مجله علوم آماری، پاییز و زمستان ۱۳۸۷
جلد ۲، شماره ۲، ص ۱۷۹–۲۰۰

آزمون نیکویی برازش توزیع‌های نرمال و نمایی بر مبنای برآوردگرهای جدید آنتروپی

احسان زمان‌زاده، ناصررضا ارقامی
گروه آمار، دانشگاه فردوسی مشهد

تاریخ دریافت: ۱۳۸۷/۸/۱۸ تاریخ آخرین بازنگری: ۱۳۸۷/۱۲/۱۸

چکیده : در این مقاله، ابتدا به معرفی دوبرآوردگر جدید آنتروپی می‌پردازیم. برآوردگرهای جدید برمبنای تصحیح برآوردگر کوریا (۱۹۹۵) در نقاط ابتدایی و انتهایی و اعمال وزن‌های متفاوت نسبت به آن برآوردگر معرفی می‌شوند. سپس به مقایسه برآوردگرهای جدید آنتروپی با برآوردگرهای آنتروپی معرفی شده توسط واسیچک (۱۹۷۶) و ابراهیمی و همکاران (۱۹۹۴) و کوریا (۱۹۹۵) می‌پردازیم. آنگاه آزمون نیکویی برازش فرضیه‌های نرمال بودن و نمایی بودن را بر مبنای برآوردگرهای جدید آنتروپی معرفی کرده و توان آن را با آزمون‌های مبتنی بر برآوردگرهای واسیچک (۱۹۷۶) و کوریا (۱۹۹۵) و آزمون شاپیرو–ویلک (۱۹۶۵) برای آزمون‌های نرمال بودن مقایسه می‌کنیم. نتایج مطالعات شبیه‌سازی نشان می‌دهد که برآوردگرهای پیشنهادی عملکرد نسبتاً خوبی نسبت به سایر برآوردگرها در برآورد آنتروپی و آزمون نیکویی برازش دارند.

واژه‌های کلیدی : آ نتروپی، برآورد آنتروپی، آزمون نیکویی برازش، اطلاع کولبک لایبلر

آدرس الکترونیک مسئول مقاله: احسان زمان‌زاده، *ehsanzamanzadeh@yahoo.com* . ∘
کد موضوع‌بندی ریاضی (۲۰۰۰): ۶۲G۱۰ و ۶۲G۳۰ . ∘

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

۱ مقدمه

فرض کنید متغیر تصادفی X دارای تابع توزیع $F(X)$ با تابع چگالی مطلقاً پیوسته $f(x)$ باشد. شانون (۱۹۴۸) آنتروپی این متغیر تصادفی را به صورت

$$H(f) = -\int_{-\infty}^{\infty} f(x)\log(f(x))dx, \qquad (۱)$$

تعریف کرد. از آنجا که مفهوم آنتروپی کاربردهای فراونی در مباحث آماری نظیر نظریه اطلاع و آزمون‌های نیکویی برازش دارد، مساله برآورد $H(f)$ بر مبنای مشاهدات $x_۱, \cdots, x_n$ توسط محققین زیادی از جمله احمد و لین (۱۹۷۶)، واسیچک (۱۹۷۶)، مک (۱۹۸۸)، دادویچ و ون‌درمولن (۱۹۸۷)، ابراهیمی و همکاران (۱۹۹۴) و کوریا (۱۹۹۵) مورد بررسی قرار گرفته است. از میان تمام این برآوردگرها، برآوردگر واسیچک (۱۹۷۶) به واسطه سادگی در محاسبات و دقت بالای آن، بسیار مورد استفاده قرار می‌گیرد. واسیچک (۱۹۷۶) نشان داد که می‌توان رابطه (۱) را با تغییر متغیر $F(x) = p$ به صورت

$$H(f) = \int_{\circ}^{۱} \log(\frac{d}{dp}F^{-۱}(p))dp \qquad (۲)$$

نوشت. آنگاه وی با جایگزینی تابع توزیع F با تابع توزیع تجربی F_n و استفاده از عملگر تفاضل به‌جای عملگر مشتق، برآورد آنتروپی خود را به صورت زیر معرفی کرد.

فرض کنید $x_۱, \cdots, x_n$ یک نمونه تصادفی از توزیع F باشند، دراین‌صورت برآورد آنتروپی واسیچک (۱۹۷۶) عبارت است از

$$HV_{mn} = \frac{۱}{n}\sum_{i=۱}^{n} \log(\frac{n}{c_i}(x_{(i+m)} - x_{(i-m)})), \qquad (۳)$$

که در آن $x_{(۱)}, \cdots, x_{(n)}$ آماره‌های مرتب، $c_i = ۲m$ و m یک عدد مثبت کوچکتر یا مساوی $\frac{n}{۲}$ است. ضمناً به‌ازای $i < ۱$ ، $x_{(i)} = x_{(۱)}$ و به‌ازای $i > n$ ، $x_{(i)} = x_{(n)}$

ابراهیمی و همکاران (۱۹۹۴) برآوردگر خود را برمبنای اصلاح ضرایب برآوردگر واسیچک (۱۹۷۶) در نقاط انتهایی (c_iها وقتی که $i \leq m$ و $i \geq n-m+۱$) به

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

صورت

$$HE_{mn} = \frac{1}{n}\sum_{i=1}^{n}\log(\frac{n}{c_i}(x_{(i+m)} - x_{(i-m)})) \qquad (۴)$$

معرفی کردند، که در آن

$$c_i = \begin{cases} m+i-1 & 1 \le i \le m \\ 2m & m+1 \le i \le n-m \\ m+n-i & n-m+1 \le i \le n \end{cases}$$

و برای $i < 1$، $x_{(i)} = x_{(1)}$ و به ازای $i > n$، $x_{(i)} = x_{(n)}$ است. مطالعات شبیه‌سازی ابراهیمی و همکاران (۱۹۹۴) نشان می‌داد که برآوردگر پیشنهادی آن‌ها اریبی و میانگین توان‌دوم خطای کمتری نسبت به برآوردگر واسیچک (۱۹۷۶) دارد. کوریا (۱۹۹۵) برآوردگر دیگری برای آنتروپی پیشنهاد داد که میانگین توان دوم خطای کمتری نسبت به برآوردگر معرفی‌شده توسط واسیچک (۱۹۷۶) داشت. وی توجه کرد که می‌توان رابطه (۳) را به صورت

$$HV_{mn} = \frac{-1}{n}\sum_{i=1}^{n}\log(\frac{\frac{i+m}{n} - \frac{i-m}{n}}{x_{(i+m)} - x_{(i-m)}}) \qquad (۵)$$

بازنویسی کرد. اما عبارت داخل لگاریتم در رابطه (۵) در واقع شیب خطی است که نقاط $(\hat{F}(X_{(i+m)}), X_{(i+m)})$ و $(\hat{F}(X_{(i-m)}), X_{(i-m)})$ را به یکدیگر متصل می‌کند. او پیشنهاد داد که این شیب را با استفاده از رگرسیون خطی موضعی برحسب $\{X_{(i-m)}, \cdots, X_{(i+m)}\}$ و استفاده از تمام $2m+1$ نقطه بجای دو نقطه برآورد کنیم. لذا با در نظر گرفتن رابطه

$$F(x_{(j)}) = \alpha + \beta x_j + \epsilon \qquad (۶)$$

و برآورد β با روش کمترین توان‌های دوم، برآوردگر خود را به صورت

$$HC_{mn} = \frac{-1}{n}\sum_{i=1}^{n}\log(b_i), \qquad (۷)$$

پیشنهاد کرد، که در آن

$$b_i = \frac{\sum_{j=i-m}^{i+m}(x_{(j)} - \bar{x}_{(i)})(\frac{j}{n} - \frac{i}{n})}{\sum_{i-m}^{i+m}(x_{(j)} - \bar{x}_{(i)})^2}$$

و

$$\bar{x}_{(i)} = \sum_{j=i-m}^{i+m} \frac{x_{(j)}}{۲m+۱}$$

و m یک عدد صحیح کوچکتریا مساوی $\frac{n}{۲}$، $x_{(۱)},\cdots,x_{(n)}$ آماره‌های مرتب و برای $j<۱$، $x_{(j)}=x_{(۱)}$ و به‌ازای $j>n$، $x_{(j)}=x_{(n)}$ می‌باشد.

برآورد آنتروپی در به دست آوردن آماره آزمون نیکویی برازش برای توزیع نرمال ابتدا توسط واسیچک (۱۹۷۶) و سپس توسط آریزونو و اوتا (۱۹۸۹) و علیزاده نوقابی و علیزاده نوقابی (۱۳۸۷)، برای توزیع یکنواخت بوسیله دادویچ و وندرمولن (۱۹۸۱) و برای توزیع نمایی بوسیله ابراهیمی و حبیب‌الله (۱۹۹۲) مورد استفاده قرار گرفت. همچنین برآورد آنتروپی در آزمون نمایی بودن توزیع برای داده‌های سانسورشده نوع–دو (پارک، ۲۰۰۵) و داده‌های سانسورفزاینده نوع–دو (بالاکریشنان و همکاران (۲۰۰۷) و یوسف‌زاده و ارقامی (۲۰۰۸) و تقارن توزیع (حبیبی‌راد و ارقامی، ۱۳۸۶) نیز مورد استفاده قرار گرفته است. نتایج شبیه‌سازی در مقالات فوق نشان می‌دهد که آزمون‌های نیکویی برازش برمبنای آنتروپی عموماً از توان بالاتری در مقایسه با سایر آزمون‌های نیکویی برخوردارند. به عنوان مثال، واسیچک (۱۹۷۶) به مقایسه آزمون نیکویی برازش نرمال‌بودن توزیع جامعه بر مبنای آنتروپی با سایر آزمون‌ها(کلموگروف–اسمیرنف[۱]، کرامر–ون میسز[۲]، کوپر[۳]، واتسون[۴]، اندرسون–دارلینگ[۵] و شاپیرو–ویلک[۶]) پرداخت و نشان داد که آزمون پیشنهادی وی عملکرد خوبی نسبت به این آزمون‌هادارد. از این‌رو آزمون‌های پیشنهاد شده در این مقاله نه با آزمون‌های فوق‌الذکر بلکه فقط با آزمون واسیچک (۱۹۷۶) مقایسه شده‌است. معذالک باتوجه به اینکه اجماع عمومی بر این است که در بین آزمون‌های فوق‌الذکر آزمون شاپیرو–ویلک (۱۹۶۵) بهتر از بقیه است. توان‌های مربوط به آزمون اخیرالذکر به نتایج شبیه‌سازی اضافه گردیده است.

