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

حسین نادب، حمزه ترابی

گروه آمار، دانشگاه یزد

چکیده: در این مقاله، یك 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 $x_1, \ldots, x_n$ مشاهدات یک نمونه تصادفی از جامعه‌ای با تابع توزیع

[۰]آدرس الکترونیک مسئول مقاله: حسین نادب، honadeb@yahoo.com

[۰]کد موضوع‌بندی ریاضی (۲۰۱۰): 62N01، 62N03.

پیوسته F و هدف انجام آزمون فرضیه‌های

$$\begin{cases} H_{0}:\ F = F_{\boldsymbol{\theta}} \\ H_{1}:\ F \neq F_{\boldsymbol{\theta}}, \end{cases}$$

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

برای ساخت آزمون نیکویی برازش روش‌های مختلفی وجود دارد. برخی از آزمون‌های نیکویی برازش بر اساس فاصله بین تابع توزیع تجربی و تابع توزیع نظری روی بازه $(0, 1)$ یا تفاضل آماره‌های ترتیبی از امید ریاضی آن‌ها، یعنی $v_i = u_{i:n} - i/(n+1)$ ساخته می‌شوند. برای جزییات بیشتر درباره این آماره‌ها می‌توان به برانک (۱۹۶۲)، استیونس (۱۹۸۶) و هگازی و گرین (۱۹۷۵) مراجعه کرد. همچنین برخی دیگر از آزمون‌های نیکویی برازش مانند آماره گرین‌وود (۱۹۶۴)، آماره کوئسنبری و میلر (۱۹۷۷) و آماره موران (۱۹۵۱) بر اساس فاصله‌ها، $D_i = u_{i:n} - u_{i-1:n}$، $i = 1, ..., n+1$ ساخته شده‌اند 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 فاصله‌ها ارائه کرده است.

در سال‌های اخیر به‌کارگیری داده‌های سانسور شده و به‌ویژه سانسور فزاینده نوع دوم به‌منظور صرفه‌جویی در وقت و هزینه بسیار مورد توجه قرار گرفته است، به‌طوری که مقالات زیادی در زمینه‌ی نیکویی برازش بر اساس این نوع داده‌ها به چاپ رسیده است. به عنوان مثال، می‌توان به بالاکریشنان و همکاران (۲۰۰۲، ۲۰۰۴)، وانگ (۲۰۰۸) و پاکیاری و بالاکریشنان (۲۰۱۲، ۲۰۱۳) اشاره کرد. برای مطالعه نمونه‌های بیش‌تر در این رابطه به آهو و همکاران (۱۹۸۳)، بالاکریشنان و همکاران (۲۰۰۷)، برات‌پور و حبیبی‌راد (۲۰۱۶)، بار و دیویدسون (۱۹۷۳)، حبیبی‌راد 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 به‌طور تصادفی از آزمایش حذف می‌شوند. سپس در زمان رخداد دومین شکست، r_2 واحد از $n - r_1 - 2$ واحد باقیمانده به‌طور تصادفی از آزمایش حذف می‌شوند؛ به همین ترتیب

آزمایش تا لحظه رخداد mامین شکست ادامه می‌یابد. در این حالت به مشاهدات $x_{1:m:n},\dots,x_{m:m:n}$ یك نمونه سانسور شده فزاینده نوع دوم با بردار حذفیات $\boldsymbol{r}=(r_1,\dots,r_m)$ گفته می‌شود که $x_{j:m:n}$ زمان رخداد jامین شکست است. این طرح سانسور، تعمیم طرح سانسور نوع دوم است به‌طوری که اگر قرار دهیم $r_1=\ldots=r_{m-1}=0$ و $r_m=n-m$، این طرح به طرح سانسور نوع دوم تبدیل می‌شود. برای اطلاعات بیشتر در مورد سانسور فزاینده نوع دوم می‌توان 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 $x_{1:m:n},\cdots,x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم با بردار حذفیات $\boldsymbol{r}=(r_1,\dots,r_m)$ از یک خانواده مکان-مقیاس با تابع توزیع

$$F_{\boldsymbol{\theta}}(x)=F_{(\mu,\sigma)}(x)=F_0(\frac{x-\mu}{\sigma}),$$

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

$$\begin{aligned} u_{i:m:n} &= F_{\hat{\boldsymbol{\theta}}}(x_{i:m:n}), \quad i=1,\dots,m,\\ v_{i:m:n} &= u_{i:m:n}-\alpha_{i:m:n}, \quad i=1,\dots,m, \end{aligned}$$

به‌طوری که $\alpha_{i:m:n}$ امید ریاضی iامین آماره ترتیبی سانسور فزاینده نوع دوم از توزیع یکنواخت بر بازه $(۰, ۱)$ است و بنابر بالاکریشنان و آگاروالا (۲۰۰۰) به‌صورت

$$\alpha_{i:m:n} = ۱ - \prod_{k=m-i+۱}^{m} \frac{k + \sum_{j=m-i+۱}^{m} r_j}{۱ + k + \sum_{j=m-i+۱}^{m} r_j}, \quad i = ۱, \dots, m,$$

