

برآوردیابی انقباضی کلاسیک و بیزی در توزیع رایلی با استفاده از حدس نقطه‌ای براساس داده‌های سانسور شده

آزاده کیاپور*؛ دانشگاه آزاد اسلامی، واحد بابل، دانشکده علوم پایه، گروه آمار

پذیرش ۹۶/۱۰/۱۷

دریافت ۹۴/۱۱/۱۰

چکیده

در این مقاله، چند آزمون انقباضی و یک برآوردگر انقباضی بیزی برای پارامتر مقیاس توزیع طول عمر رایلی براساس نمونه‌های سانسور شده تحت تابع زیان توان دوم خطای پایای مقیاس ارائه می‌شود. بهترین برآوردگر خطی نارایب و مخاطره آن تحت تابع زیان مورد بررسی محاسبه می‌شود. آزمون‌های انقباضی براساس رد یا پذیرش یک فرضیه صفر مبنی بر تساوی پارامتر واقعی با یک مقدار حدسی معرفی و مخاطره آن‌ها به دست می‌آیند. برای مقایسه عملکرد آزمون‌های انقباضی و بهترین برآوردگر خطی نارایب، کارایی نسبی بین آن‌ها محاسبه می‌شود. با استفاده از یک روش بیزی، برآوردگر انقباضی بیزی را به دست آورده و رفتار آن را نسبت به بهترین برآوردگر خطی نارایب با محاسبه کارایی نسبی مورد بررسی قرار می‌دهیم. یک مثال عددی برای تشریح برآوردگرهای پیشنهادی ارائه شده است.

واژه‌های کلیدی: برآوردگر انقباضی، تابع زیان توان دوم خطای پایای مقیاس، توزیع رایلی

مقدمه

توزیع رایلی یکی از توزیع‌های طول عمر است که به طور گسترده‌ای در حوزه‌های مختلف آمار به کار می‌رود. فرض کنید X دارای توزیع رایلی تک پارامتری با تابع چگالی

$$f(x|\sigma) = \frac{x}{\sigma} \exp\left\{-\frac{x^2}{2\sigma}\right\}, \quad x > 0, \sigma > 0, \quad (1)$$

و تابع توزیع

$$F(x|\sigma) = 1 - \exp\left\{-\frac{x^2}{2\sigma}\right\}, \quad x > 0, \sigma > 0,$$

است. تابع نرخ خطر برای مدل رایلی به صورت

$$R(t) = \frac{f(t|\sigma)}{1-F(t|\sigma)} = \frac{\frac{t}{\sigma} \exp\left\{-\frac{t^2}{2\sigma}\right\}}{\exp\left\{-\frac{t^2}{2\sigma}\right\}} = \frac{t}{\sigma},$$

محاسبه می‌شود، که نسبت به t یک تابع صعودی است، در حالی که در توزیع نمایی مقداری ثابت است. بنابراین توزیع رایلی برای مدل‌بندی توزیع طول عمر مؤلفه‌هایی که با زمان افزایش می‌یابند مناسب است. علاقه‌مندان می‌توانند به پولوگو [2] برای یافتن مثال‌های مناسب مراجعه کنند.

یک طرح سانسور نوع دوم را در نظر بگیرید که در آن فقط r تا از کوچک‌ترین مشاهده‌ها از یک نمونه تصادفی n تایی قابل مشاهده هستند. در این حالت تمام مؤلفه‌ها در آزمایش قرار می‌گیرند اما به‌جای آن که آزمایش تا پایان شکست تمام مؤلفه‌ها ادامه داده شود، این کار تا زمانی انجام می‌شود که $r(\leq n)$ تا از مؤلفه‌ها از کار بیفتند. در

این صورت نمونه سانسور شده نوع دوم به صورت $\mathbf{x} = (x_{(1)}, \dots, x_{(r)})$ با تابع درست‌نمایی

$$L(\sigma | \mathbf{x}) = \frac{n! \prod_{i=1}^r x_{(i)}}{(n-r)! \sigma^r} \exp \left\{ -\frac{\sum_{i=1}^r x_{(i)}^\gamma + (n-r)x_{(r)}^\gamma}{\gamma \sigma} \right\}, \quad 0 < x_{(1)} < x_{(r)} < \dots < x_{(r)}. \quad (2)$$

است [3]. در نتیجه، با حل معادله $\frac{\partial \ln L(\sigma | \mathbf{x})}{\partial \sigma} = 0$ ، برآوردگر بیشینه درست‌نمایی σ برابر است با

$$\hat{\sigma} = \frac{\sum_{i=1}^r X_{(i)}^\gamma + (n-r)X_{(r)}^\gamma}{\gamma r}. \quad (3)$$

در بعضی از موقعیت‌های کاربردی، محقق دارای اطلاعات پیشین درباره پارامتر نامعلوم σ به صورت یک حدس یا گمان σ_0 است. برآوردهای انقباضی برای پارامتر نامعلوم σ وقتی که مقدار حدسی σ_0 برای σ در اختیار باشد، راتامسون [4] به صورت (۴) تعریف کرده است:

$$\hat{\sigma}_S = k \hat{\sigma} + (1-k)\sigma_0, \quad 0 \leq k \leq 1, \quad 4$$

که در آن $\hat{\sigma}$ برآوردگر بیشینه درست‌نمایی σ و k ضریب انقباضی است که محقق با توجه به عقیده‌اش نسبت به مقدار حدسی σ_0 تعیین می‌کند. مقادیر k نزدیک به یک نشان‌دهنده گرایش برآوردگر به نمونه و مقادیر k نزدیک به صفر نشان‌دهنده گرایش برآوردگر به مقدار حدسی است. در صورتی که مقدار حدسی پارامتر به مقدار واقعی آن نزدیک باشد، برآوردهای انقباضی رفتار بهتری نسبت به برآوردهای معمول مانند برآوردهای بیشینه درست‌نمایی از خود نشان می‌دهند به این مفهوم که دارای مخاطره کم‌تری هستند. معمولاً برای این که تشخیص دهیم که مقدار حدسی پارامتر به مقدار واقعی آن نزدیک است یا خیر، یک آزمون مقدماتی برای آزمون فرضیه

$$\begin{cases} H_0 : \sigma = \sigma_0 \\ H_1 : \sigma \neq \sigma_0 \end{cases}$$

انجام می‌شود. با توجه به رد یا پذیرش $H_0 : \sigma = \sigma_0$ می‌توان برآوردهای جدیدی ساخت که از آن‌ها به‌عنوان آرمورد یاد می‌شود. برای بررسی بیش‌تر به پراکاش و سینگ [5] و سینگ و همکاران [6] مراجعه شود.