---
۱ Kolmogorov-Smirnov
۲ Cramer-Von Mises
۳ Kuiper
۴ Watson
۵ Anderson-Darling
۶ Shapiro-Wilk

در بخش دوم این مقاله به معرفی برآوردگرهای جدید آنتروپی که در واقع اصلاح برآوردگر معرفی شده توسط کوریا (۱۹۹۵) است، می‌پردازیم سپس با استفاده از شبیه‌سازی مونت-کارلو به مقایسه برآورگرهای پیشنهادی با برآوردگرهای پیشنهاد شده توسط واسیچک (۱۹۷۶)، ابراهیمی و همکاران (۱۹۹۴) و کوریا (۱۹۹۵) خواهیم پرداخت. در بخش ۳، آزمون نیکویی برازش را برای فرضیه‌های نمایی‌بودن و نرمال‌بودن توزیع جامعه بر اساس برآوردگرهای جدید آنتروپی معرفی خواهیم کرد و به مقایسه توان این آزمون‌ها با آزمون‌های پیشنهادی توسط واسیچک(۱۹۷۶) و کوریا (۱۹۹۵) و آزمون شاپیرو-ویلک (۱۹۶۵) درمورد نرمال خواهیم پرداخت. نتایج شبیه‌سازی حکایت از رضایت‌بخش بودن عملکرد برآوردگرهای جدید آنتروپی به منظور برآورد و آزمون نیکویی برازش دارد. بحث و نتیجه‌گیری نهایی در بخش ۴ آورده شده است.

۲ برآوردگرهای جدید آنتروپی

واضح است که مقدار b_i در رابطه (۷) (وقتی $i \leq m$ یا $i \geq n-m+1$ است) تقریب مناسبی برای شیب خط نمی‌باشد، زیرا در این نقاط بیش از یک بار از مقدار $x_{(1)}$ یا $x_{(n)}$ استفاده شده است. لذا برآورد اول آنتروپی خود را بر مبنای اصلاح این ضرایب به صورت

$$HZ1_{mn} = \frac{-1}{n}\sum_{i=1}^{n}\log(b_i^*) \qquad (8)$$

ارائه می‌دهیم، که در آن

$$b_i^* = \frac{\sum_{j=k_1(i)}^{k_2(i)}(x_{(j)} - \tilde{x}_{(i)})(\hat{F}(j) - \tilde{\hat{F}}(i))}{\sum_{j=k_1(i)}^{k_2(i)}(x_j - \tilde{x}_{(i)})^2}, i = 1, \cdots, n,$$

$$k_1(i) = \begin{cases} 1 & 1 \leq i \leq m \\ i-m & i \geq m+1 \end{cases}, \quad k_2(i) = \begin{cases} i+m & 1 \leq i \leq n-m \\ n & i \geq n-m+1 \end{cases},$$

$$\tilde{x}_{(i)} = \sum_{j=k_1(i)}^{k_2(i)} \frac{x_{(j)}}{k_2(i) - k_1(i) + 1}, \quad \tilde{\hat{F}}(i) = \sum_{j=k_1(i)}^{k_2(i)} \frac{\hat{F}_{(j)}}{k_2(i) - k_1(i) + 1},$$

$\hat{F}$ تابع توزیع تجربی و $x_{(1)},\cdots,x_{(n)}$ آماره‌های مرتب هستند. به سادگی می‌توان نشان داد رابطه

$$b_i^* = \begin{cases} \frac{\sum_{j=1}^{i+m}(x_{(j)}-\tilde{x}_{(i)})(\frac{j}{n}-\frac{m+i+1}{2n})}{\sum_{j=1}^{i+m}(x_{(j)}-\tilde{x}_{(i)})^2} & 1 \le i \le m \\ b_i & m+1 \le i \le n-m \\ \frac{\sum_{j=i-m}^{n}(x_{(j)}-\tilde{x}_{(i)})(\frac{j}{n}-\frac{n+i-m}{2n})}{\sum_{j=i-m}^{n}(x_{(j)}-\tilde{x}_{(i)})^2} & n-m+1 \le i \le n \end{cases}$$

برقرار است. دومین برآوردگر ارایه شده در این مقاله بر مبنای این ایده است که چون در برآوردگر کوریا (۱۹۹۵) برای بدست آوردن همهٔ b_iها از تعداد مساوی مشاهده استفاده شده است، در برآورد آنتروپی، همه b_iها از وزن‌های مساوی $w_i = \frac{1}{n}$ برخورداراند. اما در محاسبه $HZ1$ وقتی $1 \le i \le m$ یا $n-m+1 \le i \le n$ از مشاهدات کمتری برای بدست آوردن b_i^* استفاده شده است. لذا منطقی به نظر می‌رسد که وزن‌های کمتری به این مقادیر در برآورد آنتروپی اختصاص دهیم. بر این اساس، برآوردگر دوم آنتروپی خود را به صورت

$$HZ2_{mn} = -\sum_{i=1}^{n} w_i \log(b_i^*) \tag{۹}$$

معرفی می‌کنیم، که درآن

$$w_i = \frac{\hat{F}(x_{(i+m)}) - \hat{F}(x_{(i-m)})}{\sum_{i=1}^{n}\hat{F}(x_{(i+m)}) - \hat{F}(x_{(i-m)})}, i = 1,\cdots,n$$

و $\hat{F}$ تابع توزیع تجربی و $x_{(1)},\cdots,x_{(n)}$ آماره‌های مرتب هستند.

توجیه انتخاب وزن‌های فوق به این صورت است که هنگام تقریب یک انتگرال $\int_0^1 \log(\frac{d}{dp}F^{-1}(p))dp \approx \sum_{i=1}^{n}\log(\frac{d}{dp}F^{-1}(p_i))\Delta p_i$ به یک مجموع از رابطه استفاده می‌کنیم، لذا در رابطه (۸) به جای ضرایب مساوی $\frac{1}{n}$ از $\Delta p_i = F(x_{i+m}) - F(x_{i-m})$ استفاده می‌کنیم، که پس از تقسیم هریک بر مجموع آن‌ها همان مقادیر w_i در بالا حاصل می‌گردد. به سادگی می‌توان نشان داد

$$w_i = \begin{cases} \frac{i+m-1}{m(2n-m-1)} & 1 \le i \le m \\ \frac{2}{2n-m-1} & m+1 \le i \le n-m \\ \frac{n-i+m}{m(2n-m-1)} & n-m+1 \le i \le n \end{cases}$$

قضیه زیر را که بیان می‌کند "مقیاس متغیر تصادفی X تاثیری در دقت برآوردگر آنتروپی $HZ1_{mn}$ و $HZ2_{mn}$ ندارد" می‌توان به سادگی مشابه ابراهیمی و همکاران (۱۹۹۴) ثابت کرد.

قضیه ۱ : فرض کنید $X_1,\cdots,X_n$ یک دنباله از متغیرهای تصادفی مستقل و هم‌توزیع با آنتروپی $H^X(f)$ وبه ازای $i=1,\cdots,k, k>0$ قرار دهیم $Y_i=kX_i$. همچنین فرض کنید $HZ1^X_{mn}$ ، $HZ2^X_{mn}$ و $HZ1^Y_{mn}$، $HZ2^Y_{mn}$ به ترتیب برآوردگرهای آنتروپی $H^X(f)$ و $H^Y(g)$ باشند، که در آن g تابع چگالی احتمال $Y=kX$ است. در این صورت روابط زیر برقرار است.

الف)

$$E(HZ1^Y_{mn})=E(HZ1^Y_{mn})+\log(k) \quad \text{و} \quad E(HZ2^Y_{mn})=E(HZ2^Y_{mn})+\log(k)$$

ب)

$$Var(HZ1^Y_{mn})=MSE(HZ1^Y_{mn}) \quad \text{و} \quad Var(HZ2^Y_{mn})=Var(HZ2^Y_{mn})$$

ج)

$$MSE(HZ1^Y_{mn})=MSE(HZ1^Y_{mn}) \quad \text{و} \quad MSE(HZ2^Y_{mn})=MSE(HZ2^Y_{mn})$$

اکنون در یک مطالعه شبیه‌سازی عملکرد برآوردگرهای پیشنهادی و برآوردگرهای معرفی شده توسط واسیچک (۱۹۷۶)، ابراهیمی و همکاران (۱۹۹۴) و کوریا (۱۹۹۵)، بر حسب جذرمیانگین تواندوم خطا و اریبی، مورد مقایسه قرار می‌گیرد. برای این منظور به ازای $n=10,20,30$ و مقادیر مختلف m، تعداد $10,000$ نمونه با اندازه n تولید و مقدار اریبی و جذر میانگین توان دوم خطای آن‌ها برای سه توزیع نرمال استاندارد، نمایی با میانگین ۱ و یکنواخت $(0,1)$ محاسبه شده است و نتایج در جداول ۱ تا ۳ ارائه شده‌اند.

همان طور که در جدول‌های ۱ تا ۳ ملاحظه می‌شود، برآوردگر پیشنهادی اول ($HZ1$) عموماً اریبی کمتری نسبت به برآوردگر پیشنهادی دوم ($HZ2$) دارد، درحالی که $RMSE$های آن‌ها کم و بیش برابر هستند. ضمناً ملاحظه می‌شود که $HZ1$ چه از نظر اریبی و چه از نظر $RMSE$ به ازای mهای کوچک ($m\leq\frac{n}{4}$)

جدول ۱: اریبی و جذرمیانگین توان‌دوم خطای برآوردگرهای آنتروپی توزیع نرمال استاندارد.

