

یک توزیع چوله یکنواخت جدید: مدل و ویژگی های آن

عیسی محمودی، ریحانه لاله زاری

گروه آمار، دانشگاه یزد

تاریخ دریافت: ۱۳۹۰/۱/۳۰ تاریخ آخرین بازنگری: ۱۳۹۰/۵/۱۶

چکیده: در این مقاله یک توزیع چوله یکنواخت جدید معرفی می شود که کاملاً از آنچه قبل از این به دست آمده متمایز است. برخی از خواص مهم توزیع از جمله فرم تابع چگالی و تابع توزیع، گشتاورهای مرتبه k ام، تابع مولد گشتاور و تابع مشخصه، واریانس، ضرایب چولگی و کشیدگی، میانگین انحرافات از میانگین، میانه و مد به دست آمده و روش های مختلف برآوردیابی بررسی می شود. همچنین برای بررسی سازگاری برآوردگرهای حاصل از روش های گشتاوری و ماکسیمم درستنمایی، مطالعه ای شبیه سازی صورت می پذیرد. در پایان به مقایسه توزیع چوله یکنواخت جدید و توزیع یکنواخت پرداخته می شود.

واژه های کلیدی: آنتروپی رنی، آنتروپی شانون، توزیع چوله یکنواخت، توزیع یکنواخت، میانگین انحرافات.

۱ مقدمه

مارشال و الکین (۱۹۹۷) روشی را برای اضافه کردن یک پارامتر به خانواده ای از توزیع ها معرفی کردند و آن را برای مدل های نمایی و وایبل به کار گرفتند. اگر

آدرس الکترونیک مسئول مقاله: عیسی محمودی، emahmoudi@yazduni.ac.ir
کد موضوع بندی ریاضی (۲۰۰۰): ۶۰E۰۵، ۶۲F۱۰

$F(x)$ تابع توزیع متغیر تصادفی X و α پارامتر جدید باشد، در این صورت تابع توزیع مدل جدید با یک پارامتر اضافه به صورت

$$H(x) = \frac{F(x)}{1 - (1 - \alpha)\bar{F}(x)}, \quad -\infty < x < \infty, \quad \alpha > 0.$$

تعریف می شود، که در آن $\bar{F}(x) = 1 - F(x)$.

توزیع چوله یکنواخت علی رخم سادگی اش تا به حال به طور دقیق مورد بررسی قرار نگرفته است. به عنوان اولین کار می توان جزئیاتی از این توزیع و برخی خواص آن را در گوپتا و همکاران (۲۰۰۲) مشاهده کرد. آنها برای ساختن توزیع چوله یکنواخت از روش معروف آزالینی (۱۹۸۵) به صورت $g(x) = 2f(x)F(\lambda x)$ کمک گرفته اند، که در آن $f(x)$ و $F(x)$ به ترتیب توابع چگالی و توزیع یکنواخت $U(-\theta, \theta)$ هستند. متأسفانه بسیاری از خواص و روابط به دست آمده توسط گوپتا و همکاران (۲۰۰۲) اشتباه بوده است. سپس ناداراجا و آریال (۲۰۰۴) خواص دقیق توزیع چوله یکنواخت را به دست آوردند که عبارتند از: تابع چگالی، تابع توزیع، تابع نرخ شکست، گشتاور مرتبه k ام، گشتاور مرکزی مرتبه k ام، واریانس، ضریب چولگی و گشتاوری، تابع مولد گشتاور، آنتروپی رنی، آنتروپی شانون و انحرافات از میانگین، میانه و مد.

