



Application of Extropy in Characterization of Some Continuous Distributions

Akbari, M.¹, Kasiri, A.², Ahmadi, K.²

¹Department of Statistics, University of Mazandaran, Babolsar, Iran.

²Department of Computer Science, Shahrekord University, Shahrekord, Iran.

Corresponding author: M. Akbari, m.akbari@umz.ac.ir

Received: 27/2/2023 Revised: 13/6/2023 Accepted and Published Online: 17/6/2023.

Introduction

Shannon entropy is a widely used measure to evaluate the uncertainty associated with a random variable. So far, its various generalizations have been discussed in the literature, including dynamic cumulative residual and failure entropy. Recently, Lad et al.(2015) have proposed extropy as an alternative measure of uncertainty. Several researchers have worked on its different versions and their properties, see for example the work of Qiu (2017), Qiu et al. (2018a) and Kamari and Buono (2021). In the present paper, we focus on quantile-based dynamic cumulative residual (QDCR) and quantile-based dynamic cumulative failure (QDCF) extropies to show their applications in the characterization of some continuous distributions.

Material and Methods

Let X be a continuous random variable with quantile function (QF) $Q(p) = F^{-1}(p)$ and quantile density function $q(p) = \frac{1}{f(Q(p))}$, where $F(\cdot)$ and $f(\cdot)$ are the cumulative distribution function and probability density function, respectively. So the QCDR and QDCF extropies are given by

$$J_c(X^t; Q(p)) = -\frac{1}{p} \int_p^1 \left\{ \frac{1-u}{1-p} \right\} q(u) du, \quad (1)$$

$${}_cJ(tX; Q(p)) = -\frac{1}{p} \int_0^p \left\{ \frac{u}{p} \right\} q(u) du, \quad (2)$$

respectively. Eqs. (1) and (2) are defined based on the order statistics

$X_{(1)} \leq \dots \leq X_{(k)}$ and it is established that the following equality

$$J_c(X_{(r)}^t; Q(p)) = {}_cJ({}^tX_{(k-r+1)}; Q(1-p)), \quad (۳)$$

can characterize symmetric distributions if and only if exists $r = 1, \dots, k$ such that (3) holds for any $\frac{1}{r} \leq p \leq 1$. Moreover, it is presented a characterization of generalized Pareto (GP) distribution in terms of (1). The exponential distribution as one of the most applicable distributions in many statistical fields is a special case of GP distribution. Consequently, it is proved that a random variable X with QF $Q(p)$, is distributed as exponential with mean σ if and only if

$$J_c(X^t; Q(p)) = -\frac{\sigma}{r}. \quad (۴)$$

Results and Discussion

One of the most important applications of characterizations is in the development of goodness-of-fit tests. Hence, we focus on estimating (1) and (2) non-parametrically. Utilizing simulation results, it is concluded that the estimator based on kernel function is better than other proposed estimators. Then based on (3), a criterion for testing symmetry is proposed and its correctness is verified. Moreover, a criterion for exponentiality based on (4) is suggested.

Conclusion

In this paper, some continuous distributions have been characterized based on QDCR and QDCF extropies. Two criteria for testing symmetric and exponentiality have been proposed. For future research, they can be considered as test statistics for testing symmetric and exponentiality and can be compared with other existing tests.

Keywords: Extropy, Non-parametric estimation, Symmetric distribution, Exponential distribution, Characterization.

Mathematics Subject Classification (2010): 62G10, 62E10.



©The Author(s). The Publisher is Iranian Statistical Society.

This is an open access article distributed under the terms and conditions of [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

کاربرد اکستروپی در مشخص سازی برخی از توزیع های پیوسته

معصومه اکبری^۱، عارفه کشیری^۱ و کامبیز احمدی^۲

^۱ گروه آمار، دانشکده علوم ریاضی، دانشگاه مازندران، بابلسر

^۲ گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه شهرکرد

نویسنده مسئول: معصومه اکبری، m.akbari@umz.ac.ir

تاریخ دریافت: ۱۴۰۱/۱۲/۸ تاریخ بازنگری: ۱۴۰۲/۳/۲۳ تاریخ پذیرش و انتشار: ۱۴۰۱/۳/۲۷

چکیده: در این مقاله اندازه های اکستروپی مانده و شکست تجمعی پویای چندکی معرفی می شوند. به منظور ارائه کاربردهایی از آن ها، ابتدا با بکارگیری تکنیک شبیه سازی از میان برآوردهای مختلف، برآوردهای مناسبی برای برآورد این اندازه ها انتخاب می شود. سپس براساس تساوی دو اندازه اکستروپی برحسب آماره های ترتیبی، توزیع های پیوسته متقارن مشخص سازی می گردد. در این راستا یک معیار انحراف از تقارن معرفی و چگونگی کاربرد آن در یک مثال واقعی بیان می شود. همچنین از میان توزیع های پیوسته رایج، توزیع پارتو تعمیم یافته و در نتیجه توزیع نمایی مشخص سازی شده و براساس نتایج بدست آمده معیاری جهت تشخیص نمایی بودن یک توزیع پیشنهاد می گردد.

واژه های کلیدی: اکستروپی، برآورد ناپارامتری، توزیع متقارن، توزیع نمایی، مشخص سازی.

کد موضوع بندی ریاضی (۲۰۱۰): 62E10، 62G10.

۱ مقدمه

آنتروپی شانون و نسخه های متفاوت آن، معیارهای مناسبی برای اندازه گیری عدم قطعیت یک متغیر تصادفی است. آیرس و مارتیناس (۱۹۹۵) برای اولین بار اندازه اکستروپی به عنوان اندازه اطلاع جدید و تابع دوگان مکمل آنتروپی شانون را مطرح کردند. سپس لد و همکاران (۲۰۱۵) ضمن مطالعه جامعی بر روی روابط بین اکستروپی و آنتروپی



©نویسندگان). ناشر انجمن آمار ایران است.

این مقاله با دسترسی آزاد تحت شرایط و ضوابط (CC BY-NC 4.0) توزیع شده است.

۴ کاربرد اکستروپی در مشخص سازی توزیع ها

شانون، اکستروپی متغیر تصادفی پیوسته و نامنفی X با تابع چگالی $f(\cdot)$ را به صورت

$$J(X) = -\frac{1}{\psi} \int_0^{\infty} f^{\psi}(x) dx, \quad (1)$$

بدست آوردند. اکستروپی مانده تجمعی (جهانشاهی و همکاران، ۲۰۲۰) به صورت

$$J_c(X) = -\frac{1}{\psi} \int_0^{\infty} \bar{F}^{\psi}(x) dx, \quad (2)$$

است، که در آن $\bar{F}(x) = 1 - F(x)$ تابع بقا متغیر تصادفی X نامیده می شود و اکستروپی شکست تجمعی (کندو، ۲۰۲۱) به صورت

$${}_cJ(X) = -\frac{1}{\psi} \int_0^{\infty} F^{\psi}(x) dx, \quad (3)$$

است. اگر $X^t = (X - t | X \geq t)$ متغیر تصادفی مانده طول عمر باشد، آنگاه اکستروپی مانده تجمعی پویا (ستار و نیر، ۲۰۲۱) برابر

$$J_c(X^t) = -\frac{1}{\psi} \int_t^{\infty} \frac{\bar{F}^{\psi}(x)}{\bar{F}^{\psi}(t)} dx, \quad (4)$$

است. بطور مشابه با در نظر گرفتن متغیر تصادفی گذشته ${}^tX = (t - X | X < t)$ ، اکستروپی شکست تجمعی پویا (کندو، ۲۰۲۱) برابر

$${}_cJ({}^tX) = -\frac{1}{\psi} \int_0^t \frac{F^{\psi}(x)}{F^{\psi}(t)} dx, \quad (5)$$