تابع‌های زیان توان دوم خطا و لاینکس پایای مکان هستند و برای برآورد پارامتر مکان مناسب هستند. تابع زیان

توان دوم خطای پایای مقیاس به صورت

$$L(\sigma, \delta) = \left(\frac{\delta}{\sigma} - 1 \right)^2 \quad (5)$$

یک تابع زیان محدب و دارای نقطه کمینه در نقطه $\delta = \sigma$ و مناسب برای برآورد پارامتر مقیاس σ است. نعمت‌الهی و کیاپور [1] به برآوردیابی بییزی استوار تحت تابع زیان (δ) پرداختند. در این مقاله، آزموردهای انقباضی با توجه به رد یا پذیرش فرضیه $H_0: \sigma = \sigma_0$ معرفی و مخاطره آن‌ها به دست می‌آیند. سپس، برای مقایسه عملکرد آزموردهای انقباضی معرفی شده و بهترین برآوردگر خطی نارایب، کارایی نسبی بین آن‌ها تحت تابع زیان (δ) محاسبه می‌شود. با استفاده از یک روش بییزی، برآوردگر انقباضی بییزی را به دست آورده و رفتار آن را نسبت به بهترین برآوردگر خطی نارایب با محاسبه کارایی نسبی بین آن‌ها بررسی می‌کنیم. یک مثال عددی برای تشریح برآوردگرهای پیشنهادی ارائه شده است. در پایان، به بحث و نتیجه‌گیری می‌پردازیم.

بهترین برآوردگر نارایب خطی

یک برآوردگر خطی به صورت $c\hat{\sigma}$ برای برآورد σ در نظر بگیرید که در آن $\hat{\sigma}$ برآوردگر بیشینه درست‌نمایی σ داده شده در (۳) است. با توجه به این که $\chi_{r, \alpha}^2 \sim \chi_{r, \alpha}^2 / \sigma$ ، می‌توان نشان داد

$$E \left[\frac{\hat{\sigma}}{\sigma} \right] = \frac{r+1}{r}, \quad E \left[\left(\frac{\hat{\sigma}}{\sigma} \right)^2 \right] = 1.$$

در نتیجه، مخاطره برآوردگر $c\hat{\sigma}$ تحت تابع زیان (δ) برابر است با

$$\begin{aligned} R(\sigma, c\hat{\sigma}) &= E \left[\frac{c\hat{\sigma}}{\sigma} - 1 \right]^2 = c^2 E \left[\frac{\hat{\sigma}}{\sigma} \right]^2 - 2c E \left[\frac{\hat{\sigma}}{\sigma} \right] + 1 \\ &= c^2 \frac{r+1}{r} - 2c + 1 \end{aligned}$$

مشتق اول و دوم مخاطره برآوردگر $c\hat{\sigma}$ برابر است با

$$\frac{\partial R(\sigma, c\hat{\sigma})}{\partial \sigma} = 2c \frac{r+1}{r} - 2, \quad \frac{\partial^2 R(\sigma, c\hat{\sigma})}{\partial \sigma^2} = \frac{2(r+1)}{r} > 0.$$

بنابراین مخاطره برآوردگر $c\hat{\sigma}$ تابعی محدب نسبت به c و دارای نقطه کمینه در $C = C_1$ به صورت

$$C_1 = \frac{r}{r+1}$$

است. همچنین مخاطره برآوردگر $C_1\hat{\sigma}$ برابر است با

$$R(\sigma, C_1\hat{\sigma}) = C_1^2 \frac{r+1}{r} - 2C_1 + 1 = \frac{1}{r+1} \quad (6)$$

لازم به یادآوری است که با توجه به تعریف لهنمن (۱۹۵۱) برآوردگر δ را تحت تابع زیان (۳) برای σ مخاطره نارایب گویند اگر در شرط

$$\frac{E(\delta^*)}{E(\delta)} = \theta.$$

صدق کند. به راحتی می‌توان نشان داد

$$E(c_1 \hat{\sigma})^2 = \frac{r}{r+1} \sigma^2, \quad E(c_1 \hat{\sigma}) = \frac{r}{r+1} \sigma$$

با قراردادن روابط مذکور در شرط مخاطره نارایی، نتیجه می‌شود که برآوردگر $c_1 \hat{\sigma}$ تحت تابع زیان (۵) برای پارامتر σ مخاطره ناریب است. در نتیجه می‌توان گفت که برآوردگر $c_1 \hat{\sigma}$ یک برآوردگر مخاطره ناریب با کم‌ترین مخاطره در کلاس برآوردگرهای به فرم $c \hat{\sigma}$ یعنی بهترین برآوردگر ناریب خطی است.

آزماوردهای انقباضی

برآوردگر انقباضی تعریف شده در (۴) را در نظر بگیرید. با توجه به رد یا پذیرش فرضیه $H_0: \sigma = \sigma_0$ می‌توان آزماوردهای انقباضی را ساخت به این ترتیب که اگر $H_0: \sigma = \sigma_0$ در سطح معنی‌داری α پذیرفته شود، آزماوردهای انقباضی برابر $k \hat{\sigma} + (1-k)\sigma_0$ و در غیر این صورت برابر $c_1 \hat{\sigma}$ خواهد بود. برای آزمون فرضیه $H_0: \sigma = \sigma_0$ می‌توان از کمیت محوری $\sqrt{2r} \hat{\sigma} / \sigma \sim \chi_{2r}^2$ استفاده کرد. تحت $H_0: \sigma = \sigma_0$ داریم

$$P\left(q_1 \leq \frac{\sqrt{2r} \hat{\sigma}}{\sigma} \leq q_2\right) = 1 - \alpha,$$

که در آن $q_1 = \chi_{\frac{\alpha}{2}, 2r}^2$ و $q_2 = \chi_{\frac{1-\alpha}{2}, 2r}^2$ به ترتیب چندک‌های $\alpha/2$ و $1-\alpha/2$ توزیع کی‌دو با $2r$ درجه آزادی