جونز (۲۰۰۴) با استفاده از توزیع آماره های ترتیبی، خانواده جدیدی از چگالی ها با تعداد پارامترهای بیش تر را معرفی کرد. خانواده معروف فارلی-گامبل-مورگنسترن^۱ (FGM) از توزیع های دو متغیره با توابع توزیع کناری F_1 و F_2 به صورت

$$H(x, y) = F_1(x)F_2(y)\{1 + \lambda\bar{F}_1(x)\bar{F}_2(y)\}. \quad (1)$$

تعریف می شود، که در آن $\lambda \in [-1, 1]$ (برای جزئیات بیش تر به دروئت-ماری و کاتز (۲۰۰۱) مراجعه شود). همچنین ناداراجا و کاتز (۲۰۰۷) با استفاده از روش آزالینی (۱۹۸۵)، فرم کلی مدل های چوله را معرفی کردند و توزیع چوله یکنواخت را به عنوان یک مدل چوله مورد بررسی قرار دادند. میرحسینی و همکاران (۲۰۱۱)

^۱ Farlie-Gumbel-Morgenstern

کلاس جدیدی از توزیع‌ها را براساس آمیخته کردن مینیمم و ماکسیمم نمونه‌های تصادفی معرفی کردند و آن‌را برای توزیع‌های نمایی و لاپلاس به کار گرفتند. در این مقاله یک توزیع چوله‌یکنواخت جدید معرفی می‌شود که کاملاً با توزیع‌های چوله‌یکنواختی که قبل از این معرفی شده‌اند متفاوت است. این توزیع می‌تواند به عنوان توزیع مقادیر خطایی که تا نزدیک‌ترین k امین رقم اعشار گرد شده‌اند به کار رود. همچنین این توزیع می‌تواند کاربردهای جالبی در زمینه‌هایی مانند آزمون فرض، آزمون‌های ناپارامتری، فیزیک، تولید اعداد تصادفی، نمونه‌گیری طبقه‌بندی، آزمون تصادفی بودن مدل‌های ترافیکی و غیره داشته باشد. افزودن پارامتر چولگی باعث انعطاف‌پذیری بیشتر توزیع می‌شود و داده‌هایی که با مدل یکنواخت قابل توجیه نیستند با مدل جدید به راحتی قابل تفسیر می‌شوند.

در بخش ۲ خواص توزیع چوله‌یکنواخت جدید شامل تابع توزیع، تابع چگالی، تابع نرخ شکست، تابع مولد گشتاور، تابع مشخصه، گشتاورها مورد بررسی قرار می‌گیرد. سپس در بخش ۳ برآورد پارامترهای توزیع به دست آورده می‌شوند. در بخش ۴ مطالعه‌ای شبیه‌سازی برای نمایش نحوه برآورد پارامترها ارائه شده است و برای نشان دادن برتری توزیع چوله‌یکنواخت جدید، مقایسه‌ای بین این توزیع و توزیع یکنواخت انجام گرفته است. در بخش نهایی به بحث و نتیجه‌گیری پرداخته خواهد شد.

۲ توزیع چوله‌یکنواخت جدید

تعریف ۱: فرض کنید X_1 و X_2 نمونه‌ای تصادفی از توزیع $U(-\theta, \theta)$ باشند، همچنین برای $-1 \leq \lambda \leq 1$ فرض کنید Z یک متغیر تصادفی برنولی مستقل از X_i ها با پارامتر پیروزی $\frac{1+\lambda}{2}$ باشد، با در نظر گرفتن متغیر جدید $X = ZY_1 + (1 - Z)Y_2$ ، که در آن $Y_1 = \min(X_1, X_2)$ و $Y_2 = \max(X_1, X_2)$ ، در این صورت تابع توزیع متغیر تصادفی X برابر است با

$$F_X(x) = \left(\frac{1+\lambda}{2}\right)F_{Y_1}(x) + \left(\frac{1-\lambda}{2}\right)F_{Y_2}(x),$$

۹۰ یک توزیع چوله یکنواخت جدید

که در آن λ پارامتر چولگی توزیع است. با جایگذاری $F_{Y_1}(x)$ و $F_{Y_2}(x)$ توزیع چوله یکنواخت جدید به صورت

$$\begin{aligned} F(x) &= (1 + \lambda)G(x) - \lambda(G(x))^2 \\ &= G(x)\{1 + \lambda\bar{G}(x)\}, \quad -1 \leq \lambda \leq 1, \end{aligned} \quad (2)$$

