

آزمون نیکویی برازش برای خانواده توزیع‌های مکان-مقیاس تحت سانسور فزاینده نوع دوم

حسین نادب، حمزه ترابی

گروه آمار، دانشگاه یزد

تاریخ دریافت: ۱۳۹۵/۱۱/۰۱ تاریخ پذیرش: ۱۳۹۷/۰۷/۲۳

چکیده: در این مقاله، یک روش کلی برای انجام آزمون نیکویی برازش برای خانواده توزیع‌های مکان-مقیاس با استفاده از داده‌های سانسور شده فزاینده نوع دوم ارائه و ویژگی‌های آن بررسی می‌شود. سپس با به‌کارگیری شبیه‌سازی مونت‌کارلو، توان این آزمون با توان برخی از آزمون‌های موجود، برای فرضیه گامبل بودن مقایسه می‌شود. سرانجام از آزمون ارائه شده برای برازش توزیع به یک مجموعه از داده واقعی استفاده می‌شود.

واژه‌های کلیدی: آزمون نیکویی برازش، خانواده توزیع‌های مکان-مقیاس، سانسور فزاینده نوع دوم، شبیه‌سازی مونت‌کارلو.

۱ مقدمه

یکی از اهداف مهم آمار یافتن یک توزیع مناسب یا بررسی انطباق یا عدم انطباق یک توزیع مشخص برای یک مجموعه از داده است. فرض کنید x_1, \dots, x_n مشاهدات یک نمونه تصادفی از جامعه‌ای با تابع توزیع پیوسته F و هدف انجام آزمون فرضیه‌های

$$\begin{cases} H_0: F = F_\theta \\ H_1: F \neq F_\theta, \end{cases}$$

باشد، که در آن $\theta \in \Theta \subseteq \mathbb{R}^k$ یک بردار $k \in \mathbb{N}$ بعدی از پارامترها است. واضح است که در صورت معلوم بودن θ ، به کارگیری تبدیل انتگرال احتمال منجر به آزمون یکنواختی روی بازه $(0, 1)$ می شود. اما در صورت نامعلوم بودن θ ، ابتدا باید آن را برآورد نمود. برای اطلاعات بیشتر در مورد آزمون های نیکویی برازش می توان به دی آگوستینو و استیونس (۱۹۸۶) و هوبر-کارول و همکاران (۲۰۰۲)، مراجعه کرد.

برای ساخت آزمون نیکویی برازش روش های مختلفی وجود دارد. برخی از آزمون های نیکویی برازش بر اساس فاصله بین تابع توزیع تجربی و تابع توزیع نظری روی بازه $(0, 1)$ یا تفاضل آماره های ترتیبی از امید ریاضی آن ها، یعنی $v_i = u_{i:n} - i/(n+1)$ ساخته می شوند. برای جزئیات بیشتر درباره این آماره ها می توان به برانک (۱۹۶۲)، استیونس (۱۹۸۶) و هگازی و گرین (۱۹۷۵) مراجعه کرد. همچنین برخی دیگر از آزمون های نیکویی برازش مانند آماره گرین وود (۱۹۶۴)، آماره کوئنسبری و میلر (۱۹۷۷) و آماره موران (۱۹۵۱) بر اساس فاصله ها، $D_i = u_{i:n} - u_{i-1:n}$ ، $i = 1, \dots, n+1$ ساخته شده اند، که در آن $u_{0:n} = 0$ و $u_{n+1:n} = 1$. به علاوه ترابی (۲۰۰۶، ۲۰۰۸) یک روش کلی و جدیدی را برای برآورد و آزمون فرضیه با استفاده از این فاصله ها ارائه کرده است.

در سال های اخیر به کارگیری داده های سانسور شده و به ویژه سانسور فزاینده نوع دوم به منظور صرفه جویی در وقت و هزینه بسیار مورد توجه قرار گرفته است، به طوری که مقالات زیادی در زمینه نیکویی برازش بر اساس این نوع داده ها به چاپ رسیده است. به عنوان مثال، می توان به بالاکریشنان و همکاران (۲۰۰۲، ۲۰۰۴)، وانگ (۲۰۰۸) و پاکیاری و بالاکریشنان (۲۰۱۲، ۲۰۱۳) اشاره کرد. برای مطالعه نمونه های بیشتر در این رابطه به آهو و همکاران (۱۹۸۳)، بالاکریشنان و همکاران (۲۰۰۷)، برات پور و حبیبی راد (۲۰۱۶)، بار و دیویدسون (۱۹۷۳)، حبیبی راد و همکاران (۲۰۱۱)، لین و همکاران (۲۰۰۸)، لوری و همکاران (۱۹۷۴)، پتیت و استیونس (۱۹۷۶) و اسمیت و بین (۱۹۷۶) مراجعه شود.

فرض کنید n واحد آزمایشی در یک آزمون طول عمر قرار دارند. در زمان رخداد اولین شکست، r_1 واحد از $n-1$ واحد باقیمانده به طور تصادفی از آزمایش حذف می شوند. سپس در زمان رخداد دومین شکست، r_2 واحد از $n-r_1-2$ واحد باقیمانده به طور تصادفی از آزمایش حذف می شوند؛ به همین ترتیب آزمایش تا لحظه رخداد m امین شکست ادامه می یابد. در این حالت به مشاهدات $x_{1:m:n}, \dots, x_{m:m:n}$ یک نمونه سانسور شده فزاینده نوع دوم با بردار حذفیات $\mathbf{r} = (r_1, \dots, r_m)$ گفته می شود که $x_{j:m:n}$ زمان رخداد j امین شکست است. این طرح سانسور، تعمیم طرح سانسور نوع دوم است به طوری که اگر قرار دهیم $r_1 = \dots = r_{m-1} = 0$ و $r_m = n - m$ ، این طرح به طرح سانسور نوع دوم تبدیل می شود. برای اطلاعات بیشتر در مورد سانسور فزاینده نوع دوم می توان به بالاکریشنان و آگاروالا (۲۰۰۰)

و بالاکریشنان و کرامر (۲۰۱۴) مراجعه کرد.

در این مقاله یک رویکرد جدید برای انجام آزمون نیکویی برازش تحت سانسور فزاینده نوع دوم معرفی می‌شود. مباحث به این صورت سازماندهی شده است: در بخش ۲ برخی از آزمون‌های نیکویی برازش موجود تحت سانسور فزاینده نوع دوم مرور می‌شود. در بخش ۳ روشی برای انجام آزمون نیکویی برازش تحت سانسور فزاینده نوع دوم پیشنهاد و برخی از ویژگی‌های آن بررسی می‌شود. در بخش ۴، با استفاده از شبیه‌سازی، توان آزمون پیشنهادی با توان سایر آزمون‌ها تحت فرضیه گامبل بودن مقایسه می‌شود. سرانجام در بخش ۵، روش پیشنهادی برای یک نمونه واقعی به‌کار گرفته می‌شود.

۲ مروری بر برخی از آزمون‌های موجود

فرض کنید $x_{1:m:n}, \dots, x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم با حذفیات $r = (r_1, \dots, r_m)$ از یک خانواده مکان-مقیاس با تابع توزیع

$$F_{\theta}(x) = F_{(\mu, \sigma)}(x) = F\left(\frac{x - \mu}{\sigma}\right),$$

باشند، که در آن F_0 یک تابع توزیع استاندارد معلوم و پارامترهای مکان μ و مقیاس σ نامعلوم هستند. اگر $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$ برآورد ماکسیمم درست‌نمایی بردار پارامتر $\theta = (\mu, \sigma)$ باشد، قرار می‌دهیم:

$$u_{i:m:n} = F_{\hat{\theta}}(x_{i:m:n}), \quad i = 1, \dots, m,$$

$$v_{i:m:n} = u_{i:m:n} - \alpha_{i:m:n}, \quad i = 1, \dots, m,$$

به طوری که $\alpha_{i:m:n}$ امید ریاضی i امین آماره ترتیبی سانسور فزاینده نوع دوم از توزیع یکنواخت بر بازه $(0, 1)$ است و بنا بر بالاکریشنان و آگاروالا (۲۰۰۰) به صورت

$$\alpha_{i:m:n} = 1 - \prod_{k=m-i+1}^m \frac{k + \sum_{j=m-i+1}^m r_j}{1 + k + \sum_{j=m-i+1}^m r_j}, \quad i = 1, \dots, m,$$

است. اینک به مرور برخی از آزمون‌های موجود پرداخته می‌شود.