است. محققین زیادی به مطالعه ویژگی های اکستروپی و تعمیم های آن پرداخته اند. کیو (۲۰۱۷) اکستروپی آماره های ترتیبی را مورد مطالعه قرار داد. اکستروپی مانده آماره های ترتیبی توسط کیو و جیا (۲۰۱۸ a) مورد بررسی قرار گرفته است. کامری و بونو (۲۰۲۱) اکستروپی گذشته را تعریف و برخی از ویژگی های آن را مطالعه کرده اند. جز و عبدالستار (۲۰۱۹) ضمن معرفی اکستروپی مانده تجمعی و نسخه پویای آن، توانستند برخی از توزیع های پیوسته و مقارن را براساس آماره های ترتیبی اندازه مذکور مشخص سازی کنند. مطالعه اندازه اکستروپی براساس نمونه گیری مجموعه رتبه دار توسط رکب و کیو (۲۰۱۹) و کاظمی و همکاران (۲۰۲۱) انجام شده است. علاوه بر این تاکنون برآوردهای ناپارامتری مختلفی برای اندازه اکستروپی توسط کیو و جیا (b ۲۰۱۸)، علیزاده و جراحی (۲۰۱۹) و

هاشم‌پور و محمدی (۲۰۲۲) حاصل شده است. اگرچه اکثر توزیع‌ها در اغلب زمینه‌های آمار بر حسب تابع توزیع تجمعی یا تابع احتمال تعریف شده‌اند اما در این میان توزیع‌هایی مانند توزیع لامبدا تعمیم‌یافته (رامبرگ و اسکمیسر، ۱۹۷۴)، توزیع گوینداراجولا (گوینداراجولا، ۱۹۷۷) و توزیع لامبدا توکی تعمیم‌یافته (فریمر و همکاران، ۱۹۸۸) و موارد دیگر وجود دارند که فقط تابع چندک یا تابع چگالی چندک آنها فرم بسته دارند. از دو دهه اخیر، روش‌های مبتنی بر تابع چندک در زمینه‌های مختلف نظیر قابلیت اطمینان و کنترل کیفیت آماری (حوتی و احمدی، ۱۳۹۵) باعث ایجاد دیدگاه جایگزین و فرصت‌های جدیدی شده است که منتج به نتیجه متناظری (در صورت اتخاذ رویکرد با تابع توزیع تجمعی) نمی‌شود. جزئیات بیشتر در خصوص روش‌های استنباطی مبتنی بر چندک را می‌توان در گیلکریست (۲۰۰۰) دید. با این وجود مطالعه چندانی در خصوص اندازه اکستروپی چندکی وجود ندارد. برای اولین بار کریشنان و همکاران (۲۰۲۰) ضمن مطالعه ویژگی‌های اکستروپی مانده و اکستروپی شکست، بطور مختصر به اکستروپی مانده و شکست چندکی اشاره داشتند و اندکی از ویژگی‌های آن را بیان داشتند. سپس کریشنان و همکاران (۲۰۲۱) اکستروپی چندکی و اکستروپی مانده چندکی را مورد مطالعه بیشتری قرار دادند. آن‌ها ویژگی‌های یکنوایی این اندازه را بدست آوردند و نشان دادند که این اندازه می‌تواند برخی از توزیع‌های پیوسته و متقارن را مشخص‌سازی کند. همچنین آن‌ها، اندازه‌های معرفی شده را برای آماره‌های ترتیبی نیز تعریف، خواصشان را مطالعه و یک برآوردگر ناپارامتری برای اکستروپی چندکی ارائه کردند. تابع چندک $Q(p)$ به‌ازای $0 \leq p \leq 1$ ، برای توابع توزیع اکیداً صعودی $F(\cdot)$ به صورت $Q(p) = F^{-1}(p)$ در نظر گرفته می‌شود. با مشتق‌گیری از طرفین این تساوی نسبت به p ، تابع چگالی چندک $q(p)$ به صورت

$$q(p) = \frac{1}{f(Q(p))}, \quad 0 \leq p \leq 1,$$

خواهد بود. اگر X یک متغیر تصادفی پیوسته با تابع توزیع تجمعی اکیداً صعودی و تابع چندک $Q(p)$ و تابع چگالی چندک $q(p)$ باشد، آنگاه اکستروپی مانده تجمعی پویا بر حسب تابع چندک برابر

$$J_c(X^t; Q(p)) = -\frac{1}{p} \int_p^1 \left\{ \frac{1-u}{1-p} \right\}^p q(u) du, \quad (6)$$

است. بطور مشابه، اکستروپی شکست تجمعی پویا بر حسب تابع چندک برابر

$${}_c J({}^t X; Q(p)) = -\frac{1}{p} \int_0^p \left\{ \frac{u}{p} \right\}^p q(u) du, \quad (7)$$

است. خانواده توزیع‌های متقارن شامل تعداد زیادی از توزیع‌های معروف آماری نظیر توزیع‌های نرمال، تی‌استیودنت، لاپلاس، کوشی و لجستیک است که در بسیاری از حوزه‌های علمی نظیر مطالعات مالی و اقتصادی، پژوهش‌های بیمه و قابلیت اطمینان نقش مهمی را ایفا می‌کنند. به‌عنوان مثال، در مسائل مالی برقراری فرض تقارن توزیع بازده

دارایی در بسیاری از مدل های قیمت گذاری دارایی سرمایه، نظیر شارپ-لینتر و بلک-شولز غیر قابل اجتناب است. همچنین در رگرسیون مدل بازار، فرض تقارن خطاها یکی از فرض های ضروری است (هاجسون، ۲۰۰۲). تحقیقات متنوعی در خصوص بررسی تقارن یک توزیع ارائه شده است (شیونگ و همکاران، ۲۰۲۱؛ امیری و خالدی، ۲۰۱۵؛ کورزو و باباتیوا، ۲۰۱۳؛ باکلیزی، ۲۰۰۳، ۲۰۰۷ و ۲۰۰۸). کاتومانیل و سریدوی (۲۰۲۲) نشان دادند که متوسط طول عمر مانده یک سیستم سری و موازی با دو مولفه یکسان به ترتیب برابر هستند با $-2J_c(X^t)$ و $-2J_c(J^t(X))$. در این مقاله سعی بر آن است که کاربردهایی از روابط (۶) و (۷) در مشخص سازی برخی از توزیع های پیوسته ارائه شود. لذا ابتدا براساس تساوی دو رابطه مذکور، معیاری جدید برای بررسی تقارن یک توزیع آماری معرفی می گردد و با بیان مثال هایی، صحت پذیرش شرط متقارن بودن مورد بررسی قرار می گیرد. مزیت این روش این است که نیاز به معلوم بودن میانه توزیع نیست. سپس ضمن مشخص سازی توزیع پارتو تعمیم یافته، روشی دیگر برای تشخیص نمایی بودن یک توزیع ارائه می گردد. بدین منظور، در بخش ۲، ابتدا برآوردهای ناپارامتری مختلفی برای کمیت $J_c(X^t; Q(p))$ معرفی و سپس با استفاده از شبیه سازی، از میان برآوردهای موجود، برآوردهای مناسب انتخاب می شود. در بخش ۳، ضمن مشخص سازی توزیع های متقارن پیوسته، یک معیار انحراف از تقارن برای بررسی و کنکاش بیشتر نتایج، ارائه شده است. همچنین مشخص سازی توزیع پارتو تعمیم یافته و در نتیجه یک معیار نمایی بودن در بخش ۴ مطرح می گردد.

۲ برآوردهای ناپارامتری اکستروپی مانده تجمعی پویای چندکی

رابطه (۶) را می توان به صورت

$$J_c(X^t; Q(p)) = -\frac{1}{p} \sum_{j=1}^n \left\{ \frac{1 - \frac{j-1}{n}}{1-p} \right\}^2 I\left(\frac{j-1}{n} > p\right) \int_{\frac{j-1}{n}}^{\frac{j}{n}} q(u) du,$$

در نظر گرفت، که در آن $I(u > p)$ تابع نشانگر بوده و برابر با یک است اگر $u > p$ و در غیر این صورت، صفر است. واسیچک (۱۹۷۶)، در برآورد آنتروپی، برآوردهای تجربی $q(u)$ را به صورت $\frac{n(X_{(j+m)} - X_{(j-m)})}{2m}$ به ازای $\frac{j-1}{n} < u \leq \frac{j}{n}$ ارائه کرده است، که در آن $X_{(1)} \leq \dots \leq X_{(n)}$ آماره های ترتیبی و m یک عدد صحیح مثبت کوچکتر از $\frac{n}{4}$ است. در نتیجه اولین برآوردهای $J_c(X^t; Q(p))$ به صورت

$$J_c^{mn} = -\frac{1}{pn} \sum_{j=1}^n \left\{ \frac{1 - \frac{j-1}{n}}{1-p} \right\}^2 I\left(\frac{j-1}{n} > p\right) \left\{ \frac{n(X_{(j+m)} - X_{(j-m)})}{2m} \right\},$$

پیشنهاد می شود که در آن $X_{(j)} = X_{(1)}$ برای $j < 1$ و $X_{(j)} = X_{(n)}$ برای $j > n$.