حاصل می شود، که در آن $G(x)$ تابع توزیع تجمعی $U(-\theta, \theta)$ است. تابع توزیع معرفی شده در رابطه (۲) حالت یک بعدی تابع توزیع دو متغیره (۱) است. با جایگذاری $G(x)$ ، تابع توزیع چوله یکنواخت جدید برابر است با

$$F(x) = \left(\frac{x + \theta}{2\theta}\right)\left(1 + \lambda\frac{\theta - x}{2\theta}\right), \quad -1 \leq \lambda \leq 1. \quad (3)$$

که با مشتق گیری از آن نسبت به x ، تابع چگالی احتمال چوله یکنواخت جدید به فرم

$$f(x) = \frac{1}{2\theta}\left(1 - \frac{\lambda x}{\theta}\right), \quad x \in (-\theta, \theta), \quad -1 \leq \lambda \leq 1, \quad (4)$$

خواهد شد، که با نماد $X \sim NSU(\theta, \lambda)$ نشان داده می شود. برای حالت خاصی که مقدار پارامتر چولگی برابر صفر باشد ($\lambda = 0$)، این توزیع دقیقاً با توزیع یکنواخت $U(-\theta, \theta)$ یکی خواهد شد.

اگر $X \sim NSU(\theta, \lambda)$ ، آنگاه تابع نرخ شکست^۲ عبارت است از

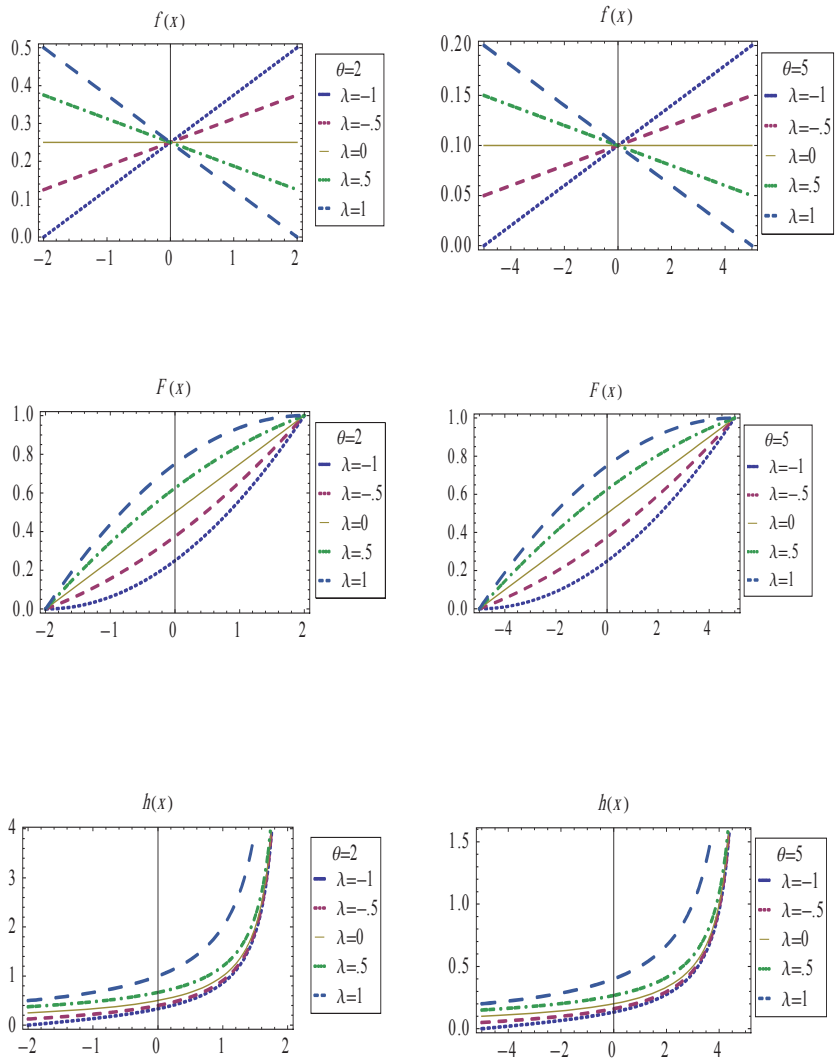
$$H(x) = \frac{2(\theta - \lambda x)}{(x - \theta)[\theta(\lambda - 2) + \lambda x]}.$$

که به ازای $0 \leq \lambda \leq 1$ تابعی نزولی و به ازای $0 \leq \lambda \leq -1$ صعودی خواهد بود. نمودارهای تابع چگالی، تابع توزیع و تابع نرخ شکست توزیع چوله یکنواخت جدید به ازای مقادیر مختلف λ و θ در شکل ۱ نشان داده شده اند.