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

$$C^{+}_{m:n} = \max_{۱\le i\le m} (v_{i:m:n}), \quad C^{-}_{m:n} = \max_{۱\le i\le m} (-v_{i:m:n}), \quad C_{m:n} = \max(C^{-}_{m:n}, C^{+}_{m:n}),$$

$$K_{m:n} = C^{-}_{m:n} + C^{+}_{m;n}, \qquad T^{(۱)}_{m:n} = \sum_{i=۱}^{m} \frac{v^{۲}_{i:m:n}}{m}, \qquad T^{(۲)}_{m:n} = \sum_{i=۱}^{m} \frac{|v_{i:m:n}|}{m},$$

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 iامین آماره ترتیبی در سانسور فزاینده نوع دوم از تابع توزیع استاندارد $F_۰$ باشد. اگر

$$G_i = \frac{x_{i:m:n} - x_{i-۱:m:n}}{\mu_{i:m:n} - \mu_{i-۱:m:n}}, \qquad i = ۲, ۳, \dots, m,$$

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 $x_{1:m:n}, \cdots, x_{m:m:n}$ داده‌ها سانسور شده فزاینده نوع دوم باشند. تابع توزیع تجربی به‌صورت

$$F^{(m)}(x) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(x_{i:m:n} \leq x) = \frac{1}{m}\#(x_{i:m:n} \leq x),$$

تعریف می‌شود.

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

الف) $\sup_{i\geq 1} r_i < \infty$؛

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

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

برقرار باشد. در این صورت رابطه

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

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

۱.۳ آماره‌ی آزمون

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

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

تعریف می‌شود استفاده می‌کنیم، که در آن

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

و $\hat{\boldsymbol{\theta}}$ برآوردگری موجه برای $\boldsymbol{\theta}$ 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 یك لم را بیان و اثبات می‌کنیم.

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

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

برهان. برای اثبات، کافی است نشان داده شود:

$$\sup_x \lim_{y\to y_0} g(x,y) \leq \lim_{y\to y_0} \sup_x g(x,y), \tag{2}$$

و

$$\lim_{y\to y_0} \sup_x g(x,y) \leq \sup_x \lim_{y\to y_0} g(x,y). \tag{3}$$

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

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

با حدگیری از طرفین نامساوی بالا، به‌ازای هر x داریم:

$$\lim_{y\to y_\circ} g(x,y) \leq \lim_{y\to y_\circ} \sup_x g(x,y).$$

چون نامساوی بالا به‌ازای هر x برقرار است، پس واضح است که

$$\sup_x \lim_{y\to y_\circ} g(x,y) \leq \lim_{y\to y_\circ} \sup_x g(x,y).$$

برای اثبات رابطه (۳) کافی است نشان داده شود که به‌ازای هر $\epsilon > \circ$، مقدار $x_\circ$ موجود است به‌گونه‌ای که

$$\lim_{y\to y_\circ} \sup_x g(x,y) - \epsilon < \lim_{y\to y_\circ} g(x_\circ,y). \qquad (۴)$$

همچنین برای اثبات رابطه (۴) کافی است نشان دهیم که به‌ازای هر y، مقدار $x_\circ$ موجود است به‌طوری که

$$\sup_x g(x,y) = g(x_\circ,y). \qquad (۵)$$

چون g تابعی کراندار است پس رابطه (۵) برقرار است و در نتیجه لم ثابت می‌شود. $\square$

اینک می‌توان ویژگی‌های آماره D را بررسی نمود که به صورت قضیه آورده شده است.

قضیه ۲. فرض کنید تحت فرض $H_\circ$ شرایط

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

ب) $\hat{\boldsymbol{\theta}} \xrightarrow{\text{a.s}} (\xrightarrow{\text{P}}) \boldsymbol{\theta}$، هنگامی که $m \to \infty$؛

پ) $F_{\boldsymbol{\theta}}$ تابعی پیوسته از $\boldsymbol{\theta}$ باشد،

برقرار باشد. در این صورت $D \xrightarrow{\text{a.s}} (\xrightarrow{\text{P}}) \circ$ ، که در آن $\xrightarrow{\text{P}}$ نشان‌دهنده همگرایی در احتمال است.

برهان.

$$\begin{aligned} D &= \sup_{x\in\mathbb{R}} |F^{(m)}(x) - (۱ - (۱ - F_{\hat{\boldsymbol{\theta}}}(x))^{\hat{r}+۱})| \\ &\leq \sup_{x\in\mathbb{R}} |F^{(m)}(x) - (۱ - (۱ - F_{\boldsymbol{\theta}}(x))^{r+۱})| \\ &\quad + \sup_{x\in\mathbb{R}} |(۱ - (۱ - F_{\hat{\boldsymbol{\theta}}}(x))^{\hat{r}+۱}) - (۱ - (۱ - F_{\boldsymbol{\theta}}(x))^{r+۱})|. \end{aligned} \qquad (۶)$$

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

$$g(x,m)=(۱-(۱-F_{\hat{\boldsymbol{\theta}}}(x))^{\hat{r}+۱})-(۱-(۱-F_{\boldsymbol{\theta}}(x))^{r+۱}).$$

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

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

بنابراین با توجه به همگرایی $\hat{\boldsymbol{\theta}}$ و $\hat{r}$ و پیوستگی تابع $F_{\boldsymbol{\theta}}$ برحسب $\boldsymbol{\theta}$، قضیه ثابت می‌شود. □

قضیه ۳. الف) اگر μ پارامتر مکان و $\hat{\mu}$ یك برآوردگر پایای مکان باشد، آنگاه توزیع D به μ بستگی ندارد.
ب) اگر σ پارامتر مقیاس و $\hat{\sigma}$ یك 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 و پایای مقیاس باشند، آنگاه توزیع D به μ و σ بستگی ندارد.

