مجله مدل‌سازی پیشرفته ریاضی
سال ۱۴۰۰، دوره ۱۱، شماره ۳، ص ۴۰۰-۴۱۴

شاپا چاپی: ۲۲۵۱-۸۰۸۸
DOI: 10.22055/JAMM.2021.35790.1877

تعمیم‌هایی از اطلاع کولبک-لیبلر بر اساس تابع بقا

هدیه افتخاری مودی، هادی علیزاده نوقابی [×]، محمد خراشادی‌زاده

گروه آمار، دانشگاه بیرجند، بیرجند، ایران

دبیر مسئول: غلامرضا محتشمی برزادران

تاریخ دریافت: ۹۹/۸/۲۸ تاریخ پذیرش: ۹۹/۱۲/۵

چکیده: در این مقاله ابتدا به برخی از تعمیم‌های اطلاع کولبک-لیبلر و خواص آن‌ها پرداخته می‌شود. سپس شرایط گشتاوری برای توزیع ماکسیمم آنتروپی مانده‌ی تجمعی بررسی و رابطه‌ی بین اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ و آنتروپی مانده‌ی تجمعی (CRE) مطالعه می‌شود. همچنین در مورد روش‌های برآورد پارامتر مقیاس توزیع رایلی بحث شده و دو روش برآورد آن ارائه می‌شود. در ادامه اطلاع کولبک-لیبلر مانده‌ی تجمعی را به عنوان یک آماره آزمون نیکویی برازش به کار برده و سپس مقادیر بحرانی و توان آزمون‌های پیشنهادی محاسبه و با توان سایر آزمون‌ها مقایسه می‌شود. در پایان، آزمون‌ها برای یک مجموعه داده‌ی واقعی به کار گرفته می‌شود

واژه‌های کلیدی: آزمون نیکویی برازش، آنتروپی مانده‌ی تجمعی، اطلاع کولبک-لیبلر، اطلاع کولبک-لیبلر مانده‌ی تجمعی، توان آزمون، شبیه سازی مونت کارلو.

رده‌بندی ریاضی: 62G10;62G20

۱ مقدمه

در قضاهای آمار و احتمال، توزیع رایلی یک توزیع احتمالی پیوسته برای متغیرهای تصادفی با مقادیر مثبت است. به طور مثال می‌توان سرعت باد را دارای این توزیع دانست.
اگر متغیر تصادفی X دارای تابع چگالی احتمال به صورت

$$f(x,\theta)=\frac{x}{\theta^{۲}}e^{\frac{-x^{۲}}{۲\theta^{۲}}} \qquad x\geq ۰,$$

باشد، گوییم X دارای توزیع رایلی با پارامتر θ است. همچنین تابع توزیع تجمعی آن به صورت

$$F(x,\theta)=۱-e^{\frac{-x^{۲}}{۲\theta^{۲}}} \qquad x\geq ۰,$$

[×] نویسنده مسئول مقاله
رایانامه: alizadehhadi@birjand.ac.ir (Hadi Alizadeh Noughabi),

است. در نظریه‌ی اطلاع مهمترین معیار اندازه گیری عدم حتمیت، آنتروپی شانون است که برای متغیرهای گسسته معرفی می‌شود[۲۳]. فرض کنید متغیر تصادفی X دارای تابع توزیع $F(x)$ و تابع چگالی پیوسته‌ی $f(x)$ باشد، تعمیم آنتروپی شانون به صورت

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

تعریف می‌شود. به دلیل تفاوت‌هایی که با تعمیم آنتروپی شانون به حالت پیوسته به وجود آمد تلاش‌های متعددی برای تعریف معیارهای جدیدی از اطلاع صورت گرفت. یکی از این معیارها، آنتروپی مانده‌ی تجمعی CRE † است که توسط رائو و همکاران[۲۱] با جایگزینی تابع بقا $(\bar{F})$ به جای تابع چگالی در آنتروپی شانون معرفی شد و به فرم

$$CRE(F) = -\int_{-\infty}^{+\infty} \bar{F}(x)\log \bar{F}(x)dx,$$

بیان شد. همچنین دیکرشنزو و لانگوباردی[۹] معیار دیگری تحت عنوان آنتروپی تجمعی CE ‡ ارائه دادند که با جایگزین کردن تابع توزیع به جای تابع چگالی در آنتروپی شانون حاصل می‌شود و به شکل

$$CE(F) = -\int_{\circ}^{+\infty} F(x)\log F(x)dx,$$

تعریف می‌شود. اطلاع کولبک-لیبلر (KL) § اندازه اطلاع اختلاف بین $g(x)$ و $f(x)$ است که به صورت

$$KL(g:f) = \int_{-\infty}^{+\infty} g(x)\log \frac{g(x)}{f(x)}dx,$$

تعریف می‌شود. از ویژگی‌های اطلاع کولبک-لیبلر می‌توان به نامنفی بودن آن اشاره کرد. همچنین این معیار برابر صفر می‌شود اگر و فقط اگر $g(x) = f(x)$ باشد.

اگر شرط

$$\int_{-\infty}^{+\infty} g(x)\log f(x)dx = \int_{-\infty}^{+\infty} f(x)\log f(x)dx, \qquad (۱.۱)$$

برقرار باشد، آنگاه می‌توان اطلاع کولبک-لیبلر را به صورت

$$KL(g:f) = H(f) - H(g), \qquad (۲.۱)$$

نوشت. رابطه‌ی (۲.۱) نشان می‌دهد که $H(f)$ بزرگتر از آنتروپی هر تابع چگالی $g(x)$ است که شرط (۱.۱) برای آن برقرار است. بنابراین $f(x)$ توزیع ماکسیمم آنتروپی (ME) در بین کلاس توزیع‌هایی است که شرط گشتاوری (۱.۱) برای آنها برقرار است.

اخیراً برخی از پژوهشگران توجه خود را روی موضوع آنتروپی تجمعی و اطلاع کولبک-لیبلر برای تابع بقا، معطوف کرده‌اند که از جمله آن‌ها می‌توان به [۱۷]، [۵]، [۱۹]، [۱۲] و [۱۱] اشاره کرد، که این دانشمندان تعمیم‌هایی را نیز در این زمینه پیشنهاد کرده‌اند. همچنین افرادی همچون پارک و پاکیاری[۲۰] و بالاکریشنان و همکاران[۴] اطلاع کولبک-لیبلر را برای داده‌های سانسور شده‌ی نوع II بررسی کردند. پارک و همکاران[۱۸] و سیومارا و پانایت[۸] ویژگی‌هایی از اطلاع کولبک-لیبلر تعمیم یافته را بیان نمودند و سپس آزمونی برای نرمال بودن ساختند. همچنین اطلاع کولبک-لیبلر بر اساس تابع چندک نیز توسط سونوج و همکاران[۲۴] بحث و بررسی شده است. در این مقاله متغیرهای تصادفی نامنفی که در آن $E(X) = \int_{\circ}^{\infty} \bar{F}(x)dx$ است، در نظر گرفته می‌شود.

براتپور و راد[۶] اخیراً تعمیمی از اطلاع کولبک-لیبلر را با استفاده از تابع بقا پیشنهاد داده‌اند که اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ ¶ نامیده می‌شود و به صورت

$$CRKL(G:F) = \int_{\circ}^{\infty} \bar{G}(x)\log \frac{\bar{G}(x)}{\bar{F}(x)}dx - (E(Y) - E(X)), \qquad (۳.۱)$$

† Cumulative Residual Entropy

‡ Cumulative Entropy

§ Kullback-Leibler Information

¶ Cumulative Residual Kullback-Leibler Information

قابل تعریف است. همچنین می‌توان مستقیماً تعمیم دیگری برای اطلاع کولبک لیبلر با استفاده از تابع توزیع تجمعی بیان کرد که اطلاع کولبک-لیبلر تجمعی CKL [۱] نامیده می‌شود و به شکل

$$CKL(G:F)=\int_{۰}^{\infty} G(x)\log\frac{G(x)}{F(x)}dx+E(Y)-E(X), \qquad (۴.۱)$$

بیان می‌شود.

