آزمون نیکویی برازش برای توزیع نمایی وزنی

محمدمهدی مقامی[۱]، نصراله ایران‌پناه

گروه آمار، دانشگاه اصفهان

تاریخ دریافت: ۱۳۹۰/۵/۳۱ تاریخ پذیرش: ۱۳۹۱/۴/۱۲

چکیده: در رده جدیدی از توزیع‌های نمایی وزنی که توسط گوپتا و کاندو [۱] ارائه شد، پارامتر چولگی به توزیع نمایی اضافه گردیده است. بنابراین توزیع نمایی وزنی دارای پارامترهای چولگی و مقیاس است. در این مقاله آزمون نیکویی برازش برای این رده با پارامترهای مجهول را بررسی می‌کنیم. آزمون بر مبنای آماره‌های معروف اندرسون و کلموگروف-اسمیرنف انجام می‌گیرد. برای یافتن چندک‌های آماره اندرسون از روش بوت استرپ اما در مورد آماره کلموگروف-اسمیرنف از روش دیگری استفاده می‌کنیم. برای برآورد پارامترها از روش ماکسیمم درست‌نمایی استفاده می‌شود. با استفاده از شبیه‌سازی مونت کارلو به بررسی اندازه و توان آزمون‌ها برای فرض‌های مقابل گوناگون و اندازه‌های نمونه متفاوت پرداخته‌ایم. نتایج نشان می‌دهد که آزمون کلموگروف-اسمیرنف دارای توان بالاتری نسبت به آزمون اندرسون است.

واژه های کلیدی: آزمون نیکویی برازش، بوت استرپ پارامتری، تابع توزیع تجربی، شبیه‌سازی مونت-کارلو.

رده‌بندی ریاضی: ۶۲G۱۰، ۶۲F۴۰

۱- مقدمه

روش‌های متنوعی برای اضافه کردن پارامتر شکل به مدل نمایی وجود دارد که منجر به توزیع-های نمایی وزنی می‌شود. برای مثال توزیع‌های گاما و نمایی تعمیم‌یافته دو تعمیم متفاوت از توزیع نمایی هستند. گوپتا و کاندو [۱] به معرفی رده توزیع‌های نمایی وزنی براساس ایده آزالینی [۲] پرداختند. این رده با اضافه نمودن یک پارامتر چولگی به توزیع نمایی به‌دست می-

۱- آدرس الکترونیکی نویسنده مسئول مقاله: محمد مهدی مقامی maghami8@gmail.com

آید. در حالت خاص اگر این پارامتر چولگی برابر یک باشد، این توزیع به توزیع نمایی تعمیم یافته‌ی گوپتا و کاندو [۳] تبدیل می‌شود.

آزالینی [۲] اولین بار توزیع چوله نرمال را با اضافه کردن پارامتر چولگی به توزیع نرمال مطرح کرد. بعد از آن تحقیقات وسیعی برای افزودن پارامتر چولگی به توزیع‌های متقارن انجام شد. برای مثال چوله تی، چوله کوشی، چوله لاپلاس و چوله لوژستیک تعریف شدند و خواص مختلف و شیوه استنباط برای آن‌ها بحث شد (برای مثال مراجع [۴، ۵ و ۶] را ببینید).

گوپتا و کاندو [۱] به بررسی برخی از خواص مهم توزیع نمایی وزنی مانند توابع توزیع و چگالی، تابع مولد گشتاور، تولید اعداد تصادفی، ارتباط این رده با توزیع‌های نمایی، گاما و نمایی تعمیم‌یافته و هم‌چنین برآورد پارامترهای آن به روش‌های گشتاوری و ماکسیمم درست‌نمایی پرداختند. سرانجام با مثال‌هایی از داده‌های واقعی نشان دادند که رده مطرح شده نسبت به توزیع‌های وایبل، گاما و نمایی تعمیم یافته دارای برتری است.

یک مسأله مهم برای برازش داده‌ها توسط یک توزیع، آزمون نیکویی برازش است. برای این منظور، آزمون‌های بر مبنای توزیع تجربی اختلاف بین تابع توزیع تحت فرض صفر و تابع توزیع تجربی را اندازه می‌گیرند. این آزمون‌ها به دو خانواده کلی کرامر و کلموگروف تقسیم می‌شوند. خانواده کرامر شامل آماره‌های کرامر-ون میزز، واتسون، اندرسون-دارلینگ و خانواده کلموگروف شامل آماره‌های D^+, D^-، گلموگروف- اسمیرنف و کیوپر است. در این مقاله از آماره‌های اندرسون و کلموگروف-اسمیرنف استفاده می‌شود.

در بخش دوم به معرفی رده نمایی وزنی و برخی از خواص مهم آن پرداخته و به بررسی برآورد پارامترهای آن با استفاده از روش‌های گشتاوری و ماکسیمم درست‌نمایی می‌پردازیم. در بخش سوم، الگوریتم آزمون نیکویی برازش برمبنای تابع توزیع تجربی را مطرح می‌کنیم. سپس آزمون اندرسون با پارامتر مقیاس معلوم انجام می‌شود. چون پارامتر شکل توزیع مجهول است، برای یافتن چندک آماره آزمون اندرسون از روش بوت استرپ استفاده می‌گردد. در ادامه به بررسی اندازه واقعی آزمون و توان آن می‌پردازیم. برای بررسی توان از برخی فرض‌های مقابل مناسب مانند لاگ نرمال، گاما، کای-دو، وایبل، نیم نرمال، نیم کوشی و یکنواخت استفاده می‌شود. در بخش چهارم آزمون اندرسون با پارامتر مقیاس مجهول بررسی می‌شود. بخش پنجم آزمون نیکویی برازش را بر مبنای آماره کلموگروف-اسمیرنف و پارامتر مقیاس دلخواه بررسی می‌کند. ولی برای یافتن چندک‌ها به‌جای استفاده از روش بوت‌استرپ از روش دیگری استفاده می‌کنیم که دارای سه مزیت نسبت به‌روش بخش چهارم است. این روش اولاً توان بالاتری دارد، ثانیا این روش نیازی به شبیه‌سازی تکراری و وقت‌گیر بوت‌استرپ نیست و سرانجام نیاز به در نظر گرفتن پارامتر مقیاس نداریم. نقاط بحرانی آزمون با استفاده از شبیه سازی‌مونت‌کارلو به‌دست

می‌آیند. بررسی اندازه واقعی آزمون کلموگروف-اسمیرنف و مقایسه توان آن با آزمون اندرسون نشان‌دهنده برتری این آزمون در اکثر فرض‌های مقابل است.

۲- رده جدیدی از توزیع های نمایی وزنی

گوییم متغیر تصادفی X دارای توزیع نمایی وزنی با پارامترهای شکل $\alpha > ۰$ و مقیاس $\lambda > ۰$ است، اگر دارای چگالی زیر باشد:

$$f_X(x;\alpha,\lambda) = \frac{\alpha+۱}{\alpha}\lambda e^{-\lambda x}\left(۱-e^{-\alpha\lambda x}\right); \quad x > ۰ \tag{۱}$$

و آن را با نماد $WE(\alpha,\lambda)$ نشان می‌دهیم [۱]. برای تولید عدد تصادفی از $WE(\alpha,\lambda)$ از لم زیر استفاده می‌شود.

لم ۱. فرض کنید $U \sim \exp(\lambda)$ و $V \sim \exp\left(\lambda(۱+\alpha)\right)$ و از هم مستقل باشند. در این-صورت $X = U + V \sim WE(\alpha,\lambda)$.

اثبات. گوپتا و کاندو [۱] را ببینید.