۱.۲ آزمون‌های مبتنی بر آماره‌های ترتیبی

آزمون‌های مختلفی بر اساس تفاضل بین آماره‌های ترتیبی و امید ریاضی آن‌ها برای داده‌های کامل توسط برانک (۱۹۶۲)، استیونس (۱۹۸۶) و هگازی و گرین (۱۹۷۵) معرفی شده است. پاکیاری و بالاکریشنان (۲۰۱۳) با پیراستن این آزمون‌ها، آماره‌های

$$C_{m:n}^+ = \max_{1 \leq i \leq m} (v_{i:m:n}), \quad C_{m:n}^- = \max_{1 \leq i \leq m} (-v_{i:m:n}), \quad C_{m:n} = \max(C_{m:n}^-, C_{m:n}^+),$$

$$K_{m:n} = C_{m:n}^- + C_{m:n}^+, \quad T_{m:n}^{(1)} = \sum_{i=1}^m \frac{v_{i:m:n}^2}{m}, \quad T_{m:n}^{(2)} = \sum_{i=1}^m \frac{|v_{i:m:n}|}{m},$$

را برای آزمون نیکویی برازش با استفاده از داده‌های سانسور شده فزاینده نوع دوم پیشنهاد دادند که مقادیر بزرگ این آماره‌ها منجر به رد فرضیه صفر می‌شود.

۲.۲ آزمون‌های مبتنی بر فاصله‌ها

آزمون‌های مختلفی بر اساس فاصله‌ها برای داده‌های کامل توسط گرین‌وود (۱۹۶۴) و کوئسنبری و میلر (۱۹۷۷) معرفی شده است. پاکیاری و بالاکریشنان (۲۰۱۳) با پیراستن این آزمون‌ها، آماره‌های جدیدی را برای آزمون نیکویی برازش با استفاده از داده‌های سانسور شده فزاینده نوع دوم پیشنهاد دادند. همچنین بالاکریشنان و همکاران (۲۰۰۴) آماره دیگری را بر اساس فاصله‌ها پیشنهاد دادند. فرض کنید $\mu_{i:m:n}$ امید ریاضی i امین آماره ترتیبی در سانسور فزاینده نوع دوم از تابع توزیع استاندارد F باشد. اگر

$$G_i = \frac{x_{i:m:n} - x_{i-1:m:n}}{\mu_{i:m:n} - \mu_{i-1:m:n}}, \quad i = 2, 3, \dots, m,$$

آن‌گاه کوچک یا بزرگ بودن آماره‌ی

$$T = \frac{\sum_{i=2}^{m-1} (m-i)G_i}{(m-2) \sum_{i=2}^m G_i},$$

منجر به رد فرضیه صفر می‌شود. توجه شود که توزیع تمام آماره‌های معرفی شده در این بخش مستقل از پارامترهای μ و σ است.

۳ آزمون پیشنهادی

در این بخش، روشی جدید برای انجام آزمون نیکویی برازش با استفاده از داده‌های سانسور شده فزاینده نوع دوم پیشنهاد می‌شود.

تعریف ۱: فرض کنید $x_{1:m:n}, \dots, x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم باشند. تابع توزیع تجربی به صورت

$$\begin{aligned} F^{(m)}(x) &= \frac{1}{m} \sum_{i=1}^m \mathbf{I}(x_{i:m:n} \leq x) \\ &= \frac{1}{m} \#(x_{i:m:n} \leq x), \end{aligned}$$

تعریف می‌شود.

قضیه ۱: (بوردس، ۲۰۰۴) فرض کنید تحت سانسور فزاینده نوع دوم شرایط

الف) $\sup_{i \geq 1} r_i < \infty$ ؛

ب) $\frac{1}{m} \sum_{i=1}^m r_i \rightarrow r$ هنگامی که $m \rightarrow \infty$ ؛

پ) برای $\tau \in \mathbb{R}$ داشته باشیم $F_{\theta}(\tau) < 1$ ،

برقرار باشد. در این صورت رابطه

$$\sup_{0 \leq x \leq \tau} |F^{(m)}(x) - (1 - (1 - F_{\theta}(x))^{r+1})| \xrightarrow{\text{a.s.}} 0,$$

برقرار است، که در آن $\xrightarrow{\text{a.s.}}$ نشان‌دهنده همگرایی با احتمال یک است.

۱.۳ آماره آزمون

اینک برای معرفی آماره آزمون از قضیه ۱ استفاده می‌شود. فرض کنید داده‌های $x_{1:m:n}, \dots, x_{m:m:n}$ و بردار حذفیات r در اختیار باشند. به منظور انجام آزمون نیکویی برازش، از آماره D که به صورت

$$D = \sup_{x \in \mathbb{R}} |F^{(m)}(x) - (1 - (1 - F_{\hat{\theta}}(x))^{\hat{r}+1})|, \quad (1)$$

تعریف می‌شود استفاده می‌کنیم، که در آن

$$\hat{r} = \frac{1}{m} \sum_{i=1}^m r_i,$$

$\hat{\theta}$ برآوردگری موجه برای θ است که می‌تواند یکی از برآوردگرهای ماکسیمم درست‌نمایی، ماکسیمم درست‌نمایی اصلاح شده، ماکسیمم درست‌نمایی وزنی، ماکسیمم حاصلضرب فاصله‌ها یا برآوردگر کمترین توان‌های دوم تحت سانسور فزاینده نوع دوم، موجود در نگ و همکاران (۲۰۱۲) باشد.

۲.۳ ویژگی‌های آماره D

در این بخش برخی از ویژگی‌های آماره D بیان و اثبات می‌شود. قبل از بیان و اثبات این ویژگی‌ها یک لم را بیان و اثبات می‌کنیم.

لم ۱: اگر $g(x, y)$ تابعی کراندار باشد، آنگاه

$$\lim_{y \rightarrow y_0} \sup_x g(x, y) = \sup_x \lim_{y \rightarrow y_0} g(x, y).$$

برهان: برای اثبات، کافی است نشان داده شود:

$$\sup_x \lim_{y \rightarrow y_0} g(x, y) \leq \lim_{y \rightarrow y_0} \sup_x g(x, y), \quad (2)$$

و

$$\lim_{y \rightarrow y_0} \sup_x g(x, y) \leq \sup_x \lim_{y \rightarrow y_0} g(x, y). \quad (۳)$$

برای اثبات (۲)، واضح است که به ازای هر x و y

$$g(x, y) \leq \sup_x g(x, y).$$

با حدگیری از طرفین نامساوی بالا، به ازای هر x داریم:

$$\lim_{y \rightarrow y_0} g(x, y) \leq \lim_{y \rightarrow y_0} \sup_x g(x, y).$$

چون نامساوی بالا به ازای هر x برقرار است، پس واضح است که

$$\sup_x \lim_{y \rightarrow y_0} g(x, y) \leq \lim_{y \rightarrow y_0} \sup_x g(x, y).$$

برای اثبات رابطه (۳) کافی است نشان داده شود که به ازای هر $\epsilon > 0$ ، مقدار x_0 موجود است به گونه‌ای که

$$\lim_{y \rightarrow y_0} \sup_x g(x, y) - \epsilon < \lim_{y \rightarrow y_0} g(x_0, y). \quad (۴)$$

همچنین برای اثبات رابطه (۴) کافی است نشان دهیم که به ازای هر y ، مقدار x_0 موجود است به طوری که

$$\sup_x g(x, y) = g(x_0, y). \quad (۵)$$

چون g تابعی کراندار است پس رابطه (۵) برقرار است و در نتیجه لم ثابت می‌شود. اینک می‌توان ویژگی‌های آماره D را بررسی نمود که به صورت قضیه آورده شده است.