معصومه اکبری، عارفه کثیری و کامبیز احمدی ۷

کیو و جیا (b ۲۰۱۸) به منظور برآورد اکستروپی، برآوردگر $q(u)$ معرفی شده توسط **واسیچک (۱۹۷۶)** را به صورت $\hat{q}(u) = \frac{n(X_{(j+m)} - X_{(j-m)})}{C_j m}$ معرفی کردند که در آن

$$C_j = \begin{cases} 1 + \frac{j-1}{m} & 1 \leq j \leq m \\ 2 & m+1 \leq j \leq n-m \\ 1 + \frac{n-j}{m} & n-m+1 \leq j \leq n. \end{cases}$$

لذا دومین برآوردگر برای $J_c(X^t; Q(p))$ به صورت

$$J_c^{mn} = -\frac{1}{n} \sum_{j=1}^n \left\{ \frac{1 - \frac{j-1}{n}}{1-p} \right\} I\left(\frac{j-1}{n} > p\right) \left\{ \frac{n(X_{(j+m)} - X_{(j-m)})}{C_j m} \right\},$$

ارائه می‌گردد، که در آن m عدد صحیح کوچکتر از $\frac{n}{4}$ است. از طرفی **پارزن (۱۹۷۹)** برآوردگر تابع چنک $Q(u)$ را به ازای $\frac{j-1}{n} \leq u \leq \frac{j}{n}$ و $j = 1, \dots, n$ ، به صورت

$$\hat{Q}(u) = n\left(\frac{j}{n} - u\right)X_{(j-1)} + n\left(u - \frac{j-1}{n}\right)X_{(j)}, \quad (8)$$

مطرح کرد. لذا یک برآوردگر دیگر از $q(u)$ با استفاده از مشتق‌گیری از (۸) به صورت

$$\hat{q}(u) = n(X_{(j)} - X_{(j-1)}); \quad \frac{j-1}{n} \leq u \leq \frac{j}{n}, \quad j = 1, \dots, n,$$

خواهد بود. بنابراین سومین برآوردگر برای $J_c(X^t; Q(p))$ به صورت

$$J_c^n = -\frac{1}{n} \sum_{j=1}^n \left\{ \frac{1 - \frac{j-1}{n}}{1-p} \right\} I\left(\frac{j-1}{n} > p\right) \left\{ n(X_{(j)} - X_{(j-1)}) \right\},$$

نتیجه می‌شود. **جز (۱۹۹۲)** یک برآوردگر هموار از $q(u)$ را به صورت

$$\hat{q}(u) = \frac{1}{f_n(Q_n(u))}, \quad (9)$$

پیشنهاد کرد که به ازای $\frac{j-1}{n} < u < \frac{j}{n}$ داریم

$$Q_n(u) = X_{(j)}, \quad j = 1, \dots, n,$$

و تابع $f_n(\cdot)$ برآوردگر هسته $f(\cdot)$ به صورت

$$f_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_{(i)}}{h}\right), \quad (10)$$

است. در ادامه، تابع هسته K را چگالی نرمال استاندارد در نظر گرفته و با استفاده از بسته *kdensity* در نرم افزار R ، مقدار h بهینه جهت ارائه یک برآوردگر دیگر، بدست می آید. از این رو چهارمین برآوردگر برای $J_c(X^t; Q(p))$ به صورت زیر پیشنهاد می شود.

$${}^4J_c^n = -\frac{1}{2n(1-p)} \sum_{j=1}^n \left(1 - \frac{j-1}{n}\right)^2 I\left(\frac{j-1}{n} > p\right) \frac{1}{f_n(X_{(j)})}.$$

جدول ۱. میانگین توان های دوم خطا و اربیی برآوردگرهای مختلف به ازای $\alpha = 0.05$.

$MSE(Bias)$				$J_c(X^t; Q(p))$	p	n
${}^1J_c^n$	${}^2J_c^n$	${}^3J_c^{mn}$	${}^4J_c^{mn}$			
۰.۱۷۹(-۰.۰۵۳)	۰.۲۴۹(-۰.۱۱۵)	۰.۸۱۱(-۰.۳۴۹)	۰.۹۰۳(-۰.۳۸۱)	-۰.۳۶	۰.۲	۵
۰.۳۱۴(-۰.۱۱۴)	۰.۴۴۱(-۰.۱۹۴)	۱.۴۲۵(-۰.۵۰۰)	۱.۵۸۵(-۰.۵۴۲)	-۰.۴۳	۰.۳	
۰.۳۳۳(-۰.۰۷۴)	۱.۰۰۷(-۰.۲۲۲)	۰.۴۷۱(-۰.۰۲۱)	۰.۶۲۶(-۰.۰۶۱)	-۰.۵۹	۰.۵	
۰.۲۹۲(-۰.۴۹۳)	۰.۷۷۷(-۰.۱۱۴)	۰.۴۲۱(-۰.۶۴۴)	۰.۳۲۵(-۰.۵۱۶)	-۰.۷۱	۰.۶	
۰.۳۷۷(-۰.۴۷۰)	۲.۴۵۷(-۰.۲۰۵)	۰.۵۶۶(-۰.۷۳۸)	۰.۴۴۶(-۰.۵۱۲)	-۰.۸۵	۰.۷	
۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	-۱.۴۰	۰.۹	
۰.۰۹۲(-۰.۰۶۹)	۰.۱۰۱(-۰.۰۷۰)	۰.۵۳۴(-۰.۳۷۳)	۰.۸۸۵(-۰.۵۳۲)	-	۰.۲	۱۰
۰.۰۹۹(-۰.۰۱۱)	۰.۱۳۶(-۰.۰۶۴)	۰.۵۸۳(-۰.۳۱۲)	۱.۰۶۸(-۰.۵۱۹)	-	۰.۳	
۰.۱۳۸(-۰.۱۴۵)	۰.۲۹۶(-۰.۰۴۴)	۰.۱۷۲(-۰.۳۲۸)	۰.۳۳۲(-۰.۰۱۱)	-	۰.۵	
۰.۱۹۰(-۰.۲۷۵)	۰.۴۷۴(-۰.۰۱۰)	۰.۳۰۳(-۰.۵۲۰)	۰.۲۴۸(-۰.۲۶۸)	-	۰.۶	
۰.۳۲۱(-۰.۴۸۳)	۰.۶۷۴(-۰.۰۸۹)	۰.۵۵۷(-۰.۷۳۸)	۰.۴۰۳(-۰.۵۷۶)	-	۰.۷	
۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	۱.۹۶۴(۱.۴۰۱)	-	۰.۹	
۰.۰۵۲(-۰.۰۴۹)	۰.۰۵۴(-۰.۰۴۸)	۰.۵۳۱(-۰.۴۶۸)	۰.۹۵۴(-۰.۶۵۸)	-	۰.۲	۱۵
۰.۰۷۰(-۰.۰۴۴)	۰.۰۸۳(-۰.۰۶۸)	۰.۷۲۵(-۰.۵۱۰)	۱.۴۰۵(-۰.۷۷۷)	-	۰.۳	
۰.۱۰۱(-۰.۰۵۹)	۰.۱۹۴(-۰.۰۸۹)	۰.۱۵۹(-۰.۱۱۹)	۰.۷۸۹(-۰.۳۹۸)	-	۰.۵	
۰.۱۳۹(-۰.۲۲۲)	۰.۲۶۷(-۰.۰۱۶)	۰.۲۳۴(-۰.۴۲۹)	۰.۲۴۷(-۰.۱۰۰)	-	۰.۶	
۰.۲۱۱(-۰.۲۵۸)	۰.۵۳۸(-۰.۱۱۳)	۰.۳۹۱(-۰.۵۸۸)	۰.۲۹۳(-۰.۲۶۵)	-	۰.۷	
۱.۰۸۲(۱.۰۰۹)	۱.۶۹۴(-۰.۱۲۱)	۱.۷۶۹(۱.۳۲۸)	۱.۵۲۹(۱.۲۲۹)	-	۰.۹	
۰.۰۲۱(-۰.۰۳۰)	۰.۰۲۲(-۰.۰۲۵)	۰.۴۱۹(-۰.۴۹۵)	۰.۹۱۵(-۰.۷۴۶)	-	۰.۲	۳۰
۰.۰۲۵(-۰.۰۰۱)	۰.۰۳۰(-۰.۰۲۵)	۰.۵۳۹(-۰.۵۲۳)	۱.۳۵۱(-۰.۸۷۹)	-	۰.۳	
۰.۰۴۸(-۰.۱۱۱)	۰.۰۶۶(-۰.۰۱۸)	۰.۱۷۷(-۰.۰۱۴)	۱.۳۵۹(-۰.۷۲۸)	-	۰.۵	
۰.۰۸۱(-۰.۱۹۱)	۰.۱۰۸(-۰.۰۰۷)	۰.۱۶۲(-۰.۲۴۹)	۰.۶۰۷(-۰.۳۲۱)	-	۰.۶	
۰.۱۵۱(-۰.۳۰۰)	۰.۱۹۲(-۰.۰۱۴)	۰.۳۳۴(-۰.۵۳۵)	۰.۲۷۲(-۰.۱۱۷)	-	۰.۷	
۱.۰۷۰(۱.۰۱۲)	۱.۱۵۹(-۰.۲۸۱)	۱.۷۸۹(۱.۳۳۷)	۱.۵۵۴(۱.۲۴۲)	-	۰.۹	