در این مقاله، ابتدا تعمیم‌هایی از اطلاع کولبک-لیبلر بیان می‌شود. سپس رابطه‌ی بین آنتروپی مانده‌ی تجمعی (CRE) و اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ را بیان کرده و شرایط گشتاوری برای توزیع ماکسیمم آنتروپی مانده‌ی تجمعی (CRE) بررسی می‌شود. سپس دو روش برای برآورد پارامتر مقیاس توزیع رایلی بر اساس ماکسیمم آنتروپی مانده‌ی تجمعی (CRE) و مینیمم اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ معرفی می‌شود. در ادامه اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ به عنوان یک آماره آزمون نیکویی برازش برای توزیع رایلی به کار برده شده و توان این آزمون محاسبه و با توان سایر آزمون‌ها مورد مقایسه قرار می‌گیرد. در پایان آماره آزمون‌ها برای یک مجموعه داده‌ی واقعی به کار گرفته می‌شود.

۲ برخی از تعمیم های اطلاع کولبک-لیبلر

در این بخش ابتدا تعمیمی از اطلاع کولبک-لیبلر بر اساس تابع بقا که توسط پارک و همکاران[۱۹] ارائه شده است معرفی می‌شود. از آنجایی که تعمیم مستقیم تابع بقا به صورت

$$\int_{۰}^{\infty}\bar{G}(x)\log\frac{G\bar{(}x)}{F\bar{(}x)}dx,$$

خاصیت نامنفی بودن را حفظ نمی‌کند، لذا از عبارت $\frac{\bar{G}(x)}{E(Y)}$ و $\frac{\bar{F}(x)}{E(X)}$، با این شرط که $\int_{۰}^{\infty}\frac{\bar{G}(x)}{E(Y)}dx=\int_{۰}^{\infty}\frac{\bar{F}(x)}{E(X)}dx=۱$ باشد استفاده می‌شود. اطلاع کولبک-لیبلر $\frac{\bar{F}(x)}{E(X)}$ و $\frac{\bar{G}(x)}{E(Y)}$ را می‌توان به صورت

$$\begin{aligned}KL\left(\frac{\bar{G}(x)}{E(Y)}:\frac{\bar{F}(x)}{E(X)}\right)&=\int_{۰}^{\infty}\frac{\bar{G}(x)}{E(Y)}\log\frac{\frac{\bar{G}(x)}{E(Y)}}{\frac{\bar{F}(x)}{E(X)}}dx\\&=\frac{۱}{E(Y)}\int_{۰}^{\infty}\bar{G}(x)\log\frac{\bar{G}(x)}{\bar{F}(x)}dx-\log\frac{E(Y)}{E(X)},\end{aligned}$$

نوشت. اگر عبارت بالا در $E(Y)$ ضرب شود آنگاه عبارت به دست آمده با رابطه‌ی (۳.۱) قابل مقایسه است که پارک و همکاران[۱۹] نام آن را $CRKL^{۱}(G:F)$ نهادند. مشابه $CRKL(G:F)$، $CRKL^{۱}(G:F)$ نیز نامنفی است و برابر صفر می‌شود اگر $F(x)=G(x)$ باشد.

براتپور و راد[۶] اخیراً حالت دیگری از اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ ارائه دادند که به شکل

$$CRKL(G:F)=\int_{۰}^{\infty}\bar{G}(x)\left(\frac{\bar{F}(x)}{\bar{G}(x)}-\log\frac{\bar{F}(x)}{\bar{G}(x)}-۱\right)dx, \qquad (۱.۲)$$

است و با تعریف ارائه شده در رابطه‌ی (۳.۱) معادل است.

از آنجایی که نامساوی $u-\log u-۱\geq ۰$ همواره برقرار است، لذا با توجه به این رابطه می‌توان دید که اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ نامنفی است و برابر صفر می‌شود اگر و فقط اگر $\bar{F}(x)=\bar{G}(x)$ باشد. همچنین از آنجا که می‌توان نشان داد نامساوی

$$E(Y)\log\frac{E(Y)}{E(X)}\geq E(Y)-E(X),$$

همواره برقرار است، لذا اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ حساس تر از $CRKL^{۱}$ است و مقیاس فاصله ای مناسب تری است.

[۱] Cumulative Kullback-Leibler Information

با توجه به رابطه‌ی (۱.۲) ، پارک و همکاران[۱۹] مستقیماً تعمیم زیر را برای تابع توزیع تجمعی ارائه دادند که اطلاع کولبک–لیبلر تجمعی نامیده می‌شود.

$$CKL(G:F)=\int_{۰}^{\infty}G(x)\left(\frac{F(x)}{G(x)}-\log\frac{F(x)}{G(x)}-۱\right)dx,$$

که با تعریف ارائه شده در رابطه‌ی (۴.۱) معادل است. همچنین می‌توانیم ببینیم که اطلاع کولبک–لیبلر تجمعی (CKL) نامنفی است و برابر صفر می‌شود اگر و فقط اگر $G(x)=F(x)$ باشد.

کولبک[۱۴] اطلاع کولبک–لیبلر را بر اساس اطلاع فیشر به صورت

$$KL(f(x;\theta);f(x;\theta+\Delta\theta))=\int_{-\infty}^{\infty}(f(x;\theta)\log\frac{(f(x;\theta)}{f(x;\theta+\Delta\theta)}dx\approx\frac{۱}{۲}(\Delta\theta)^{۲}I(\theta),$$

تقریب زد که $I(\theta)$ اطلاع فیشر است.

به همین ترتیب، پارک و همکاران[۱۹] رابطه‌ی مشابهی برای اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ ارائه کردند که به شکل

$$CRKL(F(x;\theta);F(x;\theta+\Delta\theta))\approx\frac{۱}{۲}(\Delta\theta)^{۲}I_{\bar{F}}(\theta), \tag{۲.۲}$$

بیان می‌شود و $I_{\bar{F}}(\theta)$ نیز به صورت

$$I_{\bar{F}}(\theta)=-\int_{۰}^{\infty}\bar{F}(x;\theta)\frac{\partial^{۲}}{\partial\theta^{۲}}\log\bar{F}(x;\theta)dx.$$

تعریف می‌شود. مشابه روابط فوق، رابطه‌ی

$$CKL(F(x;\theta);F(x;\theta+\Delta\theta))\approx\frac{۱}{۲}(\Delta\theta)^{۲}I_{F}(\theta), \tag{۳.۲}$$

برای اطلاع کولبک–لیبلر تجمعی (CKL) توسط پارک و همکاران[۱۹] ارائه شده است که $I_F(\theta)$ به صورت

$$I_{F}(\theta)=-\int_{۰}^{\infty}F(x;\theta)\frac{\partial^{۲}}{\partial\theta^{۲}}\log F(x;\theta)dx,$$

تعریف می‌شود. همچنین پارک و همکاران[۱۹] برای $CKRL^{۱}$ رابطه‌ی

$$CRKL^{۱}(F(x;\theta);F(x;\theta+\Delta\theta))\approx\frac{۱}{۲}(\Delta\theta)^{۲}\{I_{\bar{F}}(\theta)+E(X)\frac{\partial^{۲}}{\partial\theta^{۲}}\log E(X)\}, \tag{۴.۲}$$

را نشان دادند. با توجه به روابط (۲.۲)، (۳.۲) و (۴.۲) هرگاه هدف انتخاب یکی از معیارهای CKL، $CRKL$ و $CRKL^{۱}$ باشد، کافیست به ترتیب $I_F(\theta)$، $I_{\bar{F}}(\theta)$ و $I_{\bar{F}}(\theta)+E(X)\frac{\partial^{۲}}{\partial\theta^{۲}}\log E(X)$ با هم مقایسه شوند.

۳ رابطه بین اطلاع کولبک–لیبلر مانده‌ی تجمعی و آنتروپی مانده‌ی تجمعی

همان‌طور که در بخش ۲ در روابط (۱.۱) و (۲.۱) بیان شد، طبق شرط

$$\int_{-\infty}^{+\infty}g(x)\log f(x)dx=\int_{-\infty}^{+\infty}f(x)\log f(x)dx,$$

می‌توان اطلاع کولبک–لیبلر را به صورت

$$KL(g:f)=H(f)-H(g),$$

نوشت. مطابق رابطه‌ی فوق می‌توان رابطه بین اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ و آنتروپی مانده‌ی تجمعی (CRE) را به شرح زیر بررسی کرد. اگر شرط

$$\int_{۰}^{\infty}\bar{G}(x)\log\bar{F}(x)dx+E(Y)=\int_{۰}^{\infty}\bar{F}(x)\log\bar{F}(x)dx+E(X),$$

برقرار باشد، آنگاه

$$CRKL(G;F)=CRE(F)-CRE(G). \tag{۱.۳}$$

مثال ۱.۳. فرض کنید $F(x) = ۱ - \exp(\frac{-x^۲}{۲\theta^۲})$ آنگاه با توجه به رابطه‌ی (۱.۳) داریم،

$$\frac{-۱}{۶\theta^۲}E_G(X^۳) + E_G(X) = \frac{-۱}{۶\theta^۲}E_F(X^۳) + E_F(X). \tag{۲.۳}$$

رابطه‌ی (۲.۳) به ازای هر θ برقرار است اگر $E_G(X^۳) = E_F(X^۳)$ و $E_G(X) = E_F(X)$ باشد. توزیع رایلی با پارامتر θ، آنتروپی مانده‌ی تجمعی (CRE) را با شرط $E(X^۳) - ۶\theta^۲E(X) + ۳\theta^۳\sqrt{\frac{\pi}{۲}} = ۰$ در بین توزیع های مطلقاً پیوسته با تکیه گاه $(۰, \infty)$، ماکسیمم می‌کند. بنابراین متغیر تصادفی X، آنتروپی مانده‌ی تجمعی (CRE) را در بین تمام متغیرهای تصادفی مطلقاً پیوسته‌ی نامنفی Y، ماکسیمم می‌کند با این شرط که اگر $E(Y) = v$ و $E(Y^۳) = w$ باشد، داشته باشیم $\theta^۲ = \frac{w}{۳v}$.