تابع توزیع $WE(\alpha,\lambda)$ و تابع مولد گشتاور برای مقادیر $|t| < \lambda^{-۱}$ به‌صورت

$$F_X(x;\alpha,\lambda) = \frac{\alpha+۱}{\alpha}\left[۱-e^{-\lambda x} - \frac{\lambda}{\lambda+\alpha\lambda}\left(۱-e^{-\lambda(۱+\alpha)x}\right)\right]; \ x > ۰ \tag{۲}$$

و

$$M_X(t) = \left(۱-\frac{t}{\lambda}\right)^{-۱}\left[۱-\frac{t}{\lambda(۱+\alpha)}\right]^{-۱}; \ |t| < \frac{۱}{\lambda} \tag{۳}$$

به‌دست می‌آیند. برای $\lambda = ۱$ توزیع را با $WE(\alpha)$ نشان می‌دهیم. اگر $\bar{x}$ و $s^۲$ به‌ترتیب میانگین و واریانس نمونه مشاهده شده باشند، برآوردهای گشتاوری پارامترهای شکل و مقیاس به‌صورت زیر به‌دست می‌آیند:

$$\tilde{\alpha}=\frac{-\left(\bar{x}^{۲}-۲s^{۲}\right)+\sqrt{\left(\bar{x}^{۲}-۲s^{۲}\right)^{۲}-۲\left(\bar{x}^{۲}-s^{۲}\right)\left(\bar{x}^{۲}-۲s^{۲}\right)}}{\bar{x}^{۲}-s^{۲}}$$
$$\tilde{\lambda}=\frac{۱}{\bar{X}}\left(۱+\frac{۱}{۱+\tilde{\alpha}}\right). \qquad (۴)$$

اگر در رابطه‌ی (۴) عبارت زیر رادیکال منفی باشد، آن‌گاه برآورد گشتاوری پارامتر شکل موجود نیست و در نتیجه برآورد گشتاوری پارامتر مقیاس نیز موجود نخواهد بود. گوپتا و کاندو [۱] برای محاسبه برآوردهای ماکسیمم درست‌نمایی این توزیع از روش زیر استفاده می‌کنند:

با استفاده از تغییر پارامتر $\beta=\alpha\lambda$ تابع درست‌نمایی براساس نمونه مشاهده $\{x_{۱},...,x_{n}\}$ به‌صورت

$$l\left(x_{۱},...,x_{n}\right)=n\ln\left(\beta+\lambda\right)-n\ln\beta+n\ln\lambda-\lambda\sum_{i=۱}^{n}x_{i}+\sum_{i=۱}^{n}\ln(۱-e^{-\beta x_{i}}) \quad (۵)$$

به‌دست می‌آید. برای β ثابت با مشتق‌گیری از معادله (۵) نسبت به λ، معادله (۵) با مقدار λ

$$\hat{\lambda}(\beta)=\frac{۱}{۲\bar{x}}\left(\sqrt{(\beta\bar{x}-۲)^{۲}+۴\beta\bar{x}}-(\beta\bar{x}-۲)\right)$$

ماکسیمم می‌گردد. با جایگزینی $\hat{\lambda}(\beta)$ در معادله (۵) این معادله فقط شامل مقدار نامعلوم β خواهد بود. با ماکسیمم نمودن این رابطه با استفاده از روش‌های عددی $\hat{\beta}$ به‌دست می‌آید. بنابراین برآورد ماکسیمم درست‌نمایی β را می‌توان از حل یک معادله غیرخطی به‌دست آورد. با جایگزینی $\hat{\lambda}(\beta)$ در لگاریتم درست‌نمایی و سپس مشتق‌گیری برحسب β، معادله غیر-خطی $h(\beta)=\beta$ به‌دست می‌آید که به‌صورت زیر است:

$$h(\beta)=\left[\frac{۱}{\beta+\lambda(\hat{\beta})}+\frac{۱}{n}\sum_{i=۱}^{n}\frac{x_{i}e^{-\beta x_{i}}}{۱-e^{-\beta x_{i}}}\right]^{-۱}.$$

اما حل این معادله غیرخطی نیاز به نقطه اولیه‌ای مانند $\beta_{۰}$ برای β دارد. انتخاب این نقطه اولیه تاثیر به‌سزایی در نتیجه برآورد دارد. نتایج شبیه‌سازی نشان داد که در مواردی با نقاط ابتدایی نامناسب حتی برآورد ماکسیمم درست‌نمایی پارامترها به‌دست نمی‌آید. از طرفی چون $\beta=\alpha\lambda$، پس می‌توان حاصل‌ضرب برآوردهای گشتاوری پارامترها را به‌عنوان نقطه ابتدایی β در نظر گرفت. اما دیدیم که این برآوردها در حالت کلی موجود نیستند. بنابراین پیشنهاد می‌کنیم که به‌جای حل معادله غیر خطی $h(\beta)=\beta$، با استفاده از یک تابع بهینه، مستقیماً از ماکسیمم نمودن تابع $l(\beta)$ بر حسب β استفاده شود. در این‌صورت نیاز به نقطه اولیه

نداشته و نگران نامناسب بودن آن نیستیم. در واقع به‌طور شهودی و با رسم نمودار $l(\beta)$ نیز می‌توان $\hat{\beta}$ را تشخیص داد. در این مقاله برای برآورد پارامترها از این روش استفاده می‌شود.

۳- آزمون نیکویی برازش برای توزیع نمایی وزنی بر مبنای آماره اندرسون

آزمون‌های نیکویی برازش این فرض را که آیا نمونه تصادفی $X_۱,X_۲,...,X_n$ از جامعه‌ای با تابع توزیع $F(.)$ آمده‌اند، می‌آزماید؛ یعنی

$$H_۰ : X_۱,X_۲,...,X_n \sim F_X(.). \qquad (۶)$$

در این‌جا $F(.)$ تابع توزیع $WE(\alpha,\lambda)$ است. ابتدا حالت $\lambda = ۱$ را در نظر می‌گیریم. بنابراین $F(.)$ تابع توزیع $WE(\alpha)$ در رابطه (۲) با $\lambda = ۱$ است. در حالت کلی برای داده‌های واقعی X ابتدا برآورد پارامتر مقیاس $\hat{\lambda}$ را از روش ماکسیمم درست‌نمایی محاسبه و سپس آزمون را برای داده‌های $Y = \hat{\lambda}X$ به‌کار می‌بریم. برای آزمون در این بخش، از آماره آزمون اندرسون استفاده می‌کنیم. مقادیر بزرگ این آماره اختلاف معناداری را بین تابع توزیع تجربی و تابع توزیع مفروض نشان می‌دهد و بنابراین فرض صفر رد می‌گردد. مقدار دقیق این آماره از رابطه

$$A^۲ = n\int_{-\infty}^{\infty}\left[F_n(x)-F(x)\right]^۲\left[F(x)\left(۱-F(x)\right)\right]^{-۱} dF(x)$$

به‌دست می‌آید. برای یافتن مقدار آماره، از تقریب استفنز [۷] استفاده می‌شود که از رابطه

$$A^۲ \simeq -n-\frac{۱}{n}\sum_{i=۱}^{n}(۲i-۱)\left[\mathbf{log}\, p_{(i)}+\mathbf{log}\left(۱-p_{(n+۱-i)}\right)\right] \qquad (۷)$$

به‌دست می‌آید و در آن $p_{(i)} = F\left(x_{(i)},\alpha\right)$. چون پارامتر شکل α مجهول است، از برآورد ماکسیمم درست‌نمایی آن استفاده می‌کنیم. تابع درست‌نمایی به شکل

$$l\left(x_۱,...,x_n\right) = n\ln\left(\alpha+۱\right)-n\ln\alpha+\sum_{i=۱}^{n}\ln(۱-e^{-\alpha x_i}) \qquad (۸)$$

به‌دست می‌آید. با ماکسیمم نمودن (۸) بر حسب α برآورد ماکسیمم درست‌نمایی به‌دست می‌آید. برای یافتن چندک‌های آزمون از روش بوت‌استرپ استفاده می‌گردد. برای انجام آزمون برای یک نمونه تصادفی داده شده به اندازه n به‌طور خلاصه باید مراحل زیر را انجام داد:

۱) با داشتن نمونه $x_۱,...,x_n$ برآورد پارامتر شکل را محاسبه و با $\hat{\alpha}$ نشان می‌دهیم.