جدول ۲. میانگین توان‌های دوم خطا و اریبی برآوردگرهای مختلف به ازای $\alpha = 1$.

$MSE(Bias)$				$J_c(X^t; Q(p))$	p	n
${}^1J_c^n$	${}^2J_c^n$	${}^3J_c^{mn}$	${}^4J_c^{mn}$			
۰/۱۵(۰/۸۰)	۰/۱۸(۰/۵۵)	۰/۱۷(۰/۳۶)	۰/۱۸(۰/۲۸)	-۰/۲۵	۰/۲	۵
۰/۱۶(۰/۲۷)	۰/۲۵(-۰/۵۵)	۰/۲۸(-۰/۳۰)	۰/۳۱(-۰/۴۰)	-۰/۲۵	۰/۳	
۰/۱۶(۰/۵۲)	۰/۳۶(۰/۵۵)	۰/۱۹(۰/۸۹)	۰/۱۹(۰/۶۸)	-۰/۲۵	۰/۵	
۰/۲۹(۰/۱۶۱)	۰/۳۰(۰/۱۲۲)	۰/۵۴(۰/۲۳۴)	۰/۴۲(۰/۲۰۲)	-۰/۲۵	۰/۶	
۰/۱۶(۰/۹۲)	۰/۴۹(۰/۲۳)	۰/۴۹(۰/۲۲۱)	۰/۳۱(۰/۱۶۵)	-۰/۲۵	۰/۷	
۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	-۰/۲۵	۰/۹	
۰/۰۸(۰/۴۱)	۰/۱(۰/۲۸)	۰/۰۹(۰/۲۳)	۰/۱۲(-۰/۱۸)	-	۰/۲	۱۰
۰/۰۸(۰/۵۲)	۰/۱۱(۰/۳۳)	۰/۱۰(۰/۴۶)	۰/۱۴(-۰/۰۸)	-	۰/۳	
۰/۱۱(۰/۷۲)	۰/۱۶(۰/۴۶)	۰/۳۳(۰/۱۸۰)	۰/۱۴(۰/۹۱)	-	۰/۵	
۰/۱۴(۰/۸۹)	۰/۱۹(۰/۶۰)	۰/۴۱(۰/۲۰۲)	۰/۲۲(۰/۱۳۸)	-	۰/۶	
۰/۱۸(۰/۱۱۶)	۰/۲۳(۰/۸۳)	۰/۴۹(۰/۲۲۲)	۰/۳۴(۰/۱۸۲)	-	۰/۷	
۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	۰/۶۲(۰/۲۵۰)	-	۰/۹	
۰/۰۵(۰/۲۸)	۰/۰۷(۰/۱۹)	۰/۰۷(-۰/۰۳)	۰/۱۲(-۰/۴۶)	-	۰/۲	۱۵
۰/۰۵(۰/۱۴)	۰/۰۸(-۰/۰۲)	۰/۰۸(-۰/۱۱)	۰/۱۹(-۰/۷۶)	-	۰/۳	
۰/۰۶(۰/۲۳)	۰/۱۲(-۰/۰۲)	۰/۱۹(۰/۱۳۱)	۰/۰۹(۰/۰۲)	-	۰/۵	
۰/۰۸(۰/۶۴)	۰/۱۴(۰/۳۹)	۰/۳۴(۰/۱۸۳)	۰/۱۴(۰/۱۰۳)	-	۰/۶	
۰/۰۸(۰/۳۳)	۰/۱۸(۰/۰۳)	۰/۳۶(۰/۱۸۸)	۰/۱۶(۰/۱۱۱)	-	۰/۷	
۰/۱۳(۰/۹۶)	۰/۴۸(۰/۲۰)	۰/۵۴(۰/۲۳۴)	۰/۴۵(۰/۲۱۲)	-	۰/۹	
۰/۰۲(۰/۱۸)	۰/۰۳(۰/۰۹)	۰/۰۴(-۰/۲۹)	۰/۱۱(-۰/۷۳)	-	۰/۲	۳۰
۰/۰۳(۰/۲۸)	۰/۰۳(۰/۱۰)	۰/۰۴(-۰/۱۵)	۰/۱۵(-۰/۸۹)	-	۰/۳	
۰/۰۴(۰/۳۹)	۰/۰۵(۰/۱۵)	۰/۱۶(۰/۱۱۷)	۰/۱۰(-۰/۳۷)	-	۰/۵	
۰/۰۴(۰/۴۴)	۰/۰۷(۰/۲۱)	۰/۲۴(۰/۱۵۳)	۰/۰۶(۰/۳۲)	-	۰/۶	
۰/۰۶(۰/۵۳)	۰/۰۹(۰/۲۹)	۰/۳۴(۰/۱۸۴)	۰/۱۲(۰/۹۸)	-	۰/۷	
۰/۱۸(۰/۱۲۲)	۰/۲۲(۰/۸۱)	۰/۵۶(۰/۲۳۷)	۰/۴۸(۰/۲۱۷)	-	۰/۹	

۲.۱ مقایسه عملکرد برآوردگرها

برای ارزیابی عملکرد و سنجش کارایی برآوردگرهای معرفی شده، یک مطالعه شبیه‌سازی انجام می‌شود. براساس نمونه‌هایی به حجم $n = 5, 10, 15, 30$ ، برآوردگرها محاسبه شده‌اند و بر حسب ۱۰۰۰۰ تکرار، میانگین توان‌های دوم خطا (MSE) و اریبی ($Bias$) آن‌ها تخمین زده شده است. بدین منظور داده‌ها از توزیع وایبل با پارامترهای $\alpha = 0.5, 1, 3$ و $\beta = 1$ و چگالی

$$f_{\alpha, \beta}(x) = \alpha \beta x^{\alpha-1} \exp\{-\beta x^\alpha\}, \quad x > 0,$$

تولید شدند. نتایج شبیه‌سازی در جداول ۱-۳، بیانگر آن است که برآوردگر ${}^2J_c^{mn}$ همواره به ازای p های کوچک و $m = \lfloor \frac{n}{4} \rfloor$ دارای MSE کمتری از برآوردگر ${}^1J_c^{mn}$ است، اما برای p های بزرگ، عکس آن برقرار است. همچنین در اکثر موارد برآوردگر ${}^4J_c^n$ در مقایسه با سایر برآوردگرها دارای MSE کمتری است. علاوه‌براین با افزایش حجم نمونه MSE تمام برآوردگرها کاهش می‌یابد.