		اریبی					جذرمیانگین توان دوم خطا				
n	m	HV_{mn}	HE_{mn}	HC_{mn}	$HZ1_{mn}$	$HZ2_{mn}$	HV_{mn}	HE_{mn}	HC_{mn}	$HZ1_{mn}$	$HZ2_{mn}$
۱۰	۱	−۰/۵۹۶۷	−۰/۴۵۸۰	−۰/۴۵۸۳	−۰/۳۷۷۲	−۰/۳۹۴۳	۰/۶۷۶۴	۰/۵۵۸۰	۰/۵۵۶۷	۰/۴۹۲۱	۰/۴۹۱۸
	۲	−۰/۵۲۵۵	−۰/۳۲۹۳	−۰/۳۵۵۹	−۰/۲۳۲۵	−۰/۲۵۹۱	۰/۵۹۵۴	۰/۴۳۲۳	۰/۴۵۲۸	۰/۳۶۵۰	۰/۳۷۲۵
	۳	−۰/۵۵۵۷	−۰/۲۹۹۵	−۰/۳۷۷۱	−۰/۲۱۲۵	−۰/۲۳۵۰	۰/۶۱۶۸	۰/۴۰۱۷	۰/۴۶۳۳	۰/۳۴۵۹	۰/۳۵۲۵
	۴	−۰/۶۱۷۶	−۰/۳۰۰۸	−۰/۴۲۹۰	−۰/۲۲۴۹	−۰/۲۳۶۲	۰/۶۷۰۴	۰/۳۹۸۰	۰/۵۰۲۱	۰/۳۴۸۵	۰/۳۵۱۴
	۵	−۰/۶۶۹۸	−۰/۲۹۲۰	−۰/۴۷۱۳	−۰/۲۲۸۳	−۰/۲۳۱۸	۰/۷۱۷۱	۰/۳۸۸۳	۰/۵۳۵۳	۰/۳۴۵۳	۰/۳۴۵۷
۲۰	۱	−۰/۴۳۵۱	−۰/۳۶۵۸	−۰/۳۱۴۴	−۰/۲۷۳۸	−۰/۳۰۳۰	۰/۴۸۲۵	۰/۴۲۱۱	۰/۳۷۶۲	۰/۳۴۳۰	۰/۳۶۱۷
	۲	−۰/۳۲۴۱	−۰/۲۲۶۰	−۰/۱۸۵۳	−۰/۱۲۳۲	−۰/۱۶۷۹	۰/۳۷۲۵	۰/۲۹۱۲	۰/۲۶۰۹	۰/۲۲۱۷	۰/۲۴۵۰
	۳	−۰/۳۱۲۹	−۰/۱۸۴۸	−۰/۱۷۷۴	−۰/۰۹۳۸	−۰/۱۴۲۷	۰/۳۶۰۶	۰/۲۵۷۴	۰/۲۵۳۰	۰/۲۰۶۴	۰/۲۲۵۳
	۴	−۰/۳۲۷۹	−۰/۱۶۹۵	−۰/۱۹۲۵	−۰/۰۸۴۷	−۰/۱۳۶۲	۰/۳۷۳۶	۰/۲۴۶۵	۰/۲۶۴۴	۰/۲۰۳۵	۰/۲۲۲۱
	۵	−۰/۳۴۹۹	−۰/۱۶۱۰	−۰/۲۱۱۴	−۰/۰۸۴۸	−۰/۱۳۱۲	۰/۳۹۲۶	۰/۲۴۰۰	۰/۲۷۸۴	۰/۲۰۳۲	۰/۲۱۸۹
	۶	−۰/۳۷۸۳	−۰/۱۵۸۸	−۰/۲۳۴۱	−۰/۰۸۵۸	−۰/۱۲۷۸	۰/۴۱۹۲	۰/۲۴۰۵	۰/۲۹۷۷	۰/۲۰۷۶	۰/۲۱۹۷
	۷	−۰/۴۰۷۵	−۰/۱۵۷۵	−۰/۲۵۷۳۴	−۰/۰۸۷۸	−۰/۱۲۳۴	۰/۴۴۵۲	۰/۲۳۸۷	۰/۳۱۶۰	۰/۲۰۸۹	۰/۲۱۸۴
	۸	−۰/۴۳۴۵	−۰/۱۵۳۹	−۰/۲۷۶۵	−۰/۰۸۵۸	−۰/۱۱۴۵	۰/۴۶۹۶	۰/۲۳۵۴	۰/۳۳۰۷	۰/۲۰۷۶	۰/۲۱۴۲
	۹	−۰/۴۶۸۵	−۰/۱۵۷۲	−۰/۳۰۴۰	−۰/۰۹۳۲	−۰/۱۱۳۹	۰/۵۰۱۶	۰/۲۳۸۴	۰/۳۵۴۱	۰/۲۱۱۸	۰/۲۱۶۴
	۱۰	−۰/۴۹۹۱	۰/۱۵۷۲	−۰/۳۲۸۶	−۰/۰۹۷۴	−۰/۱۱۱۶	۰/۵۳۱۰	۰/۲۳۹۹	۰/۳۷۶۲	۰/۲۱۵۶	۰/۲۱۸۶
۳۰	۱	−۰/۳۷۴۸	−۰/۳۲۸۶	−۰/۲۶۰۶	۰/۲۳۳۶	−۰/۲۶۳۳	۰/۴۰۹۲	۰/۳۶۷۴	۰/۳۰۶۹	۰/۲۸۴۴	۰/۳۰۶۷
	۲	−۰/۲۵۸۶	−۰/۱۹۳۲	−۰/۱۳۰۲	−۰/۰۸۹۳	−۰/۱۳۲۶	۰/۲۹۸۴	۰/۲۴۴۰	۰/۱۹۸۰	۰/۱۷۳۶	۰/۱۹۶۵
	۳	−۰/۲۳۶۶	−۰/۱۵۱۲	−۰/۱۱۷۰	−۰/۰۶۱۰	−۰/۱۱۱۷	۰/۲۷۶۴	۰/۲۰۸۰	۰/۱۸۴۹۱	۰/۱۵۶۵	۰/۱۷۸۷
	۴	−۰/۲۳۶۹	−۰/۱۳۱۳	−۰/۱۲۱۱	−۰/۰۵۰۵	−۰/۱۰۵۵	۰/۲۷۷۳	۰/۱۹۴۹	۰/۱۸۹۶	۰/۱۵۶۰	۰/۱۷۶۳
	۵	−۰/۲۴۲۹	−۰/۱۱۷۰	−۰/۱۲۸۳	−۰/۰۴۲۹	−۰/۰۹۹۸	۰/۲۸۲۶	۰/۱۵۵۸	۰/۱۹۴۷	۰/۱۵۴۸	۰/۱۷۳۳
	۶	−۰/۲۵۵۶	−۰/۱۰۹۳	−۰/۱۳۹۲	−۰/۰۳۹۱	−۰/۰۹۶۴	۰/۲۹۳۳	۰/۱۸۰۷	۰/۲۰۲۰	۰/۱۵۴۴	۰/۱۷۱۱
	۷	−۰/۲۷۰۷	−۰/۱۰۴۰	−۰/۱۵۰۴	−۰/۰۳۵۵	−۰/۰۹۲۳	۰/۳۰۷۶	۰/۱۷۹۴	۰/۲۱۱۸	۰/۱۸۶۹	۰/۱۷۱۰
	۸	−۰/۲۸۵۰	−۰/۰۹۷۹	−۰/۱۶۱۳	−۰/۰۳۱۳	−۰/۰۸۵۷	۰/۳۲۰۲	۰/۱۷۵۷	۰/۲۱۹۴	۰/۱۵۶۴	۰/۱۶۷۶
	۹	−۰/۳۰۱۵	−۰/۰۹۳۹	−۰/۱۷۲۶	−۰/۰۲۷۹	−۰/۰۷۹۲۱	۰/۳۳۵۹	۰/۱۷۵۴	۰/۲۲۹۶	۰/۱۵۹۳	۰/۱۶۷۱
	۱۰	−۰/۳۲۰۲	−۰/۰۹۹۲	−۰/۱۸۵۷	−۰/۰۲۶۴	−۰/۰۷۳۴	۰/۳۵۲۲	۰/۱۷۳۲	۰/۲۳۹۱	۰/۱۵۸۸	۰/۱۶۴۴
	۱۱	−۰/۳۳۸۸	۰/۰۹۰۴	−۰/۱۹۸۵	−۰/۰۲۴۳	−۰/۰۶۶۹	۰/۳۷۰۲	۰/۱۷۴۶	۰/۲۵۰۹	۰/۱۶۲۱	۰/۱۶۵۳
	۱۲	−۰/۳۵۷۹	−۰/۰۸۹۱	−۰/۲۱۱۸	−۰/۰۲۳۳	−۰/۰۶۰۲	۰/۳۸۸۲	۰/۱۷۴۶	۰/۲۶۲۳	۰/۱۶۴۳	۰/۱۶۵۳
	۱۳	−۰/۳۸۱۱	−۰/۰۹۱۸	−۰/۲۲۹۵	−۰/۰۲۶۹	−۰/۰۵۷۸	۰/۴۱۰۳	۰/۱۷۷۷	۰/۲۷۷۶	۰/۱۶۶۹	۰/۱۶۷۷
	۱۴	− ۰/۴۰۰۱	−۰/۰۹۰۴	−۰/۲۴۳۶	−۰/۰۲۶۹	−۰/۰۵۲۰	۰/۴۲۷۳	۰/۱۷۵۲	۰/۲۸۸۱	۰/۱۶۵۵	۰/۱۶۵۵
	۱۵	−۰/۴۲۵۳	−۰/۰۹۵۲	−۰/۲۶۴۴	−۰/۰۳۴۳	−۰/۰۵۳۸	۰/۴۵۱۲	۰/۱۷۸۰	۰/۳۰۶۰	۰/۱۶۷۶	۰/۱۶۸۰

جدول ۲: اریبی و جذر میانگین توان دوم خطای برآوردگرهای آنتروپی توزیع توزیع نمایی با میانگین ۱.