تابع مولد گشتاور و تابع مشخصه:

اگر $X \sim NSU(\theta, \lambda)$ ، آنگاه تابع مولد گشتاور آن عبارت است از

^۲ Hazard rate function



شکل ۱: نمودار تابع چگالی، تابع توزیع و تابع نرخ شکست توزیع چوله‌یکنواخت جدید به ازای مقادیر مختلف λ و θ

$$M_X(t) = E(\exp(tX)) = \frac{-t\theta\lambda \cosh(t\theta) + (t\theta + \lambda) \sinh(t\theta)}{(t\theta)^2}.$$

به همین ترتیب تابع مشخصه این توزیع به صورت

$$\Phi(t) = E(\exp(itX)) = \frac{-it\theta\lambda \cosh(it\theta) + (it\theta + \lambda) \sinh(it\theta)}{(it\theta)^2},$$

است، که در آن $i = \sqrt{-1}$ عدد مختلط است.

گشتاورها:

n امین گشتاور مرکزی X عبارت است از

$$\mu'_n = E(X^n) = \begin{cases} \frac{\theta^n}{n+1} & n \text{ زوج} \\ -\frac{\lambda\theta^n}{n+2} & n \text{ فرد} \end{cases} \quad (5)$$

همانطور که ملاحظه می شود گشتاورهای مرکزی مراتب زوج به پارامتر λ بستگی ندارد درحالی که گشتاورهای مراتب فرد تابعی از λ هستند. با توجه به رابطه (5) گشتاورهای اول تا چهارم به صورت

$$\mu'_1 = -\frac{\theta\lambda}{3}, \quad \mu'_2 = \frac{\theta^2}{3}, \quad \mu'_3 = -\frac{\theta^3\lambda}{5}, \quad \mu'_4 = \frac{\theta^4}{5}.$$

محاسبه می شوند. از طرفی دیگر با توجه به بسط میانگین به صورت

$$\mu_n = \begin{cases} \theta^n \left\{ \sum_{i=0}^{[n/2]} \frac{1}{\Gamma_{i+1}} \binom{n}{2i} \left(\frac{\lambda}{\theta}\right)^{n-2i} - \sum_{i=0}^{[(n-1)/2]} \frac{1}{\Gamma_{i+2}} \binom{n}{2i+1} \left(\frac{\lambda}{\theta}\right)^{n-2i} \right\} & n \text{ زوج} \\ \theta^n \left\{ \sum_{i=0}^{[(n-1)/2]} \frac{1}{\Gamma_{i+1}} \binom{n}{2i} \left(\frac{\lambda}{\theta}\right)^{n-2i} - \sum_{i=0}^{[n/2]} \frac{1}{\Gamma_{i+2}} \binom{n}{2i+1} \left(\frac{\lambda}{\theta}\right)^{n-2i} \right\} & n \text{ فرد} \end{cases}$$

است، که در آن $[u]$ نشان دهنده جزء صحیح u است. بنابراین واریانس X برابر است با

$$\sigma^2 = \frac{\theta^2(3 - \lambda^2)}{9}.$$

گشتاورهای مرکزی مرتبه سوم و چهارم حول میانگین نیز به ترتیب عبارتند از

$$\mu_3 = \frac{2}{135}\theta^3\lambda(9 - 5\lambda^2),$$

$$\mu_4 = -\frac{1}{135}\theta^4(-27 + 6\lambda^2 + 5\lambda^4).$$

پارامتر چولگی^۳ توزیع که نشان‌دهنده میزان عدم تقارن توزیع است، برابر است با

$$\gamma_1 = \frac{\mu_3}{(\sigma^2)^{3/2}} = \frac{2\lambda(9 - 5\lambda^2)}{5(3 - \lambda^2)^{3/2}}.$$