۲) مقدار آماره آزمون $A_n^۲$ را با استفاده از $\hat{\alpha}$ و $\hat{p}_{(i)}$ و معادله (۷) به‌دست می‌آوریم.

الف) نمونه بوت استرپ به حجم n از $WE(\hat{\alpha})$ تولید می‌کنیم.

ب) با استفاده از نمونه بوت‌استرپ تولید شده در مرحله قبل، برآورد ماکسیمم درستنمایی α، یعنی $\hat{\alpha}^*$ را محاسبه می‌کنیم.

پ) با استفاده از نمونه بوت‌استرپ و $\hat{\alpha}^*$ مقدار آماره آزمون یعنی $A_n^{*۲}$ را محاسبه می‌کنیم.

۳) مراحل الف–پ را B بار تکرار کرده و مقادیر $A^{*۲}_{nj}, j=۱,۲,...,B$ را به‌دست می‌آوریم.

۴) $\tilde{C}_n(۰/۹۵)$ را با $A^{*۲}_{n(B\times ۰/۹۵)}$ که درآن $A^{*۲}_{n(j)}, j=۱,...,B$ نشان‌دهنده مقدار ترتیبی j-ام $A^{*۲}_{nj}$ است، به‌دست می‌آوریم. اگر مقدار $A_n^۲$ محاسبه شده از $\tilde{C}_n(۰/۹۵)$ بزرگ‌تر شود، در این‌صورت فرض صفر را در سطح ۰/۰۵ رد می‌کنیم.

جدول۱: برآورد اندازه آزمون با استفاده از آماره $A^۲$ که با شبیه سازی $B=۱۰۰۰$ نمونه در سطح ۰/۰۵ به‌دست آمده است.

n	α							
	۱۰۰	۲۰	۱۰	۵	۳	۱	۰/۵	۰/۲۵
۵۰	۰/۰۱۹	۰/۰۳۹	۰/۰۵	۰/۰۴۸	۰/۰۴۴	۰/۰۲۳	۰/۰۴۵	۰/۰۴۲
۱۰۰	۰/۰۴۸	۰/۰۴۳	۰/۰۳۲	۰/۰۵۳	۰/۰۳۷	۰/۰۵۳	۰/۰۵۴	۰/۰۳۹

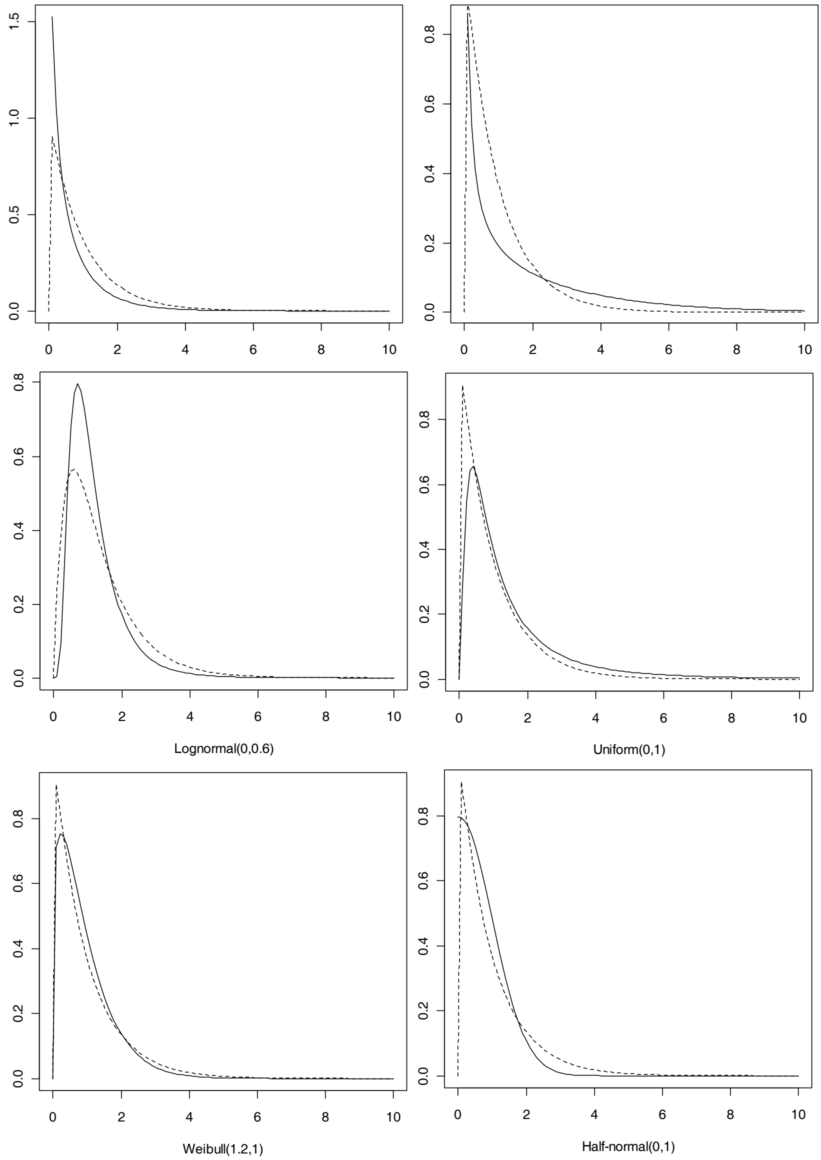
برای بررسی سطح واقعی آزمون از نمونه‌های شبیه‌سازی شده از $WE(\alpha)$ برای رده وسیعی از مقادیر پارامتر شکل، یعنی $\alpha\in\{۰/۲۵,۰/۵,۱,۳,۵,۱۰,۲۰,۱۰۰\}$ استفاده می‌کنیم. برآورد اندازه به‌وسیله شبیه‌سازی برای سطح ۰/۰۵ و $B=۱۰۰۰$ نمونه مونت‌کارلو در جدول ۱ آمده است. باتوجه به این جدول می‌بینیم که برآورد اندازه آزمون به سطح معناداری اسمی ۰/۰۵ بسیار نزدیک است.

برای بررسی توان از فرض‌های مقابل مناسب استفاده می‌کنیم. چون فرض صفر شامل پارامتر مقیاس نیست، پس توزیع فرض‌های مقابل را نیز بدون پارامتر مقیاس در نظر می‌گیریم. ملاحظه می‌شود که برای برخی فرض‌های مقابل، آزمون توان کمی دارد. به‌عبارت دیگر آزمون قابلیت تشخیص توزیع را ندارد. دلیل این مطلب این است که این توزیع‌ها شباهت زیادی با

توزیع نمایی وزنی دارند. برای مثال اگر توزیع $Gamma(۲,۱)$ را در نظر بگیریم و پارامتر شکل توزیع نمایی وزنی را از آن برآورد کنیم، به مقدار $\hat{\alpha}=۰/۰۰۰۲۲$ خواهیم رسید. اکنون با رسم چگالی توزیع‌های $Gamma(۲,۱)$ و $WE(\hat{\alpha},۱)$ درمی‌یابیم که دو توزیع دقیقاً بر هم منطبق هستند. پس انتظار داریم که آزمون قادر به تشخیص دو توزیع نباشد. در حالت کلی هرچه اختلاف بین تابع چگالی توزیع مورد نظر و نمایی وزنی برآورد شده کمتر باشد توان کمتر و هرچه این اختلاف بیشتر باشد، انتظار داشتن توان بالاتری خواهیم داشت.

جدول۲: برآورد توان آماره $A^۲$ برای برخی فرض‌های مقابل و سطح ۰/۰۵ و $B=۱۰۰۰$ شبیه‌سازی مونت‌کارلو.