قضیه ۲: فرض کنید تحت فرض H شرایط

(الف) $r \rightarrow \hat{r}$ ، هنگامی که $m \rightarrow \infty$ ؛

(ب) $\theta \xrightarrow{\text{a.s.}} \hat{\theta}$ هنگامی که $m \rightarrow \infty$ ؛

(پ) F_{θ} تابعی پیوسته از θ باشد،
برقرار باشد. در این صورت $(\xrightarrow{P})^{\circ}$ $D \xrightarrow{\text{a.s.}}$ ، که در آن \xrightarrow{P} نشان‌دهنده همگرایی در احتمال است.

برهان:

$$\begin{aligned} D &= \sup_{x \in \mathbb{R}} |F^{(m)}(x) - (\mathbb{1} - (\mathbb{1} - F_{\hat{\theta}}(x))^{\hat{r}+1})| \\ &\leq \sup_{x \in \mathbb{R}} |F^{(m)}(x) - (\mathbb{1} - (\mathbb{1} - F_{\theta}(x))^{r+1})| \\ &\quad + \sup_{x \in \mathbb{R}} |(\mathbb{1} - (\mathbb{1} - F_{\hat{\theta}}(x))^{\hat{r}+1}) - (\mathbb{1} - (\mathbb{1} - F_{\theta}(x))^{r+1})|. \end{aligned} \quad (۶)$$

بنابر قضیه ۱، اولین جمله سمت راست نابرابری (۶) به سمت صفر میل می‌کند. بنابراین کافی است نشان دهیم که جمله دوم نیز به صفر همگرا است.
بدین منظور تابع g را به صورت

$$g(x, m) = (\mathbb{1} - (\mathbb{1} - F_{\hat{\theta}}(x))^{\hat{r}+1}) - (\mathbb{1} - (\mathbb{1} - F_{\theta}(x))^{r+1}).$$

تعریف می‌کنیم. واضح است که g تابعی کراندار است؛ در نتیجه با استفاده از لم ۱ داریم:

$$\lim_{m \rightarrow \infty} \sup_{x \in \mathbb{R}} |g(x, m)| = \sup_{x \in \mathbb{R}} | \lim_{m \rightarrow \infty} g(x, m) |.$$

بنابراین با توجه به همگرایی $\hat{\theta}$ و \hat{r} و پیوستگی تابع F_{θ} بر حسب θ ، قضیه ثابت می‌شود.

قضیه ۳: الف) اگر μ پارامتر مکان و $\hat{\mu}$ یک برآوردگر پایای مکان باشد، آنگاه توزیع D به μ بستگی ندارد.

ب) اگر σ پارامتر مقیاس و $\hat{\sigma}$ یک برآوردگر پایای مقیاس باشد، آنگاه توزیع D به σ بستگی ندارد.
پ) اگر μ و σ به ترتیب پارامترهای مکان و مقیاس و $\hat{\mu}$ و $\hat{\sigma}$ برآوردگرهای پایای مکان-مقیاس باشند، آنگاه توزیع D به μ و σ بستگی ندارد.

برهان: الف) فرض کنید $x_{1:m:n}, \dots, x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم از تابع توزیع $F_{\circ}(x - \mu)$ باشند، که در آن F_{\circ} تابع توزیع مربوط به متغیر تصادفی استاندارد Z است. پس از برآورد

پارامتر μ قرار می‌دهیم

$$z_{i:m:n}^{(1)} = x_{i:m:n} - \hat{\mu}(\mathbf{x}) = z_{i:m:n} - \hat{\mu}(z), \quad i = 1, \dots, m.$$

از طرفی

$$F^{(m)}(x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(x_{i:m:n} \leq x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(1)} \leq x - \hat{\mu}(\mathbf{x})),$$

و

$$F_{\hat{\mu}}(x) = F_*(x - \hat{\mu}(\mathbf{x})).$$

بنابراین

$$\begin{aligned} D &= \sup_{x \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(1)} \leq x - \hat{\mu}(\mathbf{x})) - (1 - (1 - F_*(x - \hat{\mu}(\mathbf{x})))^{\hat{r}+1}) \right| \\ &= \sup_{z \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(1)} \leq z) - (1 - (1 - F_*(z))^{\hat{r}+1}) \right|. \end{aligned}$$

(ب) فرض کنید داده‌های سانسور شده فزاینده نوع دوم از تابع توزیع $F_*(\frac{x}{\sigma})$ باشند، که در آن F_* تابع توزیع مربوط به متغیر تصادفی استاندارد Z است. پس از برآورد پارامتر σ قرار می‌دهیم:

$$z_{i:m:n}^{(r)} = \frac{x_{i:m:n}}{\hat{\sigma}(\mathbf{x})} = \frac{z_{i:m:n}}{\hat{\sigma}(z)}, \quad i = 1, \dots, m.$$

از طرفی

$$F^{(m)}(x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(x_{i:m:n} \leq x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(r)} \leq \frac{x}{\hat{\sigma}(\mathbf{x})}),$$

و

$$F_{\hat{\sigma}}(x) = F_{\circ}\left(\frac{x}{\hat{\sigma}(\mathbf{x})}\right).$$

بنابراین

$$\begin{aligned} D &= \sup_{x \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(\nu)} \leq \frac{x}{\hat{\sigma}(\mathbf{x})}) - (1 - (1 - F_{\circ}\left(\frac{x}{\hat{\sigma}(\mathbf{x})}\right))^{\hat{r}+1}) \right| \\ &= \sup_{z \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(\nu)} \leq z) - (1 - (1 - F_{\circ}(z))^{\hat{r}+1}) \right|. \end{aligned}$$

پ) فرض کنید داده‌های سانسور شده فزاینده نوع دوم از تابع توزیع $F_{\circ}\left(\frac{x - \mu}{\sigma}\right)$ باشند، که در آن F_{\circ} تابع توزیع مربوط به متغیر تصادفی استاندارد Z است. پس از برآورد پارامترهای μ و σ قرار می‌دهیم:

$$z_{i:m:n}^{(\nu)} = \frac{x_{i:m:n} - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})} = \frac{z_{i:m:n} - \hat{\mu}(\mathbf{z})}{\hat{\sigma}(\mathbf{z})}, \quad i = 1, \dots, m.$$

از طرفی

$$F^{(m)}(x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(x_{i:m:n} \leq x) = \frac{1}{m} \sum_{i=1}^m \mathbf{I}(z_{i:m:n}^{(\nu)} \leq \frac{x - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}),$$

و

$$F_{(\hat{\mu}, \hat{\sigma})}(x) = F_{\circ}\left(\frac{x - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}\right).$$

بنابراین

$$\begin{aligned} D &= \sup_{x \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbb{I}(z_{i:m:n}^{(r)} \leq \frac{x - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}) - (1 - (1 - F_{\circ}(\frac{x - \hat{\mu}(\mathbf{x})}{\hat{\sigma}(\mathbf{x})}))^{\hat{r}+1}) \right| \\ &= \sup_{z \in \mathbb{R}} \left| \frac{1}{m} \sum_{i=1}^m \mathbb{I}(z_{i:m:n}^{(r)} \leq z) - (1 - (1 - F_{\circ}(z))^{\hat{r}+1}) \right|. \end{aligned}$$

در این مقاله روش برآورد ماکسیمم درستنمایی در نظر گرفته می‌شود؛ بنابراین شرایط قضیه ۳ برقرار است.