همان‌طور که ملاحظه می‌شود برای این توزیع میزان چولگی تنها به پارامتر λ بستگی دارد و پارامتر θ در چوله کردن توزیع بی‌تاثیر است. شکل ۲ میزان چولگی توزیع چوله‌یکنواخت جدید را به ازای λ های مختلف نشان می‌دهد. توجه شود که بیش‌ترین مقدار چولگی به ازای $\lambda = 1$ به دست می‌آید. مقدار چولگی در این نقطه برابر با $\frac{2\sqrt{2}}{5}$ است. کشیدگی^۴ این توزیع، عبارت است از

$$\gamma_2 = \frac{\mu_4}{(\sigma^2)^2} = -\frac{3(-27 + 6\lambda^2 + 5\lambda^4)}{5(\lambda^2 - 3)^2}.$$

که تنها به پارامتر λ بستگی دارد. شکل ۲ میزان کشیدگی را به ازای λ های مختلف نشان می‌دهد. همان‌طور که ملاحظه می‌شود بیش‌ترین کشیدگی نیز مانند چولگی به ازای $\lambda = 1$ اتفاق می‌افتد و برابر $\frac{12}{5}$ است. با استفاده از میانگین و انحراف معیار می‌توان به سادگی ضریب تغییرات و پراکندگی توزیع را به دست آورد.

ضریب تغییرات این توزیع برابرست با

$$C.V = \frac{\sigma}{\mu} = -\frac{\sqrt{3 - \lambda^2}}{\lambda}.$$

که همانند ضرایب چولگی و کشیدگی به پارامتر θ بستگی نداشته و تنها با λ تغییر می‌کند. ضریب پراکندگی برای توزیع چوله‌یکنواخت جدید برابر است با

$$R = \frac{\sigma^2}{\mu} = \frac{\theta(\lambda^2 - 3)}{3\lambda}.$$

^۳ Skewness

^۴ Kurtosis

۹۴ یک توزیع چوله یکنواخت جدید

شایان ذکر است که معیار پراکندگی تنها شاخصی است که به هر دو پارامتر θ و λ بستگی دارد. شکل ۲ نشان دهنده نمودارهای امید ریاضی، واریانس، چولگی، کشیدگی، ضریب تغییرات و ضریب پراکندگی به ازای $\theta = ۲$ و مقادیر مختلف λ است.

چندک‌ها و مد توزیع:

چندک p ام توزیع که در رابطه $F(x_p) = p$ صدق می‌کند با توجه به رابطه (۳) برابر است با

$$x_p = \frac{\theta[1 - \sqrt{1 + \lambda(\lambda + 2 - 4p)}]}{\lambda}. \quad (۶)$$

با قرار دادن $p = \frac{1}{2}$ در رابطه (۶) میانه توزیع چوله یکنواخت جدید به صورت

$$med = \frac{\theta(1 - \sqrt{1 + \lambda^2})}{\lambda}. \quad (۷)$$

حاصل می‌شود. با توجه به شکل ۱ مد این توزیع نیز برابر است با

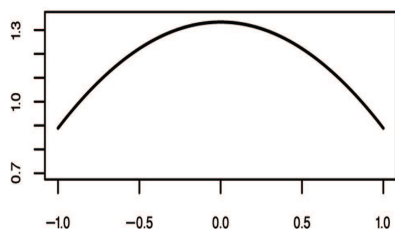
$$Mod = \begin{cases} -\theta & 0 < \lambda < 1 \\ \theta & -1 < \lambda < 0. \end{cases} \quad (۸)$$

میانگین انحرافات:

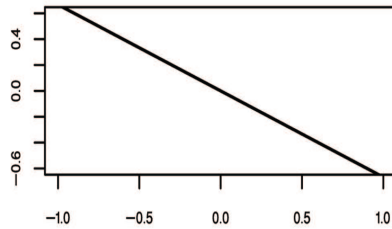
میزان پراکندگی در یک جامعه می‌تواند با استفاده از معیارهایی نظیر میانگین انحرافات از میانگین، میانه و مد اندازه گیری شود. میانگین انحرافات از میانگین یک آماره توانمند است که بیشتر برای شناسایی نقاط پرت موجود در یک مجموعه از داده‌ها مورد استفاده قرار می‌گیرد. میانگین انحرافات از میانه نسبت به معیارهایی نظیر واریانس و انحراف استاندارد نمونه توانمندتر است که حتی برای توزیع‌هایی مانند کوشی که میانگین آنها وجود ندارد نیز می‌تواند مورد استفاده قرار گیرد.

