماتریس اطلاع فشر برآورد شده برابر است با

$$\Sigma_{\hat{\alpha}, \hat{\beta}} = \begin{pmatrix} 32.89 & 7.65 \\ 7.65 & 6.32 \end{pmatrix}$$

لذا فرم درجه دوم مربوطه به صورت

$$B(\alpha, \beta) = 32.89\alpha^2 + 6.32\beta^2 + 15.309\beta\alpha - 79.642\alpha - 33.419\beta + 60.232$$

به دست می آید، لذا یک ناحیه اطمینان تقریبی 95% برای (α, β) می شود

$$\{(\alpha, \beta) : B(\alpha, \beta) < \chi_2^2(0.05) = 5.991\}$$

که یک بیضی است. شکل ۳ تصویر این ناحیه را نشان می دهد. توجه کنید که به دلیل صریح نبودن تابع توزیع تجمعی برای توزیع گاما نمی توان از روش محوری در بخش ۲ برای پیدا کردن ناحیه اطمینان توأم دقیق استفاده کرد.

۳.۴ توزیع وایبل

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه مستقل از توزیع وایبل با پارامتر مقیاس $\alpha > 0$ و پارامتر شکل $\beta > 0$ و تابع توزیع تجمعی

$$F(x; \alpha, \beta) = 1 - \exp[-(\alpha x)^\beta] \quad x > 0$$

و تابع چگالی احتمال

$$f(x; \alpha, \beta) = \beta\alpha(\alpha x)^{\beta-1} \exp[-(\alpha x)^\beta] \quad x > 0$$

باشد. اگر $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ آماره های ترتیبی متناظر باشند و

$$Y_{(i)} = -\ln[1 - F(X_{(i)}; \alpha, \beta)] = (\alpha X_{(i)})^\beta, i = 1, 2, \dots, n$$

آنگاه $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$ آماره های ترتیبی از توزیع نمایی استاندارد هستند. بنابراین با توجه به آنچه در بخش ۲ بیان شد

$$V = 2Z_1 = 2nY_{(1)} = 2n(\alpha X_{(1)})^\beta$$

دارای توزیع خبی دو با ۲ درجه آزادی و

$$U = 2 \sum_{i=2}^n Z_i = 2 \left[\sum_{i=1}^n Y_{(i)} - nY_{(1)} \right] \\ = 2\alpha^\beta \left[\sum_{i=1}^n X_{(i)}^\beta - nX_{(1)}^\beta \right]$$

دارای توزیع خبی دو با $2(n-1)$ درجه آزادی می باشد که U و V مستقل هستند. همچنین

$$T_1 = \frac{U/2(n-1)}{V/2} = \frac{\sum_{i=1}^n X_{(i)}^\beta - nX_{(1)}^\beta}{n(n-1)X_{(1)}^\beta}$$

و

$$T_2 = U + V = 2 \sum_{i=1}^n (\alpha X_{(i)})^\beta,$$

که T_1 دارای توزیع فشر با $2(n-1)$ و ۲ درجه آزادی و T_2 دارای توزیع خبی دو با $2n$ درجه آزادی می باشد. چون کمیت محوری

$$T_1(\beta) = \frac{\sum_{i=1}^n X_{(i)}^\beta - nX_{(1)}^\beta}{n(n-1)X_{(1)}^\beta}$$

برحسب پارامتر β اکیدا یکنواست [۱۷ و ۱۸]، لذا برای $t > 0$ تساوی $T_1(\beta) = t$ برای هر $\beta > 0$ دارای جواب منحصر بفرد است که آن را $T^{-1}(X_{(1)}, \dots, X_{(n)}, t)$ می نامیم. همچنین کمیت

محوری

$$T_2(\alpha, \beta) = 2 \sum_{i=1}^n Y_{(i)} = 2 \sum_{i=1}^n (\alpha X_{(i)})^\beta,$$