قضیه ۴: فرض کنید D آماره تعریف شده در رابطه (۱) باشد. در این صورت $D = \max(D^-, D^+)$ که در آن

$$\begin{aligned} D^- &= \max\{0, \max_{1 \leq i \leq m} [(1 - (1 - F_{\hat{\theta}}(x_{i:m:n}))^{\hat{r}+1}) - \frac{i-1}{m}]\}, \\ D^+ &= \max\{0, \max_{1 \leq i \leq m} [\frac{i}{m} - (1 - (1 - F_{\hat{\theta}}(x_{i:m:n}))^{\hat{r}+1})]\}. \end{aligned}$$

برهان: این رابطه مشابه با رابطه آماره آزمون کولموگروف-اسمیرنوف است و به راحتی ثابت می‌شود.

۴ مطالعات شبیه‌سازی

در این بخش با ۱۰۰۰۰ بار شبیه‌سازی، توان آزمون پیشنهادی با توان آزمون‌های موجود در بخش ۲ برای فرضیه گامبل بودن توزیع جامعه در مقابل فرضیه‌های مختلف مقایسه می‌شود. برای این منظور، ۲۷ بردار حذفیات مختلف در نظر گرفته شده است که این بردارها قبلاً توسط بالاکریشنان و همکاران (۲۰۰۴) و پاکپاری و بالاکریشنان (۲۰۱۳) استفاده شده و در جدول ۱ آورده شده‌اند. تمام شبیه‌سازی‌ها در نرم‌افزار R انجام شده است.

اینک همانند پاکپاری و بالاکریشنان (۲۰۱۳)، تحت فرضیه صفر تابع چگالی گامبل را به صورت

$$f_{(\mu, \sigma)}(x) = \frac{1}{\sigma} \exp\left\{\frac{x - \mu}{\sigma}\right\} \exp\left\{-\exp\left\{\frac{x - \mu}{\sigma}\right\}\right\}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0,$$

و تحت فرضیه مقابل، تابع چگالی لگ‌گاما را به صورت

$$f_{\kappa}(x) = \frac{\kappa^{\kappa-\frac{1}{\sqrt{\kappa}}}}{\Gamma(\kappa)} \exp\{\sqrt{\kappa}x - \kappa \exp\{\frac{x}{\sqrt{\kappa}}\}\}, \quad \kappa > 0,$$

به ازای مقادیر ۴، ۸ و ∞ برای κ در نظر می‌گیریم که اگر $\kappa = \infty$ ، آن‌گاه توزیع لگ‌گاما به توزیع نرمال استاندارد تبدیل می‌شود. برای جزئیات بیشتر در مورد این توزیع‌ها می‌توان به جانسون و همکاران (۱۹۹۴)، (۱۹۹۵) مراجعه کرد.

جدول ۲ توان‌های شبیه‌سازی شده را در سطح معنی‌داری $\alpha = 0.1$ نشان می‌دهد که مقادیر توان‌ها برای آزمون‌های موجود در بخش ۲ به‌طور مستقیم از پاک‌یاری و بالاکریشنان (۲۰۱۳) آورده شده‌اند. در جدول ۲ برای هر حالت از توزیع لگ‌گاما، بزرگترین توان به‌صورت پررنگ مشخص شده است.

۵ مثال کاربردی

پاک‌یاری و بالاکریشنان (۲۰۱۳) با اعمال یک طرح سانسور فزاینده نوع دوم بر روی داده‌های موجود در نلسون (۱۹۸۲)، صفحه ۱۰۵ جدول ۱۰.۱ و سپس لگاریتم گرفتن از آن‌ها داده‌های جدول ۳ را به‌دست آوردند. آن‌ها مقادیر آماره‌های معرفی شده و مقادیر احتمال متناظر آن‌ها را محاسبه کردند. در این مقاله نیز با در نظر گرفتن داده‌های جدول ۳، مقادیر آماره‌های T و D محاسبه و مقادیر احتمال متناظر آن‌ها به‌دست آمده است. مقادیر محاسبه شده توسط پاک‌یاری و بالاکریشنان (۲۰۱۳) برای برخی آماره‌ها، به همراه مقادیر محاسبه شده برای آماره‌های T و D در جدول ۴ آمده است. همان‌گونه که در جدول ۴ مشاهده می‌شود، آماره‌ی D موافق با آماره‌های $C_{m:n}^+$ ، $C_{m:n}^-$ ، $C_{m:n}$ ، $K_{m:n}$ ، $T_{m:n}^{(1)}$ و $T_{m:n}^{(2)}$ به‌صورت قوی فرضیه‌ی گامبل بودن داده‌های جدول ۳ (وایبول بودن توزیع داده‌های اصلی) را می‌پذیرد، اما آماره T ، دارای مقدار احتمال کمتری نسبت به سایر آزمون‌ها است.

بحث و نتیجه‌گیری

در این مقاله، یک روش کلی برای انجام آزمون نیکویی برازش برای خانواده توزیع‌های مکان-مقیاس با استفاده از داده‌های سانسور شده فزاینده نوع دوم ارائه و ویژگی‌های آن مورد بررسی قرار گرفت. سپس با استفاده از شبیه‌سازی، توان این آزمون محاسبه و همراه با توان برخی از آزمون‌های موجود برای آزمون

جدول ۱: طرح‌های سانسور فزاینده نوع دوم استفاده شده در شبیه‌سازی مونت‌کارلو.

$r = (r_1, \dots, r_m)$	m	n	طرح
$i \neq 1$ برای $r_i = 0$ و $r_1 = 12$	۸	۲۰	{۱}
$i \neq 8$ برای $r_i = 0$ و $r_8 = 12$	۸	۲۰	{۲}
$i \neq 1, 8$ برای $r_i = 0$ و $r_1 = r_8 = 6$	۸	۲۰	{۳}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 8$	۱۲	۲۰	{۴}
$i \neq 12$ برای $r_i = 0$ و $r_{12} = 8$	۱۲	۲۰	{۵}
$i \neq 3, 5, 7, 9$ برای $r_i = 0$ و $r_3 = r_5 = r_7 = r_9 = 2$	۱۲	۲۰	{۶}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 4$	۱۶	۲۰	{۷}
$i \neq 16$ برای $r_i = 0$ و $r_{16} = 4$	۱۶	۲۰	{۸}
$i \neq 5$ برای $r_i = 0$ و $r_5 = 4$	۱۶	۲۰	{۹}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 30$	۱۰	۴۰	{۱۰}
$i \neq 10$ برای $r_i = 0$ و $r_{10} = 30$	۱۰	۴۰	{۱۱}
$i \neq 1, 5, 10$ برای $r_i = 0$ و $r_1 = r_5 = r_{10} = 10$	۱۰	۴۰	{۱۲}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 20$	۲۰	۴۰	{۱۳}
$i \neq 20$ برای $r_i = 0$ و $r_{20} = 20$	۲۰	۴۰	{۱۴}
$i = 1, \dots, 20$ برای $r_i = 1$	۲۰	۴۰	{۱۵}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 10$	۳۰	۴۰	{۱۶}
$i \neq 30$ برای $r_i = 0$ و $r_{30} = 10$	۳۰	۴۰	{۱۷}
$i \neq 1, 30$ برای $r_i = 0$ و $r_1 = r_{30} = 5$	۳۰	۴۰	{۱۸}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 40$	۲۰	۶۰	{۱۹}
$i \neq 20$ برای $r_i = 0$ و $r_{20} = 40$	۲۰	۶۰	{۲۰}
$i \neq 1, 10, 20$ برای $r_i = 0$ و $r_{10} = 20$ و $r_1 = r_{20} = 10$	۲۰	۶۰	{۲۱}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 20$	۴۰	۶۰	{۲۲}
$i \neq 40$ برای $r_i = 0$ و $r_{40} = 20$	۴۰	۶۰	{۲۳}
$i = 1, \dots, 20$ برای $r_{2i} = 0$ و $r_{2i-1} = 1$	۴۰	۶۰	{۲۴}
$i \neq 1$ برای $r_i = 0$ و $r_1 = 10$	۵۰	۶۰	{۲۵}
$i \neq 50$ برای $r_i = 0$ و $r_{50} = 10$	۵۰	۶۰	{۲۶}
$i \neq 1, 50$ برای $r_i = 0$ و $r_1 = r_{50} = 5$	۵۰	۶۰	{۲۷}

جدول ۲: توان‌های برآورد شده در حالت $\alpha = ۰.۱$ با فرضیه صفر توزیع گامبل و فرضیه مقابل توزیع لگ‌گاما با پارامترهای ۴، ۸ و ∞ .