جدول ۳. میانگین توان های دوم خطا و اریبی برآوردگرهای مختلف به ازای $\alpha = ۳$.

$MSE(Bias)$				$J_c(X^t; Q(p))$	p	n
${}^۴J_c^n$	${}^۲J_c^n$	${}^۲J_c^{mn}$	${}^۱J_c^{mn}$			
$۰/۰۰۴(۰/۰۶۱)$	$۰/۰۰۴(۰/۰۵۱)$	$۰/۰۰۳(۰/۰۵۳)$	$۰/۰۰۳(۰/۰۵۲)$	$-۰/۱۳$	$۰/۲$	۵
$۰/۰۰۲(۰/۰۲۰)$	$۰/۰۰۳(۰/۰۰۷)$	$۰/۰۰۱(۰/۰۱۱)$	$۰/۰۰۱(۰/۰۰۸)$	$-۰/۱۱$	$۰/۳$	
$۰/۰۰۱(۰/۰۰۷)$	$۰/۰۰۲(۰/۰۰۸)$	$۰/۰۰۱(۰/۰۳۲)$	$۰/۰۰۱(۰/۰۲۶)$	$-۰/۰۸$	$۰/۵$	
$۰/۰۰۱(۰/۰۳۵)$	$۰/۰۰۲(۰/۰۴۱)$	$۰/۰۰۴(۰/۰۶۵)$	$۰/۰۰۳(۰/۰۵۷)$	$-۰/۰۷$	$۰/۶$	
$۰/۰۰۱(-۰/۰۰۱)$	$۰/۰۰۲(۰/۰۱۰)$	$۰/۰۰۳(۰/۰۵۳)$	$۰/۰۰۲(۰/۰۳۹)$	$-۰/۰۶$	$۰/۷$	
$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$-۰/۰۴$	$۰/۸$	
$۰/۰۰۲(۰/۰۳۷)$	$۰/۰۰۲(۰/۰۲۸)$	$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۱(۰/۰۳۰)$	$-$	$۰/۲$	۱۰
$۰/۰۰۱(۰/۰۲۸)$	$۰/۰۰۲(۰/۰۲۳)$	$۰/۰۰۲(۰/۰۳۷)$	$۰/۰۰۱(۰/۰۲۰)$	$-$	$۰/۳$	
$۰/۰۰۱(۰/۰۱۹)$	$۰/۰۰۱(۰/۰۲۱)$	$۰/۰۰۳(۰/۰۵۹)$	$۰/۰۰۱(۰/۰۳۳)$	$-$	$۰/۵$	
$۰/۰۰۱(۰/۰۱۷)$	$۰/۰۰۱(۰/۰۲۲)$	$۰/۰۰۳(۰/۰۵۶)$	$۰/۰۰۲(۰/۰۳۹)$	$-$	$۰/۶$	
$۰/۰۰۱(۰/۰۱۷)$	$۰/۰۰۱(۰/۰۲۵)$	$۰/۰۰۳(۰/۰۵۳)$	$۰/۰۰۲(۰/۰۴۳)$	$-$	$۰/۷$	
$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$۰/۰۰۲(۰/۰۴۲)$	$-$	$۰/۸$	
$۰/۰۰۱(۰/۰۲۷)$	$۰/۰۰۲(۰/۰۲۰)$	$۰/۰۰۱(۰/۰۳۱)$	$۰/۰۰۱(۰/۰۱۹)$	$-$	$۰/۲$	۱۵
$۰/۰۰۱(۰/۰۰۶)$	$۰/۰۰۱(۰/۰۰۱)$	$۰/۰۰۱(۰/۰۱۴)$	$۰/۰۰۱(-۰/۰۰۵)$	$-$	$۰/۳$	
$۰/۰۰۰(-۰/۰۰۱)$	$۰/۰۰۱(۰/۰۰۱)$	$۰/۰۰۲(۰/۰۴۳)$	$۰/۰۰۰(۰/۰۰۴)$	$-$	$۰/۵$	
$۰/۰۰۰(۰/۰۱۰)$	$۰/۰۰۱(۰/۰۱۵)$	$۰/۰۰۳(۰/۰۵۱)$	$۰/۰۰۱(۰/۰۲۹)$	$-$	$۰/۶$	
$۰/۰۰۰(-۰/۰۰۵)$	$۰/۰۰۱(۰/۰۰۳)$	$۰/۰۰۲(۰/۰۴۴)$	$۰/۰۰۱(۰/۰۲۴)$	$-$	$۰/۷$	
$۰/۰۰۱(-۰/۰۱۱)$	$۰/۰۰۱(۰/۰۰۵)$	$۰/۰۰۱(۰/۰۳۹)$	$۰/۰۰۱(۰/۰۳۴)$	$-$	$۰/۸$	
$۰/۰۰۰(۰/۰۱۳)$	$۰/۰۰۱(۰/۰۱۰)$	$۰/۰۰۰(۰/۰۰۸)$	$۰/۰۰۰(۰/۰۰۲)$	$-$	$۰/۲$	۳۰
$۰/۰۰۰(۰/۰۰۸)$	$۰/۰۰۰(۰/۰۰۸)$	$۰/۰۰۰(۰/۰۰۹)$	$۰/۰۰۰(-۰/۰۱۰)$	$-$	$۰/۳$	
$۰/۰۰۰(۰/۰۰۴)$	$۰/۰۰۰(۰/۰۰۷)$	$۰/۰۰۲(۰/۰۴۳)$	$۰/۰۰۰(۰/۰۰۲)$	$-$	$۰/۵$	
$۰/۰۰۰(۰/۰۰۳)$	$۰/۰۰۰(۰/۰۰۸)$	$۰/۰۰۲(۰/۰۴۵)$	$۰/۰۰۰(۰/۰۱۵)$	$-$	$۰/۶$	
$۰/۰۰۰(۰/۰۰۳)$	$۰/۰۰۰(۰/۰۰۹)$	$۰/۰۰۲(۰/۰۴۵)$	$۰/۰۰۱(۰/۰۲۴)$	$-$	$۰/۷$	
$۰/۰۰۰(۰/۰۰۶)$	$۰/۰۰۱(۰/۰۱۶)$	$۰/۰۰۲(۰/۰۴۰)$	$۰/۰۰۱(۰/۰۳۶)$	$-$	$۰/۸$	

قابل توجه است که کمیت ${}_cJ(tX; Q(۱-p))$ برابر است با

$${}_cJ(tX; Q(۱-p)) = -\frac{۱}{۲} \int_0^1 \left\{ \frac{۱-u}{۱-p} \right\}^۲ I(u > p) q(۱-u) du.$$

با توجه به نتایج شبیه سازی، یک برآوردگر مناسب برای ${}_cJ(tX; Q(۱-p))$ به صورت

$${}_c\hat{J}(tX; Q(۱-p)) = -\frac{۱}{۲n(۱-p)} \sum_{j=1}^n \left(۱ - \frac{j-۱}{n} \right)^۲ I\left(\frac{j-۱}{n} > p \right) \frac{۱}{f_n(X_{(n-j+۱)})},$$

است، که در آن $f_n(\cdot)$ در (۱۰) تعریف شده است.

۳ مشخص سازی کلاس توزیع های متقارن

تابع $F(\cdot)$ حول c ، متقارن است اگر و فقط اگر برای هر $\frac{1}{p} < u < 1$ ، $Q(u) = -Q(1-u) + 2c$ ، یا به عبارت دیگر $q(u) = q(1-u)$.