هستند. در نتیجه آزماوردهای انقباضی معرفی شده را می‌توان به صورت

$$\hat{\sigma}_{ST} = \begin{cases} k \hat{\sigma} + (1-k)\sigma_0 & t_1 \leq \hat{\sigma} \leq t_2 \\ c_1 \hat{\sigma} & \hat{\sigma} < t_1 \text{ or } \hat{\sigma} > t_2. \end{cases} \quad (7)$$

نوشت، که در آن $t_1 = \frac{q_1 \sigma_0}{\sqrt{2r}}$ و $t_2 = \frac{q_2 \sigma_0}{\sqrt{2r}}$. مخاطره آزماوردهای انقباضی (۷) برابر است با

$$\begin{aligned} R(\sigma, \hat{\sigma}_{ST}) &= E\left[\left(\frac{k \hat{\sigma} + (1-k)\sigma_0}{\sigma} - 1\right)^2 I(t_1 \leq \hat{\sigma} \leq t_2)\right] + E\left[\left(\frac{c_1 \hat{\sigma}}{\sigma} - 1\right)^2 I(\hat{\sigma} < t_1 \text{ or } \hat{\sigma} > t_2)\right] \\ &= \int_{w_1}^{w_2} (kw + (1-k)\delta_0 - 1)^2 g(w) dw - \int_{w_1}^{w_2} (c_1 w - 1)^2 g(w) dw + (r+1)^{-1} \end{aligned} \quad (8)$$

که در آن $\delta_0 = \frac{\sigma_0}{\sigma}$ ، $w_1 = \frac{q_1 \delta_0}{\sqrt{2r}}$ ، $w_2 = \frac{q_2 \delta_0}{\sqrt{2r}}$ ، $W = \frac{\hat{\sigma}}{\sigma} \sim \text{Gamma}(r, r^{-1})$ و $g(w)$ تابع چگالی W به صورت

$$g(w) = \frac{r^r w^{r-1} e^{-rw}}{\Gamma(r)}, \quad w > 0.$$

است. مخاطره آزماوردهای انقباضی در (۸) با استفاده از روش‌های عددی در نرم‌افزار R 3.1.2 برای مقادیر مختلف k ، α ، δ_0 و r محاسبه می‌شود.

انتخاب مقدار ضریب انقباضی k از مسائل مهم در تعیین آزماوردهای انقباضی است. در این‌جا ۴ مقدار مختلف برای k ارائه می‌دهیم که با استفاده از روش‌های متفاوتی به دست می‌آیند.

روش اول. یکی از روش‌های به دست آوردن مقداری برای ضریب انقباض k ، مقداری از k است که مخاطره برآوردگر

انقباضی $\hat{\sigma}_S$ داده شده در (۴) را کمینه کند. مخاطره برآوردگر انقباضی $\hat{\sigma}_S$ برابر است با

$$\begin{aligned}
 R(\sigma, \hat{\sigma}_s) &= E \left[\frac{k \hat{\sigma} + (1-k)\sigma}{\sigma} - 1 \right]^r \\
 &= E \left[\frac{(k \hat{\sigma})^r + (1-k)^r \sigma^r}{\sigma^r} \right] - r E \left[\frac{k \hat{\sigma} + (1-k)\sigma}{\sigma} \right] + 1 \\
 &= k^r E \left(\frac{\hat{\sigma}}{\sigma} \right)^r + (1-k)^r \left(\frac{\sigma}{\sigma} \right)^r + r k (1-k) \frac{\sigma}{\sigma} E \left(\frac{\hat{\sigma}}{\sigma} \right) - r k E \left(\frac{\hat{\sigma}}{\sigma} \right) - r (1-k) \frac{\sigma}{\sigma} + 1 \\
 &= k^r \left(\frac{r+1}{r} + \delta^r - r \delta \right) - (r k - 1)(\delta - 1)^r
 \end{aligned}$$

با مشتق‌گیری از تابع مخاطره فوق و مساوی قرار دادن آن با صفر، مقداری از k که مخاطره مذکور را کمینه می‌کند برابر است با

$$k_1 = k_{\min} = \frac{r(\delta - 1)^r}{r(\delta - 1)^r + 1} \quad (9)$$

هم‌چنین $\frac{\partial^r R}{\partial^r k} \Big|_{k=k_{\min}} > 0$ دقت کنید مقدار k_{\min} بین صفر و یک است و می‌تواند به‌عنوان ضریب انقباضی در نظر گرفته شود. توجه به این نکته، مهم است که مقدار ضریب انقباضی k_1 به پارامتر نامعلوم σ بستگی دارد و این در مسائل کاربردی امکان‌پذیر نیست. بنابراین در عمل، \hat{k}_1 با توجه به خاصیت پایایی برآوردگر بیشینه درست‌نمایی و با جای‌گذاری $\hat{\sigma}$ به‌جای σ به‌دست می‌آید.

روش دوم. تحت فرضیه $H_0: \sigma = \sigma_0$ ، رابطه

$$q_1 \leq \frac{r \hat{\sigma}}{\sigma_0} \leq q_2, \quad (10)$$

هم‌ارز با رابطه

$$0 \leq \frac{\frac{rW}{\delta_0} - q_1}{q_2 - q_1} \leq 1,$$

است، که در آن δ_0 و W قبلاً تعریف شده‌اند. بنابراین مقدار

$$k_r = \frac{\frac{rW}{\delta_0} - q_1}{q_2 - q_1}, \quad (11)$$

می‌تواند به‌عنوان ضریب انقباضی در نظر گرفته شود.

روش سوم. اگر فرضیه $H_0: \sigma = \sigma_0$ در سطح معنی‌داری α پذیرفته شود، آن‌گاه از رابطه (10) و با گرفتن امید ریاضی از طرفین آن، داریم $q_1 \leq r \leq q_2$ و در نتیجه $\frac{q_1}{r} \leq 1$. با در نظر گرفتن $\frac{q_1}{r} \approx 1$ می‌توان به ضریب انقباضی

کوچک‌تری دست یافت. از رابطه (11) داریم:

$$\frac{\frac{rW}{\delta_0} - q_1}{q_2 - q_1} = r \frac{\frac{W}{\delta_0} - \frac{q_1}{r}}{q_2 - q_1} \approx r \frac{\frac{W}{\delta_0} - 1}{q_2 - q_1}. \quad (12)$$

صورت آخرین عبارت رابطه (۱۲) ممکن است منفی باشد. بنابراین قدرمطلق آن را به‌عنوان ضریب انقباضی k_r به‌صورت (۱۳) در نظر می‌گیریم:

$$k_r = \frac{2r}{q_r - q_1} \left| \frac{W}{\delta} - 1 \right|. \quad (13)$$

روش چهارم. مقدار $c = c_1$ که مخاطره برآوردگر به فرم $c\hat{\sigma}$ را کمینه می‌کند و برابر است با

$$k_r = c_1 = \frac{r}{r+1}, \quad (14)$$

می‌تواند به‌عنوان ضریب انقباضی در نظر گرفته شود.