D	T	$T_{m:n}^{(2)}$	$T_{m:n}^{(1)}$	$K_{m:n}$	$C_{m:n}$	$C_{m:n}^-$	$C_{m:n}^+$	κ	طرح
۰/۰۴۷۰	۰/۱۲۹۹	۰/۱۱۸۲	۰/۱۱۷۰	۰/۱۲۵۹	۰/۱۱۳۹	۰/۱۷۶۳	۰/۰۴۷۵	۴	{۱}
۰/۰۳۸۶	۰/۱۵۰۲	۰/۱۳۸۱	۰/۱۳۵۸	۰/۱۳۸۵	۰/۱۲۶۹	۰/۲۰۶۹	۰/۰۳۹۶	۸	
۰/۰۲۰۲	۰/۲۱۹۹	۰/۱۹۹۵	۰/۱۹۴۱	۰/۱۸۶۱	۰/۱۷۸۲	۰/۲۹۳۲	۰/۰۲۵۶	∞	
۰/۱۳۳۵	۰/۰۸۴۹	۰/۱۰۱۵	۰/۱۰۵۰	۰/۰۹۷۴	۰/۱۱۰۸	۰/۱۲۵۱	۰/۰۷۳۵	۴	{۲}
۰/۱۳۶۱	۰/۰۸۲۴	۰/۱۰۳۷	۰/۱۱۰۴	۰/۰۹۸۷	۰/۱۱۵۳	۰/۱۳۳۷	۰/۰۶۹۳	۸	
۰/۱۵۷۷	۰/۰۸۱۰	۰/۱۱۳۵	۰/۱۲۲۳	۰/۱۰۱۴	۰/۱۲۷۶	۰/۱۵۵۰	۰/۰۵۷۴	∞	
۰/۰۶۹۷	۰/۰۸۹۱	۰/۱۰۰۶	۰/۱۰۴۲	۰/۱۰۳۴	۰/۱۰۹۰	۰/۱۳۵۹	۰/۰۶۵۴	۴	{۳}
۰/۰۶۳۴	۰/۰۹۰۱	۰/۱۰۶۱	۰/۱۰۸۰	۰/۱۰۵۴	۰/۱۱۵۱	۰/۱۴۸۳	۰/۰۵۹۰	۸	
۰/۰۵۹۸	۰/۰۹۶۰	۰/۱۲۴۵	۰/۱۲۵۷	۰/۱۱۷۵	۰/۱۳۱۲	۰/۱۸۱۰	۰/۰۴۵۳	∞	
۰/۰۴۰۰	۰/۱۴۷۱	۰/۱۲۶۵	۰/۱۲۷۰	۰/۱۰۹۹	۰/۱۲۱۱	۰/۱۸۸۴	۰/۰۴۵۵	۴	{۴}
۰/۰۲۹۴	۰/۱۷۸۹	۰/۱۴۸۶	۰/۱۵۰۴	۰/۱۲۳۸	۰/۱۴۰۴	۰/۲۲۶۹	۰/۰۳۶۶	۸	
۰/۰۱۵۸	۰/۲۹۰۴	۰/۲۲۶۶	۰/۲۲۶۳	۰/۱۷۰۵	۰/۲۱۴۱	۰/۳۳۵۸	۰/۰۲۲۱	∞	
۰/۱۴۶۷	۰/۰۹۶۵	۰/۱۰۷۶	۰/۱۰۷۳	۰/۱۰۰۰	۰/۱۱۴۳	۰/۱۴۵۳	۰/۰۶۳۳	۴	{۵}
۰/۱۶۰۰	۰/۰۹۳۹	۰/۱۱۴۴	۰/۱۱۶۷	۰/۱۰۳۷	۰/۱۲۲۳	۰/۱۵۹۸	۰/۰۵۶۵	۸	
۰/۲۰۲۰	۰/۱۱۰۸	۰/۱۳۵۸	۰/۱۴۳۱	۰/۱۱۴۳	۰/۱۴۵۲	۰/۲۰۰۴	۰/۰۴۴۴	∞	
۰/۱۱۶۲	۰/۱۴۲۸	۰/۱۲۸۷	۰/۱۳۴۹	۰/۱۴۵۵	۰/۱۳۳۰	۰/۱۷۱۹	۰/۱۰۲۷	۴	{۶}
۰/۱۳۱۵	۰/۱۶۶۶	۰/۱۴۷۳	۰/۱۵۵۰	۰/۱۶۳۶	۰/۱۴۵۸	۰/۱۹۹۵	۰/۱۰۳۳	۸	
۰/۱۸۴۸	۰/۲۶۴۶	۰/۲۰۸۷	۰/۲۱۶۶	۰/۲۱۵۵	۰/۱۹۰۴	۰/۲۸۰۱	۰/۱۰۶۱	∞	
۰/۰۳۷۸	۰/۱۶۰۱	۰/۱۳۱۸	۰/۱۲۸۵	۰/۱۱۴۶	۰/۱۲۹۵	۰/۱۸۵۷	۰/۰۴۶۷	۴	{۷}
۰/۰۲۹۱	۰/۲۰۳۵	۰/۱۵۷۵	۰/۱۵۶۸	۰/۱۲۷۹	۰/۱۵۲۹	۰/۲۲۲۸	۰/۰۳۶۸	۸	
۰/۰۱۹۵	۰/۳۴۵۲	۰/۲۵۲۳	۰/۲۵۰۶	۰/۱۸۷۹	۰/۲۳۴۲	۰/۳۳۳۶	۰/۰۲۶۱	∞	
۰/۱۶۶۴	۰/۱۰۷۸	۰/۱۲۱۶	۰/۱۲۱۷	۰/۱۱۲۸	۰/۱۱۹۸	۰/۱۶۲۷	۰/۰۵۸۹	۴	{۸}
۰/۱۸۷۱	۰/۱۲۰۹	۰/۱۳۴۱	۰/۱۳۷۴	۰/۱۲۰۷	۰/۱۳۳۸	۰/۱۸۶۰	۰/۰۵۰۱	۸	
۰/۲۵۹۹	۰/۱۷۵۴	۰/۱۸۸۶	۰/۱۹۰۰	۰/۱۵۱۶	۰/۱۷۶۲	۰/۲۵۴۱	۰/۰۳۶۹	∞	
۰/۰۵۰۵	۰/۱۵۵۸	۰/۱۲۱۰	۰/۱۲۷۶	۰/۱۱۴۳	۰/۱۲۳۲	۰/۱۹۱۷	۰/۰۵۱۹	۴	{۹}
۰/۰۵۲۸	۰/۱۹۷۳	۰/۱۴۳۹	۰/۱۵۰۴	۰/۱۳۱۳	۰/۱۴۱۸	۰/۲۳۰۳	۰/۰۴۳۲	۸	
۰/۰۵۹۴	۰/۳۳۸۲	۰/۲۳۱۴	۰/۲۳۱۰	۰/۱۸۷۰	۰/۲۰۸۱	۰/۳۳۳۹	۰/۰۳۰۳	∞	
۰/۰۴۰۵	۰/۱۵۹۱	۰/۱۵۳۲	۰/۱۴۴۸	۰/۱۳۹۲	۰/۱۲۹۹	۰/۲۱۹۱	۰/۰۴۰۰	۴	{۱۰}
۰/۰۲۸۷	۰/۱۹۶۳	۰/۱۸۹۶	۰/۱۷۹۱	۰/۱۶۸۲	۰/۱۶۰۲	۰/۲۶۹۷	۰/۰۳۰۷	۸	
۰/۱۱۴	۰/۳۰۸۰	۰/۳۰۹۱	۰/۲۹۲۷	۰/۲۵۱۵	۰/۲۵۹۰	۰/۴۰۷۳	۰/۰۱۷۶	∞	

ادامه جدول ۲. توان‌های برآورد شده در حالت $\alpha = ۰/۱$ با فرضیه صفر توزیع گامبل و فرضیه مقابل توزیع لگ‌گاما با پارامترهای ۴، ۸ و ∞ .