برهان. الف) فرض کنید $x_{۱:m:n},\cdots,x_{m:m:n}$ 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:

$$z^{(۱)}_{i:m:n}=x_{i:m:n}-\hat{\mu}(\boldsymbol{x})=z_{i:m:n}-\hat{\mu}(\boldsymbol{z}),\qquad i=۱,\cdots,m.$$

از طرفی

$$F^{(m)}(x)=\frac{۱}{m}\sum_{i=۱}^{m}\mathrm{I}(x_{i:m:n}\leq x)=\frac{۱}{m}\sum_{i=۱}^{m}\mathrm{I}(z^{(۱)}_{i:m:n}\leq x-\hat{\mu}(\boldsymbol{x})),$$

و

$$F_{\hat{\mu}}(x)=F_{\circ}(x-\hat{\mu}(\boldsymbol{x})).$$

بنابراین

$$\begin{aligned}D&=\sup_{x\in\mathbb{R}}|\frac{۱}{m}\sum_{i=۱}^{m}\mathrm{I}(z^{(۱)}_{i:m:n}\leq x-\hat{\mu}(\boldsymbol{x}))-(۱-(۱-F_{\circ}(x-\hat{\mu}(\boldsymbol{x}))^{\hat{r}+۱})|\\&=\sup_{z\in\mathbb{R}}|\frac{۱}{m}\sum_{i=۱}^{m}\mathrm{I}(z^{(۱)}_{i:m:n}\leq z)-(۱-(۱-F_{\circ}(z))^{\hat{r}+۱})|.\end{aligned}$$

ب) فرض کنید $x_{1:m:n}, \cdots, x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم از تابع توزیع $F_\circ(\frac{x}{\sigma})$ باشند، که در آن $F_\circ$ تابع توزیع مربوط به متغیر تصادفی استاندارد Z است. پس از برآورد پارامتر σ قرار می‌دهیم:

$$z^{(2)}_{i:m:n} = \frac{x_{i:m:n}}{\hat{\sigma}(\boldsymbol{x})} = \frac{z_{i:m:n}}{\hat{\sigma}(\boldsymbol{z})}, \quad i = 1, \cdots, m.$$

از طرفی

$$F^{(m)}(x) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(x_{i:m:n} \leq x) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(z^{(2)}_{i:m:n} \leq \frac{x}{\hat{\sigma}(\boldsymbol{x})}),$$

و

$$F_{\hat{\sigma}}(x) = F_\circ(\frac{x}{\hat{\sigma}(\boldsymbol{x})}).$$

بنابراین

$$\begin{aligned} D &= \sup_{x\in\mathbb{R}} \Big|\frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(z^{(2)}_{i:m:n} \leq \frac{x}{\hat{\sigma}(\boldsymbol{x})}) - (1 - (1 - F_\circ(\frac{x}{\hat{\sigma}(\boldsymbol{x})}))^{\hat{r}+1})\Big| \\ &= \sup_{z\in\mathbb{R}} \Big|\frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(z^{(2)}_{i:m:n} \leq z) - (1 - (1 - F_\circ(z))^{\hat{r}+1})\Big|. \end{aligned}$$

پ) فرض کنید $x_{1:m:n}, \cdots, x_{m:m:n}$ داده‌های سانسور شده فزاینده نوع دوم از تابع توزیع $F_\circ(\frac{x-\mu}{\sigma})$ باشند، که در آن $F_\circ$ تابع توزیع مربوط به متغیر تصادفی استاندارد Z است. پس از برآورد پارامترهای μ و σ قرار می‌دهیم:

$$z^{(3)}_{i:m:n} = \frac{x_{i:m:n} - \hat{\mu}(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})} = \frac{z_{i:m:n} - \hat{\mu}(\boldsymbol{z})}{\hat{\sigma}(\boldsymbol{z})}, \quad i = 1, \cdots, m.$$

از طرفی

$$F^{(m)}(x) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(x_{i:m:n} \leq x) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{I}(z^{(3)}_{i:m:n} \leq \frac{x - \hat{\mu}(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})}),$$

و

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

بنابراین

$$\begin{aligned} D &= \sup_{x\in\mathbb{R}} |\frac{۱}{m}\sum_{i=۱}^{m} \mathrm{I}(z_{i:m:n}^{(۳)} \leq \frac{x-\hat{\mu}(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})}) - (۱-(۱-F_{\circ}(\frac{x-\hat{\mu}(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})}))^{\hat{r}+۱})| \\ &= \sup_{z\in\mathbb{R}} |\frac{۱}{m}\sum_{i=۱}^{m} \mathrm{I}(z_{i:m:n}^{(۳)} \leq z) - (۱-(۱-F_{\circ}(z))^{\hat{r}+۱})|. \end{aligned}$$