فرض کنید $G(x)$ و $F(x)$ به ترتیب تابع توزیع تجربی $(F_n(x))$ و یک تابع توزیع پارامتری $(F_\theta(x))$ باشند، آنتروپی مانده‌ی تجمعی بر اساس این دو تابع به صورت

$$CRE(F_n) = -\sum_{i=۰}^{n} \frac{n-i}{n} \log \frac{n-i}{n}(x_{i+۱} - x_i), \tag{۳.۳}$$

به دست می‌آید. همچنین اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ بین F_n و F_θ به صورت

$$CRKL(F_n, F_\theta) = -CRE(F_n) - \sum_{i=۰}^{n} \frac{n-i}{n} \int_{x_i}^{x_{i+۱}} \log \bar{F}_\theta(x)dx - (\bar{x} - E_{F_\theta}(X)), \tag{۴.۳}$$

است. رابطه‌ی (۴.۳) می‌تواند با انتخاب یک برآورگر پارامتری مناسب برای θ، برآورد شود. برای برآورد پارامتر دو روش زیر در نظر گرفته شده است:

- شرایط گشتاوری برای ماکسیمم آنتروپی مانده‌ی تجمعی (CRE) برقرار باشد.
- مینیمم کردن اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$

برای مینیمم کردن $(CRKL)$ می‌توان به کمک روش مینیمم کردن اطلاع تشخیص MDI [xx] از رابطه‌ی

$$\begin{aligned}\hat{\theta} &= \underset{\theta}{arg\min}\, CRKL(F_n, F_\theta)\\ &= \underset{\theta}{arg\max}\{\sum_{i=۰}^{n} \frac{n-i}{n} \int_{x_i}^{x_{i+۱}} \log \bar{F}_\theta(x)dx - E_{F_\theta(X)}\},\end{aligned}$$

استفاده کرد.

مثال ۲.۳. برای توزیع رایلی با تابع بقا به شکل $\bar{F}(x) = e^{\frac{-x^۲}{۲\theta^۲}}$، $\hat{\theta}$ به دست آمده از روش ماکسیمم آنتروپی مانده‌ی تجمعی (CRE) برابر

$$\hat{\theta}_m = \left(\frac{\sum_{i=۱}^{n} {x_i}^۳}{۳\sum_{i=۱}^{n} x_i}\right)^{\frac{۱}{۲}},$$

است.

از آنجا که اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ بین F_n و F_θ برای توزیع رایلی طبق رابطه‌ی (۴.۳) به صورت

$$CRKL(F_n, F_\theta) = \frac{۱}{۶n\theta^۲}\sum_{i=۱}^{n} {x_i}^۳ - \left(\bar{x} - \theta\sqrt{\frac{\pi}{۲}}\right) - CRE(F_n),$$

به دست می‌آید. لذا $\hat{\theta}$ به دست آمده از روش مینیمم کردن اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ برابر

$$\hat{\theta}_{MDI} = \left(\frac{\sum_{i=۱}^{n} {x_i}^۳}{۳n\sqrt{\frac{\pi}{۲}}}\right)^{\frac{۱}{۳}},$$

است.

[xx] Minimum Discrimination Information

توجه داریم که برآوردگر گشتاوری $\hat{\theta}_m$ با برآوردگر مینیمم کننده‌ی اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ متفاوت است و داریم:

$$CRKL(F_n, F_{\hat{\theta}_{MDI}}) \leq CRKL(F_n, F_{\hat{\theta}_m}).$$

۴ کاربرد اطلاع کولبک–لیبلر مانده‌ی تجمعی برای آزمون نیکویی برازش توزیع رایلی

پارک و همکاران[۱۹] و زهره‌وند و همکاران[۲۶] براساس اطلاع کولبک–لیبلر مانده‌ی تجمعی آزمون‌هایی برای نیکویی برازش توزیع نمایی ارائه دادند و توان آزمون‌های پیشنهادی را با دیگر آزمون‌ها مقایسه کردند. در این بخش ما قصد داریم آزمون‌های مشابهی برای توزیع رایلی به دست آوریم.
از آنجایی که اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ به پارامتر مقیاس وابسته است لذا اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ نمی‌تواند به عنوان یک آماره آزمون نیکویی برازش استفاده شود. بنابراین باید آماره آزمون را به گونه ای در نظر گرفت که در مقیاس پایا باشد. می‌توان اطلاع کولبک–لیبلر مانده‌ی تجمعی $(CRKL)$ برای توزیع رایلی را با تقسیم بر پارامتر مقیاسی برآورد شده به صورت

$$\begin{aligned} T_۱ &= \frac{۱}{\hat{\theta}_۱} CRKL(F_n, F_{\hat{\theta}_۱}) \\ &= \frac{۱}{\sqrt{\frac{\sum_{i=۱}^n x^۲}{۲n}}} \left(\frac{\sum_{i=۱}^n {x_i}^۳}{۳\sum_{i=۱}^n {x_i}^۲} - \bar{x} + (\sqrt{\frac{\sum_{i=۱}^n {x_i}^۲}{۲n}} \sqrt{\frac{\pi}{۲}}) - CRE(F_n) \right), \end{aligned} \tag{۱.۴}$$

نوشت. اما رابطه ی (۱.۴) در واقع برآورد $\frac{CRKL(F_n, F_\theta)}{\theta}$ است که θ یک پارامتر مقیاسی است، لذا بهتر است برآوردگری را در نظر بگیریم که $\frac{CRKL(F_n, F_\theta)}{\theta}$ را مینیمم کند. برآوردگر مینیمم کننده برای توزیع رایلی به صورت

$$\hat{\theta} = \left(\frac{\sum_{i=۱}^n {x_i}^۳}{۲n(\bar{x} + CRE(F_n))} \right)^{\frac{۱}{۲}},$$

حاصل می‌شود و آماره آزمونی که براساس برآوردگر فوق به دست می‌آید برابر

$$T_۲ = \frac{-۲(\overline{x} + CRE(F_n))}{۳\left(\frac{\sum_{i=۱}^n x_i^۳}{۲n(\overline{x} + CRE(F_n))}\right)^{\frac{۱}{۲}}} + \sqrt{\frac{\pi}{۲}}, \tag{۲.۴}$$

است.
با توجه به اینکه توزیع دقیق آماره آزمون‌های پیشنهاد شده به راحتی به دست نمی‌آیند، مقادیر بحرانی آماره آزمون‌ها را به کمک شبیه سازی مونت کارلو با ۵۰۰۰۰ تکرار به دست آورده‌ایم. برای این منظور نمونه‌ای به حجم n از توزیع رایلی با پارامتر $\theta = ۱$ در سطح $\alpha = ۰/۰۱, ۰/۰۵, ۰/۱$ تولید کرده و چندک مرتبه $\alpha = ۰/۹۹, ۰/۹۵, ۰/۹۰$ام آماره آزمون‌ها را محاسبه می‌کنیم. آماره آزمون‌ها نسبت به تبدیلات مقیاسی پایا هستند و بنابراین مقادیر بحرانی به مقدار پارامتر θ بستگی ندارند. جدول ۱ مقادیر بحرانی آماره‌های $T_۱$ و $T_۲$، را به ازای حجم نمونه n و سطوح معناداری ذکر شده، نشان می‌دهد.