D	T	$T_{m:n}^{(2)}$	$T_{m:n}^{(1)}$	$K_{m:n}$	$C_{m:n}$	$C_{m:n}^-$	$C_{m:n}^+$	k	طرح
۰/۱۳۳۹	۰/۰۸۷۶	۰/۱۰۵۸	۰/۱۰۷۵	۰/۰۹۵۷	۰/۱۱۳۶	۰/۱۲۶۸	۰/۰۷۳۷	۴	{۱۱}
۰/۱۴۲۶	۰/۰۸۱۰	۰/۱۰۷۱	۰/۱۱۱۴	۰/۰۹۵۵	۰/۱۲۰۴	۰/۱۳۶۳	۰/۰۶۵۲	۸	
۰/۱۴۹۳	۰/۰۸۱۷	۰/۱۱۸۷	۰/۱۲۳۴	۰/۰۹۸۶	۰/۱۳۶۴	۰/۱۵۹۰	۰/۰۵۴۷	∞	
۰/۰۷۵۵	۰/۰۹۴۸	۰/۱۱۷۳	۰/۱۱۶۳	۰/۱۱۹۴	۰/۱۱۳۰	۰/۱۴۴۲	۰/۰۸۳۶	۴	{۱۲}
۰/۰۷۰۷	۰/۰۹۳۵	۰/۱۲۳۸	۰/۱۲۵۶	۰/۱۲۷۴	۰/۱۱۸۷	۰/۱۵۶۴	۰/۰۸۰۵	۸	
۰/۰۶۹۰	۰/۱۱۰۴	۰/۱۴۴۴	۰/۱۴۷۳	۰/۱۴۸۹	۰/۱۳۷۵	۰/۱۹۴۱	۰/۰۷۳۵	∞	
۰/۰۲۶۱	۰/۲۱۱۰	۰/۱۷۸۲	۰/۱۷۶۰	۰/۱۳۵۷	۰/۱۶۹۴	۰/۲۵۵۸	۰/۰۳۴۱	۴	{۱۳}
۰/۰۱۹۳	۰/۲۷۸۵	۰/۲۳۱۸	۰/۲۳۴۳	۰/۱۶۵۰	۰/۲۱۷۳	۰/۳۱۹۶	۰/۰۲۵۲	۸	
۰/۰۰۶۸	۰/۴۹۰۳	۰/۴۱۵۸	۰/۴۱۱۰	۰/۲۶۹۹	۰/۳۷۷۱	۰/۵۰۲۸	۰/۰۱۳۰	∞	
۰/۱۳۹۰	۰/۰۹۶۵	۰/۱۱۷۸	۰/۱۲۰۱	۰/۱۰۱۹	۰/۱۲۱۱	۰/۱۶۶۲	۰/۰۵۷۵	۴	{۱۴}
۰/۱۴۷۳	۰/۱۰۱۹	۰/۱۳۳۵	۰/۱۳۵۰	۰/۱۰۶۳	۰/۱۳۶۴	۰/۱۸۹۸	۰/۰۴۸۴	۸	
۰/۱۸۶۳	۰/۱۳۶۲	۰/۱۷۵۶	۰/۱۷۹۸	۰/۱۲۴۶	۰/۱۷۷۵	۰/۲۵۲۴	۰/۰۳۴۴	∞	
۰/۱۳۶۶	۰/۱۴۰۳	۰/۱۵۳۴	۰/۱۷۰۹	۰/۱۹۳۵	۰/۱۷۹۳	۰/۱۸۳۵	۰/۱۷۳۷	۴	{۱۵}
۰/۱۵۴۶	۰/۱۷۲۸	۰/۱۸۶۷	۰/۲۰۶۵	۰/۲۲۹۱	۰/۲۱۶۶	۰/۲۱۷۳	۰/۲۰۸۸	۸	
۰/۲۲۹۶	۰/۲۹۳۵	۰/۲۷۹۵	۰/۳۰۶۶	۰/۳۳۷۰	۰/۳۱۰۶	۰/۳۱۵۸	۰/۲۹۷۱	∞	
۰/۰۲۷۱	۰/۲۵۰۹	۰/۱۸۵۴	۰/۱۸۲۹	۰/۱۳۶۹	۰/۱۷۳۰	۰/۲۶۰۸	۰/۰۳۵۸	۴	{۱۶}
۰/۰۱۲۳	۰/۳۴۵۵	۰/۲۴۷۳	۰/۲۴۳۱	۰/۱۷۶۱	۰/۲۲۳۶	۰/۳۳۵۰	۰/۰۲۹۵	۸	
۰/۰۱۲۳	۰/۶۱۵۷	۰/۴۵۳۱	۰/۴۵۱۱	۰/۳۱۰۹	۰/۴۰۰۳	۰/۵۳۲۳	۰/۰۳۴۹	∞	
۰/۲۷۹۰	۰/۱۲۸۸	۰/۱۴۱۹	۰/۱۴۶۳	۰/۱۱۷۹	۰/۱۴۲۰	۰/۲۰۶۴	۰/۰۴۵۸	۴	{۱۷}
۰/۱۹۲۶	۰/۱۵۷۲	۰/۱۷۳۳	۰/۱۷۵۱	۰/۱۳۴۵	۰/۱۶۶۴	۰/۲۵۱۱	۰/۰۳۷۷	۸	
۰/۲۷۹۰	۰/۲۸۰۹	۰/۲۷۳۸	۰/۲۷۳۷	۰/۱۹۴۴	۰/۲۵۲۴	۰/۳۵۹۰	۰/۰۲۵۸	∞	
۰/۰۷۷۳	۰/۱۵۳۰	۰/۱۶۲۴	۰/۱۶۰۴	۰/۱۲۹۳	۰/۱۵۳۳	۰/۲۲۷۱	۰/۰۳۹۷	۴	{۱۸}
۰/۰۹۶۶	۰/۱۹۸۸	۰/۲۰۲۰	۰/۲۰۴۸	۰/۱۵۵۶	۰/۱۸۸۰	۰/۲۸۰۷	۰/۰۳۲۱	۸	
۰/۱۵۴۰	۰/۳۶۸۴	۰/۳۴۱۴	۰/۳۳۷۸	۰/۲۴۲۵	۰/۲۹۷۷	۰/۴۲۶۵	۰/۰۲۱۹	∞	
۰/۰۲۵۱	۰/۲۲۶۰	۰/۲۱۰۸	۰/۲۰۹۷	۰/۱۵۷۹	۰/۱۸۸۱	۰/۲۹۰۵	۰/۰۲۸۵	۴	{۱۹}
۰/۰۱۴۸	۰/۳۰۴۲	۰/۲۸۶۰	۰/۲۷۹۴	۰/۱۹۶۰	۰/۲۵۶۱	۰/۳۶۷۳	۰/۰۲۰۴	۸	
۰/۰۰۴۶	۰/۵۲۴۷	۰/۵۰۹۴	۰/۵۰۳۸	۰/۳۳۴۶	۰/۴۴۶۵	۰/۵۸۲۵	۰/۰۰۸۴	∞	
۰/۱۲۹۲	۰/۰۹۲۸	۰/۱۱۳۵	۰/۱۱۸۲	۰/۱۰۰۰	۰/۱۱۹۹	۰/۱۵۷۳	۰/۰۵۹۱	۴	{۲۰}
۰/۱۴۴۸	۰/۰۹۴۰	۰/۱۲۵۷	۰/۱۳۳۹	۰/۱۰۳۳	۰/۱۳۱۶	۰/۱۷۹۱	۰/۰۵۰۲	۸	
۰/۱۶۷۲	۰/۱۱۳۳	۰/۱۶۲۹	۰/۱۷۰۶	۰/۱۱۴۴	۰/۱۶۸۷	۰/۲۳۱۷	۰/۰۳۵۳	∞	