در ادامه روشی برای تشخیص متقارن بودن یک توزیع ارائه می شود. براساس متغیرهای تصادفی پیوسته، مستقل و هم توزیع $X_1, \dots, X_k$ با تابع چگالی چندک $q(\cdot)$ ، فرم کلی روابط (۶) و (۷) برحسب r امین آماره ترتیبی متغیر تصادفی مانده و گذشته به ترتیب عبارتند از

$$J_c(X_{(r)}^t; Q(p)) = \frac{-1}{2(\sum_{j=0}^{r-1} \binom{k}{j} p^j (1-p)^{k-j})^2} \int_p^1 \left\{ \sum_{j=0}^{r-1} \binom{k}{j} u^j (1-u)^{k-j} \right\}^2 q(u) du, \quad (11)$$

$${}_c J({}^t X_{(r)}; Q(p)) = \frac{-1}{2(\sum_{j=r}^k \binom{k}{j} p^j (1-p)^{k-j})^2} \int_0^p \left\{ \sum_{j=r}^k \binom{k}{j} u^j (1-u)^{k-j} \right\}^2 q(u) du. \quad (12)$$

قضیه ۱. فرض کنید $X_1, \dots, X_k$ متغیرهای تصادفی پیوسته با تابع توزیع تجمعی اکیداً صعودی $F(\cdot)$ و تابع چندک $Q(\cdot)$ باشند. $F(\cdot)$ متقارن است اگر و فقط اگر، $r = 1, \dots, k$ وجود داشته باشد که برای هر $\frac{1}{p} \leq p \leq 1$ ، رابطه زیر برقرار باشد.

$$J_c(X_{(r)}^t; Q(p)) = {}_c J({}^t X_{(k-r+1)}; Q(1-p)). \quad (13)$$

برهان: با فرض متقارن بودن توزیع $F(\cdot)$ ، رابطه (۱۳) به سادگی حاصل می شود. بنابراین با فرض آن که (۱۳) برقرار باشد، با استفاده از روابط (۱۱) و (۱۲) داریم

$$\int_p^1 \left\{ \sum_{j=0}^{r-1} \binom{k}{j} u^j (1-u)^{k-j} \right\}^2 q(u) du = \int_0^{1-p} \left\{ \sum_{j=0}^{r-1} \binom{k}{j} (1-v)^j v^{k-j} \right\}^2 q(v) dv.$$

با تغییر متغیر u به $1-u$ ، تساوی

$$\int_0^{1-p} \left\{ \sum_{j=0}^{r-1} \binom{k}{j} (1-u)^j u^{k-j} \right\}^2 [q(1-u) - q(u)] du = 0,$$

برای هر $\frac{1}{p} < p < 1$ ، بدست می آید. از آنجایی که برای هر $\frac{1}{p} < u < 1$ ، $q(u) \leq -q(1-u)$ ، عبارت $[q(1-u) - q(u)]$ در انتگرال فوق همواره بزرگتر و یا مساوی صفر است و با توجه به مثبت بودن ضریب این عبارت، انتگرال فوق زمانی برابر صفر خواهد بود اگر و فقط اگر تساوی، $q(1-u) = q(u)$ برقرار باشد. در نتیجه اثبات کامل می شود.

۱۲ کاربرد اکستروپی در مشخص سازی توزیع ها

براساس قضیه ۱، یک آماره آزمون برای تشخیص متقارن بودن توزیع، به صورت

$$S_{r,p} = \hat{J}_c(X_{(r)}^t; Q(p)) - {}_c\hat{J}(X_{(k-r+1)}^t; Q(1-p)),$$

خواهد بود. بر مبنای نتایج شبیه سازی بخش ۲ و برای سادگی با در نظر گرفتن $r = 1$ ، داریم:

$$S_{1,p} = \frac{1}{2kn(1-p)^{2k}} \sum_{j=1}^{kn} \left(1 - \frac{j-1}{kn}\right)^{2k} I\left(\frac{j-1}{kn} > p\right) \left[\frac{1}{f_{kn}(X_{(kn-j+1)})} - \frac{1}{f_{kn}(X_{(j)})} \right]. \quad (14)$$

فرع ۱. اگر $k = 1$ باشد، شرط تقارن (۱۳) به رابطه $J_c(X^t; Q(p)) = {}_cJ(X^t; Q(1-p))$ برای هر $\frac{1}{4} \leq p \leq 1$ ، تقلیل می یابد و در نتیجه $S_{1,p}$ برابر است با

$$S_{1,p} = \frac{1}{2n(1-p)^2} \sum_{j=1}^n \left(1 - \frac{j-1}{n}\right)^2 I\left(\frac{j-1}{n} > p\right) \left[\frac{1}{f_n(X_{(n-j+1)})} - \frac{1}{f_n(X_{(j)})} \right]. \quad (15)$$

مقادیر بزرگ آماره $|S_{1,p}|$ منجر به رد متقارن بودن توزیع آماری خواهد شد.

۳.۱ آماره $|S_{1,p}|$

برای بررسی صحت پذیرش شرط تقارن، در یک مطالعه شبیه سازی، ۴۸ داده ($n = 48, k = 1$) و ($n = 16, k = 3$) از توزیع های نرمال استاندارد (N)، یکنواخت استاندارد (U) و توزیع های کای دو با دو درجه آزادی ($\chi_{(2)}^2$) و یک درجه آزادی ($\chi_{(1)}^2$) تولید شده است. سپس به ازای p های مختلف، آماره $|S_{1,p}|$ محاسبه شده است. جدول ۴ مقادیر بزرگ $|S_{1,p}|$ را برای توزیع های نامتقارن و مقادیر کوچک را برای توزیع های متقارن نشان می دهد.

جدول ۴. مقادیر $|S_{1,p}|$ در توزیع های مختلف.

$ S_{1,p} $						
$\chi_{(1)}^2$	$\chi_{(2)}^2$	U	N	p	k	n
۰/۱۳۲	۰/۲۲۳	۰/۰۱۰	۰/۰۰۱	۰/۵	۱	۴۸
۰/۱۸۴	۰/۳۱۴	۰/۰۱۱	۰/۰۰۱	۰/۶		
۰/۲۳۲	۰/۴۲۳	۰/۰۱۰	۰/۰۰۰	۰/۷		
۰/۳۴۴	۰/۶۲۲	۰/۰۱۰	۰/۰۱۰	۰/۸		
۰/۴۹۲	۰/۵۱۴	۰/۰۱۱	۰/۰۲۲	۰/۹		
۰/۰۳۹	۰/۰۵۸	۰/۰۰۱	۰/۰۰۳	۰/۵	۳	۱۶
۰/۰۵۶	۰/۰۸۵	۰/۰۰۲	۰/۰۰۶	۰/۶		
۰/۰۶۴	۰/۰۹۷	۰/۰۰۳	۰/۰۰۶	۰/۷		
۰/۰۷۴	۰/۰۹۲	۰/۰۰۴	۰/۰۱۰	۰/۸		
۰/۰۹۳	۰/۰۹۲	۰/۰۰۱	۰/۰۱۲	۰/۹		

مثال ۱. داده‌های زیر، زمان‌های تعمیر (برحسب ساعت) برای یک گیرنده ارتباطی هوابرد هستند.

۰/۲, ۰/۳, ۰/۵, ۰/۵, ۰/۵, ۰/۵, ۰/۶, ۰/۶, ۰/۷, ۰/۷, ۰/۷, ۰/۸, ۰/۸, ۱/۰, ۱/۰, ۱/۵,
۱/۵, ۱/۵, ۱/۵, ۲/۰, ۲/۰, ۲/۲, ۲/۵, ۳/۰, ۳/۰, ۳/۳, ۳/۳, ۴/۰, ۴/۰, ۴/۵, ۴/۷,
۵/۰, ۵/۴, ۵/۴, ۷/۰, ۷/۵, ۸/۸, ۹/۰, ۱۰/۳, ۲۲/۰, ۲۴/۵, ۱/۰, ۱/۰, ۱/۱, ۱/۳

کیو و جیا (۲۰۱۸ b) نشان داده‌اند این داده‌ها از توزیع چوله-گوسی معکوس با چگالی

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, \quad x > 0,$$

پیروی می‌کنند، که در آن $\mu = 3$ و $\lambda = 1/588$. در این مثال، $k = 1$ و $n = 45$ در نظر گرفته شده و آماره $S_{1,p}$ در (۱۵) به ازای مقادیر مختلف p در جدول ۵ آمده است که اعداد بدست آمده انحراف قابل توجهی از صفر دارند و چوله بودن توزیع داده‌ها را تایید می‌کند.