		اریبی					جذرمیانگین توان دوم خطا				
n	m	HV_{mn}	HE_{mn}	HC_{mn}	$HZ۱_{mn}$	$HZ۲_{mn}$	HV_{mn}	HE_{mn}	HC_{mn}	$HZ۱_{mn}$	$HZ۲_{mn}$
۱۰	۱	−۰/۵۳۱۶	−۰/۳۹۳۰	−۰/۳۹۱۳	−۰/۳۱۰۲	−۰/۳۰۵۷	۰/۶۵۵۹	۰/۵۴۹۶	۰/۵۴۶۲	۰/۴۹۱۳	۰/۴۷۸۰
	۲	−۰/۴۴۸۸	−۰/۲۵۲۶	−۰/۲۷۱۶	−۰/۱۵۱۳	−۰/۱۵۷۴	۰/۵۸۳۱	۰/۴۴۹۹	۰/۴۶۲۱	۰/۴۰۴۰	۰/۳۹۷۸
	۳	−۰/۴۳۴۵	−۰/۱۷۸۳	−۰/۲۴۰۴	−۰/۰۸۱۶	−۰/۰۸۴۶	۰/۵۶۹۸	۰/۴۰۹۴	۰/۴۴۳۲	۰/۳۸۳۳	۰/۳۷۸۴
	۴	−۰/۴۶۶۰	−۰/۱۴۹۱	−۰/۲۵۴۳	−۰/۰۵۸۳	−۰/۰۵۲۴	۰/۵۹۱۹	۰/۳۹۴۳	۰/۴۵۱۹	۰/۳۷۹۸	۰/۳۷۹۵
	۵	−۰/۴۶۶۱	−۰/۰۸۸۳	−۰/۲۳۰۹	−۰/۰۰۰۰۱	−۰/۰۱۴۱	۰/۶۰۲۰	۰/۳۹۱۱	۰/۴۵۵۱	۰/۳۹۶۳	۰/۴۰۱۸
۲۰	۱	−۰/۴۱۳۲	−۰/۳۴۳۹	−۰/۲۹۲۵	−۰/۲۵۱۹	−۰/۲۶۳۰	۰/۴۸۹۸	۰/۴۳۳۰	۰/۳۹۲۶	۰/۳۶۳۴	۰/۳۶۸۰
	۲	−۰/۳۰۳۷	−۰/۲۰۵۶	−۰/۱۶۱۲	−۰/۱۰۰۲	−۰/۱۲۴۹	۰/۳۹۶۹	۰/۳۲۸۰	۰/۳۰۲۹	۰/۲۷۵۶	۰/۲۸۲۵
	۳	−۰/۲۵۳۴	−۰/۱۲۵۳	−۰/۱۱۱۷	−۰/۰۳۱۳	−۰/۰۵۸۸	۰/۳۴۹۴	۰/۲۷۱۲	۰/۲۶۷۳	۰/۲۴۵۲	۰/۲۴۶۲
	۴	−۰/۲۵۷۹	−۰/۰۹۹۵	−۰/۱۱۴۲	−۰/۰۱۴۵	−۰/۰۳۹۶	۰/۳۵۴۴	۰/۲۶۲۶	۰/۲۷۱۱	۰/۲۴۷۰	۰/۲۴۵۷
	۵	−۰/۲۶۴۳	−۰/۰۷۵۴	−۰/۱۱۳۶	۰/۰۰۵۵	−۰/۰۱۸۸	۰/۳۶۱۸	۰/۲۵۸۲	۰/۲۷۵۷	۰/۲۵۲۸	۰/۲۴۶۷
	۶	−۰/۲۶۲۸	−۰/۰۴۳۴	−۰/۱۰۳۹	۰/۰۳۴۱	−۰/۰۱۵۹	۰/۳۵۴۸	۰/۲۴۲۳	۰/۲۶۵۵	۰/۲۴۸۶	۰/۲۴۱۰
	۷	−۰/۲۵۹۵	−۰/۰۰۹۵	−۰/۰۹۱۴	۰/۰۶۵۶	۰/۰۵۲۲	۰/۳۵۷۱	۰/۲۴۵۴	۰/۲۶۹۹	۰/۲۶۳۸	۰/۲۵۶۶
	۸	−۰/۲۶۲۵	−۰/۰۱۸۱	−۰/۰۸۲۹	۰/۰۹۳۹	۰/۰۸۷۷	۰/۳۵۷۱	۰/۲۴۲۷	۰/۲۶۳۶	۰/۲۷۱۳	۰/۲۶۵۲
	۹	−۰/۲۸۴۲	۰/۰۲۷۰	−۰/۰۹۲۷	۰/۱۰۱۶	۰/۱۰۴۲	۰/۳۸۵۱	۰/۲۶۱۳	۰/۲۸۵۵	۰/۲۹۳۰	۰/۲۹۳۶
	۱۰	−۰/۲۸۳۸	۰/۰۵۸۰	−۰/۰۷۹۳	۰/۱۳۳۳	۰/۱۴۲۹	۰/۳۸۱۱	۰/۲۶۰۸	۰/۲۷۵۹	۰/۲۹۹۷	۰/۳۰۶۷
۳۰	۱	−۰/۳۷۱۵	−۰/۳۲۵۳	−۰/۲۵۷۲	۰/۲۳۰۲	−۰/۲۴۲۷	۰/۴۲۹۴	۰/۳۹۰۱	۰/۳۳۴۳	۰/۳۱۴۰	۰/۳۲۰۹
	۲	−۰/۲۳۶۳	−۰/۱۷۰۹	−۰/۱۰۷۱	۰/۰۶۶۶	−۰/۰۹۰۹	۰/۳۰۷۲	۰/۲۶۰۲	۰/۲۲۳۹	۰/۲۰۷۷	۰/۲۱۴۴
	۳	−۰/۲۰۳۶	−۰/۱۱۸۲	−۰/۰۸۰۸	۰/۰۲۷۱	−۰/۰۵۷۲	۰/۲۸۵۴	۰/۲۳۲۳	۰/۲۱۷۱	۰/۲۰۳۶	۰/۲۰۶۸
	۴	−۰/۱۹۳۷	−۰/۰۸۸۱	−۰/۰۷۳۰	۰/۰۰۶۸	−۰/۰۳۷۷	۰/۲۷۷۰	۰/۲۱۶۷	۰/۲۱۳۳	۰/۲۰۱۰	۰/۱۹۹۹
	۵	−۰/۱۹۷۹	−۰/۰۷۱۹	−۰/۰۷۶۹	۰/۰۰۲۵	−۰/۰۳۰۴	۰/۲۷۷۵	۰/۲۰۷۴	۰/۲۱۱۴	۰/۱۹۸۲	۰/۱۹۶۸
	۶	−۰/۱۸۸۰	−۰/۰۴۱۷	−۰/۰۶۵۱	۰/۰۲۷۷	−۰/۰۰۴۸	۰/۲۶۶۴	۰/۱۹۳۳	۰/۲۰۲۵	۰/۱۹۴۳	۰/۱۸۸۰
	۷	−۰/۱۹۰۶	۰/۰۲۳۹	−۰/۰۶۰۹	۰/۰۶۵۲	۰/۰۱۲۹	۰/۲۷۲۶	۰/۱۹۴۶	۰/۲۰۷۴	۰/۲۰۴۶	۰/۱۹۵۲
	۸	−۰/۱۸۶۷	۰/۰۰۰۳	−۰/۰۵۱۰	۰/۰۶۸۴	۰/۰۳۸۱	۰/۲۷۶۸	۰/۲۰۴۳	۰/۲۱۴۱	۰/۲۱۹۸	۰/۲۰۷۸
	۹	−۰/۲۰۷۰	۰/۰۲۰۹	−۰/۰۵۶۷	۰/۰۸۸۸	۰/۰۶۶۱	۰/۲۸۷۲	۰/۲۰۰۱	۰/۲۱۲۲	۰/۲۲۵۵	۰/۲۱۱۹
	۱۰	−۰/۲۰۷۰	۰/۰۲۰۹	−۰/۰۵۶۷	۰/۰۸۸۸	۰/۰۶۶۱	۰/۲۸۷۲	۰/۲۰۰۱	۰/۲۱۲۲	۰/۲۲۵۵	۰/۲۱۱۹
	۱۱	−۰/۲۰۷۰	۰/۰۴۱۳	−۰/۰۴۹۷	۰/۱۰۸۵	۰/۰۹۱۳	۰/۲۹۴۲	۰/۲۱۳۰	۰/۲۲۲۱	۰/۲۴۵۲	۰/۲۳۳۱
	۱۲	−۰/۲۰۳۹	۰/۰۶۴۸	−۰/۰۳۷۹	۰/۱۳۲۸	۰/۱۲۱۱	۰/۲۹۴۷	۰/۲۲۲۴	۰/۲۲۲۶	۰/۲۵۹۲	۰/۲۵۰۵۰
	۱۳	−۰/۲۰۷۹	۰/۰۸۱۲	−۰/۰۳۳۴	۰/۱۴۹۸	۰/۱۴۴۵	۰/۳۰۳۱	۰/۲۳۵۰	۰/۲۳۱۰	۰/۲۷۷۲	۰/۲۷۱۹
	۱۴	− ۰/۲۵۵۰	۰/۱۰۴۶	−۰/۰۲۱۷	۰/۱۷۳۱	۰/۱۷۴۵	۰/۲۹۲۷	۰/۲۳۳۷	۰/۲۱۷۱	۰/۲۸۰۹	۰/۲۸۰۷
	۱۵	۰/۲۱۶۰	۰/۱۱۴۱	−۰/۰۲۴۸	۰/۱۸۳۲	۰/۱۸۲۵	۰/۳۱۴۱	۰/۲۵۵۰	۰/۲۳۹۸	۰/۳۰۴۹	۰/۳۰۹۶

جدول ۳: اریبی و جذرمیانگین توان‌دوم خطای برآوردگرهای آنتروپی توزیع یکنواخت (۰,۱).