ادامه جدول ۲. توان‌های برآورد شده در حالت $\alpha = 0.1$ با فرضیه صفر توزیع گامبل و فرضیه مقابل توزیع لگ‌گاما

D	T	$T_{m:n}^{(2)}$	$T_{m:n}^{(1)}$	$K_{m:n}$	$C_{m:n}$	$C_{m:n}^-$	$C_{m:n}^+$	κ	طرح
۰.۰۷۰۶	۰.۱۱۶۰	۰.۱۳۱۸	۰.۱۳۷۶	۰.۱۴۶۸	۰.۱۲۵۵	۰.۱۶۹۰	۰.۰۸۱۱	۴	{۲۱}
۰.۰۷۱۲	۰.۱۳۳۹	۰.۱۴۷۷	۰.۱۵۵۳	۰.۱۶۶۳	۰.۱۳۶۷	۰.۱۹۵۴	۰.۰۷۸۷	۸	
۰.۰۷۲۳	۰.۲۱۵۰	۰.۲۰۴۲	۰.۲۰۸۰	۰.۲۲۲۱	۰.۱۷۰۸	۰.۲۶۶۵	۰.۰۷۱۲	∞	
۰.۰۱۹۴	۰.۳۲۱۸	۰.۲۲۵۳	۰.۲۲۰۹	۰.۱۶۰۴	۰.۲۰۲۳	۰.۳۰۲۸	۰.۰۳۰۸	۴	{۲۲}
۰.۰۱۱۸	۰.۴۵۱۱	۰.۳۱۵۸	۰.۳۱۱۶	۰.۲۱۱۵	۰.۲۷۵۴	۰.۴۰۰۷	۰.۰۲۶۷	۸	
۰.۰۰۴۳	۰.۷۵۶۴	۰.۵۹۵۱	۰.۵۸۷۳	۰.۴۰۷۵	۰.۵۱۳۵	۰.۶۵۲۶	۰.۰۴۴۳	∞	
۰.۱۳۰۰	۰.۱۳۵۲	۰.۱۵۲۲	۰.۱۵۴۳	۰.۱۲۲۷	۰.۱۴۲۱	۰.۲۲۲۱	۰.۰۳۹۹	۴	{۲۳}
۰.۱۵۰۸	۰.۱۷۵۳	۰.۱۸۸۱	۰.۱۸۷۶	۰.۱۴۱۹	۰.۱۷۴۴	۰.۲۶۵۴	۰.۰۳۲۰	۸	
۰.۲۰۳۸	۰.۳۲۱۱	۰.۳۰۶۶	۰.۳۰۴۶	۰.۲۱۲۰	۰.۲۶۴۸	۰.۳۸۹۵	۰.۰۲۱۳	∞	
۰.۱۵۸۲	۰.۲۵۴۵	۰.۱۹۱۶	۰.۱۹۹۱	۰.۲۰۱۰	۰.۱۸۷۰	۰.۲۶۰۰	۰.۰۸۵۶	۴	{۲۴}
۰.۲۰۲۳	۰.۳۵۶۵	۰.۲۵۵۷	۰.۲۶۱۴	۰.۲۵۰۵	۰.۲۳۸۴	۰.۳۲۸۰	۰.۰۹۳۶	۸	
۰.۳۵۶۴	۰.۶۴۵۸	۰.۴۶۲۷	۰.۴۶۶۸	۰.۴۱۷۴	۰.۳۹۲۵	۰.۵۱۱۸	۰.۱۳۰۱	∞	
۰.۰۲۲۰	۰.۳۵۴۳	۰.۲۴۳۱	۰.۲۴۱۸	۰.۱۸۱۴	۰.۲۱۵۲	۰.۳۰۹۴	۰.۰۴۰۷	۴	{۲۵}
۰.۰۱۵۸	۰.۵۱۰۰	۰.۳۴۳۷	۰.۳۳۲۹	۰.۲۴۵۳	۰.۲۹۴۹	۰.۴۰۷۸	۰.۰۴۶۶	۸	
۰.۰۲۹۶	۰.۸۱۸۴	۰.۶۳۸۶	۰.۶۲۶۵	۰.۴۸۹۹	۰.۵۳۸۳	۰.۶۵۶۹	۰.۱۲۷۸	∞	
۰.۱۹۶۵	۰.۱۸۹۶	۰.۱۹۴۲	۰.۱۹۵۵	۰.۱۵۲۹	۰.۱۷۷۱	۰.۲۵۹۵	۰.۰۳۳۹	۴	{۲۶}
۰.۲۳۸۳	۰.۲۷۷۲	۰.۲۵۴۳	۰.۲۵۵۱	۰.۱۹۱۸	۰.۲۳۰۶	۰.۳۲۷۰	۰.۰۳۰۵	۸	
۰.۳۷۱۳	۰.۵۲۶۸	۰.۴۵۰۴	۰.۴۴۴۴	۰.۳۲۴۵	۰.۳۸۰۹	۰.۵۰۵۲	۰.۰۳۵۶	∞	
۰.۱۰۲۵	۰.۲۳۰۶	۰.۲۱۲۲	۰.۲۱۳۲	۰.۱۶۴۷	۰.۱۹۴۳	۰.۲۸۷۶	۰.۰۳۲۵	۴	{۲۷}
۰.۱۴۵۹	۰.۳۳۹۱	۰.۲۸۸۱	۰.۲۸۳۸	۰.۲۱۱۵	۰.۲۵۴۱	۰.۳۶۷۳	۰.۰۲۹۵	۸	
۰.۲۷۱۶	۰.۶۲۷۵	۰.۵۲۲۶	۰.۵۱۲۹	۰.۳۷۴۹	۰.۴۴۰۰	۰.۵۷۱۰	۰.۰۴۳۳	∞	

جدول ۳: لگاریتم داده‌های به دست آمده از اجرای طرح سانسور روی داده‌های موجود در نلسون (۱۹۸۲).

i	۱	۲	۳	۴	۵
$x_{i:m:n}$	-۱,۳۰۹۳	-۰,۹۱۶۳	-۰,۳۷۱۱	-۰,۲۳۵۷	۱,۰۱۱۶
r_i	۰	۱	۰	۱	۰
i	۶	۷	۸	۹	۱۰
$x_{i:m:n}$	۱,۳۶۳۵	۲,۷۶۸۲	۳,۳۲۵۰	۳,۹۷۴۸	۵,۳۷۱۱
r_i	۱	۰	۱	۰	۱

جدول ۴: مقادیر آماره‌ها و مقادیر احتمال متناظر آن‌ها برای داده‌های جدول ۳ تحت فرض گامبل بودن.