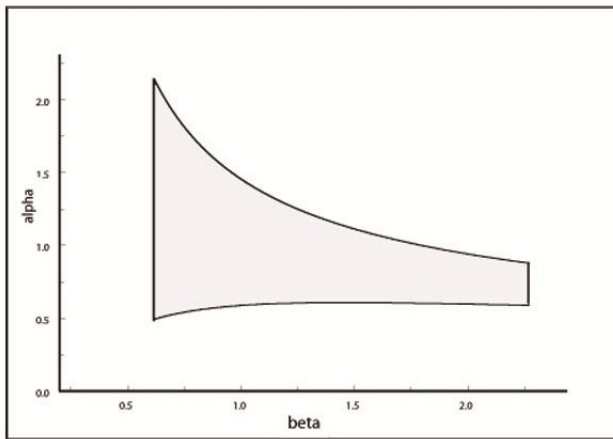
به هر دو پارامتر وابسته است بنابراین از استقلال T_1 و T_2 داریم

$$P \left[F_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n-2, 2)} < T_1 < F_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n-2, 2)}, \right. \\ \left. \chi_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n)}^2 < T_2 < \chi_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n)}^2 \right] = 1 - \gamma$$

با جایگذاری خواهیم داشت

$$P \left[F_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n-2, 2)} < \frac{\sum_{i=1}^n X_{(i)}^\beta - nX_{(1)}^\beta}{n(n-1)X_{(1)}^\beta} \right. \\ \left. < F_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n-2, 2)}, \quad \chi_{\left(\frac{1+\sqrt{1-\gamma}}{2}\right), (2n)}^2 \right. \\ \left. < 2 \sum_{i=1}^n (\alpha X_{(i)})^\beta < \chi_{\left(\frac{1-\sqrt{1-\gamma}}{2}\right), (2n)}^2 \right] = 1 - \gamma$$

که معادل است با



شکل ۴. ناحیه اطمینان دقیق در توزیع وایبل

داده های زیر یک نمونه ۲۵ تایی از توزیع وایبل با پارامترهای

$\alpha = 1$ و $\beta = 1.5$ را نشان می دهد که مرتب شده اند

0.0492 , 0.0823 , 0.1411 , 0.1521 , 0.1997 ,

0.2986 , 0.4246 , 0.4340 , 0.4512 , 0.5531 ,

0.6672 , 0.6714 , 0.7112 , 1.1088 , 1.1574 ,

1.2214 , 1.4718 , 1.4924 , 1.5256 , 1.7750 ,

1.8174 , 1.8761 , 2.0883 , 2.3444 , 3.0492

یک ناحیه اطمینان توأم دقیق برای پارامترهای α و β را می توان

بصورت نامساوی های زیر نشان داد:

$$0.6171 < \beta < 2.2711$$

$$\left\{ \frac{30.351}{2 \sum_{i=1}^{25} X_{(i)}^{\beta}} \right\}^{\frac{1}{\beta}} < \alpha < \left\{ \frac{74.974}{2 \sum_{i=1}^{25} X_{(i)}^{\beta}} \right\}^{\frac{1}{\beta}}$$

شکل (۴) ناحیه اطمینان دقیق ۹۵ درصد برای پارامترهای α و β

را نشان می دهد.

حال ناحیه اطمینان تقریبی را در نظر می گیریم. با استفاده از

روش های عددی برآوردهای درستیابی ماکزیمم پارامترهای α و

β عبارتند از $\hat{\alpha} = 0.91$ و $\hat{\beta} = 1.20$. بنابراین ماتریس اطلاع فیشر