ضرایب انقباضی k_i ، $i = 1, 2, 3, 4$ ، آزمایه‌های انقباضی $\hat{\sigma}_{STi}$ ، $i = 1, 2, 3, 4$ را فراهم می‌کنند. مخاطره آزمایه‌های انقباضی $\hat{\sigma}_{STi}$ نیز با قراردادن مقادیر ضرایب انقباضی k_i متناظر در رابطه (۸) محاسبه می‌شوند. برای مقایسه آزمایه‌های انقباضی معرفی شده و برآوردگر مخاطره ناریب با کم‌ترین مخاطره $c_1\hat{\sigma}$ ، کارایی نسبی بین آن‌ها به‌صورت (۱۵) محاسبه می‌شود:

$$RE(\hat{\sigma}_{STi}, c_1\hat{\sigma}) = \frac{R(\sigma, c_1\hat{\sigma})}{R(\sigma, \hat{\sigma}_{STi})}, \quad i = 1, 2, 3, 4. \quad (15)$$

جدول‌های ۱ تا ۴ کارایی نسبی (۱۵) را به‌ازای مقادیر مختلف $r = 4, 6, 8$ ، $\delta = 0.25(0.25)1/75$ و $\alpha = 0.01, 0.05, 0.1$ نشان می‌دهد. آزمایه‌های $\hat{\sigma}_{ST4}$ و $\hat{\sigma}_{ST3}$ به‌ازای $0.25 \leq \delta \leq 1$ ، آزمایه‌های $\hat{\sigma}_{ST1}$ به‌ازای $0.25 \leq \delta \leq 1/75$ و آزمایه‌های $\hat{\sigma}_{ST2}$ به‌ازای $0.25 \leq \delta \leq 1/75$ دارای کارایی بیش‌تری نسبت به $c_1\hat{\sigma}$ هستند. کارایی تمام آزمایه‌ها در نقطه یک به بیشینه خود می‌رسد. تقریباً در اکثر موارد کارایی نسبی با افزایش r کاهش می‌یابد. برای $0.25 \leq \delta \leq 1/75$ و مقادیر کوچک r ، آزمایه‌های $\hat{\sigma}_{ST1}$ نسبت به دیگر آزمایه‌ها دارای مخاطره کم‌تری است.

جدول ۱. کارایی نسبی بین $\hat{\sigma}_{ST1}$ و $c_1\hat{\sigma}$

δ								r
۱/۷۵	۱/۵	۱/۲۵	۱	۰/۷۵	۰/۵	۰/۲۵	α_{\downarrow}	
۱/۰۴۴۵	۱/۴۴۲۶	۳/۳۲۵۲	۱۳/۴۲۹۲	۲/۳۶۹۱	۱/۱۹۴۹	۱/۰۵۰۳	۰/۰۱	۴
۰/۸۶۰۷	۱/۱۵۴۹	۲/۳۱۶۶	۴/۳۹۷۷	۱/۹۱۷۲	۱/۲۰۸۶	۱/۰۴۳۴	۰/۰۵	
۰/۷۶۵۱	۰/۹۹۳۹	۱/۶۹۵۴	۲/۸۵۶۸	۱/۷۶۸۴	۱/۳۱۸۶	۱/۰۳۷۰	۰/۱	
۰/۹۵۵۷	۱/۲۴۳۱	۲/۶۰۸۷	۱۲/۹۳۰۵	۱/۹۳۵۷	۱/۰۹۱۳	۱/۰۲۳۲	۰/۰۱	۶
۰/۷۸۰۶	۰/۹۸۸۴	۱/۸۰۹۷	۴/۱۳۱۰	۱/۶۲۱۳	۱/۱۱۵۴	۱/۰۱۷۲	۰/۰۵	
۰/۷۱۸۷	۰/۸۶۸۲	۱/۴۳۴۵	۲/۶۶۴۲	۱/۵۲۷۷	۱/۱۲۷۴	۱/۰۱۳۴	۰/۱	
۰/۸۹۴۱	۱/۱۲۴۶	۲/۱۹۸۷	۱۲/۶۶۹۷	۱/۶۸۳۶	۱/۰۴۳۵	۱/۰۱۱۰	۰/۰۱	۸
۰/۷۳۹۹	۰/۸۹۷۷	۱/۵۶۷۸	۳/۹۹۵۸	۱/۴۴۸۵	۱/۰۷۱۴	۱/۰۰۶۹	۰/۰۵	
۰/۶۹۲۷	۰/۸۰۴۹	۱/۲۷۸۱	۲/۵۶۷۵	۱/۳۸۶۷	۱/۰۸۲۲	۱/۰۰۵۰	۰/۱	