جدول ۱: مقادیر بحرانی آماره‌های T_1 و T_2

	T_1			T_2		
	α			α		
n	۰/۰۱	۰/۰۵	۰/۱	۰/۰۱	۰/۰۵	۰/۱
۵	۰/۲۱۲	۰/۱۶۳	۰/۱۴۲	۰/۲۰۰	۰/۱۵۸	۰/۱۳۷
۱۰	۰/۱۴۵	۰/۱۰۴	۰/۰۸۶	۰/۱۳۰	۰/۰۹۸	۰/۰۸۳
۱۵	۰/۱۱۲	۰/۰۷۷	۰/۰۶۴	۰/۱۰۰	۰/۰۷۲	۰/۰۶۱
۲۰	۰/۰۹۲	۰/۰۶۲	۰/۰۵۱	۰/۰۸۳	۰/۰۵۸	۰/۰۴۹
۲۵	۰/۰۷۹	۰/۰۵۲	۰/۰۴۳	۰/۰۷۲	۰/۰۴۹	۰/۰۴۱
۳۰	۰/۰۶۹	۰/۰۴۶	۰/۰۳۷	۰/۰۶۲	۰/۰۴۳	۰/۰۳۶
۴۰	۰/۰۵۶	۰/۰۳۶	۰/۰۳۰	۰/۰۵۰	۰/۰۳۴	۰/۰۲۸
۵۰	۰/۰۴۶	۰/۰۳۰	۰/۰۲۵	۰/۰۴۲	۰/۰۲۹	۰/۰۲۴
۶۰	۰/۰۴۰	۰/۰۲۶	۰/۰۲۱	۰/۰۳۶	۰/۰۲۵	۰/۰۲۰
۷۰	۰/۰۳۶	۰/۰۲۳	۰/۰۱۹	۰/۰۳۳	۰/۰۲۲	۰/۰۱۸
۸۰	۰/۰۳۲	۰/۰۲۱	۰/۰۱۷	۰/۰۲۹	۰/۰۱۹	۰/۰۱۶
۹۰	۰/۰۲۹	۰/۰۱۹	۰/۰۱۵	۰/۰۲۷	۰/۰۱۸	۰/۰۱۴
۱۰۰	۰/۰۲۷	۰/۰۱۷	۰/۰۱۴	۰/۰۲۵	۰/۰۱۶	۰/۰۱۳

در ادامه توزیع‌های زیر را به عنوان توزیع‌های جانشین در نظر گرفته و توان آزمون‌ها محاسبه شده است:

- توزیع وایبل $W(\theta)$ با تابع چگالی

$$f(x;\theta) = \theta x^{\theta-1}\exp\{-x^{\theta}\}, \qquad \theta > 0, x \geq 0.$$

- توزیع گاما $\Gamma(\theta)$ با تابع چگالی

$$f(x;\theta) = \frac{x^{\theta-1}\exp\{-x\}}{\Gamma(\theta)}, \qquad \theta > 0, x \geq 0.$$

- توزیع لگ-نرمال $LN(\theta)$ با تابع چگالی

$$f(x;\theta) = \frac{1}{\theta x\sqrt{2\pi}}\exp\left\{-\frac{1}{2\theta^2}(\log x)^2\right\}, \qquad \theta > 0, x > 0.$$

- توزیع نیم‌نرمال HN با تابع چگالی

$$f(x) = \Gamma(\frac{2}{\pi})^{1/2}\exp\left\{-\frac{x^2}{2}\right\}, \qquad x \geq 0.$$

- توزیع یکنواخت U با تابع چگالی

$$f(x) = 1, \qquad 0 \leq x \leq 1.$$

- توزیع مقادیر غایی تعدیل یافته $EV(\theta)$ با تابع چگالی

$$f(x;\theta) = \frac{1}{\theta}\exp\left\{x + \frac{1}{\theta}(1-e^{x})\right\}, \qquad \theta > 0, x \geq 0.$$

- توزیع نرخ خطر از کار افتادگی صعودی خطی $LF(\theta)$ با تابع چگالی

$$f(x;\theta) = (1+\theta x)\left(-x - \frac{\theta x^2}{2}\right), \qquad \theta > 0, x \geq 0.$$

• توزیع دهیلون $DL(\theta)$ با تابع چگالی

$$f(x;\theta) = \frac{\theta+1}{x+1}(\log(x+1))^{\theta}\exp\{-(\log(x+1))^{\theta+1}\}, \qquad \theta \geq 0, x \geq 0.$$

• توزیع چن $CH(\theta)$ با تابع چگالی

$$f(x;\theta) = 2\theta x^{\theta-1}\exp\{x^{\theta} + 2(1-e^{x^{\theta}})\}, \qquad \theta \geq 0, x \geq 0.$$

• توزیع بتا $B(\alpha,\beta)$ با تابع چگالی

$$f(x;\alpha,\beta) = \frac{1}{\beta(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \qquad \alpha > 0, \beta > 0, 0 < x < 1.$$

با استفاده از شبیه سازی مونت کارلو و تعداد ۵۰۰۰۰ بار تکرار، توان آزمون‌های رایلی بودن را تحت توزیع‌های جانشین متفاوت به دست می‌آوریم. برای این منظور ابتدا نمونه‌هایی به حجم‌های $n = 10, 20, 30$ تحت هریک از توزیع‌های جانشین تولید کرده و سپس توان هریک از این آزمون‌ها را با استفاده از نسبت تعداد دفعاتی که آماره آزمون‌ها از مقادیر بحرانی‌شان بیشتر هستند به کل تعداد تکرارها، برآورد می‌کنیم. مقادیر توان برای آماره‌های T_1، T_2 و توزیع‌های جانشین در جدول‌های ۲، ۳ و ۴ ارائه شده است. همچنین توان آزمون‌های رقیب در این جدول‌ها نشان داده شده‌اند که در ادامه به آماره‌های آنها اشاره شده است:

• آماره آزمون کرامر-ون میسز

$$W^2 = \sum_{i=1}^{n}(Z_{(i)} - \frac{2i-1}{2n})^2 + \frac{1}{12n}.$$

• آماره آزمون کولموگروف [۱۳]

$$D^+ = \max_{1\leq i\leq n}\left\{\frac{i}{n} - Z_{(i)}\right\}, \quad D^- = \max_{1\leq i\leq n}\left\{Z_{(i)} - \frac{i-1}{n}\right\},$$

$$D = max(D^+, D^-).$$

• آماره آزمون اندرسون-دارلینگ [۲]

$$A^2 = -\frac{1}{n}\sum_{i=1}^{n}((2i-1)\{\log(Z_{(i)}) + \log(1-Z_{(i)})\}) - n.$$

• آماره آزمون واتسون [۲۵]

$$U^2 = W^2 - n\left(\overline{Z} - \frac{1}{2}\right)^2,$$

که W^2 آماره آزمون کرامر-وان میسز است.

• آماره آزمون کوپر [۱۵]

$$D^+ = \max_{1\leq i\leq n}\left\{\frac{i}{n} - Z_{(i)}\right\}, \quad D^- = \max_{1\leq i\leq n}\left\{Z_{(i)} - \frac{i-1}{n}\right\},$$

$$V = D^+ + D^-.$$

در اینجا $\overline{Z} = (1/n)\sum_{i=1}^{n} Z_{(i)}$ و $Z_{(i)} = F(x_{(i)})$ بوده و $F(.)$ تابع توزیع رایلی است.

- آماره آزمون پیشنهاد شده توسط منتانیس و ایلیئوپولوس[۱۶] بر اساس تبدیل لاپلاس:

$$L = \frac{n}{b} + \frac{\sqrt{۲}}{n} \sum_{j=۱}^{n} \sum_{k=۱}^{n} \left\{ \frac{۱}{(\hat{Y}_j + \hat{Y}_k + b)} + \frac{(\hat{Y}_j + \hat{Y}_k)}{(\hat{Y}_j + \hat{Y}_k + b)^{۲}} + \frac{۲(\hat{Y}_j \hat{Y}_k + ۲)}{(\hat{Y}_j + \hat{Y}_k + b)^{۳}} + \frac{۶(\hat{Y}_j + \hat{Y}_k)}{(\hat{Y}_j + \hat{Y}_k + b)^{۴}} + \frac{۲۴}{(\hat{Y}_j + \hat{Y}_k + b)^{۵}} \right\} - ۲\sqrt{۲} \sum_{j=۱}^{n} \left\{ \frac{۱}{(Y_j \hat{+} b)} + \frac{\hat{Y}_j}{(Y_j \hat{+} b)^{۲}} + \frac{۲}{(Y_j \hat{+} b)^{۳}} \right\}$$

که $b = ۲\sqrt{۲}$، $\hat{Y}_j = X_j/\hat{\theta}$ و $\hat{\theta}$ برآوردگر درستنمایی ماکسیمم θ است.