جدول ۵. مقادیر $S_{1,p}$.					
p	۰/۵	۰/۶	۰/۷	۰/۸	۰/۹
$S_{1,p}$	-۰/۴۱۶	-۰/۵۳۸	-۰/۶۷۷	-۰/۷۳۳	-۰/۸۶۵

۴ مشخص‌سازی توزیع پارتو تعمیم‌یافته

فرض کنید X ، دارای توزیع پارتو تعمیم‌یافته با تابع توزیع تجمعی

$$F(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - \exp\left\{-\frac{(x-\mu)}{\sigma}\right\} & \xi = 0, \end{cases} \quad (16)$$

باشد، که در آن $x \geq \mu$ برای $\xi \geq 0$ و $\mu - \frac{\sigma}{\xi} \leq x \leq \mu$ به ازای $\xi < 0$. پارامترهای $\mu \in \mathbb{R}$ ، $\sigma > 0$ و $\xi \in \mathbb{R}$ به ترتیب پارامترهای مکان، مقیاس و شکل هستند و با نماد $X \sim GPD(\mu, \sigma, \xi)$ نشان داده می‌شود. این توزیع یکی از پرکاربردترین توزیع‌های آماری است که در زمینه‌های مختلفی از جمله، مدل‌سازی مطالبات بیمه‌ای بزرگ مقدار، تحلیل مقادیر فرین و در مطالعات قابلیت اطمینان به عنوان توزیع شکست کاربرد دارد. از طرفی این توزیع، سه توزیع نمایی، پارتو و یکنواخت را نیز تحت شرایطی شامل می‌شود.

- به ازای $\xi = 0$ و $\mu = 0$ ، متغیر X دارای توزیع نمایی با میانگین σ است.

• به ازای $\xi > 0$ و $\mu = \frac{\sigma}{\xi}$ ، متغیر X دارای توزیع پارتو با پارامترهای α و β و چگالی

$$f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}, \quad x \geq \beta, \quad (17)$$

که $\alpha = \frac{1}{\xi}$ و $\beta = \frac{\sigma}{\xi}$ است.

• به ازای $\xi = -1$ و $\mu = 0$ ، متغیر X دارای توزیع یکنواخت روی بازه $(0, \sigma)$ است.

تابع چگالی چندک توزیع $GPD(\mu, \sigma, \xi)$ برابر است با

$$q(u) = \begin{cases} \sigma(1-u)^{-\xi-1} & \xi \neq 0 \\ \frac{\sigma}{1-u} & \xi = 0. \end{cases}$$

قضیه ۲. فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی چندک $q(u)$ باشد. در این صورت $X \sim GPD(\mu, \sigma, \xi)$ است اگر و فقط اگر برای هر $0 < p < 1$ ، رابطه

$$J_c(X^t; Q(p)) = \begin{cases} h(p, \sigma, \xi) & \xi \neq 0 \\ -\frac{\sigma}{p} & \xi = 0, \end{cases} \quad (18)$$

برقرار باشد، که در آن

$$h(p, \sigma, \xi) = \frac{-\sigma}{p(1-p)^2} \left\{ -\frac{1}{\xi}(1-p^{-\xi}) + \frac{1}{2-\xi}(1-p^{2-\xi}) - \frac{2}{1-\xi}(1-p^{1-\xi}) \right\}.$$

برهان: اگر $X \sim GPD(\mu, \sigma, \xi)$ باشد، به راحتی (۱۸) بدست می آید. اثبات شرط لازم در دو حالت $\xi \neq 0$ و $\xi = 0$ یکسان است. لذا فرض می شود رابطه (۱۸) برای $\xi \neq 0$ برقرار است. در این صورت با استفاده از (۶) نتیجه می شود

$$\int_p^1 (1-u)^2 q(u) du = \sigma \left\{ -\frac{1}{\xi}(1-p^{-\xi}) + \frac{1}{2-\xi}(1-p^{2-\xi}) - \frac{2}{1-\xi}(1-p^{1-\xi}) \right\}.$$

با مشتق گیری از طرفین تساوی فوق نسبت به p ، داریم:

$$-(1-p)^2 q(p) = \sigma p^{-\xi} (-p^{-1} - p + 2).$$

در نتیجه $q(p) = \sigma p^{-\xi-1}$ برای $\xi \neq 0$ بدست می آید و اثبات کامل می شود.

فرع ۲. بنابر قضیه ۲، X دارای توزیع نمایی با میانگین σ است اگر و فقط اگر به ازای هر $0 < p < 1$ ،

$$J_c(X^t; Q(p)) = -\frac{\sigma}{p}, \quad (19)$$

X دارای توزیع پارتو با چگالی (۱۷) است اگر و فقط اگر به ازای هر $0 < p < 1$ ،

$$J_c(X^t; Q(p)) = \frac{\beta(1-p)^{-\frac{1}{\alpha}}}{2(1-2\alpha)},$$

و X دارای توزیع یکنواخت روی بازه $(0, \sigma)$ است اگر و فقط اگر به ازای هر $0 < p < 1$ ،

$$J_c(X^t; Q(p)) = -\frac{\sigma}{p}(1-p).$$

توزیع نمایی یکی از معروفترین توزیعهای طول عمر است که به دلیل سادگی آن دارای کاربردهای فراوان در زمینه‌های مختلف است و مجموعه داده‌های زیادی وجود دارد که از این توزیع پیروی می‌کنند. از آنجایی که یکی از کاربردهای مهم نتایج مشخص‌سازی، استفاده از آن‌ها در آزمون‌های نیکویی برازش است، در مثال زیر با استفاده از (۱۹) یک معیار نمایی بودن معرفی می‌شود.

مثال ۲. با توجه به رابطه (۱۹) و برآوردگر J_c^n ، متغیر تصادفی X دارای توزیع نمایی است اگر و فقط اگر به ازای هر $0 < p < 1$

$$J_c^n T = -\frac{1}{2n(1-p)^2} \sum_{j=1}^n \left(1 - \frac{j-1}{n}\right)^2 I\left(\frac{j-1}{n} > p\right) \frac{1}{f_n(X_{(j)})} + \frac{\bar{X}}{p},$$

دارای مقداری نزدیک صفر باشد. در نتیجه اگر به ازای یک نمونه تصادفی، کمیت $J_c^n T$ ، دارای مقادیر بسیار کوچکی در اطراف صفر باشد، می‌توان فرض کرد که داده‌ها از توزیع نمایی پیروی می‌کنند. زمان‌های بین ورود ۲۵ مشتری به یک مرکز خرید (وادزورس، ۱۹۹۰) را در نظر بگیرید.

$$180, 3743, 398, 423, 465, 289, 348, 406, 434, 484, 293, 357, 411 \\ 437, 491, 303, 385, 413, 453, 499, 315, 392, 416, 462, 517$$

علیزاده (۲۰۱۵) براساس چند آماره آزمون نشان داد که این داده‌ها از توزیع نمایی پیروی نمی‌کنند. مقادیر $J_c^n T$ به ازای p های مختلف برای مجموعه داده فوق محاسبه و در جدول ۶ گزارش شده است. همان‌طور که مشاهده می‌شود این مقادیر به ازای همه p ها به‌ویژه مقادیر بزرگ p ، انحراف قابل توجهی از صفر دارند. بنابراین مشاهدات از توزیع نمایی پیروی نمی‌کنند. از طرفی قابل ذکر است که کمیت $|J_c^n T|$ به نوعی آماره آزمون نمایی بودن است که مقادیر بزرگ آن منجر به رد فرضیه نمایی بودن یک توزیع می‌گردد.