فرض مقابل	$n=۵۰$	$n=۱۰۰$	$n=۱۵۰$
$Gamma(۰/۶,۱)$	۰/۹۴۳	۰/۹۵۶	۰/۹۸۸
$Gamma(۱,۱)$	۰/۰۵۰	۰/۰۶۳	/۰۶۸
$Gamma(۲,۱)$	۰/۰۵۲	۰/۱۵۳	۰/۱۶۶
$Gamma(۵,۱)$	/۹۴۸	/۹۵۶	۰/۹۶۹
$Chi\ squared\ (۱)$	۰/۹۲۷	۰/۹۵۳	۰/۹۸۴
$Chi\ squared\ (۲)$	۰/۹۱۵	۰/۹۳۶	۰/۹۷۲
$Chi\ squared\ (۱۰)$	۹۴۲	۰/۹۷۰	۰/۹۷۸
$Weibull\ (۰/۵,۱)$	۰/۹۷۷	۹۸۳۶	۰/۹۹۰
$Weibull\ (۱/۲,۱)$	۰/۱۱۵	۰/۱۲۳	۰/۱۸۲
$Weibull\ (۲,۱)$	۰/۹۵۸	۰/۹۸۲	۰/۹۵۹
$Lognormal\ (۰/۰,۰/۶)$	۰/۷۶۳	۰/۹۵۰	۰/۹۹۶
$Lognormal\ (۰/۰,۱)$	۰/۴۳۲	۰/۷۹۵	۰/۸۸۶
$Lognormal\ (۰/۰,۵)$	۰/۹۲۵	۰/۹۴۶	۰/۹۷۳
$Uniform\ (۰/۰,۱)$	۰/۹۵۰	۰/۹۴۸	۰/۹۸۷
$Half-normal\ (۰/۰,۱)$	۰/۲۲۵	۰/۵۲۱	۰/۸۳۷
$Half-Cauchy\ (۰/۰,۱)$	۰/۹۷۱	۰/۹۷۰	۰/۹۹۱



شکل۱: نمودار چگالی برخی از فرض‌های مقابل (نمودار پررنگ) و برآورد آن توسط توزیع $WE\,(\alpha)$ (نمودار نقطه چین).

در شکل ۱ نمودار برخی از چگالی‌های مقابل به همراه چگالی $WE(\hat{\alpha})$ برآورد شده رسم شده است. در این نمودارها چگالی اصلی را به‌صورت پررنگ و چگالی نمایی وزنی برآورد شده را با نقطه چین نشان داده‌ایم.

۴- آزمون برای توزیع شامل پارامتر مقیاس

در این بخش آزمون را در حضور پارامتر مقیاس بررسی می‌کنیم. بنابراین در فرض صفر (۶)، $F(.)$ تابع توزیع $WE(\alpha,\lambda)$ است. روش یافتن چندک‌ها مشابه بخش سوم است با این تفاوت که چون پارامتر مقیاس نامعلوم است از برآورد ماکسیمم درست‌نمایی آن استفاده می‌گردد. بنابراین با داشتن نمونه $\{x_۱,...,x_n\}$ ابتدا با استفاده از روش پیشنهاد شده در بخش دوم، برآورد ماکسیمم درست‌نمایی پارامتر مقیاس را محاسبه و با $\hat{\lambda}$ نشان می‌دهیم. اکنون با ضرب این برآورد در نمونه، مقادیر $\left\{y_۱ = \hat{\lambda}x_۱,...., y_n = \hat{\lambda}x_n\right\}$ به‌دست می‌آیند. برای nهای بزرگ، داده‌های $\{y_۱,...,y_n\}$ به‌طور مجانبی و تقریبی دارای پارامتر مقیاس $\lambda = ۱$ است، پس برای انجام آزمون برای این داده‌ها می‌توان از الگوریتم بخش سوم استفاده نمود. جداول ۳ و ۴ برآورد اندازه و توان این آزمون را نشان می‌دهند. همان‌طور که انتظار داریم توان آزمون برای برخی از فرض‌های مقابل کمتر از بخش سوم است. در عوض آزمون محافظه‌کاری بیشتری دارد. یعنی سطح معناداری آن از سطح اسمی ۰/۰۵ کمتر است.

جدول۳: برآورد اندازه آزمون با استفاده از آماره $A^۲$ که با شبیه‌سازی $B = ۱۰۰۰$ نمونه برای سطح ۰/۰۵ به‌دست آمده‌اند.

n	(α,λ)							
	(۳و۱۰۰)	(۹و۲۰)	(۶و۱۰)	(۳و۵)	(۲و۳)	(۵و۱)	(۴و۰/۵)	(۲و۰/۲)
۵۰	۰/۰۱۵	۰/۰۰۶	۰/۰۰۸	۰/۰۰۱	۰/۰۰۲	۰/۰۱۲	۰/۰۰۶	۰/۰۱۶
۱۰۰	۰/۰۱۲	۰/۰۱	۰/۰۰۳	۰/۰۰۵	۰/۰۱۲	۰/۰۱۴	۰/۰۱۱	۰/۰۲۲

با ملاحظه جدول ۴ متوجه می‌شویم که توان این آزمون برای برخی فرض‌های مقابل مانند *Chi squared* (۲) حتی به اندازه سطح معناداری آزمون یعنی ۰/۰۵ نیز نیست.

جدول۴: برآورد توان آماره $A^{۲}$ برای برخی فرض‌های مقابل و سطح ۰/۰۵ و $B=$۱۰۰۰ نمونه مونت‌کارلو.

فرض مقابل	$n=$۵۰	$n=$۱۰۰	$n=$۱۵۰
$Gamma$(۰/۶,۱)	۰/۵۵۳	۰/۸۳۲	۰/۹۷۴
$Gamma$(۰/۶,۲)	۰/۵۸۳	۰/۸۷۲	۰/۹۳۴
$Gamma$(۳,۲)	۰/۰۸	۰/۴۱۳	۰/۶۲۷
$Gamma$(۵,۱)	۰/۸۴۱	۰/۸۹۳	۰/۹۳۲
$Chi\ squared$ (۲)	۰/۰۰۲	۰/۰۰۴	۰/۰۰۵
$Chi\ squared$ (۱۰)	۰/۸۹۲	۰/۸۹۸	۰/۹۶۳
$Weibull$ (۳,۲)	۰/۹۹۳	۰/۹۸۵	۰/۹۹۹
$Weibull$ (۲,۱)	۰/۳۶۵	۰/۷۵۱	۰/۹۳۴
$Lognormal$ (۰/۰,۰/۶)	۰/۱۵۲	۰/۵۸۵	۰/۶۵۲
$Lognormal$ (۰/۰,۱)	۰/۱۴۴	۰/۲۸۴	۰/۳۸۲
$Lognormal$ (۰/۰,۵)	۰/۸۵	۰/۹۶۵	۰/۹۸۲
$Uniform$ (۰/۰,۱)	۰/۴۳۲	۰/۸۷۲	۰/۹۳۸
$Half-normal$ (۰/۰,۱)	۰/۰۱۵	۰/۱۱۸	۰/۱۲۳
$Half-Cauchy$ (۰/۰,۱)	۰/۸۶۲	۰/۸۹۶	۰/۹۶۳

۵- آزمون نیکویی برازش بر مبنای آماره کلموگروف-اسمیرنف

در این بخش آزمون (۷) که در آن $F(.)$ تابع توزیع $WE(\alpha,\lambda)$ است، بررسی می‌شود. برای این منظور از آماره کلموگروف-اسمیرنف استفاده می‌کنیم. مقدار دقیق آماره عبارتست از $D=\mathbf{max}\left(D^{+},D^{-}\right)$ که در آن

$$D^{-}=\sup_{x}\left(F(x)-F_n(x)\right)\ ,\ D^{+}=\sup_{x}\left(F_n(x)-F(x)\right).$$

تقریب استفنز [۷] برای آماره به‌صورت $D\simeq\mathbf{max}\left(D^{+},D^{-}\right)$ است. چنانچه $D^{+}\simeq\max_{i}\left(\frac{i}{n}-p_{(i)}\right)$ و $D^{-}\simeq\max_{i}\left(p_{(i)}-\frac{i-۱}{n}\right)$، آن‌گاه فرض صفر در سطح معناداری $۱-\gamma$ وقتی رد می‌شود که $D\geq C_n(\gamma)$، که در آن $C_n(\gamma)$ چندک γ-ام تابع

توزیع تجربی D است. درحقیقت چون برای مقادیر بزرگ آماره، فرض صفر را رد می‌کنیم و می‌خواهیم به سطح معناداری $\gamma-۱$ برسیم، پس باید چندک γ-ام را در نظر بگیریم.