جدول ۲. کارایی نسبی بین $\hat{\sigma}_{ST_2}$ و $\hat{\sigma}_1$

δ								r
۱/۷۵	۱/۵	۱/۲۵	۱	۰/۷۵	۰/۵	۰/۲۵	α_{\downarrow}	
۰/۴۱۱۵	۰/۸۰۹۷	۱/۷۲۱۰	۲/۳۷۰۷	۱/۶۹۲۳	۱/۲۸۴۸	۱/۰۹۱۳	۰/۰۱	۴
۰/۴۰۰۷	۰/۷۰۸۶	۱/۳۱۱۶	۱/۸۹۳۷	۱/۷۶۱۵	۱/۳۸۹۸	۱/۰۶۹۲	۰/۰۵	
۰/۴۰۳۹	۰/۶۵۹۲	۱/۱۳۸۷	۱/۷۱۹۶	۱/۷۹۳۶	۱/۳۹۰۸	۱/۰۵۴۷	۰/۱	
۰/۳۰۸۱	۰/۶۳۵۳	۱/۵۶۰۹	۲/۴۴۷۵	۱/۵۷۳۴	۱/۱۷۳۲	۱/۰۳۷۰	۰/۰۱	۶
۰/۳۱۶۰	۰/۵۷۳۴	۱/۱۸۹۵	۱/۸۹۰۹	۱/۶۲۰۹	۱/۲۴۳۶	۱/۰۲۴۵	۰/۰۵	
۰/۳۳۷۱	۰/۵۵۲۴	۱/۰۳۰۱	۱/۶۷۱۷	۱/۶۴۰۱	۱/۲۴۴۲	۱/۰۱۸۲	۰/۱	
۰/۲۴۶۷	۰/۵۱۷۲	۱/۴۰۲۳	۲/۴۸۳۸	۱/۴۷۴۸	۱/۱۰۸۲	۱/۰۱۵۳	۰/۰۱	۸
۰/۲۶۹۵	۰/۴۸۳۹	۱/۰۸۰۵	۱/۸۸۶۳	۱/۵۱۸۲	۱/۱۶۰۸	۱/۰۰۸۹	۰/۰۵	
۰/۳۰۳۷	۰/۴۸۳۸	۰/۹۴۵۱	۱/۶۴۴۷	۱/۵۳۵۶	۱/۱۵۸۸	۱/۰۰۶۲	۰/۱	

جدول ۳. کارایی نسبی بین $\hat{\sigma}_{ST_3}$ و $\hat{\sigma}_1$

δ								r
۱/۷۵	۱/۵	۱/۲۵	۱	۰/۷۵	۰/۵	۰/۲۵	α_{\downarrow}	
۰/۵۲۸۹	۱/۰۶۶۱	۲/۷۹۷۸	۴/۹۱۰۹	۲/۱۴۶۷	۱/۱۰۴۳	۰/۹۶۲۳	۰/۰۱	۴
۰/۵۳۴۵	۰/۹۵۷۲	۱/۹۷۴۸	۲/۹۵۳۸	۱/۸۸۹۶	۱/۱۶۴۳	۱/۰۰۰۵	۰/۰۵	
۰/۵۳۴۱	۰/۸۷۳۵	۱/۵۷۹۷	۲/۲۸۴۵	۱/۷۶۷۲	۱/۱۸۹۹	۰/۰۱۱۳	۰/۱	
۰/۴۰۱۸	۰/۸۰۸۲	۲/۳۲۵۶	۵/۲۳۳۲	۱/۸۴۲۷	۰/۹۳۰۳	۰/۹۴۸۸	۰/۰۱	۶
۰/۴۲۲۱	۰/۷۴۸۶	۱/۶۸۵۶	۲/۹۰۸۸	۱/۶۲۹۷	۱/۰۱۳۵	۰/۹۸۶۲	۰/۰۵	
۰/۴۴۲۹	۰/۷۱۰۶	۱/۳۷۳۶	۲/۱۸۶۳	۱/۵۴۵۶	۱/۰۵۶۶	۰/۹۹۶۲	۰/۱	
۰/۳۰۰۷	۰/۶۵۴۲	۱/۹۶۳۰	۵/۳۹۹۷	۱/۶۱۲۴	۰/۸۳۷۸	۰/۹۶۲۲	۰/۰۱	۸
۰/۳۶۴۶	۰/۶۲۵۹	۱/۴۷۱۵	۲/۸۷۹۱	۱/۴۵۱۸	۰/۹۳۹۱	۰/۹۸۹۵	۰/۰۵	
۰/۴۰۰۶	۰/۶۱۶۰	۱/۲۲۹۷	۲/۱۳۳۶	۱/۳۹۸۸	۰/۹۹۱۲	۰/۹۹۶۰	۰/۱	

جدول ۴. کارایی نسبی بین $\hat{\sigma}_{ST_4}$ و $\hat{\sigma}_1$

δ								r
۱/۷۵	۱/۵	۱/۲۵	۱	۰/۷۵	۰/۵	۰/۲۵	α_{\downarrow}	
۱/۰۱۶۱	۱/۱۲۳۴	۱/۲۰۵۸	۱/۲۶۱۹	۱/۳۱۴۸	۱/۳۲۳۱	۱/۰۹۴۸	۰/۰۱	۴
۰/۸۶۸۴	۱/۰۰۳۸	۱/۱۴۲۴	۱/۲۸۱۵	۱/۴۰۰۷	۱/۳۲۸۳	۱/۰۶۲۱	۰/۰۵	
۰/۷۸۳۸	۰/۹۲۲۳	۱/۰۹۲۰	۱/۳۸۸۹	۱/۴۳۴۲	۱/۳۰۳۵	۱/۰۴۷۳	۰/۱	
۰/۹۵۴۱	۱/۰۵۳۳	۱/۱۲۶۱	۱/۱۷۴۲	۱/۲۳۲۸	۱/۲۵۳۳	۱/۰۴۵۰	۰/۰۱	۶
۰/۸۰۷۴	۰/۹۳۱۵	۱/۰۵۹۳	۱/۱۸۶۴	۱/۳۰۷۸	۱/۲۳۹۲	۱/۰۲۵۱	۰/۰۵	
۰/۷۴۱۴	۰/۸۵۹۱	۱/۰۰۹۷	۱/۱۹۱۰	۱/۳۳۷۰	۱/۲۱۲۱	۱/۰۱۷۶	۰/۱	
۰/۹۰۹۳	۱/۰۱۱۵	۱/۰۸۵۰	۱/۱۳۰۴	۱/۱۹۲۷	۱/۲۱۰۰	۱/۰۲۱۰	۰/۰۱	۸
۰/۷۷۵۲	۰/۸۹۱۹	۱/۰۱۷۰	۱/۱۳۹۳	۱/۲۶۰۵	۱/۱۸۳۰	۱/۰۱۰۱	۰/۰۵	
۰/۷۲۷۹	۰/۸۲۸۹	۰/۹۶۹۱	۱/۱۴۲۶	۱/۲۸۵۹	۱/۱۵۵۶	۱/۰۰۶۵	۰/۱	

برآوردگر انقباضی بیزی

برخلاف آمار کلاسیک که محقق براساس اطلاعات نمونه‌ای به برآورد پارامتر نامعلوم می‌پردازد، در روش‌های بیزی براساس ترکیب اطلاعات نمونه‌ای و توزیع پیشین برای پارامتر به برآوردیابی پرداخته می‌شود. پراکاش و سینگ [7] به برآوردیابی انقباضی بیزی پارامتر مقیاس توزیع وایبول و بر اساس داده‌های سانسور شده تحت تابع لاینکس پرداختند.