- آماره آزمون ارائه شده توسط براتپور و خدادادی[۵] بر اساس آنتروپی مانده‌ی تجمعی:

$$CK_n = \frac{\sum_{i=۱}^{n-۱} ((n-i)/n)(ln((n-i)/n))\{X_{(i+۱)} - X_{(i)}\} + \sqrt{\pi/۲}\sqrt{\sum_{i=۱}^{n} X_i^{۳}/۳ \sum_{i=۱}^{n} X_i}}{\overline{X}}.$$

- آماره آزمون پیشنهاد شده توسط علیزاده و همکاران[۱] بر اساس اطلاع کولبک-لیبلر:

$$KL_{mn} = -H_{mn} + ۲\log(\hat{\theta}) - \frac{۱}{n}\sum_{i=۱}^{n} \log(X_i) + ۱,$$

که $\hat{\theta}$ برآوردگر درستنمایی ماکسیمم θ بوده و H_{mn} به صورت

$$H_{mn} = \frac{۱}{n}\sum_{i=۱}^{n} \log\left\{\frac{n}{۲m}(X_{(i+m)} - X_{(i-m)})\right\} \qquad m < n/۲,$$

است.

جدول ۲: مقایسه‌ی توان آزمون‌ها برای اندازه نمونه‌ی $n = ۱۰$ و سطح معنی داری $\alpha = ۰/۰۵$

D	V	$W^{۲}$	$U^{۲}$	$A^{۲}$	L	CK_n	KL_{mn}	$T_۱$	$T_۲$	توزیع جانشین
۰/۷۷۴	۰/۷۰۶	۰/۸۱۲	۰/۷۰۶	۰/۹۲۵	۰/۹۲۵	۰/۸۳۶	۰/۷۶۴	۰/۷۶۳	۰/۷۲۲	W(۰/۸)
۰/۱۹۴	۰/۱۵۸	۰/۲۲۴	۰/۱۵۶	۰/۳۲۳	۰/۳۴۲	۰/۲۷۴	۰/۱۰۱	۰/۲۱۴	۰/۱۷۸	W(۱/۴)
۰/۹۳۰	۰/۸۹۸	۰/۹۴۹	۰/۹۰۱	۰/۹۹۲	۰/۹۹۱	۰/۹۵۵	۰/۹۶۷	۰/۹۱۹	۰/۸۹۹	Γ(۰/۴)
۰/۵۵۷	۰/۴۷۸	۰/۶۱۳	۰/۴۹۹	۰/۷۶۵	۰/۷۸۱	۰/۶۵۴	۰/۴۷۴	۰/۵۵۹	۰/۵۱۰	Γ(۱)
۰/۱۸۱	۰/۱۴۰	۰/۲۰۹	۰/۱۴۱	۰/۲۶۵	۰/۲۶۰	۰/۲۶۵	۰/۰۷۳	۰/۲۰۷	۰/۱۷۵	Γ(۲)
۰/۳۸۰	۰/۳۱۳	۰/۴۲۳	۰/۳۱۴	۰/۴۵۳	۰/۴۵۱	۰/۴۶۱	۰/۲۱۵	۰/۴۲۷	۰/۳۸۹	LN(۰/۸)
۰/۸۹۰	۰/۸۴۷	۰/۹۱۰	۰/۸۵۱	۰/۹۵۶	۰/۹۶۴	۰/۹۲۰	۰/۸۵۵	۰/۸۹۱	۰/۸۶۹	LN(۱/۵)
۰/۲۴۴	۰/۲۱۵	۰/۲۸۲	۰/۲۲۶	۰/۴۶۶	۰/۴۷۷	۰/۳۲۰	۰/۲۰۶	۰/۲۴۳	۰/۲۰۳	HN
۰/۱۰۳	۰/۱۶۳	۰/۱۱۹	۰/۱۵۲	۰/۲۶۸	۰/۱۹۷	۰/۱۲۳	۰/۱۷۲	۰/۰۹۱	۰/۱۰۰	U
۰/۹۰۷	۰/۸۶۵	۰/۹۲۹	۰/۸۶۶	۰/۹۸۴	۰/۹۸۴	۰/۹۳۳	۰/۹۴۱	۰/۸۸۷	۰/۸۶۱	CH(۰/۵)
۰/۳۱۲	۰/۲۶۴	۰/۳۵۳	۰/۲۷۵	۰/۵۴۵	۰/۵۵۶	۰/۳۷۱	۰/۲۵۲	۰/۲۸۵	۰/۲۴۰	CH(۱)
۰/۰۵۴	۰/۰۶۴	۰/۰۵۹	۰/۰۶۰	۰/۰۹۹	۰/۰۹۹	۰/۰۵۷	۰/۰۴۱	۰/۰۴۹	۰/۰۴۷	CH(۱/۵)
۰/۲۰۵	۰/۱۷۷	۰/۲۳۷	۰/۱۸۸	۰/۴۰۲	۰/۴۰۹	۰/۲۸۳	۰/۱۶۱	۰/۱۹۸	۰/۱۶۵	LF(۲)
۰/۱۴۲	۰/۱۳۳	۰/۱۶۷	۰/۱۳۴	۰/۳۰۲	۰/۳۰۸	۰/۱۹۵	۰/۱۱۳	۰/۱۴۱	۰/۱۱۶	LF(۴)
۰/۳۰۹	۰/۲۷۱	۰/۳۵۰	۰/۲۷۵	۰/۵۴۲	۰/۵۴۷	۰/۳۸۵	۰/۲۴۸	۰/۲۸۴	۰/۲۳۶	EV(۰/۵)
۰/۱۳۲	۰/۱۳۸	۰/۱۵۰	۰/۱۳۴	۰/۳۱۰	۰/۲۹۷	۰/۱۵۲	۰/۱۲۶	۰/۱۰۴	۰/۰۸۵	EV(۱/۵)
۰/۳۶۳	۰/۲۹۵	۰/۴۱۰	۰/۲۹۶	۰/۴۸۲	۰/۴۸۶	۰/۴۵۵	۰/۲۰۲	۰/۴۰۳	۰/۳۶۰	DL(۱)
۰/۱۴۴	۰/۱۰۸	۰/۱۶۵	۰/۱۱۰	۰/۱۸۵	۰/۱۸۱	۰/۲۰۶	۰/۰۶۰	۰/۱۸۲	۰/۱۵۷	DL(۱/۵)
۰/۲۸۲	۰/۳۵۵	۰/۳۶۳	۰/۳۴۲	۰/۳۰۷	۰/۲۵۱	۰/۳۹۶	۰/۴۵۰	۰/۴۶۵	۰/۵۰۶	B(۲, ۱)
۰/۳۳۹	۰/۴۸۴	۰/۳۹۵	۰/۴۹۹	۰/۷۴۹	۰/۵۷۹	۰/۴۵۲	۰/۶۷۹	۰/۳۵۰	۰/۳۶۳	B(۰/۵, ۰/۵)
۰/۰۸۱	۰/۰۹۱	۰/۰۹۲	۰/۰۹۳	۰/۰۷۵	۰/۰۷۰	۰/۰۶۵	۰/۱۰۷	۰/۰۹۰	۰/۱۰۷	B(۲, ۲)
۰/۲۹۰	۰/۳۱۵	۰/۳۶۴	۰/۳۲۴	۰/۳۰۱	۰/۲۸۳	۰/۲۸۵	۰/۴۱۴	۰/۳۵۹	۰/۳۹۴	B(۳, ۲)
۰/۶۶۷	۰/۸۲۵	۰/۸۰۴	۰/۷۹۶	۰/۷۵۹	۰/۵۵۸	۰/۸۹۷	۰/۹۱۵	۰/۹۲۲	۰/۹۳۶	B(۲, ۰/۵)

جدول ۳: مقایسه‌ی توان آزمون‌ها برای اندازه نمونه‌ی $n = ۲۰$ و سطح معنی داری $\alpha = ۰/۰۵$