برای به‌دست آوردن نقاط بحرانی $C_n(\gamma)$ برای (α,λ) ثابت از الگوریتم زیر استفاده می‌گردد.

۱. (α,λ) و n را ثابت در نظر می‌گیریم.

۲. نمونه‌ای تصادفی به حجم n از $WE(\alpha,\lambda)$ شبیه‌سازی می‌کنیم.

۳. برآورد ماکسیمم درست‌نمایی پارامترهای (α,λ) را به‌دست می‌آوریم.

۴. D_n، با استفاده از تقریب استفنز [۷] محاسبه می‌شود.

۵. مراحل ۱-۴ را B بار تکرار می‌کنیم.

با اتمام فرآیند شبیه سازی، برای (α,λ) داده شده B مقدار D_n داریم. بنابراین مقدار ثابت بحرانی $C_n(\gamma)$ با یافتن چندک ۱۰۰γ-ام از تابع توزیع تجربی D_n محاسبه می‌شود.

شبیه‌سازی نشان می‌دهد که چندک $C_n(\gamma)$ به‌پارامتر مقیاس λ وابسته نیست. جدول ۵ نمونه‌ای از این شبیه‌سازی را برای مقادیر مختلف (α,λ) با $n=۵۰$ و $B=۱۰۰۰۰$ نمونه مونت‌کارلو نشان می‌دهد. توجه کنید که این مساله در مورد آماره اندرسون درست نیست. پس برای راحتی کار در الگوریتم بالا می‌توان از $\lambda=۱$ استفاده نمود. از طرفی چون می‌خواهیم حداکثر اندازه آزمون برابر $۱-\gamma$ شود، بایستی $C_n(\gamma)$ طوری مشخص شود که داشته باشیم:

$$\begin{aligned} ۱-\gamma &= \max_{(\alpha,\lambda=۱)} P(\text{Reject } H_\circ \mid H_\circ) \\ &= \max_{(\alpha,\lambda=۱)} P(D \geq C_n(\gamma)). \end{aligned} \tag{۹}$$

شکل ۲ نمودارهای $C_n(\gamma)$ را به‌عنوان تابعی از α برای $\gamma=\{۰/۹۵,۰/۹۷۵,۰/۹۹\}$ و $n=۵۰$ نشان می‌دهد. این نمودارها نشان می‌دهند که توزیع آماره آزمون D_n تحت $H_\circ$ وابسته به مقدار پارامتر نامعلوم α است. از شکل‌های ۲ و ۹ مقادیر ثابت بحرانی $C_n(\gamma)$ با یافتن چندک‌های ۱۰۰γ از تابع توزیع تجربی D_n به‌وسیله شبیه‌سازی برای $\alpha=۳۰$ به‌دست می‌آیند؛ زیرا با توجه به شکل ۲ چندک‌های $C_n(\gamma)$ برای $\alpha=۳۰$ دارای حداکثر اندازه‌ای برابر $۱-\gamma$ خواهند بود.

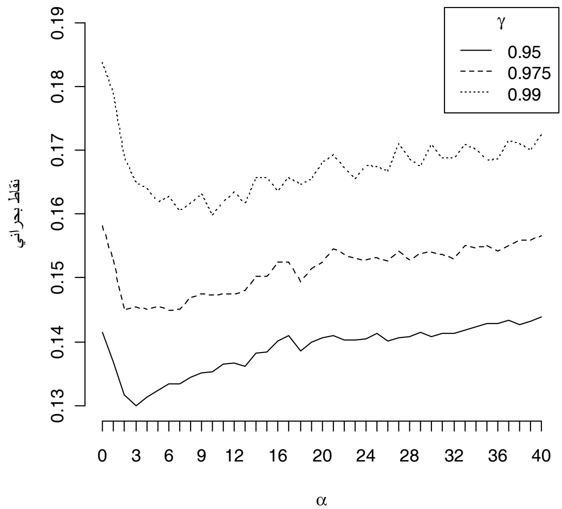
جدول۵: برآورد چندک‌های آماره کلموگروف-اسمیرنف برای مقادیر مختلف (α,λ) با $n=$ ۵۰ و $B=$ ۱۰۰۰۰ نمونه مونت‌کارلو.