□

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

$$\begin{aligned} D^- &= \max\{\circ, \max_{۱\leq i\leq m}[(۱-(۱-F_{\hat{\boldsymbol{\theta}}}(x_{i:m:n}))^{\hat{r}+۱}) - \frac{i-۱}{m}]\}, \\ D^+ &= \max\{\circ, \max_{۱\leq i\leq m}[\frac{i}{m} - (۱-(۱-F_{\hat{\boldsymbol{\theta}}}(x_{i:m:n}))^{\hat{r}+۱})]\}. \end{aligned}$$

برهان. این رابطه مشابه با رابطه آماره آزمون کولموگروف–اسمیرنف است و به‌راحتی 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 قبلا توسط بالاکریشنان و همکاران (۲۰۰۴) و پاکیاری و بالاکریشنان (۲۰۱۳) استفاده شده و در جدول ۱ آورده شده‌اند. تمام شبیه‌سازی‌ها در نرم‌افزار R انجام شده است.

اینک همانند پاکیاری و بالاکریشنان (۲۰۱۳)، تحت فرضیه صفر تابع چگالی گامبل را به‌صورت

$$f_{(\mu,\sigma)}(x) = \frac{۱}{\sigma}\exp\{\frac{x-\mu}{\sigma}\}\exp\{-\exp\{\frac{x-\mu}{\sigma}\}\}, \quad x\in\mathbb{R}, \mu\in\mathbb{R}, \sigma>\circ,$$

و تحت فرضیه مقابل، تابع چگالی لگ‌گاما را به‌صورت

$$f_{\kappa}(x) = \frac{\kappa^{\kappa-\frac{۱}{۲}}}{\Gamma(\kappa)}\exp\{\sqrt{\kappa}x - \kappa\exp\{\frac{x}{\sqrt{\kappa}}\}\}, \quad \kappa>\circ,$$

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 $\alpha = ۰/۱$ 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 پر رنگ 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 ۱.۱ 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 $C^+_{m:n}$، $C^-_{m:n}$، $C_{m:n}$، $K_{m:n}$، $T^{(1)}_{m:n}$ و $T^{(2)}_{m:n}$ 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، یك 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 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 $\{۲\}, \{۵\}, \{۸\}, \{۱۱\}$) آزمون D عملکرد بهتری نسبت

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

طرح	n	m	$\boldsymbol{r} = (r_۱, ..., r_m)$
{۱}	۲۰	۸	$r_۱ = ۱۲$ و $r_i = ۰$ برای $i \neq ۱$
{۲}	۲۰	۸	$r_۸ = ۱۲$ و $r_i = ۰$ برای $i \neq ۸$
{۳}	۲۰	۸	$r_۱ = r_۸ = ۶$ و $r_i = ۰$ برای $i \neq ۱,۸$
{۴}	۲۰	۱۲	$r_۱ = ۸$ و $r_i = ۰$ برای $i \neq ۸$
{۵}	۲۰	۱۲	$r_{۱۲} = ۸$ و $r_i = ۰$ برای $i \neq ۱۲$
{۶}	۲۰	۱۲	$r_۳ = r_۵ = r_۷ = r_۹ = ۲$ و $r_i = ۰$ برای $i \neq ۳, ۵, ۷, ۹$
{۷}	۲۰	۱۶	$r_۱ = ۴$ و $r_i = ۰$ برای $i \neq ۱$
{۸}	۲۰	۱۶	$r_{۱۶} = ۴$ و $r_i = ۰$ برای $i \neq ۱۶$
{۹}	۲۰	۱۶	$r_۵ = ۴$ و $r_i = ۰$ برای $i \neq ۵$
{۱۰}	۴۰	۱۰	$r_۱ = ۳۰$ و $r_i = ۰$ برای $i \neq ۱$
{۱۱}	۴۰	۱۰	$r_{۱۰} = ۳۰$ و $r_i = ۰$ برای $i \neq ۱۰$
{۱۲}	۴۰	۱۰	$r_۱ = r_۵ = r_{۱۰} = ۱۰$ و $r_i = ۰$ برای $i \neq ۱, ۵, ۱۰$
{۱۳}	۴۰	۲۰	$r_۱ = ۲۰$ و $r_i = ۰$ برای $i \neq ۱$
{۱۴}	۴۰	۲۰	$r_{۲۰} = ۲۰$ و $r_i = ۰$ برای $i \neq ۲۰$
{۱۵}	۴۰	۲۰	$r_i = ۱$، برای $r_i = ۱, ۲, ..., ۲۰$
{۱۶}	۴۰	۳۰	$r_۱ = ۱۰$ و $r_i = ۰$ برای $i \neq ۱$
{۱۷}	۴۰	۳۰	$r_{۳۰} = ۱۰$ و $r_i = ۰$ برای $i \neq ۳۰$
{۱۸}	۴۰	۳۰	$r_۱ = r_{۳۰} = ۵$ $r_i = ۰$ برای $i \neq ۱, ۳۰$
{۱۹}	۶۰	۲۰	$r_۱ = ۴۰$ و $r_i = ۰$ برای $i \neq ۱$
{۲۰}	۶۰	۲۰	$r_{۲۰} = ۴۰$ و $r_i = ۰ i \neq ۲۰$
{۲۱}	۶۰	۲۰	$r_۱ = r_{۲۰} = ۱۰$ و $r_i = ۰$ $r_{۱۰} = ۲۰$ برای $i \neq ۱, ۱۰, ۲۰$

ادامه جدول ۱.