دی و همکاران [8] برآوردگرهای انقباضی بی‌زی را برای پارامتر مقیاس توزیع نمایی بر اساس داده‌های سانسور شده تحت تابع زیان آنتروپی به‌دست آوردند. با در نظر گرفتن توزیع گامای وارون، $IGamma(a, b)$ ، به‌عنوان توزیع پیشین θ با تابع چگالی

$$g(\sigma | a, b) = \frac{a^b \exp(-\frac{a}{\sigma})}{\sigma^b \Gamma(b) \sigma^{b+1}}, \quad a, b, \sigma > 0, \quad (16)$$

تابع چگالی پسین θ به شرط داده‌های سانسور شده $\mathbf{X} = (X_{(1)}, \dots, X_{(r)})$ با ترکیب تابع درست‌نمایی (۳) و توزیع پیشین (۱۶) به‌صورت (۱۷) به‌دست می‌آید:

$$g(\sigma | \mathbf{x}; a, b) = \frac{(2r\hat{\sigma} + a)^{r+b} \exp(-\frac{2r\hat{\sigma} + a}{\sigma})}{\sigma^{r+b} \Gamma(r+b) \sigma^{r+b+1}}, \quad \theta > 0, \quad (17)$$

که در واقع همان توزیع گامای وارون با پارامترهای $2r\hat{\sigma} + a$ و $r+b$ ، یعنی $IGamma(2r\hat{\sigma} + a, r+b)$ است. مخاطره پسین تحت تابع زیان (۵) برابر با

$$\begin{aligned} \rho(\delta, \sigma) &= E[L(\delta, \sigma) | \mathbf{x}] = E\left[\frac{\delta^r}{\sigma^r} - 2\frac{\delta}{\sigma} + 1 | \mathbf{x}\right] \\ &= \delta^r E\left[\frac{1}{\sigma^r} | \tilde{x}\right] - 2\delta E\left[\frac{1}{\sigma} | \mathbf{x}\right] + 1, \end{aligned}$$

است. مشتق اول و دوم تابع مخاطره پسین نسبت به δ عبارتند از:

$$\frac{\partial \rho(\delta, \sigma)}{\partial \delta} = 2\delta E\left[\frac{1}{\sigma^r} | \tilde{x}\right] - 2E\left[\frac{1}{\sigma} | \mathbf{x}\right], \quad \frac{\partial^2 \rho(\delta, \sigma)}{\partial \delta^2} = 2E\left[\frac{1}{\sigma^r} | \tilde{x}\right] > 0.$$

است. بنابراین تابع مخاطره پسین نسبت به δ تابعی محدب است و به‌ازای

$$\delta^\pi(\tilde{x}) = \frac{E\left[\frac{1}{\sigma} | \tilde{x}\right]}{E\left[\frac{1}{\sigma^r} | \tilde{x}\right]}. \quad (18)$$

کمینه می‌شود. برآوردگر (۱۸) برآوردگر بی‌زی تحت تابع زیان (۵) است. با توجه به توزیع پسین (۱۶) داریم:

$$E\left[\frac{1}{\sigma} | \tilde{x}\right] = \frac{2(r+b)}{2r\hat{\sigma} + a}, \quad E\left[\frac{1}{\sigma^r} | \tilde{x}\right] = \frac{2(r+b)(r+b+1)}{(2r\hat{\sigma} + a)^2}$$

بنابراین برآوردگر بی‌زی تحت تابع زیان (۵) برابر است با

$$\delta^\pi(\tilde{x}) = \frac{2r\hat{\sigma} + a}{2(r+b+1)}. \quad (19)$$

اگر مقدار حدسی σ در مورد پارامتر σ در اختیار باشد آن‌گاه داریم $2r\hat{\sigma} / \sigma \sim \chi_{2r}^2$. با استفاده از روش پراکاش و سینگ [6] و قرار دادن $E[\delta^\pi(\tilde{x})] = \sigma$ داریم:

$$E[\delta^\pi(\tilde{x})] = \frac{2rE[\hat{\sigma}] + a}{2(r+b+1)} = \frac{2r\sigma + a}{2(r+b+1)} = \sigma. \quad (20)$$

با حل رابطه (۲۰) نتیجه می‌شود $a = 2(b+1)\sigma$. با قراردادن این مقدار در برآوردگر بیزی (۱۹)، برآوردگر بیزی بدین صورت است:

$$\delta_S^\pi(x) = \frac{2r\hat{\sigma} + 2(b+1)\sigma}{2(r+b+1)} = \lambda\hat{\sigma} + (1-\lambda)\sigma, \quad (21)$$

که در آن $\lambda = \frac{r}{r+b+1}$. برآوردگر بیزی (۲۱) یک برآوردگر انقباضی بیزی نامیده می‌شود که مخاطره آن تحت تابع زیان (۵) برابر است با

$$\begin{aligned} R(\sigma, \delta_S^\pi(x)) &= E \left[\frac{\delta_S^\pi(x)}{\sigma} - 1 \right]^2 \\ &= E \left[\frac{\lambda\hat{\sigma} + (1-\lambda)\sigma}{\sigma} - 1 \right]^2 \\ &= \lambda^2 \left(\frac{1}{r} + (\delta - 1)^2 \right) - (2\lambda - 1)(\delta - 1) \end{aligned}$$

برای مقایسه برآوردگر انقباضی بیزی (۲۱) و برآوردگر مخاطره ناریب با کم‌ترین مخاطره $c_1\hat{\sigma}$ ، کارایی نسبی بین آن‌ها به صورت (۲۲) محاسبه می‌شود:

$$RE(\delta_S^\pi(x), c_1\hat{\sigma}) = \frac{R(\sigma, c_1\hat{\sigma})}{R(\sigma, \delta_S^\pi(x))}. \quad (22)$$

جدول ۵ کارایی نسبی (۲۲) را به‌ازای مقادیر مختلف $r = 4, 6, 8$ ، $\delta = 0/25, 0/25, 1/75$ و $b = 1, 5, 10, 15$ نشان می‌دهد که از گام‌های $0/25$ برای δ استفاده شده است. در بیش‌تر حالت‌ها برآوردگر انقباضی بیزی نسبت به برآوردگر $c_1\hat{\sigma}$ دارای مخاطره کم‌تری است. کارایی نسبی به‌ازای r و b ثابت تابعی مقعر نسبت به δ است که در نقطه $\delta = 1$ به بیشینه خود می‌رسد. به‌ازای r ثابت، کارایی نسبی به‌ازای $0/75 \leq \delta \leq 1/25$ با افزایش b افزایش می‌یابد. این نشان می‌دهد که وقتی σ به σ نزدیک است، برآوردگر انقباضی بیزی برای مقادیر بزرگ‌تر b دارای مخاطره کم‌تری نسبت به برآوردگر $c_1\hat{\sigma}$ است.