D	V	$W^{۲}$	$U^{۲}$	$A^{۲}$	L	CK_n	KL_{mn}	$T_۱$	$T_۲$	توزیع جانشین
۰/۹۷۶	۰/۹۴۵	۰/۹۸۵	۰/۹۵۱	۰/۹۹۷	۰/۹۹۷	۰/۹۷۹	۰/۹۷۶	۰/۹۶۱	۰/۹۵۰	W(۰/۸)
۰/۳۶۸	۰/۲۷۴	۰/۴۱۷	۰/۲۹۲	۰/۵۲۴	۰/۵۶۵	۰/۴۲۷	۰/۲۴۱	۰/۳۷۳	۰/۳۳۳	W(۱/۴)
۰/۹۹۹	۰/۹۹۵	۰/۹۹۹	۰/۹۹۶	۱/۰۰۰	۱/۰۰۰	۰/۹۹۸	۱/۰۰۰	۰/۹۹۷	۰/۹۹۶	Γ(۰/۴)
۰/۸۶۷	۰/۷۷۶	۰/۸۹۶	۰/۷۹۵	۰/۹۵۱	۰/۹۶۲	۰/۸۸۷	۰/۸۳۷	۰/۸۳۷	۰/۸۰۳	Γ(۱)
۰/۳۲۴	۰/۲۳۰	۰/۳۶۷	۰/۲۵۰	۰/۴۲۱	۰/۴۵۲	۰/۴۰۴	۰/۱۷۲	۰/۳۶۸	۰/۳۳۱	Γ(۲)
۰/۶۶۴	۰/۵۴۵	۰/۷۰۲	۰/۵۷۴	۰/۷۱۷	۰/۷۱۴	۰/۷۴۱	۰/۵۰۱	۰/۷۰۳	۰/۶۷۲	LN(۰/۸)
۰/۹۹۶	۰/۹۸۹	۰/۹۹۷	۰/۹۹۲	۰/۹۹۹	۰/۹۹۹	۰/۹۹۶	۰/۹۹۳	۰/۹۹۴	۰/۹۹۱	LN(۱/۵)
۰/۴۶۹	۰/۳۸۴	۰/۵۲۳	۰/۴۱۲	۰/۷۰۳	۰/۷۲۸	۰/۴۹۵	۰/۴۴۴	۰/۴۰۳	۰/۳۵۷	HN
۰/۱۵۴	۰/۲۶۲	۰/۱۸۱	۰/۲۸۱	۰/۳۸۵	۰/۲۶۶	۰/۲۱۰	۰/۴۵۳	۰/۱۸۴	۰/۲۱۴	U
۰/۹۹۸	۰/۹۹۲	۰/۹۹۹	۰/۹۹۵	۱/۰۰۰	۱/۰۰۰	۰/۹۹۸	۰/۹۹۹	۰/۹۹۳	۰/۹۹۰	CH(۰/۵)
۰/۵۵۳	۰/۴۶۱	۰/۶۰۳	۰/۵۰۶	۰/۷۷۷	۰/۸۰۳	۰/۵۶۹	۰/۵۳۸	۰/۴۸۰	۰/۴۲۵	CH(۱)
۰/۰۶۹	۰/۰۷۶	۰/۰۶۸	۰/۰۷۶	۰/۱۱۹	۰/۱۱۱	۰/۰۴۵	۰/۰۷۱	۰/۰۴۳	۰/۰۴۲	CH(۱/۵)
۰/۳۷۹	۰/۳۰۴	۰/۴۲۷	۰/۳۳۷	۰/۶۰۶	۰/۶۳۳	۰/۴۰۸	۰/۳۵۷	۰/۳۳۵	۰/۲۸۸	LF(۲)
۰/۲۵۹	۰/۲۱۴	۰/۲۹۲	۰/۲۲۵	۰/۴۶۷	۰/۴۹۳	۰/۲۸۲	۰/۲۳۹	۰/۲۲۵	۰/۱۹۰	LF(۴)
۰/۵۵۷	۰/۴۷۱	۰/۶۰۹	۰/۴۹۷	۰/۷۷۷	۰/۸۰۲	۰/۵۷۷	۰/۵۴۱	۰/۴۸۱	۰/۴۲۸	EV(۰/۵)
۰/۲۲۸	۰/۲۲۱	۰/۲۵۱	۰/۲۴۰	۰/۴۵۸	۰/۴۵۶	۰/۱۹۹	۰/۲۷۴	۰/۱۳۶	۰/۱۱۶	EV(۱/۵)
۰/۶۴۷	۰/۵۳۰	۰/۶۹۶	۰/۵۴۱	۰/۷۴۴	۰/۷۴۸	۰/۷۱۲	۰/۴۷۵	۰/۶۷۴	۰/۶۳۶	DL(۱)
۰/۲۵۵	۰/۱۷۱	۰/۲۸۴	۰/۱۸۰	۰/۳۰۴	۰/۳۰۸	۰/۳۵۲	۰/۱۲۳	۰/۳۲۷	۰/۲۹۵	DL(۱/۵)
۰/۵۱۰	۰/۶۶۲	۰/۶۶۴	۰/۶۰۳	۰/۶۰۷	۰/۴۶۸	۰/۳۹۶	۰/۸۵۳	۰/۸۶۳	۰/۸۸۴	B(۲, ۱)
۰/۵۹۳	۰/۷۸۶	۰/۶۹۱	۰/۸۰۹	۰/۹۲۸	۰/۷۷۹	۰/۷۹۰	۰/۹۶۵	۰/۷۰۳	۰/۷۳۲	B(۰/۵, ۰/۵)
۰/۱۱۳	۰/۱۱۶	۰/۱۳۴	۰/۱۱۸	۰/۱۰۷	۰/۰۹۸	۰/۱۵۱	۰/۱۷۴	۰/۱۷۱	۰/۱۹۴	B(۲, ۲)
۰/۵۳۷	۰/۵۹۱	۰/۶۸۸	۰/۵۸۵	۰/۶۳۴	۰/۶۴۹	۰/۷۰۵	۰/۶۸۳	۰/۷۴۰	۰/۷۶۶	B(۳, ۲)
۰/۹۳۸	۰/۹۹۳	۰/۹۸۷	۰/۹۸۳	۰/۹۸۵	۰/۸۷۴	۱/۰۰۰	۰/۹۹۹	۰/۹۹۹	۱/۰۰۰	B(۲, ۰/۵)

جدول ۴: مقایسه‌ی توان آزمون‌ها برای اندازه نمونه‌ی $n = ۳۰$ و سطح معنی داری $\alpha = ۰/۰۵$