برآورد شده می شود

$$\Sigma_{\hat{\alpha}, \hat{\beta}} = \begin{pmatrix} 31.661 & 11.615 \\ 11.615 & 43.473 \end{pmatrix}.$$

$$P \left[T^{-1}(X_{(1)}, \dots, X_{(n)}, F_{(\frac{1+\sqrt{1-\gamma}}{2}, (2n-2, 2)})} < \beta < T^{-1}(X_{(1)}, \dots, X_{(n)}, F_{(\frac{1-\sqrt{1-\gamma}}{2}, (2n-2, 2)})} \right. \\ \left. \left\{ \frac{\chi_{(\frac{1+\sqrt{1-\gamma}}{2}, (2n))}^2}{2 \sum_{i=1}^n X_{(i)}^{\beta}} \right\}^{\frac{1}{\beta}} < \alpha < \left\{ \frac{\chi_{(\frac{1-\sqrt{1-\gamma}}{2}, (2n))}^2}{2 \sum_{i=1}^n X_{(i)}^{\beta}} \right\}^{\frac{1}{\beta}} \right] \\ = 1 - \gamma.$$

توجه کنید که چون هر دو کمیت های محوری U و V به هر دو پارامتر α و β وابسته می باشند بنابراین نمی توان از آنها برای بدست آوردن ناحیه اطمینان توأم استفاده کرد.

توزیع وایبل از جمله توزیع هایی می باشد که شرایط نظم در آن برقرار است. بنابراین می توان ناحیه اطمینان تقریبی را در این توزیع بدست آورد. داریم

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \simeq N_2 \left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \Sigma_{\alpha, \beta}^{-1} \right)$$

به طوری که $([12])$,

$$\Sigma_{\alpha, \beta} = \begin{pmatrix} \frac{n}{\beta}(\psi'(1) + \psi^2(2)) & \frac{n}{\alpha}(1 - \psi(1)) \\ \frac{n}{\alpha}(1 - \psi(1)) & \frac{n\beta^2}{\alpha^2} \end{pmatrix},$$

که در آن $\psi(t) = \frac{d}{dt}\Gamma(t)$ و $\psi'(t) = \frac{d}{dt}\psi(t)$. لذا برای n های بزرگ، فرم درجه دوم مرتبط یعنی

$$A(\alpha, \beta) = (\hat{\alpha} - \alpha, \hat{\beta} - \beta) \Sigma_{\hat{\alpha}, \hat{\beta}} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix}$$

تقریباً دارای توزیع خبی دو با دو درجه آزادی است. در نتیجه یک ناحیه اطمینان تقریبی $\gamma\%(1 - \gamma)$ برای پارامترها عبارت است از:

$$\{(\alpha, \beta) : A(\alpha, \beta) < \chi_2^2(\gamma)\}$$

نتیجه گیری

در نتیجه فرم درجه دوم مربوط عبارت است از

$$A(\alpha, \beta) = (0.91 - \alpha, 1.20 - \beta) \times \begin{pmatrix} 31.661 & 11.615 \\ 11.615 & 43.473 \end{pmatrix} \times \begin{pmatrix} 0.91 - \alpha \\ 1.20 - \beta \end{pmatrix}$$

که معادل است با

$$A(\alpha, \beta) = 31.661\alpha^2 + 43.473\beta^2 + 23.231\alpha\beta - 85.500\alpha - 125.475\beta + 114.188$$

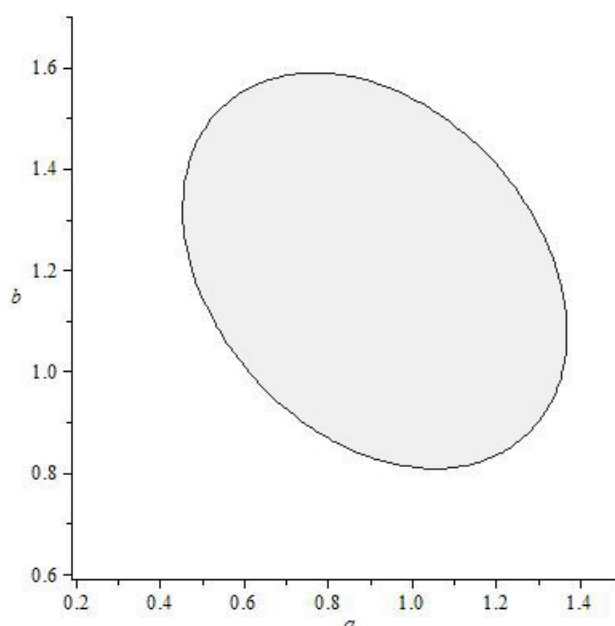
لذا ناحیه اطمینان توأم تقریبی ۹۵٪ برای (α, β) می شود