(α,λ)	γ						
	۰/۵	۰/۷۵	۰/۸۵	۰/۹	۰/۹۵	۰/۹۷۵	۰/۹۹
(۱,۱)	۰/۰۸۵۶۷۳۶۲	۰/۱۰۱۷۰۵۹۲	۰/۱۱۱۳۸۰۹۵	۰/۱۱۹۱۲۹۸۳	۰/۱۳۳۴۲۰۹۸	۰/۱۴۷۸۶۵۶۹	۰/۱۶۷۰۸۹۰۲
(۱,۳)	۰/۰۸۴۶۸۳۰۷	۰/۱۰۱۷۴۱۰۵	۰/۱۱۱۵۷۱۷۳	۰/۱۱۹۵۵۳۲۴	۰/۱۳۱۴۰۵۳۶	۰/۱۵۱۷۵۵۶۲	۰/۱۶۸۱۷۴۰۱
(۱,۵)	۰/۰۸۶۳۹۵۴۸	۰/۱۰۱۸۹۷۶۳	۰/۱۱۱۶۱۷۹۸	۰/۱۱۸۹۷۳۹۹	۰/۱۳۴۹۶۴۶۶	۰/۱۵۱۳۹۲۳۶	۰/۱۸۲۱۴۳۶۳
(۱,۷)	۰/۰۸۵۲۸۳۸۱	۰/۱۰۱۰۴۷۹۲	۰/۱۱۰۹۸۹۱۴	۰/۱۱۹۹۱۲۸۳	۰/۱۳۲۴۸۵۰۵	۰/۱۴۸۶۰۸۹	۰/۱۶۹۰۵۹۵۱
(۱,۱۰)	۰/۰۸۵۱۳۱۹۷	۰/۱۰۲۷۴۱۱۱	۰/۱۱۲۹۲۱۱۳	۰/۱۱۹۳۷۰۷۳	۰/۱۳۳۹۷۲۶	۰/۱۴۸۷۶۴۰۷	۰/۱۶۴۴۳۲۷
(۱,۱۵)	۰/۰۸۶۷۲۸۰۳	۰/۱۰۴۵۸۳۶۵	۰/۱۱۴۵۰۶۴۹	۰/۱۲۲۳۷۴۳۵	۰/۱۳۶۱۵۸۳۹	۰/۱۵۲۸۶۴۷۷	۰/۱۸۶۷۲۴۳۵
(۱,۲۰)	۰/۰۸۷۰۵۴۲۵	۰/۱۰۲۶۰۲۰۳	۰/۱۱۴۱۵۴۱۷	۰/۱۲۳۵۳۶۶۱	۰/۱۳۸۰۵۲۶۶	۰/۱۶۲۱۷۲۴۷	۰/۱۸۳۱۷۵۱
(۱,۳۰)	۰/۰۸۵۵۵۰۷۴	۰/۱۰۲۲۷۴۴	۰/۱۱۳۴۹۲۷۲	۰/۱۲۰۶۴۵۰۲	۰/۱۳۱۳۰۹۳۶	۰/۱۴۹۳۴۳۱۳	۰/۱۷۴۷۲۳۵۶
(۳,۱)	۰/۰۸۴۱۲۸۷۲	۰/۱۰۰۴۴۳۷۳	۰/۱۱۰۱۵۹۹۵	۰/۱۲۰۰۸۲۴۳	۰/۱۳۱۸۰۴۶۱	۰/۱۴۴۴۰۶۲۲	۰/۱۶۲۷۹۷۱۷
(۳,۳)	۰/۰۸۲۸۳۵۰۴	۰/۰۹۸۲۶۵۳۶	۰/۱۰۹۵۷۳۶۶	۰/۱۱۸۱۸۰۷۱	۰/۱۳۳۲۷۵۵۴	۰/۱۴۴۱۹۰۴۴	۰/۱۶۱۲۸۶۵۸
(۳,۵)	۰/۰۸۳۳۷۷۴۶	۰/۰۹۹۲۶۳۲۸	۰/۱۱۱۱۳۰۱۲	۰/۱۱۹۱۹۰۶	۰/۱۳۳۳۴۹۸۵	۰/۱۴۶۳۶۳۴۹	۰/۱۷۳۴۴۷۷۹
(۳,۷)	۰/۰۸۲۹۵۲۰۶	۰/۰۹۸۰۲۱۶۸	۰/۱۰۹۷۹۹۷۴	۰/۱۱۷۷۳۹۴۹	۰/۱۲۷۷۸۳۵۵	۰/۱۳۸۴۳۱۱۸	۰/۱۵۵۰۴۳۹۹
(۳,۱۰)	۰/۰۸۴۰۵۳۶	۰/۰۹۹۴۴۱۵۲	۰/۱۰۸۷۸۸۹۵	۰/۱۱۶۴۳۴۴۴	۰/۱۲۹۱۴۵۰۵	۰/۱۴۱۹۵۳۶۶	۰/۱۶۵۰۶۷۰۳
(۳,۱۵)	۰/۰۸۳۰۰۳۳۴	۰/۰۹۸۹۶۸۸۷	۰/۱۰۸۹۷۲۰۱	۰/۱۱۸۴۸۲۱۶	۰/۱۳۴۴۴۵۷۷	۰/۱۴۶۷۲۳۱۲	۰/۱۶۵۴۱۲۸۶
(۳,۲۰)	۰/۰۸۳۳۴۳۱۸	۰/۰۹۶۶۲۴۶۳	۰/۱۰۷۳۳۴۶	۰/۱۱۴۲۳۴۶۹	۰/۱۲۸۷۲۷۴۲	۰/۱۳۹۳۰۹۳۲	۰/۱۵۰۶۲۴۴
(۳,۳۰)	۰/۰۸۴۲۰۱۳۹	۰/۱۰۱۱۶۱۵۳	۰/۱۱۲۷۸۸۴	۰/۱۱۹۳۲۱۵۴	۰/۱۳۱۶۴۷۴۴	۰/۱۴۶۹۸۶۲۳	۰/۱۶۱۸۶۹۶۶
(۵,۱)	۰/۰۸۴۹۳۵۸۹	۰/۱۰۰۴۸۴۱۵	۰/۱۱۱۴۳۷۸۲	۰/۱۱۸۵۶۲۰۸	۰/۱۳۱۴۵۳۷۴	۰/۱۴۳۱۱۵۲۸	۰/۱۵۷۷۷۹۰۲
(۵,۳)	۰/۰۸۴۴۹۴۰۴	۰/۱۰۰۳۸۹	۰/۱۱۰۲۶۵۰۳	۰/۱۱۸۶۰۹۷۴	۰/۱۲۹۵۷۳۳۸	۰/۱۴۶۳۲۰۲۹	۰/۱۶۴۳۵۳۵۲
(۵,۵)	۰/۰۸۴۴۶۰۴۵	۰/۱۰۰۲۷۸۲۹	۰/۱۰۸۷۹۲۷۵	۰/۱۱۸۴۷۸۴	۰/۱۳۴۱۳۶۰۹	۰/۱۴۶۳۷۳۴	۰/۱۶۱۱۷۴۸۲
(۵,۷)	۰/۰۸۵۶۹۱۹۳	۰/۱۰۱۲۵۴۱۲	۰/۱۱۱۴۸۵۷۵	۰/۱۱۹۰۴۲۵۴	۰/۱۳۵۶۹۲۹۶	۰/۱۵۰۶۱۴۳۱	۰/۱۶۰۴۸۴۶۴
(۵,۱۰)	۰/۰۸۵۸۹۳۴۵	۰/۱۰۰۸۶۱۵۲	۰/۱۰۹۶۳۰۸۲	۰/۱۱۶۷۹۳۱۳	۰/۱۲۶۴۹۴۵۵	۰/۱۴۰۲۴۲۲	۰/۱۵۹۹۱۷۰۴
(۵,۱۵)	۰/۰۸۴۰۴۹۸۱	۰/۰۹۹۷۷۷۱۹	۰/۱۱۰۴۱۸۳۱	۰/۱۱۷۷۱۰۱۲	۰/۱۲۷۹۴۹۹۵	۰/۱۴۱۷۲۵۳	۰/۱۶۴۸۷۵۲۶
(۵,۲۰)	۰/۰۸۵۳۷۲۷۴	۰/۱۰۱۳۷۷۳۷	۰/۱۱۳۴۳۵۲۵	۰/۱۲۲۰۹۰۶	۰/۱۳۵۲۷۸	۰/۱۴۷۹۰۳۶۵	۰/۱۶۲۸۲۸۵۳
(۵,۳۰)	۰/۰۸۳۴۸۹۵۹	۰/۰۹۹۶۸۱۱۸	۰/۱۰۹۷۵۱۶۶	۰/۱۱۷۷۹۷۸۸	۰/۱۲۹۱۵۳۶۹	۰/۱۳۹۳۶۳۰۶	۰/۱۵۹۹۲۷۵

جدول۵ (ادامه): برآورد چندک‌های آماره کلموگروف-اسمیرنف برای مقادیر مختلف (α,λ) با $n=$ ۵۰ و $B=$ ۱۰۰۰۰ نمونه مونت‌کارلو.