n	m	اریبی HV_{mn}	اریبی HE_{mn}	اریبی HC_{mn}	اریبی $HZ1_{mn}$	اریبی $HZ2_{mn}$	جذرمیانگین توان دوم خطا HV_{mn}	جذرمیانگین توان دوم خطا HE_{mn}	جذرمیانگین توان دوم خطا HC_{mn}	جذرمیانگین توان دوم خطا $HZ1_{mn}$	جذرمیانگین توان دوم خطا $HZ2_{mn}$
۱۰	۱	−۰/۵۱۰۵	−۰/۳۷۱۸	−۰/۳۷۱۷	−۰/۲۹۰۷	−۰/۲۵۳۹	۰/۵۶۱۸	۰/۴۳۹۷	۰/۴۳۸۱	۰/۳۷۱۷	۰/۳۲۵۸
	۲	−۰/۴۰۹۹	−۰/۲۱۳۸	−۰/۲۳۹۵	−۰/۱۲۲۲	−۰/۱۰۲۹	۰/۴۴۸۹	۰/۲۸۱۴	۰/۳۰۱۱	۰/۲۱۸۷	۰/۲۰۳۱
	۳	−۰/۴۲۷۱	−۰/۱۷۰۹	−۰/۲۴۷۶	−۰/۰۹۶۷	−۰/۰۸۷۹	۰/۴۵۹۴	۰/۲۴۰۵	۰/۲۹۹۰	۰/۱۹۰۸	۰/۱۹۰۷
	۴	−۰/۴۶۸۸	−۰/۱۵۱۹	−۰/۲۷۴۱	−۰/۰۸۹۹	−۰/۰۸۱۹	۰/۴۹۸۳	۰/۲۲۷۳	۰/۳۲۲۷	۰/۱۸۶۷	۰/۱۸۶۳
	۵	−۰/۵۰۳۸	−۰/۱۲۵۹	−۰/۲۹۴۲	−۰/۰۷۶۷	−۰/۰۷۲۳	۰/۵۲۹۲	۰/۲۰۵۳	۰/۳۳۸۵	۰/۱۷۵۱	۰/۱۷۴۷
	۲	−۰/۲۶۸۳	−۰/۱۷۰۳	−۰/۱۳۲۰	−۰/۰۷۳۴	−۰/۰۶۶۲	۰/۴۱۹۰	۰/۳۵۴۵	۰/۳۰۸۱	۰/۲۷۲۵	۰/۲۴۹۸
	۳	−۰/۲۵۰۳	−۰/۱۲۲۴	−۰/۱۱۸۹	−۰/۰۴۳۶	−۰/۰۳۸۳	۰/۲۶۵۸	۰/۱۵۱۴	۰/۱۴۹۰	۰/۰۹۸۵	۰/۰۹۹۵
	۴	−۰/۲۵۹۵	−۰/۱۰۱۱	−۰/۱۲۷۲	−۰/۰۳۵۴	−۰/۰۲۸۶	۰/۲۷۳۶	۰/۱۳۳۱	۰/۱۵۵۳	۰/۰۹۳۸	۰/۰۹۲۸
	۵	−۰/۲۷۴۵	−۰/۰۸۵۶	−۰/۱۴۰۲	−۰/۰۳۱۳	−۰/۰۲۶۹	۰/۲۸۵۹	۰/۱۱۷۱	۰/۱۶۲۷	۰/۰۸۴۷	۰/۰۸۸۵
	۶	−۰/۲۹۸۷	−۰/۰۷۹۳	−۰/۱۵۸۹	−۰/۰۳۳۸	−۰/۰۲۹۸	۰/۳۱۰۳	۰/۱۱۵۵	۰/۱۸۰۷	۰/۰۸۸۸	۰/۰۹۱۵
	۷	−۰/۳۲۰۲	−۰/۰۷۰۲	−۰/۱۷۲۸	−۰/۰۳۱۱	−۰/۰۲۷۵	۰/۳۲۹۹	۰/۱۰۶۱	۰/۱۹۱۶	۰/۰۸۳۵	۰/۰۸۷۴
	۸	−۰/۳۴۷۳	−۰/۰۶۶۷	−۰/۱۹۰۴	−۰/۰۳۱۳	−۰/۰۲۸۳	۰/۳۵۶۷	۰/۱۰۵۲	۰/۲۰۷۹	۰/۰۸۲۹	۰/۰۸۷۰
	۹	−۰/۳۷۴۹	−۰/۰۶۳۶	−۰/۲۰۸۱	−۰/۰۳۲۷	۰/۰۲۹۷	۰/۳۵۷۱	۰/۲۴۲۷	۰/۲۶۳۶	۰/۲۷۱۳	۰/۲۶۵۲
	۱۰	−۰/۴۰۳۲	−۰/۰۶۱۳	−۰/۲۲۵۲	۰/۰۳۳۳	−۰/۰۳۱۱	۰/۴۱۲۴	۰/۱۰۶۲	۰/۲۴۳۲	۰/۰۸۸۱	۰/۰۹۰۲
۳۰	۱	−۰/۳۵۲۵	−۰/۳۰۶۳	−۰/۲۳۷۸	۰/۲۱۰۷	−۰/۱۹۸۱	۰/۳۶۸۵	۰/۳۲۴۵	۰/۲۵۹۵	۰/۲۳۴۹	۰/۲۲۰۵
	۲	−۰/۲۲۷۰	−۰/۱۶۱۷	−۰/۱۰۲۰	۰/۰۶۲۸	−۰/۰۵۵۰	۰/۲۴۰۲	۰/۱۷۹۶	۰/۱۲۷۴	۰/۰۹۸۶	۰/۰۹۱۰
	۳	−۰/۲۰۰۱	−۰/۱۱۴۷	−۰/۰۸۳۱	−۰/۰۳۲۸	−۰/۰۲۷۷	۰/۲۱۱۵	۰/۱۳۳۶	۰/۱۰۸۷	۰/۰۷۶۵	۰/۰۷۲۵
	۴	−۰/۱۹۷۰	−۰/۰۹۱۳	−۰/۰۸۵۸	−۰/۰۲۴۴	−۰/۰۲۰۹	۰/۲۰۶۴	۰/۱۱۰۳	۰/۱۰۶۳	۰/۰۶۶۱	۰/۰۶۶۴
	۵	−۰/۲۰۴۷	−۰/۰۷۸۸	−۰/۰۹۵۲	−۰/۰۲۲۶	−۰/۰۱۹۴	۰/۲۱۳۵	۰/۰۹۹۴	۰/۱۱۳۷	۰/۰۶۴۳	۰/۰۶۵۰
	۶	−۰/۲۱۲۳	−۰/۰۶۶۰	−۰/۱۰۲۰	−۰/۰۱۸۶	−۰/۰۱۵۰	۰/۲۱۹۸	۰/۰۸۷۰	۰/۱۱۸۲	۰/۰۵۹۶	۰/۰۶۱۷
	۷	−۰/۲۲۵۹	−۰/۰۵۹۲	−۰/۱۱۲۵	−۰/۰۱۸۵	−۰/۰۱۴۶	۰/۲۳۲۶	۰/۰۸۰۹	۰/۱۲۶۹	۰/۰۵۸۴	۰/۰۶۱۱
	۸	−۰/۲۴۴۶	−۰/۰۵۷۵	−۰/۱۲۷۰	−۰/۰۲۱۰	−۰/۰۱۸۰	۰/۲۵۰۸	۰/۰۷۹۷	۰/۱۳۹۹	۰/۰۵۸۹	۰/۰۶۲۳
	۹	−۰/۲۶۰۱	۰/۰۵۲۵	−۰/۱۳۷۶	−۰/۰۲۰۰	−۰/۰۱۸۱	۰/۲۶۶۰	۰/۰۷۶۷	۰/۱۴۹۹	۰/۰۵۸۳	۰/۰۶۳۶
	۱۰	−۰/۲۷۹۵	−۰/۰۵۱۵	−۰/۱۵۰۶	−۰/۰۲۲۳	−۰/۰۲۰۲	۰/۲۸۵۳	۰/۰۷۷۳	۰/۱۶۲۱	۰/۰۵۹۸	۰/۰۶۴۶
	۱۱	−۰/۲۹۳۶	−۰/۰۶۵۲	−۰/۱۵۷۵	−۰/۰۱۸۷	−۰/۰۱۶۴	۰/۲۹۹۱	۰/۰۷۲۸	۰/۱۶۸۶	۰/۰۵۷۴	۰/۰۶۲۴
	۱۲	−۰/۳۱۵۶	−۰/۰۴۶۸	−۰/۱۷۲۲	−۰/۰۲۱۷	−۰/۰۲۰۰	۰/۳۲۱۳	۰/۰۷۶۳	۰/۱۸۳۵	۰/۰۵۸۹	۰/۰۶۵۸
	۱۳	−۰/۳۳۲۹	−۰/۰۴۳۶	−۰/۱۸۲۸	−۰/۰۲۱۳	−۰/۰۱۹۹	۰/۳۳۷۸	۰/۰۷۲۱	۰/۱۹۲۸	۰/۰۵۳۷	۰/۰۶۱۸
	۱۴	− ۰/۳۴۷۲	−۰/۰۳۷۵	−۰/۱۸۸۶	−۰/۰۱۷۳	−۰/۰۱۵۰	۰/۳۵۲۳	۰/۰۷۰۵	۰/۱۹۹۱	۰/۰۵۷۵	۰/۰۶۱۶
	۱۵	− ۰/۳۷۱۹	۰/۰۴۱۸	−۰/۲۰۷۰	−۰/۰۲۳۲	−۰/۰۲۲۰	۰/۳۷۶۶	۰/۰۷۲۵	۰/۲۱۶۵	۰/۰۵۹۰	۰/۰۶۱۷

تحت هرسه توزیع نرمال، نمایی و یکنواخت و به ازای هر سه حجم نمونه ۲۰،۱۰ و ۳۰ از HV، HE و HCO بهتراست. با توجه به اینکه اولاً شرط $m \leq \frac{n}{۴}$ فقط در مورد توزیع نمایی لازم است و ثانیاً جداول نشان می‌دهد که m بهینه کمتر از $\frac{n}{۴}$ می‌باشد، نتیجه می‌گیریم $HZ۱$ به طور یکنواخت از همه برآوردگرهای مورد بررسی از نظر اریبی و $RMSE$ بهتر است.

فرض کنید که مشاهدات $x_۱, \cdots, x_n$ داده شده اند، نحوهٔ انتخاب $m(m \leq \frac{n}{۴})$ که به ازای آن بهترین برآورد آنتروپی حاصل گردد، هنوز به عنوان مساله‌ای باز مطرح است. ویزورکوسکی و گورزگورزسکی (۱۹۹۹) رابطه $m = [\sqrt{n} + ۰/۵]$ را برای انتخاب m هنگامی که n داده شده‌است، پیشنهاد دادند، که در آن $[\cdot]$ علامت جزء صحیح است. همان‌طور که از جداول ۱ تا ۳ مشخص است، برآوردگرهای پیشنهادی به ازای این مقدار m نیز عملکردی بهتر نسبت به سایر برآوردگرها دارند.

۳ آزمون‌های نیکویی برازش

در این بخش به معرفی آزمون‌های نیکویی برازش برای فرضیه‌های نمایی بودن و نرمال بودن توزیع جامعه بر اساس برآوردگرهای جدید آنتروپی معرفی شده می‌پردازیم. از آنجا که رابطه

$$HE_{mn} = HV_{mn} + \frac{۲}{n}[m \log(۲m) + \log(\frac{(m-۱)!}{(۲m-۱)!})],$$

میان برآوردگرهای واسیچک (۱۹۷۶) و ابراهیمی و همکاران (۱۹۹۴) برقرار است، توان آزمون‌های نیکویی برازش برمبنای این دو برآوردگر دقیقاً یکسان می‌باشد. لذا دراین مقاله توان آزمون‌های نیکویی برازش بر مبنای برآوردگرهای جدید را تنها با آزمون‌های نیکویی برازش بر مبنای برآوردگر واسیچک (۱۹۷۶) و برآوردگر کوریا (۱۹۹۵) مقایسه می‌کنیم.

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

۱.۳ آزمون نمایی بودن

فرض کنید $X_۱, \cdots, X_n$ یک نمونه تصادفی مستقل و هم‌توزیع از توزیعی با تابع چگالی مطلقاً پیوسته f باشد. می‌خواهیم فرضیه $H_۰ : f = f^\circ$ را در مقابل نقیض آن آزمون کنیم، که در آن $f^\circ(x, \theta) = \frac{۱}{\theta} e^{-\frac{x}{\theta}}$ و θ مقداری مثبت و نامعلوم است. فاصله نامتقارن کولبک–لایبلر f از f° عبارت است از

$$\begin{aligned} D(f^\circ, f) &= \int_{-\infty}^{\infty} f(x,\theta) \ln(\frac{f(x,\theta)}{f^\circ(x,\theta)}) dx \\ &= -H(X) + \ln(\theta) + \frac{۱}{\theta} E_f(X). \end{aligned}$$

به سادگی به کمک مشتق‌گیری معلوم می‌شود که

$$D_{inf} = -H(X) + \ln E_f(X) + ۱,$$

بدیهی است که $D_{\inf} = \inf_\theta D(f, f^\circ) = ۰$ اگر و تنها اگر $H_۰$ درست باشد. لذا می‌توان آماره آزمون را به صورت‌های

$$TZ_۱ = ۱ + \ln(\bar{X}) - HZ۱_{mn}, \quad TZ_۲ = ۱ + \ln(\bar{X}) - HZ۲_{mn}$$

پیشنهاد کرد، که بر مبنای آن‌ها فرضیه صفر را به ازای مقادیر بزرگ $TZ_۱$ و $TZ_۲$ رد می‌کنیم. توزیع این آماره‌ها به روش تحلیلی قابل محاسبه نمی‌باشد، بنابراین برای بدست آوردن نقاط بحرانی آزمون‌ها از روش شبیه سازی مونت کارلو استفاده می‌کنیم. قابل ذکر است که نقاط بحرانی به خاطر اینکه آماره آزمون نسبت به تبدیلات مقیاسی ناوردا می‌باشد، به پارامتر مجهول بستگی ندارد. نقاط بحرانی به این صورت محاسبه می‌شود، که ابتدا از توزیع نمایی با میانگین یک، نمونه‌ای به حجم n تولید و مقدار آماره آزمون را محاسبه می‌کنیم و این کار را ۱۰,۰۰۰ بار تکرار می‌کنیم. مقدار بحرانی آزمون با استفاده از چندک $۱ - \alpha$ام توزیع تجربی آماره آزمون بدست می‌آید. جداول ۴ و ۵ حاوی نتایج ۱۰,۰۰۰ دفعه شبیه‌سازی (حجم نمونه ۱۰ و ۲۰) است که برای بدست آوردن توان آزمون‌های پیشنهادی و توان آزمون‌های نمایی بودن بر مبنای برآوردگر واسیچک (۱۹۷۶) و کوریا (۱۹۹۵) در