D	T	$T_{m:n}^{(2)}$	$T_{m:n}^{(1)}$	$K_{m:n}$	$C_{m:n}$	$C_{m:n}^-$	$C_{m:n}^+$	آماره
۰,۱۷۹۸	۰,۳۵۵۴	۰,۰۴۷۸	۰,۰۰۳۹	۰,۲۰۶۲	۰,۱۰۳۲	۰,۱۰۳۰	۰,۱۰۳۲	مقدار آماره
۰,۴۹۰۳	۰,۱۳۵۴	۰,۶۸۸۴	۰,۵۵۸۵	۰,۵۱۸۹	۰,۷۰۹۶	۰,۳۹۰۰	۰,۶۱۰۶	مقدار احتمال

فرضیه گامبل بودن در جدول ۲ آورده شد. با بررسی جدول ۲ ملاحظه می‌شود که وقتی تعداد شکست‌ها (m) کم و سانسور در آخرین مرحله انجام می‌شود (مانند طرح‌های $\{2\}$, $\{5\}$, $\{8\}$, $\{11\}$) آزمون D عملکرد بهتری نسبت به سایر آزمون‌ها دارد. در حالتی که m زیاد و سانسور در اولین مرحله انجام می‌شود (مانند طرح‌های $\{16\}$, $\{22\}$, $\{24\}$) آزمون T عملکرد بهتری نسبت به سایر آزمون‌ها دارد و در سایر موارد آزمون $C_{m:n}^-$ بهتر از سایرین عمل می‌کند. بنابراین پیشنهاد می‌شود که در حالت m کم و طرح سانسور نوع دوم از آماره D ، در حالتی که m زیاد و سانسور در اولین مرحله انجام می‌شود از آماره T و در سایر حالات از آماره $C_{m:n}^-$ استفاده شود.

تقدیر و تشکر

نویسندگان مقاله بر خود لازم می‌دانند از سردبیر محترم مجله و داوران گرامی که نظراتشان موجب بهبودی مقاله گردید، تشکر کنند.

مراجع

- Aho, M., Bain, L. J. and Engelhardt, M. (1983), Goodness-of-Fit Tests for the Weibull Distribution with Unknown Parameters and Censored Sampling, *Journal of Statistical Computation and Simulation*, **18** (1), 59-69.
- Balakrishnan, N. and Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods, and Applications*, Birkhäuser, Boston.
- Balakrishnan, N. and Cramer, E. (2014), *The Art of Progressive Censoring: Application to Reliability and Quality*, Birkhäuser, New York.
- Balakrishnan, N., Habibi Rad, A. and Arghami, N. R. (2007), Testing Exponentiality Based on Kullback-Leibler Information with Progressively Type-II Censored Data, *IEEE Transactions on Reliability*, **56** (2), 301-307.
- Balakrishnan, N., Ng, H. K. T. and Kannan, N. (2002), A Test of Exponentiality Based on Spacings for Progressively Type-II Censored Data, in Huber-Carol, C., Balakrishnan, N., Nikulin, M. S. and Mesbah, M. (eds), *Goodness-of-Fit Tests and Model Validity*, Birkhäuser, Boston.
- Balakrishnan, N., Ng, H. K. T. and Kannan, N. (2004), Goodness-of-Fit Tests Based on Spacings for Progressively Type-II Censored Data from a General Location-Scale Distribution, *IEEE Transaction on Reliability*, **53** (3), 349-356.
- Baratpour, S. and Habibi Rad, A. (2016), Exponentiality Test Based on the Progressive Type-II Censoring via Cumulative Entropy, *Communications in Statistics-Simulation and Computation*, **45** (7), 2625-2637.
- Barr, D. R. and Davidson, T. (1973), A Kolmogorov-Smirnov Test for Censored Samples, *Technometrics*, **15** (4), 739-757.
- Bordes, L. (2004), Non-parametric Estimation Under Progressive Censoring, *Journal of Statistical Planning and Inference*, **119** (1), 171-189.
- Brunk, H. D. (1962), On the Range of Difference Between Hypothetical Distribution Function and Pyke's Modified Empirical Distribution Function, *Annals of Mathematical Statistics*, **33** (2), 525-532.
- D'Agostino, R. B. and Stephens, M. A. (eds), (1986), *Goodness-of-Fit Techniques*, Marcel Dekker, New York.
- Greenwood, M. (1964), The Statistical Study of Infectious Diseases, *Journal of the Royal Statistical Society*, **109** (2), 85-110.

- Habibi Rad, A., Yousefzadeh, F. and Balakrishnan, N. (2011), Goodness-of-Fit Test Based on Kullback-Leibler Information for Progressively Type-II Censored Data, *IEEE Transactions on Reliability*, **60** (3), 570-579.
- Hegazy, Y. A. S. and Green, J. R. (1975), Some New Goodness-of-Fit Tests Using Order Statistics, *Journal of Applied Statistics*, **24** (3), 299-308.
- Huber-Carol, C., Balakrishnan, N., Nikulin, M. S. and Mesbah, M. (eds), (2002), *Goodness-of-Fit Tests and Model Validity*, Birkhäuser, Boston.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distributions*, 2nd, Vol. 1, John Wiley & Sons, New York.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distributions*, 2nd, Vol. 2, John Wiley & Sons, New York.
- Lin, C. T., Huang, Y. L. and Balakrishnan, N. (2008), A New Method for Goodness-of-Fit Testing Based on Type-II Right Censored Sample, *IEEE Transactions on Reliability*, **57** (4), 633-642.
- Lurie, D., Hartley, H. O. and Stroud, M. R. (1974), A Goodness-of-Fit test for Censored Data, *Communications in Statistics*, **3** (8), 745-753.
- Moran, P. A. P. (1951), The Random Division of an Interval-Part II, *Journal of the Royal Statistical Society, Series B*, **13** (1), 147-150.
- Nelson, W. (1982), *Applied Life Data Analysis*, John Wiley & Sons, New York.
- Ng, H. K. T., Luo, L., Hu, Y. and Duan, F. (2012), Parameter Estimation of Three-parameter Weibull Distribution Based on Progressively Type-II Censored Samples, *Journal of Statistical Computation and Simulation*, **82** (11), 1661-1678.
- Pakyari, R. and Balakrishnan, N. (2012), A General Purpose Approximate Goodness-of-Fit for Progressively Type-II Censored Data, *IEEE Transactions on Reliability*, **61** (1), 238-244.
- Pakyari, R. and Balakrishnan, N. (2013), Goodness-of-Fit Tests for Progressively Type-II Censored Data from Location-Scale Distributions, *Journal of Statistical Computation and Simulation*, **83** (1), 167-178.
- Pettitt, A. N. and Stephens, M. A. (1976), Modified Cramér-von Mises Statistics for Censored Data, *Biometrika*, **63** (2), 291-298.
- Quesenberry, C. P. and Miller Jr, F. L. (1977), Power Studies of Some Tests for Uniformity, *Journal of Statistical Computation and Simulation*, **5** (3), 169-191.

- Smith, R. M. and Bain, L. J. (1976), Correlation Type Goodness-of-Fit Statistics with Censored Samples, *Communications in Statistics-Theory and Methods*, **5** (2), 119-132.
- Stephens, M. A. (1986), Tests Based on EDF Statistics, in D'Agostino, R. B. and Stephens, M. A. (eds), *Goodness-of-Fit Techniques*, Marcel Dekker, New York.
- Torabi, H. (2006), A New Method for Hypotheses Testing Using Spacing, *Statistics and Probability Letters*, **76** (13), 1345-1347.
- Torabi, H. (2008), A General Method for Estimating and Hypotheses Testing Using Spacing, *Journal of Statistical Theory and Applications*, **8** (2), 163-168.
- Wang, B. (2008), Goodness-of-Fit Test for the Exponential Distribution Based on Progressively Type-II Censored Sample, *Journal of Statistical Computation and Simulation*, **78** (2), 125-132.