طرح	n	m	$\boldsymbol{r}=(r_۱,r_۲,...,r_m)$
{۲۲}	۶۰	۴۰	$r_۱=۲۰$ و $r_i=۰$ برای $i\neq ۱$
{۲۳}	۶۰	۴۰	$r_{۴۰}=۲۰$ و $r_i=۰$ برای $i\neq ۴۰$
{۲۴}	۶۰	۴۰	$r_{۲i-۱}=۱$ و $r_{۲i}=۰$ برای $i=۱,۲,...,۲۰$
{۲۵}	۶۰	۵۰	$r_۱=۱۰$ و $r_i=۰$ برای $i\neq ۱$
{۲۶}	۶۰	۵۰	$r_۱=۱۲$ و $r_i=۰$ برای $i\neq ۱$
{۲۷}	۶۰	۵۰	$r_۱=r_{۵۰}=۵$ و $r_i=۰$ برای $i\neq ۱,۵۰$

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 ∞.

طرح	κ	$C^+_{m:n}$	$C^-_{m:n}$	$C_{m:n}$	$K_{m:n}$	$T^{(۱)}_{m:n}$	$T^{(۲)}_{m:n}$	T	D
{۱}	۴	۰٫۰۴۷۵	**۰٫۱۷۶۳**	۰٫۱۱۳۹	۰٫۱۲۵۹	۰٫۱۱۷۰	۰٫۱۱۸۲	۰٫۱۲۹۹	۰٫۰۴۷۰
	۸	۰٫۰۳۹۶	**۰٫۲۰۶۹**	۰٫۱۲۶۹	۰٫۱۳۸۵	۰٫۱۳۵۸	۰٫۱۳۸۱	۰٫۱۵۰۲	۰٫۰۳۸۶
	∞	۰٫۰۲۵۶	**۰٫۲۹۳۲**	۰٫۱۷۸۲	۰٫۱۸۶۱	۰٫۱۹۴۱	۰٫۱۹۹۵	۰٫۲۱۹۹	۰٫۰۲۰۲
{۲}	۴	۰٫۰۷۳۵	۰٫۱۲۵۱	۰٫۱۱۰۸	۰٫۰۹۷۴	۰٫۱۰۵۰	۰٫۱۰۱۵	۰٫۰۸۴۹	**۰٫۱۳۳۵**
	۸	۰٫۰۶۹۳	۰٫۱۳۳۷	۰٫۱۱۵۳	۰٫۰۹۸۷	۰٫۱۱۰۴	۰٫۱۰۳۷	۰٫۰۸۲۴	**۰٫۱۳۶۱**
	∞	۰٫۰۵۷۴	۰٫۱۵۵۰	۰٫۱۲۷۶	۰٫۱۰۱۴	۰٫۱۲۲۳	۰٫۱۱۳۵	۰٫۰۸۱۰	**۰٫۱۵۷۷**
{۳}	۴	۰٫۰۶۵۴	**۰٫۱۳۵۹**	۰٫۱۰۹۰	۰٫۱۰۳۴	۰٫۱۰۴۲	۰٫۱۰۰۶	۰٫۰۸۹۱	۰٫۰۶۹۷
	۸	۰٫۰۵۹۰	**۰٫۱۴۸۳**	۰٫۱۱۵۱	۰٫۱۰۵۴	۰٫۱۰۸۰	۰٫۱۰۶۱	۰٫۰۹۰۱	۰٫۰۶۳۴
	∞	۰٫۰۴۵۳	**۰٫۱۸۱۰**	۰٫۱۳۱۲	۰٫۱۱۷۵	۰٫۱۲۵۷	۰٫۱۲۴۵	۰٫۰۹۶۰	۰٫۰۵۹۸

ادامه جدول ۲.

D	T	$T_{m:n}^{(۲)}$	$T_{m:n}^{(۱)}$	$K_{m:n}$	$C_{m:n}$	$C_{m:n}^{-}$	$C_{m:n}^{+}$	κ	طرح
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	
٫۰۱۱۴	۰٫۳۰۸۰	۰٫۳۰۹۱	۰٫۲۹۲۷	۰٫۲۵۱۵	۰٫۲۵۹۰	**۰٫۴۰۷۳**	۰٫۰۱۷۶	∞	

ادامه جدول ۲.