D	V	$W^{۲}$	$U^{۲}$	$A^{۲}$	L	CK_n	KL_{mn}	$T_۱$	$T_۲$	توزیع جانشین
۰/۹۹۸	۰/۹۹۳	۰/۹۹۹	۰/۹۹۳	۱/۰۰۰	۱/۰۰۰	۰/۹۹۷	۰/۹۹۹	۰/۹۹۵	۰/۹۹۳	W(۰/۸)
۰/۵۲۴	۰/۴۱۲	۰/۵۸۴	۰/۴۳۹	۰/۶۹۲	۰/۷۲۸	۰/۵۷۲	۰/۴۱۳	۰/۴۹۰	۰/۴۴۶	W(۱/۴)
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	Γ(۰/۴)
۰/۹۶۴	۰/۹۲۷	۰/۹۷۸	۰/۹۳۳	۰/۹۹۱	۰/۹۹۶	۰/۹۶۸	۰/۹۶۱	۰/۹۴۵	۰/۹۲۸	Γ(۱)
۰/۴۶۰	۰/۳۴۰	۰/۵۱۱	۰/۳۶۴	۰/۵۶۰	۰/۵۸۷	۰/۵۴۲	۰/۲۷۹	۰/۴۹۵	۰/۴۵۶	Γ(۲)
۰/۸۳۰	۰/۷۳۴	۰/۸۶۱	۰/۷۵۴	۰/۸۶۴	۰/۸۵۸	۰/۸۶۶	۰/۶۸۹	۰/۸۴۷	۰/۸۲۸	LN(۰/۸)
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۹۷	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	LN(۱/۵)
۰/۶۴۹	۰/۵۵۸	۰/۷۰۵	۰/۵۹۸	۰/۸۴۷	۰/۸۶۵	۰/۶۲۶	۰/۶۳۵	۰/۵۳۱	۰/۴۸۲	HN
۰/۲۲۲	۰/۳۹۱	۰/۲۸۰	۰/۴۲۱	۰/۵۲۲	۰/۳۲۲	۰/۳۶۷	۰/۷۱۲	۰/۳۳۴	۰/۳۸۷	U
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	CH(۰/۵)
۰/۷۴۹	۰/۶۶۱	۰/۷۹۵	۰/۶۹۸	۰/۹۰۸	۰/۹۲۰	۰/۷۲۳	۰/۷۵۶	۰/۶۲۵	۰/۵۷۵	CH(۱)
۰/۰۷۹	۰/۰۹۴	۰/۰۸۱	۰/۰۹۹	۰/۱۴۳	۰/۱۳۷	۰/۰۴۸	۰/۰۹۳	۰/۰۴۱	۰/۰۴۲	CH(۱/۵)
۰/۵۴۴	۰/۴۵۳	۰/۵۹۷	۰/۴۸۴	۰/۷۶۶	۰/۷۸۳	۰/۵۱۴	۰/۵۳۳	۰/۴۴۰	۰/۳۹۳	LF(۲)
۰/۳۷۹	۰/۳۱۲	۰/۴۳۳	۰/۳۴۰	۰/۶۱۸	۰/۶۳۵	۰/۳۸۲	۰/۳۷۹	۰/۲۹۶	۰/۲۵۷	LF(۴)
۰/۷۴۹	۰/۶۵۵	۰/۷۹۶	۰/۶۹۶	۰/۹۰۹	۰/۹۲۳	۰/۷۱۹	۰/۷۵۲	۰/۶۲۶	۰/۵۷۴	EV(۰/۵)
۰/۳۳۲	۰/۳۲۳	۰/۳۶۲	۰/۳۵۵	۰/۵۸۹	۰/۵۹۲	۰/۲۲۷	۰/۴۳۳	۰/۱۶۷	۰/۱۴۲	EV(۱/۵)
۰/۸۱۸	۰/۷۰۸	۰/۸۵۶	۰/۷۲۳	۰/۸۸۵	۰/۸۹۱	۰/۸۵۳	۰/۶۸۱	۰/۸۲۲	۰/۷۹۵	DL(۱)
۰/۳۶۱	۰/۲۵۱	۰/۴۰۶	۰/۲۷۰	۰/۴۱۹	۰/۴۱۶	۰/۴۷۰	۰/۲۰۰	۰/۴۴۰	۰/۴۱۰	DL(۱/۵)
۰/۷۲۵	۰/۸۶۹	۰/۸۶۵	۰/۸۰۱	۰/۸۵۱	۰/۶۶۰	۰/۹۷۵	۰/۹۷۳	۰/۹۸۲	۰/۹۸۷	B(۲, ۱)
۰/۷۸۷	۰/۹۳۰	۰/۸۸۲	۰/۹۴۷	۰/۹۸۶	۰/۸۸۸	۰/۳۴۴	۰/۹۹۸	۰/۹۱۷	۰/۹۳۳	B(۰/۵, ۰/۵)
۰/۱۶۳	۰/۱۵۴	۰/۱۹۱	۰/۱۵۱	۰/۱۶۴	۰/۱۳۳	۰/۲۱۹	۰/۲۲۶	۰/۲۸۴	۰/۳۱۷	B(۲, ۲)
۰/۷۴۲	۰/۷۸۸	۰/۸۶۸	۰/۷۶۹	۰/۸۶۰	۰/۸۴۶	۰/۲۱۸	۰/۸۶۷	۰/۹۳۴	۰/۹۴۲	B(۳, ۲)
۰/۹۹۵	۱/۰۰۰	۱/۰۰۰	۰/۹۹۹	۱/۰۰۰	۰/۹۷۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	B(۲, ۰/۵)

با توجه به جدول‌های (۲) و (۳) برای حجم نمونه $n = ۱۰, ۲۰$ آماره آزمون پیشنهادی $T_۲$ در مقابل توزیع‌های جانشین $B(۲, ۱)$، $B(۲, ۲)$ و $B(۲, ۰/۵)$ دارای بیشترین توان است و با توان سایر آزمون‌ها اختلاف زیادی دارد. همچنین برای $n = ۲۰$ آزمون $T_۲$ در مقابل توزیع $B(۳, ۲)$ دارای بیشترین توان است. در مقابل سایر توزیع‌های جانشین توان آزمون پیشنهادی با توان آزمون‌های رقیب اختلاف چندانی ندارد.

از جدول (۴) مشاهده می‌شود که برای حجم نمونه $n = ۳۰$ آماره آزمون پیشنهادی $T_۲$ در مقابل توزیع‌های جانشین $B(۲, ۱)$، $B(۲, ۲)$، $B(۲, ۰/۵)$ و $B(۳, ۲)$ دارای بیشترین توان است و با توان سایر آزمون‌ها اختلاف زیادی دارد. به طور کلی از جدول‌های شبیه سازی توان‌ها مشاهده می‌شود، آزمون‌های پیشنهادی در مقایسه با آزمون‌های رقیب عملکرد نسبتاً خوبی دارند و برای فرضیه‌های جانشین بتا از توان بالاتری برخوردارند.

۵ کاربرد برای داده‌های واقعی

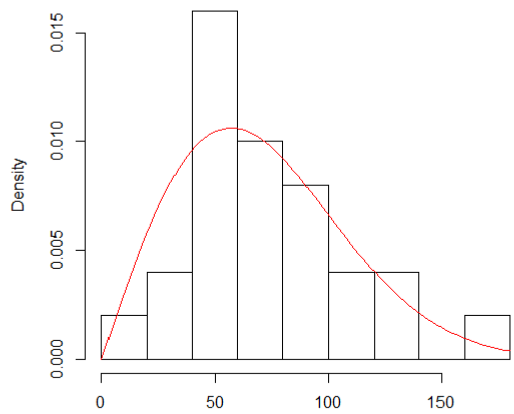
در این بخش یک مجموعه داده‌ی واقعی در نظر گرفته شده و به کمک آزمون‌های پیشنهادی این مجموعه داده تجزیه و تحلیل شده است.

مثال ۱.۵. داده‌های ارائه شده در جدول ۵، را در نظر می‌گیریم. این داده‌ها مربوط به تعداد دفعات شکست ۲۵بلبرینگ در آزمون استقامت است که توسط کارونی[۷] ارائه و تحلیل شده است. نویسندگانی مانند براتپور و خدادادی[۵]، جهانشاهی و همکاران[۱۰] و صفوی‌نژاد و همکاران[۲۲] نیز این داده‌ها را تجزیه و تحلیل کرده‌اند. به علت استفاده زیاد نویسندگان از این داده‌ها ما نیز این داده‌ها را به کار می‌بریم.

هیستوگرام مجموعه داده‌ها، در شکل ۱ نشان داده شده است.

جدول ۵: تعداد دفعات شکست ۲۵ بلبرینگ در آزمون استقامت

۱۷/۸۸	۲۸/۹۲	۳۳/۰۰	۴۱/۵۲	۴۲/۱۲	۴۵/۶۰	۴۸/۴۸	۵۱/۸۴	۵۱/۹۶	۵۴/۱۲
۵۵/۵۶	۶۷/۸۰	۶۷/۸۰	۶۷/۸۰	۶۸/۶۴	۸۶/۶۴	۶۸/۸۸	۸۴/۱۲	۹۳/۱۲	۹۸/۶۴
۱۰۵/۱۲	۱۰۵/۸۴	۱۲۷/۹۲	۱۲۸/۰۴	۱۷۳/۴۰					



شکل ۱: هیستوگرام تعداد دفعات شکست بلبرینگ‌ها

در اینجا از آماره آزمون‌های (۱.۴) و (۲.۴) برای بررسی این که آیا داده‌ها از توزیع رایلی پیروی می‌کنند یا خیر استفاده شده است. برآورد درستنمایی ماکسیمم پارامتر θ به صورت

$$\hat{\theta} = ۵۷/۰۷۶,$$

است. مقادیر بحرانی، مقادیر آماره آزمون‌ها و همچنین p–مقدار تعریف شده به ازای آماره آزمون‌های (۱.۴)، (۲.۴)، آزمون کولموگروف–اسمیرنوف (D)، آزمون کوپر (V)، آزمون کرامر–وان میسز $(W^۲)$، آزمون واتسون $(U^۲)$، و آزمون اندرسون–دارلینگ $(A^۲)$ در جدول ۶ ارائه شده‌اند.