(α,λ)	γ						
(۱۵,۱)	۰/۰۸۸۰۶۷۴۸	۰/۱۰۵۹۴۶۵۴	۰/۱۱۷۸۰۵۴۱	۰/۱۲۵۸۴۶۲۴	۰/۱۳۹۲۷۳۳۳	۰/۱۵۰۳۵۸۱	۰/۱۶۵۴۰۹۶۲
(۱۵,۳)	۰/۰۸۸۴۰۱۳۹	۰/۱۰۵۸۷۲۰۲	۰/۱۱۶۶۸۹۲۷	۰/۱۲۵۱۲۷۹۱	۰/۱۳۸۶۱۷۸۶	۰/۱۵۱۹۳۹۷۲	۰/۱۶۶۰۷۳۴۶
(۱۵,۵)	۰/۰۸۷۸۱۷۲۷	۰/۱۰۵۸۰۸۰۶	۰/۱۱۶۹۵۶۷۶	۰/۱۲۵۰۶۸۸۵	۰/۱۳۷۹۹۶۹۶	۰/۱۵۰۰۰۷	۰/۱۶۴۷۰۲۰۳
(۱۵,۷)	۰/۰۸۸۲۴۵۳	۰/۱۰۵۹۵۶۴	۰/۱۱۷۳۶۲۷	۰/۱۲۵۵۹۷۶	۰/۱۳۸۹۳۹۶	۰/۱۵۰۹۴۶۹	۰/۱۶۵۶۸۰۴
(۱۵,۱۰)	۰/۰۸۸۴۶۱۳۵	۰/۱۰۶۳۴۳۳۶	۰/۱۱۸۱۰۴۸۲	۰/۱۲۶۰۹۳۷۱	۰/۱۳۹۱۰۵۴۵	۰/۱۴۹۷۶۹۵۱	۰/۱۶۴۳۲۰۸۸
(۱۵,۱۵)	۰/۰۸۸۵۸۲۷۴	۰/۱۰۶۲۷۱۱۱	۰/۱۱۷۱۸۸۸۴	۰/۱۲۵۶۸۶۹۲	۰/۱۳۹۰۰۰۸۸	۰/۱۵۰۰۰۳۸۲	۰/۱۶۶۷۰۰۱
(۱۵,۲۰)	۰/۰۸۷۷۱۸۱۲	۰/۱۰۵۷۱۴۹۶	۰/۱۱۶۳۷۰۳	۰/۱۲۴۹۷۷۸۸	۰/۱۳۷۵۹۳۳۲	۰/۱۵۰۵۰۰۲۴	۰/۱۶۵۲۱۱
(۱۵,۳۰)	۰/۰۸۸۲۶۳۱۵	۰/۱۰۵۶۷۱۴۵	۰/۱۱۷۴۱۵	۰/۱۲۵۶۰۸۰۴	۰/۱۳۸۳۳۰۶۹	۰/۱۵۰۵۴۱۳۳	۰/۱۶۵۴۳۷۶
(۲۰,۱)	۰/۰۸۸۸۳۲۴۶	۰/۱۰۷۰۴۴۵۶	۰/۱۱۷۸۹۲۰۹	۰/۱۲۶۱۴۵۸۲	۰/۱۳۸۹۹۶۵	۰/۱۴۸۷۶۹۴۹	۰/۱۶۲۹۸۵۶
(۲۰,۳)	۰/۰۸۹۰۰۸۶۵	۰/۱۰۷۱۷۱۸۳	۰/۱۱۸۲۳۴۱۱	۰/۱۲۵۴۸۰۵۸	۰/۱۳۸۲۰۳۸	۰/۱۴۸۸۱۹۶۱	۰/۱۶۳۸۸۳۸۷
(۲۰,۵)	۰/۰۸۹۱۹۵۰۶	۰/۱۰۷۶۶۷۵۶	۰/۱۱۹۰۵۷۱۶	۰/۱۲۶۹۷۴۸۷	۰/۱۳۹۸۸۹۰۱	۰/۱۵۲۰۸۶۸	۰/۱۶۴۵۲۷۱۹
(۲۰,۷)	۰/۰۸۸۶۲۹۴۷	۰/۱۰۶۹۰۰۸۱	۰/۱۱۸۳۶۲۵۶	۰/۱۲۶۳۵۱۰۶	۰/۱۳۹۱۳۶۷۶	۰/۱۵۱۸۵۶۸۷	۰/۱۶۷۸۷۵۴۸
(۲۰,۱۰)	۰/۰۸۹۳۲۵۱۵	۰/۱۰۶۹۶۳۶۱	۰/۱۱۸۱۶۷۴	۰/۱۲۶۶۶۸۷۵	۰/۱۳۸۸۰۵۹۷	۰/۱۵۱۵۰۴۸۷	۰/۱۶۳۷۸۰۵۲
(۲۰,۱۵)	۰/۰۸۹۰۱۱۰۵	۰/۱۰۷۲۵۳۶	۰/۱۱۸۳۲۸۷۱	۰/۱۲۶۳۷۸۲۴	۰/۱۳۹۵۳۲۸۴	۰/۱۵۲۶۰۲۷	۰/۱۶۴۶۲۶
(۲۰,۲۰)	۰/۰۸۹۶۷۰۱۶	۰/۱۰۷۷۰۱۲۶	۰/۱۱۸۶۵۳۷۲	۰/۱۲۶۳۳۷۶	۰/۱۴۰۴۲۰۵۳	۰/۱۵۲۰۳۵۲۲	۰/۱۶۶۸۳۷۹
(۲۰,۳۰)	۰/۰۸۹۵۲۶۱۱	۰/۱۰۷۵۰۲۷	۰/۱۱۸۹۸۰۹۱	۰/۱۲۷۱۹۲۸۴	۰/۱۳۹۵۷۶۹۱	۰/۱۵۱۱۹۵۳۱	۰/۱۶۸۷۹۶۴۷
(۳۰,۱)	۰/۰۹۰۷۳۷۲۳	۰/۱۰۹۲۹۵۱۸	۰/۱۲۰۷۲۳۶۸	۰/۱۲۹۵۰۰۴۲	۰/۱۴۲۶۴۴۷۷	۰/۱۵۴۳۹۱۳	۰/۱۶۹۵۸۴۷۷
(۳۰,۳)	۰/۰۹۰۹۴۳۳۵	۰/۱۰۸۹۳۸۴۱	۰/۱۲۰۱۲۵۵۴	۰/۱۲۹۲۷۰۳۴	۰/۱۴۲۷۳۶۳۸	۰/۱۵۴۸۲۴۸۳	۰/۱۶۹۴۱۶۲۵
(۳۰,۵)	۰/۰۹۰۳۳۱۰۲	۰/۱۰۸۹۶۱۵۵	۰/۱۲۰۴۹۶۶	۰/۱۲۹۴۸۶۲۸	۰/۱۴۲۳۴۶۰۴	۰/۱۵۴۵۱۹۹۶	۰/۱۷۰۱۱۱۸۶
(۳۰,۷)	۰/۰۹۰۴۸۵۱۹	۰/۱۰۸۶۷۶۵۵	۰/۱۲۰۲۰۵۶۲	۰/۱۲۹۰۱۹۵۶	۰/۱۴۲۴۱۹۳۹	۰/۱۵۴۶۴۵۱	۰/۱۶۹۱۴۹۲۵
(۳۰,۱۰)	۰/۰۹۰۱۶۰۲	۰/۱۰۸۶۶۲۳	۰/۱۱۹۹۸۸۹	۰/۱۲۸۳۵۶۲	۰/۱۴۱۱۳۷	۰/۱۵۲۸۰۶۶	۰/۱۶۷۵۱۰۵
(۳۰,۱۵)	۰/۰۹۰۴۸۹۲۵	۰/۱۰۸۵۲۲۶۸	۰/۱۲۰۷۱۱۷۱	۰/۱۲۹۶۰۸۷	۰/۱۴۲۷۹۶۸۴	۰/۱۵۵۱۸۱۰۱	۰/۱۶۹۹۲۲۲۷
(۳۰,۲۰)	۰/۰۹۰۲۷۰۵۳	۰/۱۰۸۴۲۶۷۹	۰/۱۲۰۶۸۰۶۷	۰/۱۲۸۷۳۴۳۶	۰/۱۴۱۱۳۲۲	۰/۱۵۳۸۶۰۹۱	۰/۱۷۰۶۷۲۱۵
(۳۰,۳۰)	۰/۰۹۰۳۷۳۱۹	۰/۱۰۸۷۲۶۹۱	۰/۱۲۰۰۹۲۲۳	۰/۱۲۸۰۸۱۰۱	۰/۱۴۲۰۵۹۶۴	۰/۱۵۴۸۰۶۱۸	۰/۱۶۹۷۷۱۴۴

جدول ۶، نقاط بحرانی آزمون D_n را که با شبیه‌سازی $B=$ ۱۰۰۰۰ نمونه مونت‌کارلو و مراحل ۱-۵ برای $(\alpha=۳۰,\lambda=۱)$ به‌دست آمده‌اند، نشان می‌دهد.



شکل۲: نقاط بحرانی به‌عنوان تابعی از α با $n=$ ۵۰ و $B=$ ۱۰۰۰۰ نمونه برای آماره D_n.

جدول۶: نقاط بحرانی برای آزمون D_n که با استفاده از شبیه‌سازی مونت‌کارلو با $B=$ ۱۰۰۰۰ نمونه به حجم n به‌دست آمده اند.