جدول ۶. مقادیر $J_c^n T$.

p	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵	۰/۶	۰/۷	۰/۸	۰/۹
$J_c^n T$	۰.۵۵۹	۰.۷۰۰	۰.۷۳۷	۰.۷۹۴	۰.۸۰۳	۰.۸۴۷	۰.۸۴۶	۰.۸۹۲	۰.۹۰۲

بحث و نتیجه‌گیری

اکستروپی‌های مانده و شکست تجمعی پویای چندکی معرفی و برآورد ناپارامتری آن‌ها مورد بررسی قرار گرفت. سپس براساس این اندازه‌ها، توزیع‌های متقارن پیوسته و توزیع پارتو تعمیم‌یافته مشخص‌سازی شدند. براساس [آرنولد و همکاران \(۱۹۹۸\)](#)، یکی از مهم‌ترین کاربردهای نتایج مشخص‌سازی، استفاده از آنها در توسعه آزمون‌های نیکویی برازش است. از این‌رو دو معیار نمایی بودن و انحراف از تقارن معرفی و کاربرد آنها با استفاده از مثال‌های عددی تشریح شد. به‌علاوه، این پژوهش را به گونه‌ای می‌توان گسترش داد که با در نظر گرفتن معیارهای معرفی شده به عنوان آماره‌های آزمون، به مقایسه توان آزمون‌های پیشنهادی با آزمون‌های متداول و معمول نمایی بودن و متقارن بودن پرداخت. همچنین ارائه آزمون‌های پارتو بودن و یکنواخت بودن و مقایسه آنها با دیگر آزمون‌های معمول، می‌تواند یک موضوع دیگر برای مطالعات آینده باشد.

تقدیر و تشکر

نویسندگان مقاله از داوران گرانقدر و سردبیر محترم مجله علوم آماری که با رهنمودهای ارزنده خود باعث بهبود سطح کیفی و افزایش غنای مقاله شده‌اند، کمال تشکر و قدردانی را دارند.

مراجع

حوتی، ف. و احمدی، ج. (۱۳۹۵)، آنتروپی مانده تجمعی پویای چندکی و مشخص‌سازی توزیع‌های یکنواخت، نمایی و پارتو، مجله علوم آماری، ۱۰، ۸۰-۶۷.

Alizadeh Noughabi, H. (2015), Testing Exponentiality Based on the Likelihood Ratio and Power Comparison, *Annals of Data Science*, **2**, 195-204.

Alizadeh Noughabi, H., and Jarrahiferiz, J. (2019), On the Estimation of Extropy, *Journal of Nonparametric Statistics*, **31**, 88-99.

Amiri, M., and Khaledi, B. (2015), A New Test for Symmetry against Right Skewness, *Journal of Statistical Computation and Simulation*, **86**, 1479-1496.

- Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998), *Records*, New York: John Wiley.
- Ayres, R. U., and Martinas, K. (1995), Waste Potential Entropy: The Ultimate Ecotoxic, *Économie Appliquée*, **48**, 95–120.
- Baklizi, A. (2003), A Conditional Distribution Free Runs Test for Symmetry, *Journal of Nonparametric Statistics*, **15**, 713–718.
- Baklizi, A. (2007), Testing Symmetry Using a Trimmed Longest Run Statistics, *Australian & New Zealand Journal of Statistics*, **49**, 339–347.
- Baklizi, A. (2008), Improving the Power of the Hybrid Test, *International Journal of Contemporary Mathematical Sciences*, **3**, 497–499.
- Corzo, J., and Babativa, G. (2013), A Modified Runs Test for Symmetry, *Journal of Statistical Computation and Simulation*, **83**, 984–991.
- Freimer, M., Kollia, G., Mudholkar, G. S., and Lin, C. T. (1988), A Study of the Generalized Tukey Lambda Family, *Communications in Statistics -Theory and Methods*, **17**, 3547-3567.
- Gilchrist, W. (2000), *Statistical Modelling with Quantile Functions*, Chapman and Hall/CRC, Boca Raton, FL.
- Govindarajula, Z. (1977), A Class of Distributions Useful in Life Testing and Reliability, *IEEE Transactions on Reliability*, **26**, 67-69.
- Hashempour, M., and Mohammadi, M. (2022), On Dynamic Cumulative Past Inaccuracy Measure Based on Extropy, *Communications in Statistics - Theory and Methods*, doi: 10.1080/03610926.2022.2098335.
- Hodgson, D. j., Linton, O., and Vorkink, K. (2002), Testing the Capital Asset Pricing Model Efficiently under Elliptical Symmetry: A Semiparametric Approach, *Journal of Applied Econometrics*, **17**, 617-639.

- Jahanshahi, S. M. A., Zarei, H., and Khammar, A. H. (2020), On Cumulative Residual Extropy, *Probability in the Engineering and Informational Sciences*, **34**, 605-625.
- Jones, M. C. (1992), Estimating Densities, Quantiles, Quantile Densities and Density Quantiles. *Annals of the Institute of Statistical Mathematics*, **44**, 721–27.
- Jose, J., and Abdul-Sathar, E. I. (2019), Residual Extropy of K-Record Values, *Statistics and Probability Letters*, doi: 10.1016/j.spl.2018.10.019.
- Kamari, O., and Buono, F. (2021), On Extropy of Past Lifetime Distribution, *Ricerche di Matematica*, **70**, 505-515.
- Kattumannil, S. K., and Sreedevi, E. P. (2022), Non-Parametric Estimation of Cumulative (Residual) Extropy. *Statistics and Probability Letter*, **185**, 109-434.
- Kazemi, M. R., Tahmasebi, S., Cali, C., and Longobardi, M. (2021), Cumulative Residual Extropy of Minimum Ranked Set Sampling with Unequal Samples, *Results in Applied Mathematics*, **10**, doi: 10.1016/j.rinam.2021.100156.
- Krishnan, A. S., Sunoj, S. M., and Unnikrishnan Nair, N. (2020), Some Reliability Properties of Extropy for Residual and Past Lifetime Random Variables, *Journal of the Korean Statistical Society*, **49**, 457-474.
- Krishnan, A. S., Sunoj, S. M., and Sankaran, P. G. (2021), Some Reliability Properties of Extropy and Its Related Measures Using Quantile Function, *Statistica*, **80**, 413-437.
- Kundo, C. (2021), On Cumulative Residual (Past) Extropy of Extreme Order Statistics, *Communications in Statistics - Theory and Methods*, doi: 10.1080/03610926.2021.2021238.
- Lad, F., Sanfilippo, G., and Agro, G. (2015), Extropy: Complementary Dual of Entropy, *Statistical Science*, **30**, 40–58.

- Parzen, E. (1979), Nonparametric Statistical Data Modeling, *Journal of the American Statistical Association*, **74**, 105-21.
- Qiu, G. (2017), The Extropy of Order Statistics and Record Values, *Statistics and Probability Letters*, **120**, 52-60.
- Qiu, G., and Jia, K. (2018a), The Residual Extropy of Order Statistics, *Statistics and Probability Letters*, **133**, 15-22.
- Qiu, G., and Jia, K. (2018b), Extropy Estimators with Applications in Testing Uniformity, *Journal of Nonparametric Statistics*, **30**, 182-196.
- Ramberg, J. S., and Schmeiser, B. W. (1974), An Approximate Method for Generating Asymmetric Random Variables, *Communications of the ACM*, **17**, 78-82.
- Raqab, M. Z., and Qiu, G. (2019), On Extropy Properties of Ranked Set Sampling, *Statistics*, **53**, 210-26.
- Sathar, E. A. and Nair, R. D. (2021), On Dynamic Survival Extropy, *Communications in Statistics - Theory and Methods*, **50**, 1295-1313.
- Vasicek, O. (1976), A Test for Normality Based on Sample Entropy, *Journal of Royal Statistical Society*, **38**, 730-737.
- Wadsworth, H. M. (1990), *Handbook of Statistical Methods for Engineers and Scientists*, McGraw-Hill, NewYork.
- Xiong, P., Zhuang, W. and Qiu, G. (2021), Testing Symmetry Based on the Extropy of Record Values, *Journal of Nonparametric Statistics*, **33**, 134-155.