طرح	κ	$C^{+}_{m:n}$	$C^{-}_{m:n}$	$C_{m:n}$	$K_{m:n}$	$T^{(۱)}_{m:n}$	$T^{(۲)}_{m:n}$	T	D
{۱۱}	۴	۰/۰۷۳۷	۰/۱۲۶۸	۰/۱۱۳۶	۰/۰۹۵۷	۰/۱۰۷۵	۰/۱۰۵۸	۰/۰۸۷۶	**۰/۱۳۳۹**
	۸	۰/۰۶۵۲	۰/۱۳۶۳	۰/۱۲۰۴	۰/۰۹۵۵	۰/۱۱۱۴	۰/۱۰۷۱	۰/۰۸۱۰	**۰/۱۴۲۶**
	∞	۰/۰۵۴۷	**۰/۱۵۹۰**	۰/۱۳۶۴	۰/۰۹۸۶	۰/۱۲۳۴	۰/۱۱۸۷	۰/۰۸۱۷	۰/۱۴۹۳
{۱۲}	۴	۰/۰۸۳۶	**۰/۱۴۴۲**	۰/۱۱۳۰	۰/۱۱۹۴	۰/۱۱۶۳	۰/۱۱۷۳	۰/۰۹۴۸	۰/۰۷۵۵
	۸	۰/۰۸۰۵	**۰/۱۵۶۴**	۰/۱۱۸۷	۰/۱۲۷۴	۰/۱۲۵۶	۰/۱۲۳۸	۰/۰۹۳۵	۰/۰۷۰۷
	∞	۰/۰۷۳۵	**۰/۱۹۴۱**	۰/۱۳۷۵	۰/۱۴۸۹	۰/۱۴۷۳	۰/۱۴۴۴	۰/۱۱۰۴	۰/۰۶۹۰
{۱۳}	۴	۰/۰۳۴۱	**۰/۲۵۵۸**	۰/۱۶۹۴	۰/۱۳۵۷	۰/۱۷۶۰	۰/۱۷۸۲	۰/۲۱۱۰	۰/۰۲۶۱
	۸	۰/۰۲۵۲	**۰/۳۱۹۶**	۰/۲۱۷۳	۰/۱۶۵۰	۰/۲۳۴۳	۰/۲۳۱۸	۰/۲۷۸۵	۰/۰۱۹۳
	∞	۰/۰۱۳۰	**۰/۵۰۲۸**	۰/۳۷۷۱	۰/۲۶۹۹	۰/۴۱۱۰	۰/۴۱۵۸	۰/۴۹۰۳	۰/۰۰۶۸
{۱۴}	۴	۰/۰۵۷۵	**۰/۱۶۶۲**	۰/۱۲۱۱	۰/۱۰۱۹	۰/۱۲۰۱	۰/۱۱۷۸	۰/۰۹۶۵	۰/۱۳۹۰
	۸	۰/۰۴۸۴	**۰/۱۸۹۸**	۰/۱۳۶۴	۰/۱۰۶۳	۰/۱۳۵۰	۰/۱۳۳۵	۰/۱۰۱۹	۰/۱۴۷۳
	∞	۰/۰۳۴۴	**۰/۲۵۲۴**	۰/۱۷۷۵	۰/۱۲۴۶	۰/۱۷۹۸	۰/۱۷۵۶	۰/۱۳۶۲	۰/۱۸۶۳
{۱۵}	۴	۰/۱۷۳۷	۰/۱۸۳۵	۰/۱۷۹۳	**۰/۱۹۳۵**	۰/۱۷۰۹	۰/۱۵۳۴	۰/۱۴۰۳	۰/۱۳۶۶
	۸	۰/۲۰۸۸	۰/۲۱۷۳	۰/۲۱۶۶	**۰/۲۲۹۱**	۰/۲۰۶۵	۰/۱۸۶۷	۰/۱۷۲۸	۰/۱۵۴۶
	∞	۰/۲۹۷۱	۰/۳۱۵۸	۰/۳۱۰۶	**۰/۳۳۷۰**	۰/۳۰۶۶	۰/۲۷۹۵	۰/۲۹۳۵	۰/۲۲۹۶
{۱۶}	۴	۰/۰۳۵۸	**۰/۲۶۰۸**	۰/۱۷۳۰	۰/۱۳۶۹	۰/۱۸۲۹	۰/۱۸۵۴	۰/۲۵۰۹	۰/۰۲۷۱
	۸	۰/۰۲۹۵	۰/۳۳۵۰	۰/۲۲۳۶	۰/۱۷۶۱	۰/۲۴۳۱	۰/۲۴۷۳	**۰/۳۴۵۵**	۰/۰۱۲۳
	∞	۰/۰۳۴۹	۰/۵۳۲۳	۰/۴۰۰۳	۰/۳۱۰۹	۰/۴۵۱۱	۰/۴۵۳۱	**۰/۶۱۵۷**	۰/۰۱۲۳
{۱۷}	۴	۰/۰۴۵۸	**۰/۲۰۶۴**	۰/۱۴۲۰	۰/۱۱۷۹	۰/۱۴۶۳	۰/۱۴۱۹	۰/۱۲۸۸	۰/۲۷۹۰
	۸	۰/۰۳۷۷	**۰/۲۵۱۱**	۰/۱۶۶۴	۰/۱۳۴۵	۰/۱۷۵۱	۰/۱۷۳۳	۰/۱۵۷۲	۰/۱۹۲۶
	∞	۰/۰۲۵۸	**۰/۳۵۹۰**	۰/۲۵۲۴	۰/۱۹۴۴	۰/۲۷۳۷	۰/۲۷۳۸	۰/۲۸۰۹	۰/۲۷۹۰

ادامه جدول ۲.