n	γ						
	۰/۵	۰/۷۵	۰/۸۵	۰/۹	۰/۹۵	۰/۹۷۵	۰/۹۹
۱۰	۰/۱۹۱۱۵۶۳	۰/۲۲۹۹۵۱۵	۰/۲۵۴۴۲۷۴	۰/۲۷۲۱۲۲	۰/۲۹۹۶۱۵۲	۰/۳۲۵۶۳۴۱	۰/۳۵۸۱۷۷۲
۲۰	۰/۱۳۷۷۳۱۲	۰/۱۶۶۲۰۹۵	۰/۱۸۵۴۱۰۷	۰/۱۹۷۸۸۶	۰/۲۱۹۷۳۶۱	۰/۲۳۸۶۶۲	۰/۲۶۴۸۷۰۴
۳۰	۰/۱۱۴۸۶۵۳	۰/۱۳۸۰۸۶۲	۰/۱۵۲۹۶۶۷	۰/۱۶۴۸۴۱	۰/۱۸۰۹۵۵	۰/۱۹۵۰۵۲۳	۰/۲۱۷۳۳۵۵
۵۰	۰/۰۹۰۶۸۱۹۹	۰/۱۰۸۸۳۱۷	۰/۱۲۰۴۲۶۹۲	۰/۱۲۹۰۷۴	۰/۱۴۲۳۵۱۱۶	۰/۱۵۴۵۶۶۷	۰/۱۶۸۸۵۰۸
۱۰۰	۰/۰۶۵۲۷۵۵۹	۰/۰۷۸۱۰۸۹	۰/۰۸۶۶۲۵۶۵	۰/۰۹۲۷۲۸	۰/۱۰۱۴۸۸۷۳	۰/۱۰۹۷۰۹۵	۰/۱۱۹۱۱۶۷
۱۵۰	۰/۰۵۳۶۷۸۴۲	۰/۰۶۴۴۸۴۷	۰/۰۷۱۱۶۷۹۶	۰/۰۷۵۹۲۶	۰/۰۸۳۲۰۵۵۴	۰/۰۹۱۵۷۷۵	۰/۰۹۹۳۳۸
۲۰۰	۰/۰۴۶۹۵۵۴۶	۰/۰۵۶۵۵۸۲	۰/۰۶۲۳۴۷۷۸	۰/۰۶۶۶۴۹۷	۰/۰۷۳۵۳۱۲۳	۰/۰۷۹۶۳۵۰	۰/۰۸۷۹۰۴۷
۳۰۰	۰/۰۳۸۸۴۵۷۲	۰/۰۴۶۴۱۸۳	۰/۰۵۱۰۳۱۵۴	۰/۰۵۴۶۸۸	۰/۰۶۰۴۴۲۸۲	۰/۰۶۵۸۷۰۴	۰/۰۷۲۳۲۱۹
۵۰۰	۰/۰۳۰۰۶۱۴۹	۰/۰۳۶۰۰۷۴	۰/۰۳۹۸۰۳۹۷	۰/۰۴۲۴۵۶	۰/۰۴۶۷۰۸۰۱	۰/۰۵۰۸۰۲۶	۰/۰۵۵۶۷۰۲

در جداول ۷ و ۸ به‌ترتیب به بررسی سطح معناداری و توان آزمون کلموگروف-اسمیرنف با $B=$ ۱۰۰۰ نمونه مونت‌کارلو پرداخته‌ایم. جدول ۷ نشان می‌دهد که سطح معناداری آزمون نزدیک سطح اسمی آن است. مقایسه جدول ۸ با جدول ۴ نشان می‌دهد که این آزمون نسبت به آزمون برمبنای آماره اندرسون به مراتب تواناتر است.

جدول۷: برآورد اندازه آزمون با استفاده از آماره D_n که با شبیه‌سازی $B=$ ۱۰۰۰ نمونه برای $1-\gamma=$ ۰/۰۵ به‌دست آمده است.

n	(α,λ)							
	(۳و۱۰۰)	(۹و۲۰)	(۶و۱۰)	(۳و۵)	(۲و۳)	(۵و۱)	(۴و۰/۵)	(۲و۰/۲)
۵۰	۰/۰۵۷	۰/۰۵۴	۰/۰۳۸	۰/۰۳۲	۰/۰۲۷	۰/۰۵	۰/۰۴۱	۰/۰۵۷
۱۰۰	۰/۰۵۸	۰/۰۴۳	۰/۰۴۵	۰/۰۲۴	۰/۰۲۴	۰/۰۲۶	۰/۰۳۳	۰/۰۴۲

جدول۸: برآورد توان آماره D_n برای برخی فرض‌های مقابل، $1-\gamma=$ ۰/۰۵ و $B=$ ۱۰۰۰ شبیه‌سازی مونت‌کارلو.

فرض مقابل	$n=$ ۵۰	$n=$ ۱۰۰	$n=$ ۱۵۰
$Gamma(۰/۶,۱)$	۰/۶۸۳	۰/۹۱۶	۰/۹۸۲
$Gamma(۰/۶,۲)$	۰/۶۲۱	۰/۹۲۱	۰/۹۸۷
$Gamma(۳,۲)$	۰/۳۳۲	۰/۶۱۳	۰/۷۹۹
$Gamma(۵,۱)$	۰/۹۶۵	۰/۹۹۶	۱
$Chi\ squared\ (۲)$	۰/۰۷۶	۰/۰۷۸	۰/۰۷۵
$Chi\ squared\ (۱۰)$	۰/۹۴۱	۰/۹۸۶	۰/۹۹۵
$Weibull\ (۳,۲)$	۰/۹۹۸	۰/۹۹۹	۱
$Weibull\ (۲,۱)$	۰/۵۳۱	۰/۸۴۶	۰/۹۴۴
$Lognormal\ (۰/۰,۰/۶)$	۰/۳۳۴	۰/۶۶۱	۰/۸۷۷
$Lognormal\ (۰/۰,۱)$	۰/۲۹۵	۰/۵۰۹	۰/۶۲۸
$Lognormal\ (۰/۰,۵)$	۰/۹۹۲	۰/۹۹۷	۱
$Uniform\ (۰/۰,۱)$	۰/۷۰۶	۰/۹۴۷	۰/۹۹۳
$Half-normal\ (۰/۰,۱)$	۰/۲۴۹	۰/۵۰۵	۰/۶۷۳
$Half-Cauchy\ (۰/۰,۱)$	۰/۹۳	۰/۹۹۵	۰/۹۹۹

۶- نتیجه‌گیری

در این مقاله آزمون نیکویی برازش برای رده نمایی وزنی براساس آماره‌های اندرسون و کلموگروف-اسمیرنف مورد بررسی قرار گرفت. برای محاسبه چندک‌های آماره آزمون اندرسون از روش بوت‌استرپ و جهت آزمون کلموگروف-اسمیرنف از یک روش دیگر استفاده شد. در ادامه توان این آزمون‌ها برای فرض‌های مقابل مناسب بررسی شد. نتایج شبیه‌سازی نشان داد که چندک‌های آماره آزمون کلموگروف-اسمیرنف با تقریب بسیار خوبی به پارامتر مقیاس وابسته نیست. هم‌چنین این آزمون به مراتب از آزمون اندرسون تواناتر است.

مراجع

[1] Gupta, R. D. and Kundu, D. (2009), A new class of weighted exponential distributions, *Statistics*, **43**, 621 - 643.

[2] Azzalini, A. (1985), A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

[3] Gupta, R.D. and Kundu, D. (1999), Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, **41(2)**, 173-188.

[4] Arnold, B.C. and Beaver, R.J. (2000), The skew-Cauchy distribution, *Statistics and probability letters*, **49**, 285–290.

[5] Gupta, R.D. and Kundu, D. (2007), Skew-Logistic distribution, Tech. Rep., *Indian Institute of Technology Kanpur, Kanpur, India*.

[6] Genton, M.G. (2004), *Skew-elliptical distributions and their applications: a journey beyond normality*, Chapman and Hall/CRC.

[7] Stephens, M.A. (1986), Tests based on EDF statistics, *Goodness-of-fit Techniques*, **68**, 97-193.

Goodness-of-fit tests for the weighted exponential distribution

Mohammad Mehdi Maghami, Nasrollah Iranpanah

Department of Statistics, University of Isfahan, Isfahan, Iran

Abstract

In the new class of weighted exponential distributions was presented by Gupta and Kundu [1], the skewness parameter has been added to the exponential distribution. Therefore the weighted exponential distribution has the skewness and scale parameters. In this paper, we first study Anderson and Kolmogorov-Smirnov goodness of fit tests for this class with unknown parameters. Then, we apply bootstrap method for estimation of Anderson's quantile and another method for Kolmogorov-Smirnov. We use the maximum likelihood method for estimation of parameters. Finally, we compare Kolmogorov-Smirnov and Anderson tests in a Monte Carlo simulation study. The results show that the Kolmogorov–Smirnov test has greater power than Anderson test.

Keywords: Goodness of fit test, Parametric bootstrap, Empirical distribution function, Monte Carlo simulation.

Mathematics Subject Classification (2000): 62G10, 62F40