جدول ۶: مقادیر بحرانی، مقادیر آماره و p-مقدار آزمون‌ها

	مقدار بحرانی	آماره آزمون	p-مقدار
D	۰/۲۰۹	۰/۱۲۳	۰/۶۲۰
V	۰/۳۱۸	۰/۲۳۵	۱/۰۰۰
W^2	۰/۲۱۷	۰/۰۲۵	۰/۷۹۹
U^2	۰/۱۵۸	۰/۰۵۱	۰/۶۱۹
A^2	۱/۲۸۹	۰/۳۴۲	۰/۸۲۱
T_1	۰/۰۵۲۱	۰/۰۱۳۰۹	۰/۹۲۶
T_2	۰/۰۴۹۲	۰/۰۱۳۰۸	۰/۹۲۱

همان‌طور که در جدول ۶ مشاهده می‌شود، تمام آزمون‌های ذکر شده در آزمون نیکویی برازش با سطح معناداری ۵ درصد نتیجه‌ی یکسانی را می‌دهند. از آنجا که p-مقدار آزمون‌های T_1 و T_2 به ترتیب برابر با ۰/۹۲۶ و ۰/۹۲۱ بوده و از ۰/۰۵ بیشتر هستند، لذا فرضیه‌ی صفر مبنی بر این که تعداد دفعات شکست بلبرینگ‌ها از توزیع رایلی پیروی می‌کنند رد نمی‌شود و این آزمون‌ها نتیجه‌ی ارائه شده توسط سایر آزمون‌ها را تایید می‌کنند.

۶ نتیجه‌گیری

در این مقاله برخی از تعمیم‌های اطلاع کولبک-لیبلر برای تابع توزیع تجمعی و تابع بقا بررسی شد. رابطه‌ی بین اطلاع کولبک-لیبلر مانده‌ی تجمعی $(CRKL)$ و آنتروپی مانده‌ی تجمعی (CRE) ما را قادر ساخت تا با استفاده از برآورد گشتاوری، توزیع ماکسیمم آنتروپی مانده‌ی تجمعی را ارائه دهیم. همچنین در رابطه با دو روش برآورد پارامتر براساس اطلاع کولبک-لیبلر مانده‌ی تجمعی و آنتروپی مانده‌ی تجمعی بحث شد. سپس، کاربرد اطلاع کولبک-لیبلر مانده‌ی تجمعی به عنوان یک آماره آزمون نیکویی برازش برای توزیع رایلی بررسی شد. پس از آن توزیع‌های جانشین را به همراه آماره‌های رقیب ارائه داده و به کمک شبیه سازی مونت کارلو توان آزمون‌ها محاسبه شد. به طور کلی از شبیه سازی توان آزمون‌ها مشاهده شد که آزمون‌های پیشنهادی در مقایسه با آزمون‌های رقیب عملکرد نسبتاً خوبی دارند و برای فرضیه‌های جانشین بتا دارای بیشترین توان هستند. در آخر یک مثال واقعی بیان کرده و نتایج آن برای آماره‌های پیشنهاد شده ارائه شد.

فهرست منابع

[1] Alizadeh Noughabi, R., & Alizadeh Noughabi, H., & Ebrahimi Moghaddam Behabadi, A., (2014) . An entropy test for the Rayleigh distribution and power comparison. *Journal of Statistical Computation and Simulation*, 84 , 151-158 .

[2] Anderson, T.W., & Darling, D.A., (1954) . A test of goodness of fit. *Journal of American Statistical Association*, 49 , 765-769 .

[3] Arizono, I., & Ohta, H., (1989) . A test for normality based on Kullback-Leibler information. *The American Statistician*, 43 , 20-22 .

[4] Balakreshnan, N., & Rad, A.H., & Arghami, N.R., (2007) . Testing exponentiality based on Kullback-Leibler information with peogressively Type-II censored data. *IEEE Transactions on Reliability*, 56 , 349-356 .

[5] Baratpour, S., & Khodadadi, F., (2012) . A cumulative residual entropy characterization of the Rayleigh distribution and related goodness-of-fit test. *Journal of Statistical Research of Iran*, 9 , 115-131 .

[6] Baratpour, S., & Rad, A.H., (2012) . Testing goodness-of fit for exponential distribution based on cumulative residual entropy. *Communications in Statistics Theory and Methods*, 41 , 1387-1396 .

[7] Caroni, C., (2002) . The correct ball bearing data. *Lifetime Data Analysis*, 8 , 395-399 .

[8] Ciumara, R., & Panait, I.I., (2018) . On Generalized Cumulative Information of Kullback-Leibler Type. *Order*, 2 , 1 .

[9] Di Crescenzo, A., & Longobardi, M., (2009) . On cumulative entropies. *Journal of Statistical Planning and Inference*, 139 , 4072-4087 .

[10] Jahanshahi, S.M.A., & Habibi Rad, A., & Fakoori, V., (2016) . A goodness of fit test Rayleigh distribution based on Hellinger distance. *Annals of Data Science*, 3 , 401-411 .

[11] Khorashadizadeh, M., (2018) . More results on dynamic cumulative inaccuracy measure. *JIRSS-Journal of The Iranian Statistical Society*, 17(1) , 89-108 .

[12] Khorashadizadeh, M., & Roknabadi, A.R., & Borzadaran, G.M., (2016) . Discrete dynamic cumulative residual entropy. *International Journal of Reliability and Safety*, 10(3) , 210-226 .

[13] Kolmogorov, A.N., (1933) . Sulla Determinazione Empirica di une legge di Distribuzione. *Giornale dell'Intituto Italiano degli Attuari*, 4 , 83-91 .

[14] Kullback, S., (1959) . Information Theory and Statistics. *Wiley, NY*.

[15] Kuiper, N.H., (1960) . Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A*, 63 , 38-47 .

[16] Meintanis, S., & Iliopoulos, G., (2003) . Test of fit for the Rayleigh distribution based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics*, 55 , 137-151 .

[17] Navarro, J., & Aguila, Y., & Asadi, M., (2010) . Some new result on the cumulative residual entropy. *Journal of Statistical Planning and Inference*, 140 , 310-322 .

[18] Park, S., & Noughabi, H.A., & Kim, I., (2018) . General cumulative Kullback-Leibler information. *Communications in Statistics-Theory and Methods*, 47(7) , 1551-1560 .

[19] Park, S., & Rao, M., & Shin, D.W., (2012) . On cumulative residual Kullback-Leibler information. *Statistics and Probability Letters*, 82 , 2025-2032 .

[20] Park, S., & Pakyari, R., (2015) . Cumulative residual Kullback-Leibler information with the progressively Type-II censored data. *Statistics and Probability Letters*, 106 , 287-294 .

[21] Rao, M., & Chen, Y., & Vemuri, B.C., & Wang, F., (2004) . Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information theory*, 50 , 1220-1228 .

[22] Safavinejad, M., & Jomhoori, S., & Alizadeh Noughabi, H., (2015) . A devsity based empirical likelihood ratio goodness of fit test for the Rayleigh distribution and power comparison. *Jornal of Statistical Computation and Simulation*, 85 , 3322-3334 .

[23] Shannon, C.E., (1948) . A Mathematical of Communication. *Bell System Technical Journal*, 27 , 379-423 .

[24] Sunoj, S.M., & Sankaran, P.G., & Unnikrishnan Nair, N., (2018) . Quantile-based cumulative Kullback-Leibler divergence. *Statistics*, 52(1) , 1-17 .

[25] Watson, G.S., (1961) . Goodness of fit tests on a circle. *Biometrika*, 48 , 109-114 .

[26] Zohrevand, Y., & Hashemi, R., & Asadi, M., (2020) . An adjusted cumulative Kullback-Leibler information with application to test of exponentiality. *Communications in Statistics-Theory and Methods*, 49(1) , 44-60 .

Some extensions of Kullback-Leibler information based on survival function

Hedieh Eftekhari Moodi, Hadi Alizadeh Noughabi [††] **, Mohammad Khorashadizadeh**

Department of Statistics, University of Birjand, Birjand , Iran.

Communicated by: G.R. Mohtashami Borzadaran

Received: 2020/11/18 Accepted: 2021/2/23

Abstract: In this article, we first investigate some extensions of the Kullback-Leibler information and their properties. Then, we consider the moment constraints for the maximum distribution of cumulative residual entropy and investigate the relationship between the cumulative residual Kullback-Leibler information (CRKL) and cumulative residual entropy (CRE). We also discuss the methods for estimating the scale parameter of Rayleigh distribution and provide two estimation methods. In the following we use cumulative residual Kullback-Leibler information as a goodness of fit test statistic. Then we compute the critical values and the power of proposed tests and compare the power values with the power of other tests. Finally, we apply the tests for a real data set.

Keywords: Goodness of fit test, Cumulative residual entropy, Kullback-Leibler information, Cumulative residual Kullback-Leibler information, Power of test, Monte Carlo simulation.
Mathematics Subject Classification (2010): 62G10, 62G20.

[††]**Corresponding author.**
E-mail addresses: alizadehhadi@birjand.ac.ir **(Hadi Alizadeh Noughabi)**