طرح	κ	$C^{+}_{m:n}$	$C^{-}_{m:n}$	$C_{m:n}$	$K_{m:n}$	$T^{(۱)}_{m:n}$	$T^{(۲)}_{m:n}$	T	D
{۱۸}	۴	۰/۰۳۹۷	**۰/۲۲۷۱**	۰/۱۵۳۳	۰/۱۲۹۳	۰/۱۶۰۴	۰/۱۶۲۴	۰/۱۵۳۰	۰/۰۷۷۳
	۸	۰/۰۳۲۱	**۰/۲۸۰۷**	۰/۱۸۸۰	۰/۱۵۵۶	۰/۲۰۴۸	۰/۲۰۲۰	۰/۱۹۸۸	۰/۰۹۶۶
	∞	۰/۰۲۱۹	**۰/۴۲۶۵**	۰/۲۹۷۷	۰/۲۴۲۵	۰/۳۳۷۸	۰/۳۴۱۴	۰/۳۶۸۴	۰/۱۵۴۰
{۱۹}	۴	۰/۰۲۸۵	**۰/۲۹۰۵**	۰/۱۸۸۱	۰/۱۵۷۹	۰/۲۰۹۷	۰/۲۱۰۸	۰/۲۲۶۰	۰/۰۲۵۱
	۸	۰/۰۲۰۴	**۰/۳۶۷۳**	۰/۲۵۶۱	۰/۱۹۶۰	۰/۲۷۹۴	۰/۲۸۶۰	۰/۳۰۴۲	۰/۰۱۴۸
	∞	۰/۰۰۸۴	**۰/۵۸۲۵**	۰/۴۴۶۵	۰/۳۳۴۶	۰/۵۰۳۸	۰/۵۰۹۴	۰/۵۲۴۷	۰/۰۰۴۶
{۲۰}	۴	۰/۰۵۹۱	**۰/۱۵۷۳**	۰/۱۱۹۹	۰/۱۰۰۰	۰/۱۱۸۲	۰/۱۱۳۵	۰/۰۹۲۸	۰/۱۲۹۲
	۸	۰/۰۵۰۲	**۰/۱۷۹۱**	۰/۱۳۱۶	۰/۱۰۳۳	۰/۱۳۳۹	۰/۱۲۵۷	۰/۰۹۴۰	۰/۱۴۴۸
	∞	۰/۰۳۵۳	**۰/۲۳۱۷**	۰/۱۶۸۷	۰/۱۱۴۴	۰/۱۷۰۶	۰/۱۶۲۹	۰/۱۱۳۳	۰/۱۶۷۲
{۲۱}	۴	۰/۰۸۱۱	**۰/۱۶۹۰**	۰/۱۲۵۵	۰/۱۴۶۸	۰/۱۳۷۶	۰/۱۳۱۸	۰/۱۱۶۰	۰/۰۷۰۶
	۸	۰/۰۷۸۷	**۰/۱۹۵۴**	۰/۱۳۶۷	۰/۱۶۶۳	۰/۱۵۵۳	۰/۱۴۷۷	۰/۱۳۳۹	۰/۰۷۱۲
	∞	۰/۰۷۱۲	**۰/۲۶۶۵**	۰/۱۷۰۸	۰/۲۲۲۱	۰/۲۰۸۰	۰/۲۰۴۲	۰/۲۱۵۰	۰/۰۷۲۳
{۲۲}	۴	۰/۰۳۰۸	۰/۳۰۲۸	۰/۲۰۲۳	۰/۱۶۰۴	۰/۲۲۰۹	۰/۲۲۵۳	**۰/۳۲۱۸**	۰/۰۱۹۴
	۸	۰/۰۲۶۷	۰/۴۰۰۷	۰/۲۷۵۴	۰/۲۱۱۵	۰/۳۱۱۶	۰/۳۱۵۸	**۰/۴۵۱۱**	۰/۰۱۱۸
	∞	۰/۰۴۴۳	۰/۶۵۲۶	۰/۵۱۳۵	۰/۴۰۷۵	۰/۵۸۷۳	۰/۵۹۵۱	**۰/۷۵۶۴**	۰/۰۰۴۳
{۲۳}	۴	۰/۰۳۹۹	**۰/۲۲۲۱**	۰/۱۴۲۱	۰/۱۲۲۷	۰/۱۵۴۳	۰/۱۵۲۲	۰/۱۳۵۲	۰/۱۳۰۰
	۸	۰/۰۳۲۰	**۰/۲۶۵۴**	۰/۱۷۴۴	۰/۱۴۱۹	۰/۱۸۷۶	۰/۱۸۸۱	۰/۱۷۵۳	۰/۱۵۰۸
	∞	۰/۰۲۱۳	**۰/۳۸۹۵**	۰/۲۶۴۸	۰/۲۱۲۰	۰/۳۰۴۶	۰/۳۰۶۶	۰/۳۲۱۱	۰/۲۰۳۸
{۲۴}	۴	۰/۰۸۵۶	**۰/۲۶۰۰**	۰/۱۸۷۰	۰/۲۰۱۰	۰/۱۹۹۱	۰/۱۹۱۶	۰/۲۵۴۵	۰/۱۵۸۲
	۸	۰/۰۹۳۶	۰/۳۲۸۰	۰/۲۳۸۴	۰/۲۵۰۵	۰/۲۶۱۴	۰/۲۵۵۷	**۰/۳۵۶۵**	۰/۲۰۲۳
	∞	۰/۱۳۰۱	۰/۵۱۱۸	۰/۳۹۲۵	۰/۴۱۷۴	۰/۴۶۶۸	۰/۴۶۲۷	**۰/۶۴۵۸**	۰/۳۵۶۴