$$\{(\alpha, \beta) : A(\alpha, \beta) < \chi_2^2(0.05) = 5.991\}$$

در انتها قابل ذکر است که روش های ارائه شده در این مقاله

را می توان به سادگی برای داده های سانسور شده و رکوردی نیز

که یک بیضی است، شکل (۵) تصویری از ناحیه اطمینان تقریبی تعمیم داد. برای مطالعه بیشتر در این خصوص می توان به مراجع [۷-۹] و [۱۴-۱۷] مراجعه کرد.



شکل ۵. ناحیه اطمینان تقریبی در توزیع وایبل

مراجع

[۱] آرنولد، بی، سی،، بالاکریشنان، ان. و ناگارجا، اچ، ان، (۱۳۷۹)، نخستین درس در آماره های ترتیبی، چاپ اول، ترجمه حسنعلی

آذرنوش، انتشارات دانشگاه فردوسی مشهد.

- [۲] بهبودیان، جواد. چاپ اول، (۱۳۸۵)، آمار ریاضی، انتشارات امیرکبیر.
- [۳] پارسیان، احمد، چاپ سوم، (۱۳۸۶)، مبانی آمار ریاضی، مرکز نشر دانشگاه صنعتی اصفهان.
- [۴] مود، ا. گریبیل، ف. و بوز، د. (۱۳۸۳)، مقدمه ای بر نظریه آمار، چاپ پنجم، مشکانی علی، دانشگاه فردوسی مشهد.
- [5] Arnold, B. C. and Shavelle, R. M. (1998). Joint confidence sets for the mean and variance of a normal distribution. *The American Statistician*, **52**, 133-140.
- [6] Asgharzadeh, A., Abdi, M. (2008). Point and Interval Estimation for the Burr Type III Distribution, *Journal of Statistical Research of Iran*, **5**, 147-159.
- [7] Asgharzadeh, A., Abdi, M. (2011). Confidence intervals and joint confidence regions for the two-parameter exponential distribution based on records. *Communications of the Korean Statistical Society*, **18(1)**, 103-110.
- [8] Asgharzadeh, A., Abdi, M. and Kus, C. (2011). Interval Estimation for the Two-Parameter Pareto Distribution Based on Records, *Selçuk Journal of Applied Mathematics, Special Issue*, 149-161.
- [9] Asgharzadeh, A., Abdi, M. and Wu, S-J. (2015). Interval Estimation for the Two-Parameter Bathtub-Shaped Lifetime Distribution Based on Records, *Hacettepe Journal of Mathematics and Statistics*. **44(2)**, 399-416.
- [10] Chen, Z. (1996). Joint confidence region for the parameters of Pareto distribution. *Metrika*, **44**, 191-197.
- [11] Chen, Z. (1998). Joint estimation for the parameters of weibull distribution. *Journal of Statistical Planning and Inference*, **66**, 113-120.
- [12] Gupta, P. L. and Gupta, C. R. (1998). Numerical methods for maximum likelihood estimation of weibull parameters. *Journal of Statistical Computation and Simulation*, **62**, 1-7.
- [13] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley Inc.
- [14] 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(1)**, 94-106.
- [15] Wu, S-J., (2003). Estimation for the two-parameter Pareto distribution under progressive censoring with uniform removals. *Journal of Statistical Computation and Simulation*, **73**, 125-134.
- [16] Wu, S-F., (2010). Interval estimation for the Pareto distribution based on a progressive Type II censored sample, *Journal of Statistical Computation and Simulation*, **80(4)**, 463-474.
- [17] Wu, S-J. (2002). Estimation of the parameters of the Weibull distribution with progressive censored data, *Journal Japan Statistics Society*, **32(2)**, 155-163.