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

سطح معنی‌داری $\alpha = ۰/۰۵$ مورد استفاده قرار گرفته‌اند. آماره‌های آزمون برمبنای برآوردگر واسیچک (۱۹۷۶) و کوریا (۱۹۹۵) به صورت زیر است.

$$TV = ۱ + \ln(\bar{X}) - HV_{mn}, \quad TC = ۱ + \ln(\bar{X}) - HC_{mn}.$$

در حالت کلی در آزمون‌های مبتنی بر آنتروپی مقدار بهینه m علاوه بر حجم نمونه به توزیع فرضیه مقابل نیز بستگی دارد و نمی‌توان، در حجم نمونه ثابت n، مقدار m را به قسمی تعیین کرد که توان آزمون به ازای تمام توزیع‌های جانشین ماکسیمم شود. ابراهیمی و همکاران (۱۹۹۲) به ازای حجم نمونه n، مقداری از m را پیشنهاد دادند که آزمون مربوطه دارای توان نسبتاً خوبی به ازای تمام توزیع‌های جانشین است. این مطلب برای آزمون‌های پیشنهادی ما نیز صادق می‌باشد لذا ما نیز مقدار m را برابر ۳ و ۴ به ترتیب برای حجم نمونه ۱۰ و ۲۰ درنظر می‌گیریم. در مقایسه توان‌ها از توزیع‌های زیر به عنوان فرضیه جانشین استفاده شده‌است.
الف) توزیع وایبل با تابع چگالی

$$f(x;\lambda;\beta) = \beta\lambda^{\beta}x^{\beta-۱}\exp(-(\lambda x)^{\beta}), \beta > ۰, \lambda > ۰, x \geq ۰$$

ب) توزیع گاما با تابع چگالی

$$f(x;\lambda;\beta) = \frac{\lambda^{\beta}x^{\beta-۱}\exp(-\lambda x)}{\Gamma(\beta)}, \beta > ۰, \lambda > ۰, x \geq ۰$$

ج) توزیع لگ‌نرمال با تابع چگالی

$$f(x;\nu;\sigma^{۲}) = \frac{۱}{x\sigma\sqrt{۲\pi}}\exp(\frac{۱}{۲\sigma^{۲}}(\ln(x) - \nu)^{۲}), -\infty \leq \nu \leq \infty, \sigma > ۰, x > ۰$$

در این مقاله مقادیر سیاه در جدول‌ها نشان‌دهندهٔ آن است که توان آزمون، در حجم نمونه در نظر گرفته شده و تحت توزیع جانشین از توان سایر آزمون‌های بر مبنای آنتروپی بیشتر است.

از آنجا که برای بدست آوردن نقاط بحرانی آزمون‌ها از توزیع نمایی با میانگین یک داده تولید کردیم، در تمامی توزیع‌های بالا پارامترهای توزیع طوری انتخاب شده‌اند که امید ریاضی متغیر تصادفی برابر یک شود. بنابراین برای توزیع وایبل

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

$\lambda = \Gamma(۱ + \frac{۱}{\beta})$، برای توزیع گاما $\lambda = \beta$ و برای توزیع لگ‌نرمال $\nu = -\frac{\sigma^۲}{۲}$ در نظر گرفته می‌شود. البته واضح است که همهٔ آزمون‌های معرفی شده آزمون‌هایی دقیق می‌باشند و نقاط بحرانی این آزمون‌ها به مقدار پارامتر λی توزیع نمایی بستگی ندارند، یعنی اگر از توزیع نمایی با میانگینی غیر از یک، داده تولید شود و مقادیر بحرانی آزمون‌ها را محاسبه کنیم مقادیر بحرانی تغییری نمی‌کنند.

ما در اینجا از فرضیه‌های جانشینی استفاده کردیم که ابراهیمی و همکاران (۱۹۹۲) از آن‌ها برای مقایسه توان آزمون خود با آزمون‌های دیگر استفاده کرده‌اند. البته می‌توان فرضیه‌های جانشینی را در نظر گرفت که میانگین توزیع جانشین یک نباشد. مجدداً تاکید می‌شود که در توزیع‌های گاما، وایبل و لگ‌نرمال تنها پارامتر شکل در توان آزمون موثر و پارامتر مقیاس در توان آزمون تاثیری ندارد.

جدول ۴: توان آزمون‌های نمایی بودن بر مبنای آنتروپی، $\alpha = ۰/۰۵$ و $n = ۱۰$.

$Alternatives$	TV	TC	$TZ_۱$	$TZ_۲$
$Gamma(۲, ۲)$	۰/۳۲۵۵	۰/۳۳۲۵	۰/۳۱۹۴	**۰/۳۵۰۰**
$Gamma(۳, ۳)$	۰/۶۳۴۵	۰/۶۳۶۰	۰/۶۱۶۲	**۰/۶۶۹۳**
$Gamma(۴, ۴)$	۰/۸۲۷۴	۰/۸۳۰۰	۰/۸۱۱۴	**۰/۸۵۸۰**
$Weibull(۲, ۱/gamma(۱ + ۱/۲))$	۰/۶۹۰۴	۰/۷۰۲۰	۰/۶۸۲۶	**۰/۷۳۲۳**
$Weibull(۳, ۱/gamma(۱ + ۱/۳))$	۰/۹۸۲۵	۰/۹۸۳۱	۰/۹۷۷۶	**۰/۹۸۷۸**
$Weibull(۴۱/gamma(۱ + ۱/۴))$	۰/۹۹۹۸	۰/۹۹۹۸	۰/۹۹۹۷	**۰/۹۹۹۹**
$Lognormal(-۴/۵, ۳)$	**۰/۲۰۴۱**	۰/۱۴۱۸	۰/۱۳۷۹	۰/۱۳۳۰
$Lognormal(-۴/۵, ۳)$	**۰/۵۹۱۰**	۰/۵۰۱۸	۰/۴۹۹۳	۰/۷۱۷۶
$Lognormal(-۸, ۴)$	**۰/۸۱۵۵**	۰/۷۵۷۳	۰/۷۵۷۹	۰/۷۱۷۶

جدول ۵: توان آزمون‌های نمایی بودن بر مبنای آنتروپی، $\alpha = ۰/۰۵$ و $n = ۲۰$.

$Alternatives$	TV	TC	$TZ_۱$	$TZ_۲$
$Gamma(۲, ۲)$	۰/۵۰۴۶	۰/۴۹۷۱	۰/۴۷۷۶	**۰/۵۷۸۸**
$Gamma(۳, ۳)$	۰/۸۹۰۰	۰/۸۷۷۶	۰/۸۵۸۷	**۰/۹۲۸۱**
$Gamma(۴, ۴)$	۰/۹۸۳۵	۰/۹۷۷۶	۰/۹۷۰۶	**۰/۹۹۱۷**
$Weibull(۲, ۱/gamma(۱ + ۱/۲))$	۰/۹۳۲۳	۰/۹۲۹۶	۰/۹۱۸۵	**۰/۹۶۱۱**
$Weibull(۳, ۱/gamma(۱ + ۱/۳))$	۰/۹۹۹۹	۰/۹۹۹۹	۰/۹۹۹۸	**۱**
$Weibull(۴, ۱/gamma(۱ + ۱/۴))$	**۱**	**۱**	**۱**	**۱**
$Lognormal(-۴/۵, ۳)$	**۰/۶۵۴۹**	۰/۵۸۳۰	۰/۵۷۸۱	۰/۶۱۲۲
$Lognormal(-۴/۵, ۳)$	**۰/۹۶۷۳**	۰/۹۴۸۹	۰/۹۴۹۲	۰/۹۴۸۸
$Lognormal(-۸, ۴)$	**۰/۹۹۹۶**	۰/۹۹۵۱	۰/۹۹۵۲	۰/۹۹۳۹

همان‌طور که در جداول ۴ و ۵ ملاحظه می‌شود، هیچ کدام از آزمون‌ها به طور کامل بر دیگری تسلط ندارد. به این معنی که هیچ کدام به ازای هر سه توزیع

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

جانشین توان بیشتری ندارد، اما آزمون‌های مبتنی بر برآوردگر واسیچک (۱۹۷۶) و TZ۲ در توزیع‌های جانشین مختلف دارای ماکسیمم توان هستند. همچنین ملاحظه می‌شود که در یکی از سه توزیع جانشین (توزیع لگ‌نرمال) توان آزمون مبتنی بر TV و در دو توزیع وایبل و گاما توان آزمون برمبنای TZ۲ بیشتر است. آزمون مبتنی بر TZ۲، در بیشتر موارد توان بهتری نسبت به سایر آزمون‌ها دارد.

۲.۳ آزمون نرمال بودن

فرض کنید $X_1, \cdots, X_n$ یک نمونه تصادفی مستقل و هم‌توزیع از توزیعی با تابع چگالی مطلقاً پیوسته f باشد. می‌خواهیم فرضیه $H_0 : f = f^0$ را در مقابل نقیض آن آزمون کنیم، که در آن $f^0(x; \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$ و μ و $\sigma > 0$ مقادیر نامعلوم هستند. بنابر قضیه‌ای در نظریه اطلاع (شانون، ۱۹۴۸) توزیع نرمال دارای بیشترین مقدار آنتروپی در رده توزیع‌های با واریانس معلوم σ^2 است. واسیچک (۱۹۷۶) بر مبنای این خاصیت توزیع نرمال، آماره آزمون خود را برای فرضیه نرمال بودن به صورت

$$TV_{mn} = \frac{\exp(HV_{mn})}{S_x},$$

معرفی کرد، که در آن S_x انحراف معیار نمونه است. مشابه واسیچک (۱۹۷۶) می‌توان برمبنای دیگر برآوردهای آنتروپی آماره‌های

$$TC_{mn} = \frac{\exp(HC_{mn})}{S_x},\ \ TZ1_{mn} = \frac{\exp(HZ1_{mn})}{S_x},\ \ TZ2_{mn} = \frac{\exp(HZ2_{mn})}{S_x}.$$