ادامه جدول ۲.

طرح	κ	$C^{+}_{m:n}$	$C^{-}_{m:n}$	$C_{m:n}$	$K_{m:n}$	$T^{(۱)}_{m:n}$	$T^{(۲)}_{m:n}$	T	D
{۲۵}	۴	۰/۰۴۰۷	۰/۳۰۹۴	۰/۲۱۵۲	۰/۱۸۱۴	۰/۲۴۱۸	۰/۲۴۳۱	**۰/۳۵۴۳**	۰/۰۲۲۰
	۸	۰/۰۴۶۶	۰/۴۰۷۸	۰/۲۹۴۹	۰/۲۴۵۳	۰/۳۳۲۹	۰/۳۴۳۷	**۰/۵۱۰۰**	۰/۰۱۵۸
	∞	۰/۱۲۷۸	۰/۶۵۶۹	۰/۵۳۸۳	۰/۴۸۹۹	۰/۶۲۶۵	۰/۶۳۸۶	**۰/۸۱۸۴**	۰/۰۲۹۶
{۲۶}	۴	۰/۰۳۳۹	**۰/۲۵۹۵**	۰/۱۷۷۱	۰/۱۵۲۹	۰/۱۹۵۵	۰/۱۹۴۲	۰/۱۸۹۶	۰/۱۹۶۵
	۸	۰/۰۳۰۵	**۰/۳۲۷۰**	۰/۲۳۰۶	۰/۱۹۱۸	۰/۲۵۵۱	۰/۲۵۴۳	۰/۲۷۷۲	۰/۲۳۸۳
	∞	۰/۰۳۵۶	۰/۵۰۵۲	۰/۳۸۰۹	۰/۳۲۴۵	۰/۴۴۴۴	۰/۴۵۰۴	**۰/۵۲۶۸**	۰/۳۷۱۳
{۲۷}	۴	۰/۰۳۲۵	**۰/۲۸۷۶**	۰/۱۹۴۳	۰/۱۶۴۷	۰/۲۱۳۲	۰/۲۱۲۲	۰/۲۳۰۶	۰/۱۰۲۵
	۸	۰/۰۲۹۵	**۰/۳۶۷۳**	۰/۲۵۴۱	۰/۲۱۱۵	۰/۲۸۳۸	۰/۲۸۸۱	۰/۳۳۹۱	۰/۱۴۵۹
	∞	۰/۰۴۳۳	۰/۵۷۱۰	۰/۴۴۰۰	۰/۳۷۴۹	۰/۵۱۲۹	۰/۵۲۲۶	**۰/۶۲۷۵**	۰/۲۷۱۶

جدول ۳: لگاریتم داده‌های به‌دست 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.

آماره	$C^{+}_{m:n}$	$C^{-}_{m:n}$	$C_{m:n}$	$K_{m:n}$	$T^{(۱)}_{m:n}$	$T^{(۲)}_{m:n}$	T	D
مقدار آماره	۰/۱۰۳۲	۰/۱۰۳۰	۰/۱۰۳۲	۰/۲۰۶۲	۰/۰۰۳۹	۰/۰۴۷۸	۰/۳۵۵۴	۰/۱۷۹۸
مقدار احتمال	۰/۶۱۰۶	۰/۳۹۰۰	۰/۷۰۹۶	۰/۵۱۸۹	۰/۵۵۸۵	۰/۶۸۸۴	۰/۱۳۵۴	۰/۴۹۰۳

به سایر آزمون‌ها دارد. در حالاتی که m زیاد و سانسور در اولین مرحله انجام می‌شود (مانند طرح‌های $\{۱۶\}, \{۲۲\}, \{۲۴\}$) 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.

A Goodness of Fit Test for Location-Scale Family of Distributions Under Type-II Progressive Censoring

Nadeb, H., Torabi, H.

Department of Statistics, Yazd University, Yazd , Iran.

Abstract: In this paper, a general method for goodness of fit test for the location-scale family of distributions under Type-II progressive censoring is presented and its properties are investigated. Then, using Monte Carlo simulation studies, the power of this test is compared with the powers of some existing tests for testing the Gumbel distribution. Finally the proposed test is used for fitting a distribution to a real data set.

Keywords: Goodness of fit test, location-scale distributions family, Type-II progressive censoring, Monte Carlo simulation.

Mathematics Subject Classification (2010): 62N01, 62N03.