مثال کاربردی

مجموعه داده‌های زیر (لاولس [۹]) مدت زمان بهبودی ۲۰ بیمار سرطان خون را که با مصرف دارو درمان شده‌اند نشان می‌دهد.

۱/۰۱۳	۱/۰۳۴	۱/۱۰۹	۱/۱۶۹	۱/۲۶۶	۱/۵۰۹	۱/۵۳۳
۱/۵۶۳	۱/۷۱۶	۱/۹۲۹	۱/۹۶۵	۲/۰۶۱	۲/۳۴۴	۲/۵۴۶
۲/۶۲۶	۲/۷۷۸	۲/۹۵۱	۳/۴۱۳	۴/۱۱۸	۵/۱۲۶	

فرض کنید مدت زمان بهبودی ۱۰ بیمار اول را در نظر بگیریم یعنی $r = 10$. یک آزمون کلموگروف-اسمیرنوف با مقدار آماره آزمون $0/1591$ و مقدار معنی‌داری $0/6352$ حاکی از برآزش توزیع رایلی با پارامتر $2/8820$ است. برآوردگر مخاطره ناریب با کم‌ترین مخاطره برابر است با

$$c_1\hat{\sigma} = \frac{r}{r+1} \hat{\sigma} = \frac{10}{11} (2/8620) = 2/6018$$

جدول ۵. کارایی نسبی بین $c_1 \hat{\sigma}$ و $\delta_s^\pi(x)$

b					r
۱۵	۱۰	۵	۱	$\delta_s \downarrow$	
۰/۵۴۰۵	۰/۶۲۴۴	۰/۸۲۴۷	۱/۱۵۲۰	۰/۲۵	۴
۱/۱۷۶۴	۱/۳۱۳۸	۱/۵۳۸۴	۱/۴۴۰۰	۰/۵	
۴/۰۰۰۰	۳/۸۹۱۸	۳/۲۰۰۰	۱/۶۹۴۱	۰/۷۵	
۲۰/۰۰۰۰	۱۱/۲۵۰۰	۵/۰۰۰۰	۱/۸۰۰۰	۱	
۴/۰۰۰۰	۳/۸۹۱۸	۳/۲۰۰۰	۱/۶۹۴۱	۱/۲۵	
۱/۱۷۶۴	۱/۳۱۳۸	۱/۵۳۸۴	۱/۴۴۰۰	۱/۵	
۰/۵۴۰۵	۰/۶۲۴۴	۰/۸۲۴۷	۱/۱۵۲۰	۱/۷۵	
۰/۳۰۷۶	۰/۳۶۰۰	۰/۵۰۰۰	۰/۹۰۰۰	۲	
۰/۴۶۰۹	۰/۵۵۷۴	۰/۷۸۳۶	۱/۱۰۸۲	۰/۲۵	۶
۰/۹۸۷۷	۱/۱۳۸۹	۱/۳۷۱۴	۱/۳۰۶۱	۰/۵	
۳/۱۴۲۸	۳/۰۴۴۱	۲/۴۹۳۵	۱/۴۶۲۸	۰/۷۵	
۱۱/۵۲۳۸	۶/۸۸۰۹	۳/۴۲۸۵	۱/۵۲۳۸	۱	
۳/۱۴۲۸	۳/۰۴۴۱	۲/۴۹۳۵	۱/۴۶۲۸	۱/۲۵	
۰/۹۸۷۷	۱/۱۳۸۹	۱/۳۷۱۴	۱/۳۰۶۱	۱/۵	
۰/۴۶۰۹	۰/۵۵۷۴	۰/۷۸۳۸	۱/۱۰۸۲	۱/۷۵	
۰/۲۶۳۹	۰/۳۲۵۰	۰/۴۸۹۷	۰/۹۱۴۲	۲	
۰/۴۲۱۰	۰/۵۲۷۳	۰/۷۷۰۸	۱/۰۸۴۰	۰/۲۵	۸
۰/۸۸۸۸	۰/۰۴۸۹	۱/۲۸۱۰	۱/۲۳۴۵	۰/۵	
۲/۶۶۶۶	۲/۵۷۷۴	۲/۱۲۴۶	۱/۳۴۶۸	۰/۷۵	
۸/۰۰۰۰	۵/۰۱۳۸	۲/۷۲۲۲	۱/۳۸۸۸	۱	
۲/۶۶۶۶	۲/۵۷۷۴	۲/۱۲۴۶	۱/۳۴۶۸	۱/۲۵	
۰/۸۸۸۸	۱/۰۴۸۶	۱/۲۸۱۰	۱/۲۳۴۵	۱/۵	
۰/۴۲۱۰	۰/۵۲۷۳	۰/۷۷۰۸	۱/۰۸۴۰	۱/۷۵	
۰/۲۴۲۴	۰/۳۱۰۹	۰/۴۹۴۹	۰/۹۲۵۹	۲	

اگر مقدار حدسی $\sigma = ۲$ باشد آن‌گاه برای آزمون فرض

$$H_0: \sigma = ۲ \quad \text{v.s.} \quad H_1: \sigma \neq ۲$$

مقدار آماره آزمون برابر است با

$$U = \frac{۲r\hat{\sigma}}{\sigma_0} = \frac{۲(۱۰)(۲/۸۶۲۰)}{۲} = ۲۸/۶۲۰$$

از طرفی با در نظر گرفتن سطح معنی‌داری $۰/۰۵$ داریم $q_1 = \chi_{۰/۰۲۵, ۲۰}^۲ = ۹/۵۹۰۸$ و

$$q_2 = \chi_{۰/۹۷۵, ۲۰}^۲ = ۳۴/۱۶۹۶$$

از آن‌جاکه $U \in (۹/۵۹۰۸, ۳۴/۱۶۹۶)$ ، در نتیجه فرض صفر در سطح معنی‌داری $۰/۰۵$ پذیرفته می‌شود.