را برای آزمون فرضیه نرمال بودن پیشنهاد داد، که فرضیه صفر به ازای مقادیر کوچک آماره‌های فوق رد می‌شود. ناوردا بودن این آماره‌ها نسبت به تبدیلات مکانی و مقیاسی واضح است. توزیع این آماره‌ها نیز به روش تحلیلی قابل محاسبه نمی‌باشد، بنابراین برای بدست آوردن نقاط بحرانی آزمون‌ها از روش شبیه سازی مونت کارلو استفاده می‌کنیم. بدین صورت که ابتدا از توزیع نرمال استاندارد، نمونه‌ای به حجم n تولید و مقدار آماره آزمون را محاسبه می‌کنیم و این کار را ۱۰,۰۰۰ بار تکرار می‌کنیم. مقدار بحرانی آزمون را با استفاده از چندک αام توزیع

تجربی آماره آزمون بدست می‌آوریم. به منظور مقایسه توان این آزمون‌ها با یکدیگر، شبیه‌سازی مونت–کارلو با ۱۰,۰۰۰ بار تکرار، برای حجم نمونه $n =$ ۱۰,۲۰ تحت ۲۰ توزیع در فرضیه مقابل انجام می‌شود. در اینجا نیز همانند آزمون نمایی بودن، مقدار بهینه m علاوه بر حجم نمونه به توزیع جانشین نیز بستگی دارد، واسیچک (۱۹۷۶) به ازای مقادیر مختلف n مقادیری از m را پیشنهاد داد که آزمون مربوطه دارای توان نسبتاً خوبی به ازای تمام توزیع‌های جانشین است. این مطلب برای آزمون‌های پیشنهادی ما نیز صادق می‌باشد لذا ما نیز مقدار m را برابر ۲ و ۳ به ترتیب برای ۱۰ $= n$ و ۲۰ $= n$ در نظر می‌گیریم. توزیع‌های جانشین مورد بررسی را بنابر شکل و تکیه‌گاه‌شان می‌توان به ۴ گروه مختلف تقسیم کرد. این تقسیم‌بندی می‌تواند دیدگاه بهتری نسبت به رفتار آزمون ارائه دهد. این سبک تقسیم‌بندی برای مقایسه توان آزمون‌های نرمال‌بودن برای نخستین بار توسط استبان و همکاران (۲۰۰۱) انجام شد.

گروه I: خانواده توزیع‌های متقارن با تکیه‌گاه $(-\infty, \infty)$

- توزیع t استیودنت با یک درجه آزادی(کوشی استاندارد)
- توزیع t استیودنت با سه درجه آزادی
- توزیع لوژستیک استاندارد
- توزیع نمایی دوگانه استاندارد

گروه II: خانواده توزیع‌های نامتقارن با تکیه‌گاه $(-\infty, \infty)$

- توزیع گامبل با پارامترهای ۰ $= \alpha$(مکان) و ۱ $= \beta$(مقیاس)
- توزیع نرمال چوله با پارامترهای ۰ $= \mu$(مکان)، ۱ $= \sigma$(مقیاس) و ۲ $= \alpha$(شکل)
- توزیع نمایی دوگانه چوله (ترکیب توزیع نمایی با میانگین ۲ $= \beta$ و قرینه توزیع نمایی با میانگین ۱ $= \alpha$)

گروهIII: خانواده توزیع‌های با تکیه‌گاه $(۰,\infty)$

- توزیع نمایی با میانگین ۱
- توزیع گاما با پارامترهای $۱=\beta$(مقیاس) و $۲=\alpha$(شکل)
- توزیع گاما با پارامترهای $۱=\beta$(مقیاس) و $\frac{۱}{۲}=\alpha$(شکل)
- توزیع لگ‌نرمال با پارامترهای $۰=\mu$(مقیاس) و $۱=\sigma$(شکل)
- توزیع لگ‌نرمال با پارامترهای $۰=\mu$(مقیاس) و $۲=\sigma$(شکل)
- توزیع لگ‌نرمال با پارامترهای $۰=\mu$(مقیاس) و $\frac{۱}{۲}=\sigma$(شکل)
- توزیع وایبل با پارامترهای $۱=\beta$(مقیاس) و $\frac{۱}{۲}=\alpha$(شکل)
- توزیع لگ‌نرمال با پارامترهای $۱=\beta$(مقیاس) و $۲=\alpha$(شکل)

گروهIV: خانواده توزیع‌های با تکیه‌گاه $(۰,۱)$

- توزیع یکنواخت $(۰,۱)$
- توزیع بتا$(۲,۲)$
- توزیع بتا$(۰/۵,۰/۵)$
- توزیع بتا$(۳,۱/۵)$
- توزیع بتا$(۲,۱)$

جداول ۶ تا ۹ نتایج شبیه‌سازی را نشان می‌دهند. مجدداً متذکر می‌شویم که دراین جداول مقادیر سیاه نشان دهنده این است که در حجم نمونه و توزیع جانشین مربوطه، توان آزمون مربوطه نسبت به دیگر آزمون‌های بر مبنای آنتروپی بیشتر است. همان‌طور که در این جداول ملاحظه می‌شود، آمارهٔ آزمون مبتنی بر $TZ۲_{mn}$ درهمه گروه‌های توزیع‌های مورد بررسی بجز گروه IV دارای بیشترین توان نسبت به سایر آزمون‌های مبتنی بر آنتروپی است و مقدار اختلاف توان‌ها نیز قابل توجه

جدول ۶: توان آزمون‌های نرمال بودن بر مبنای آنتروپی و آزمون شاپیرو–ویلک تحت خانواده‌های توزیع‌های جانشین گروه I و $\alpha = ۰/۰۵$.

Alternatives	n	TV	TC	$TZ_۱$	$TZ_۲$	SW
$t_۱$	۱۰	۰/۴۳۳۱	۰/۴۰۲۹	۰/۳۹۵۲	۰/۵۰۸۶	۰/۵۹۶۸
	۲۰	۰/۷۳۶۰	۰/۶۸۹۴	۰/۶۷۶۶	**۰/۸۱۲۱**	۰/ ۸۵۶۸
$t_۳$	۱۰	۰/۰۹۵۸	۰/۰۸۸۵	۰/۰۸۶۷	**۰/۱۳۳۹**	۰/۱۸۹۳
	۲۰	۰/۱۵۷۷	۰/۱۲۸۹	۰/۱۲۱۳	**۰/۲۲۳۸**	۰/۳۳۶۵
$Double\ exponential(۰, ۱)$	۱۰	۰/۰۷۰۶	۰/۰۶۳۸	۰/۰۶۱۳	**۰/۱۰۳۴**	۰/۱۵۳۹
	۲۰	۰/۰۸۹۱	۰/۰۶۹۷	۰/۰۶۵۱	**۰/۱۶۶۵**	۰/۲۶۰۲
$Logistic(۰, ۱)$	۱۰	۰/۰۵۷۲	۰/۰۵۴۱	۰/۰۵۴۶	**۰/۰۶۴۷**	۰/۰۸۰۸
	۲۰	۰/۰۵۰۳	۰/۰۴۶۱	۰/۰۴۴۲	**۰/۰۷۳۸**	۰/۱۰۹۹

جدول ۷: توان آزمون‌های نرمال بودن بر مبنای آنتروپی و آزمون شاپیرو–ویلک تحت خانواده‌های توزیع‌های جانشین گروه II و $\alpha = ۰/۰۵$.

Alternatives	n	TV	TC	$TZ_۱$	$TZ_۲$	SW
$Gumbel(۰, ۱)$	۱۰	۰/۱۰۳۶	۰/۰۹۹۴	۰/۰۹۷۵	**۰/۱۲۸۳**	۰/۱۵۲۳
	۲۰	۰/۱۹۲۱	۰/۱۸۵۱	۰/۱۷۷۲	**۰/۲۴۷۱**	۰/۳۰۳۶
$Skew\ Normal(۰, ۱, ۲)$	۱۰	۰/۰۶۲۴	۰/۰۶۳۸	۰/۰۶۲۷	**۰/۰۶۹۳**	۰/۰۷۱۰
	۲۰	۰/۰۶۶۹	۰/۰۶۶۷	۰/۰۶۶۰	**۰/۰۸۰۸**	۰/۱۰۳۵
$Skew\ Double\ Exponential(۱, ۲)$	۱۰	۰/۱۲۴۴	۰/۱۱۷۳	۰/۱۱۵۱	**۰/۱۶۸۴**	۰/۲۲۱۸
	۲۰	۰/۲۱۵۹	۰/۱۹۲۳	۰/۱۸۱۱	**۰/۳۰۵۳**	۰/۴۰۴۴

جدول ۸: توان آزمون‌های نرمال بودن بر مبنای آنتروپی و آزمون شاپیرو–ویلک تحت خانواده‌های توزیع‌های جانشین گروه III و $\alpha = ۰/۰۵$.

Alternatives	n	TV	TC	$TZ_۱$	$TZ_۲$	SW
$Exponentail(۱)$	۱۰	۰/۴۲۷۹	۰/۴۲۵۰	۰/۴۲۴۵	**۰/۴۷۶۶**	۰/۴۴۱۷
	۲۰	۰/۸۴۷۷	۰/۸۳۸۴	۰/۸۳۳۰	**۰/۸۶۴۰**	۰/۸۳۱۳
$Gamma(۲)$	۱۰	۰/۱۹۲۸	۰/۱۹۰۳	۰/۱۸۸۰	**۰/۲۳۲۵**	۰/۲۴۶۹*
	۲۰	۰/۴۴۰۰	۰/۴۳۳۵	۰/۴۲۲۲	**۰/۵۰۳۹**	۰/۵۳۴۵*
$Gamma(\frac{۱}{۲})$	۱۰	۰/۷۹۲۳	۰/۷۸۰۳	۰/۷۸۴۶	**۰/۸۰۱۱**	۰/۷۲۸۹
	۲۰	۰/۹۹۰۷	۰/۹۹۱۰	۰/۹۹۰۴	**۰/۹۹۱۳**	۰/۸۹۴۸
$Lognormal(۰, ۱)$	۲۰	۰/۵۷۴۹	۰/۵۶۹۸	۰/۵۶۶۶	**۰/۶۲۴۵**	۰/۶۰۶۲
	۲۰	۰/۹۲۳۸	۰/۹۱۸۱	۰/۹۱۴۵	**۰/۹۳۶۸**	۰/۹۲۹۹
$Lognormal(۰, ۲)$	۱۰	۰/۹۴۱۵	۰/۹۳۷۶	۰/۹۳۸۹	**۰/۹۴۵۹**	۰/۹۱۹۲
	۲۰	**۰/۹۹۹۸**	۰/۹۷۹۷	۰/۹۹۹۷	**۰/۹۹۹۸**	۰/۹۹۹۳
$Lognormal(۰, \frac{۱}{۲})$	۱۰	۰/۱۷۴۳	۰/۱۶۷۶	۰/۱۶۶۷	**۰/۲۱۵۰**	۰/۲۴۹۶*
	۲۰	۰/۴۰۲۸	۰/۳۹۳۲	۰/۳۷۸۲	**۰/۴۷۵۵**	۰/۵۲۴۲*
$Weibull(\frac{۱}{۲})$	۱۰	۰/۹۳۲۹	۰/۹۲۸۱	۰/۹۲۹۲	**۰/۹۳۲۹**	۰/۸۹۶۳
	۲۰	**۰/۹۹۹۶**	**۰/۹۹۹۹۶**	**۰/۹۹۹۶**	**۰/۹۹۹۶**	۰/۹۹۸۹
$Weibull(۲)$	۱۰	۰/۰۸۱۵	۰/۰۹۷۲	۰/۰۸۰۲	**۰/۰۸۱۵**	۰/۰۸۴۹*
	۲۰	۰/۱۲۴۸	۰/۱۲۸۵	۰/۱۲۶۸	**۰/۱۴۴۴**	۰/۱۵۰۴*

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

جدول ۹: توان آزمون‌های نرمال بودن بر مبنای آنتروپی و آزمون شاپیرو–ویلک تحت خانواده‌های توزیع‌های جانشین گروه IV و $\alpha = ۰/۰۵$.