آزماوردهای پیشنهادی در سطر اول جدول ۶ خلاصه شده‌اند. به‌طور مشابه، برای مقدار حدسی $\sigma = ۳$ نیز فرض

$H_0: \sigma = ۳ \quad \text{v.s.} \quad H_1: \sigma \neq ۳$ در سطح معنی‌داری $۰/۰۵$ پذیرفته می‌شود و آزماوردهای پیشنهادی در سطر

دوم جدول ۶ محاسبه شده‌اند.

اگر مقدار حدسی $\sigma = ۵$ باشد آن‌گاه فرض $H_0: \sigma = ۵ \quad \text{v.s.} \quad H_1: \sigma \neq ۵$ در سطح معنی‌داری $۰/۰۵$ رد

می‌شود، از این رو، آزماوردهای پیشنهادی برابر با برآورد مخاطره ناریب با کم‌ترین مخاطره هستند.

جدول ۶. برآورد پارامتر برای داده‌های مثال کاربردی

$\hat{\sigma}_{ST4}$	$\hat{\sigma}_{ST3}$	$\hat{\sigma}_{ST2}$	$\hat{\sigma}_{ST1}$	$c_1 \hat{\sigma}$	σ
۲/۷۸۳۶	۲/۳۰۲۳	۲/۶۶۷۳	۲/۴۱۰۰	۲/۶۰۱۸	۲
۲/۸۷۴۵	۲/۹۹۴۸	۲/۹۴۶۷	۲/۹۹۶۸	۲/۶۰۱۸	۳
۲/۶۰۱۸	۲/۶۰۱۸	۲/۶۰۱۸	۲/۶۰۱۸	۲/۶۰۱۸	۵

حال فرض کنید اطلاعات پیشین برای پارامتر σ در اختیار داشته باشیم. اگر مقدار حدسی $\sigma = 2$ ، $b = 2$ و در

نتیجه $a = 2(b+1)\sigma = 12$ باشند آن‌گاه

$$\lambda = \frac{r}{r+b+1} = \frac{10}{13} = 0.7692$$

در نتیجه برآوردگر انقباضی بیزی برابر است با:

$$\delta_S^\pi(x) = \lambda \hat{\sigma} + (1-\lambda)\sigma = 0.7692(2/8620) + 0.2308(2) = 2/6630$$

مقادیر برآورد انقباضی بیزی برای مقادیر مختلف $\sigma = 2, 3, 5$ و $b = 2, 5, 10, 15$ در جدول ۷ محاسبه شده‌اند.

چنان‌که ملاحظه می‌شود برای مقدار ثابت b و با افزایش مقدار حدسی σ ، مقدار برآورد افزایش می‌یابد.

جدول ۷. برآورد انقباضی بیزی پارامتر برای داده‌های مثال

b				σ
۱۵	۱۰	۵	۲	
۲/۳۳۱۵	۲/۴۱۰۵	۲/۵۳۸۷	۲/۶۶۳۰	۲
۲/۹۴۶۹	۲/۹۳۴۳	۲/۹۱۳۷	۲/۸۹۳۸	۳
۴/۱۷۷۷	۳/۹۸۱۹	۳/۶۶۳۷	۳/۳۵۵۴	۵

بحث و نتیجه‌گیری

در این مقاله، با در نظر گرفتن توزیع طول عمر رایلی براساس داده‌های سانسور شده نوع دوم، آزمودهای انقباضی و یک برآوردگر انقباضی بیزی تحت تابع زیان توان دوم خطای پایای مقیاس ارائه شد. آزمودهای انقباضی با استفاده از ۴ روش مختلف، معرفی و مخاطره آن‌ها به‌دست آمد. برای مقایسه عملکرد آزمودهای انقباضی و بهترین برآوردگر خطی نارایب، از ملاک کارایی نسبی بین آن‌ها استفاده شد. نتایج نشان داد که آزمودهای انقباضی وقتی σ به σ نزدیک است دارای مخاطره کم‌تری نسبت به برآوردگر مخاطره نارایب با کم‌ترین مخاطره $c_1 \hat{\sigma}$ است. در ادامه، یک برآوردگر انقباضی بیزی را معرفی و رفتار آن را نسبت به $c_1 \hat{\sigma}$ با محاسبه کارایی نسبی بین آن‌ها بررسی کردیم. محاسبات عددی نشان می‌دهد که وقتی σ به σ نزدیک است، برآوردگر انقباضی بیزی برای مقادیر بزرگ‌تر b دارای مخاطره کم‌تری نسبت به برآوردگر $c_1 \hat{\sigma}$ است. با توجه به این که $a = 2(b+1)\sigma$ ، اگر σ به σ نزدیک باشد، مقادیر بزرگ‌تر a و b برای انتخاب ابرپارامترهای توزیع پیشین پیشنهاد می‌شود.

منابع

۱. نعمت‌الهی ن.، کیاپور آ.، "تحلیل بیزی استوار و کاربرد آن در برآورد حق بیمه"، نشریه علوم خوارزمی، دوره ۱۳، شماره ۳، ۸۴۳-۸۶۰ (۱۳۹۲).
2. Polovko A. M., "Fundamentals of Reliability Theory", Academic Press, SanDiego (1968).
3. Arnold B. C., Balakrishnan N., Nagaraja H. N., "A first course in order statistics", John Wiley & Sons, New York (2008).
4. Thompson J. R., "Some shrunken techniques for estimating the Mean", Journal of the American Statistical Association 63 (1968) 113-122.
5. Prakash G., Singh D. C., "Shrinkage estimation in exponential type-II censored data under LINEX loss", Journal of the Korean Statistical Society 37 (2008) 53-61.
6. Singh D. C., Prakash G., Singh P., "Shrinkage estimator for the shape parameter of pareto distribution using LINEX loss function", Communication in Statics Theory and Methods (2007) 741-753.
7. Prakash G., Singh D. C., "A Bayesian shrinkage approach in Weibull type-II censored data using prior point information", REVSTAT-Statistical Journal 7(2) (2009) 171-187.
8. Day S., Day T., Maiti S. S., "Bayes shrinkage estimation of the parameter of Rayleigh distribution for progressively type-II censored data", Austrian Journal of Statistics 44 (2015) 3-15.
9. Lawless J. F., "Statistical Models and Methods for Lifetime Data", Wiley, New York (1982).