Alternatives	n	TV	TC	$TZ_۱$	$TZ_۲$	SW
$Uniform(۰, ۱)$	۱۰	/۱۷۶۶	۰/۱۷۴۴	۰/**۱۸۱۰**	۰/۱۳۳۶	۰/۰۸۱۹
	۲۰	۰/۴۰۸۶	۰/۴۱۸۳	۰/**۴۳۱۵**	۰/۳۳۲۶	۰/۱۹۵۰
$Beta(۲, ۲)$	۱۰	۰/۰۸۱۲	۰/۰۸۴۵	۰/۰**۸۵۹**	۰/۰۶۴۳	۰/۰۴۱۹
	۲۰	۰/۱۲۸۸	۰/۱۳۳۲	۰/**۱۳۵۴**	۰/۰۹۸۹	۰/۰۵۲۸
$Beta(\frac{۱}{۲}, \frac{۱}{۲})$	۱۰	۰/۵۲۰۴	۰/۵۱۰۸	۰/**۵۲۴۶**	۰/۴۴۴۷	۰/۲۹۱۴
	۲۰	۰/۹۰۸۹	۰/۹۰۰۶	۰/**۹۱۲۹**	۰/۸۵۵۱	۰/۷۱۶۲
$Beta(۳, \frac{۱}{۲})$	۱۰	۰/۱۰۶۶	۰/۱۰۵۵	۰/**۱۰۶۸**	۰/۰۹۹۵	۰/۰۸۶۸
	۲۰	۰/۲۲۳۱	۰/۲۲۸۳	۰/**۲۲۶۹**	۰/۲۱۹۱	۰/۱۶۸۱
$Beta(۲, ۱)$	۱۰	۰/۱۷۷۹	۰/۱۷۸۲	۰/**۱۸۴۲**	۰/۱۶۹۱	۰/۱۲۸۹
	۲۰	۰/۴۲۳۳	۰/۴۲۹۲	۰/**۴۳۱۳**	۰/۴۱۱۷	۰/۳۰۳۷

است. اما در خانواده توزیع‌های گروه IV، از آزمون مبتنی بر $TZ۱$، توان بیشتری نسبت به سایر آزمون‌های بر مبنای آنتروپی دارد، اما اختلاف توان آن باتوان آزمون مبتنی بر TV قابل ملاحظه نیست. از طرفی مشاهده می‌شود که گرچه آزمون مبتنی بر $TZ۲$ در مقایسه با آزمون شاپیرو–ویلک (۱۹۶۵) در گروه‌های III و IV، به استثنای مواردی که با ستاره مشخص شده‌اند، از توان بیشتری برخوردار است ولی در گروه‌های I و II توان کمتری دارد(در مورد حجم نمونه‌های بررسی شده). این عدم کارایی از سه جهت جبران می‌شود، نخست اینکه انجام این آزمون بسیار ساده‌تر از آزمون شاپیرو–ویلک (۱۹۶۵) است، که به جداول زیادی نیاز دارد. دوم اینکه آزمون شاپیرو–ویلک (۱۹۶۵) به ازای حجم نمونه‌های بیش از ۲۰ تقریبی است و ضرایب لازم برای محاسبهٔ آماره آزمون برای هر حجم نمونه تغییر می‌کنند. سوم اینکه توان آزمون نرمال بودن توزیع جامعه،در گروه I (از توزیع‌های جانشین) خیلی اهمیت ندارد، زیرا اکثر روش‌های آماری مبتنی بر توزیع نرمال، نسبت به توزیع‌های متقارن بسیار استوار هستند و مادام که توزیع داده‌ها متقارن باشد، ولو اینکه نرمال نباشد، دقت روش‌های آماری از اعتبار ساقط نخواهد شد. ضمناً متذکر می‌شویم که در مورد بیش از نیمی از ۲۰ توزیع جانشین آزمون مبتنی بر $TZ۲$ تواناتر از آزمون شاپیرو–ویلک (۱۹۶۵) است.

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

۴ بحث و نتیجه‌گیری

در این مقاله دو برآوردگر جدید آنتروپی مبتنی بر اصلاح ضرایب و تخصیص وزن‌های متفاوت از برآوردگر کوریا (۱۹۹۵) ارایه گردید، سپس در مطالعه شبیه‌سازی عملکرد برآوردگرهای پیشنهادی با برآوردگرهای واسیچک (۱۹۷۶)، ابراهیمی و همکاران (۱۹۹۴) و کوریا (۱۹۹۵) مقایسه شد. نتایج بیانگر آن است که برآوردگر پیشنهادی اول به طور یکنواخت (یعنی به‌ازای هر سه توزیع بررسی شده و هرسه حجم نمونه ۱۰، ۲۰ و ۳۰ و به ازای $m \leq \frac{n}{۴}$) عملکرد بهتری نسبت به سایر برآوردگرهای آنتروپی دارد. مقایسه آزمون نیکویی برازش برای فرضیه‌های نرمال‌بودن و نمایی‌بودن توزیع جامعه بر مبنای برآوردگرهای جدید آنتروپی و برآوردگرهای واسیچک (۱۹۷۶) و کوریا (۱۹۹۵) و شاپیرو–ویلک (۱۹۶۵) در مورد نرمال، بیانگر آن است که در آزمون فرضیه نمایی‌بودن توزیع، توان آزمون مبتنی بر $TZ۲$ در بیشتر موارد از توان دیگر آزمون‌های مبتنی بر آنتروپی بیشتر است. در آزمون فرضیه نرمال بودن توزیع، تنها هنگامی که تکیه‌گاه توزیع متغیر تصادفی تحت فرضیه جانشین $(۰, ۱)$ است، آزمون مبتنی بر $TZ_۱$ دارای بیشترین توان و در غیر این‌صورت آزمون مبتنی بر $TZ_۲$ دارای ماکسیمم توان درمیان سایر آزمون‌های برمبنای آنتروپی است. توان آزمون نرمال بودن مبتنی بر $TZ۲$ حداقل به خوبی توان آزمون شاپیرو–ویلک (۱۹۶۵) می‌باشد. قدرمسلم این است که هیچ‌یک دیگری را تحت تسلط ندارد. ولی ساده‌تر بودن آزمون‌های برمبنای آنتروپی امتیازی بر آزمون‌های معرفی شده محسوب می‌شود.

تقدیر و تشکر

نویسندگان از پیشنهادات ارزندهٔ داوران و هیئت تحریریه محترم مجله که باعث اصلاحات سازنده در این مقاله شده کمال تشکر و سپاسگزاری را دارند. در ضمن از حمایت مالی قطب داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد نیز قدردانی و تشکر می‌شود.

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

مراجع

حبیبی‌راد، آ.، ارقامی، ن. ر.، (۱۳۸۶)، آزمون متقارن توزیع بر اساس آنتروپی، مجله علوم آماری، جلد **۱**، شماره ۲، ۱۲۰-۱۰۹.

علیزاده نوقابی، هـ، علیزاده نوقابی، ر.، (۱۳۸۷)، مقایسه توان آزمونهای نیکویی برازش بر مبنای آنتروپی با سایر روشها، مجله علوم آماری، جلد **۲**، شماره ۱، ۱۱۳-۹۷.

Ahmad, I. A. and Lin, P. E (1976), A Nonparametric Estimation of the Entropy of the Absolutely Continuous Distributions, *IEEE Transacation on Information Theory*, IT-22, 327-375.

Arizono, I. and Ohta, H. (1989), A Test for Normality Based on Kullback-Leibler Information, *The American Statisticain*, **43**, 20-23.

Balakrishnan, N., Habibi Rad, A. and Arghami, N. R. (2007), Testing Exponentiality Based on the Kullback-Leibler Information with Progressively Type-II Cencored Data, *IEEE Transacation on Reliability*, **56**, 301-307.

Corea, J. C., (1995), A New Estimator of Entropy, *Communications in Statistics-Theory and Methods*, **24**, 2439-2449.

Dudewicz, E. S. and Van der Meulen, E. C. (1981), Entropy-based Tests of Uniformity, *Journal of American Statistical Association*, **76**, 967-974.

Ebrahimi, N. and Habibullah, M.,(1992), Testing Exponentiality Based on Kullback-Leibler Information, *Journal of Royal Statistical Society*, Ser. B **54**, 739-748.

[DOR: 20.1001.1.17358183.1387.2.2.4.0]

Ebrahimi, N. ,Pflughoeft, K. and Soofi, E. S. (1994), Two Measures of Sample Entropy, *Probability and Statistics Letter*, **20**, 225-234.

Esteban, M. D. ,Castellanos, M. E., Morales, D. and Vajda I., (2001), Monte Carlo Comparison of Four Normality Tests Using Different Entropy Estimates, *Communications in Statistics-Simulation and computation*, **30**, 761-285.

Park, S. (2005), Testing Exponentiality Based on the Kullback-Leibler Information With the Type II Censored Data, *IEEE Transaction On Reliability*, **54**, 22-26.

Shanon, C. E.(1948), Mathematical Theory of Communications, *Bell Sysyem Technical Journal*, **27**, 379-423; 623-656.

Shapiro, S. S. and Wilk, M. B. (1965), An Analysis of Variance Test for Normality (Complete Sample), *Biometrika*, **52**, 591-611.

Mack, S. P. (1988), A Comparative Study of Entropy Estimators and Entropy-Based Goodness of-fit Test, *Ph.D. Dissertation, Univeristy of California, Riverside.*

Vasicek, O. (1976), A Test for Normality Based on Sample Entropy, *Jouranl of Royal Statistical Society*, Ser. B, **38**, 730-737.

Wieczorkowski, P., and Grzegorzewsky, P. (1999), Entropy Estimators Improvements and Comparisons, *Communication in Statistics-Computation and Simulation*, **28**, 541-567.

Yousefzadeh, F. and Arghami, N. R. (2008), Testing Exponentiality Based on Tyoe II Censored Data and a New cdf Estimator, *Communications in Statistics-Computation and Simulation*, **37**, 1479-1499.