الف- میانگین انحراف از میانگین: میانگین انحراف از میانگین توزیع چوله یکنواخت جدید به صورت

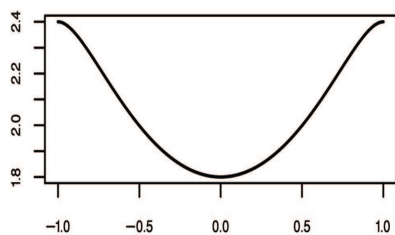
$$\delta_1(X) = \int_{-\theta}^{\theta} |x - \mu| f(x) dx$$



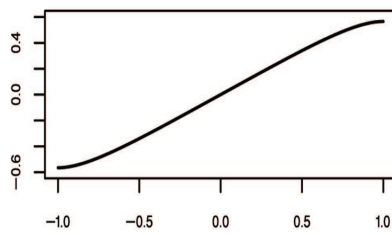
(ب)



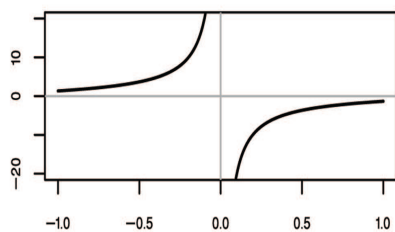
(الف)



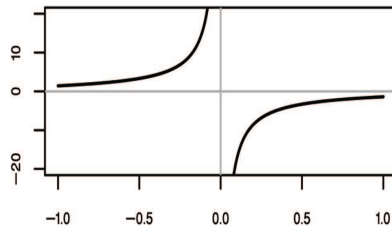
(د)



(ج)



(و)



(ه)

شکل ۲: نمودارهای (الف) امید ریاضی، (ب) واریانس، (ج) چولگی، (د) کشیدگی، (ه) ضریب تغییرات و (و) ضریب پراکندگی توزیع چوله‌یکنواخت جدید به‌ازای $\theta = 2$ و مقادیر مختلف λ

۹۶ یک توزیع چوله یکنواخت جدید

$$\begin{aligned}
 &= \int_{-\theta}^{\mu} (\mu - x)f(x)dx + \int_{\mu}^{\theta} (x - \mu)f(x) dx \\
 &= \mu F(\mu) - \int_{-\theta}^{\mu} xf(x) dx + \int_{\mu}^{\theta} xf(x) dx - \mu(1 - F(\mu)) \\
 &= 2\mu F(\mu) - 2 \int_{-\theta}^{\mu} xf(x) dx.
 \end{aligned}$$

محاسبه می شود، به طوری که $\mu = E(X)$ چون

$$\int_{-\theta}^{\mu} xf(x) dx = \frac{\theta}{324} (2\lambda^4 + 9(\lambda - 3) - 162),$$

$$\delta_1 = \frac{\theta(\lambda^2 - 9)^2}{162} \text{ بنابراین}$$

ب- میانگین انحراف از میانه: با توجه به مقدار به دست آمده برای میانه از رابطه (۷) خواهیم داشت:

$$\begin{aligned}
 \delta_2 &= \int_{-\theta}^{\theta} |x - med|f(x)dx \\
 &= \frac{1}{48} \theta \left(\sqrt{1 + \lambda^2} (\lambda^2 - 8\lambda - 12) - 3\lambda^2 + 6\lambda^2 + 8\lambda + 36 \right).
 \end{aligned}$$

ج- میانگین انحراف از مد: با استفاده از رابطه (۸) داریم:

$$\begin{aligned}
 \delta_3 &= \int_{-\theta}^{\theta} |x - Mod|f(x)dx \\
 &= \begin{cases} \theta - \frac{\lambda\theta}{4} & 0 < \lambda < 1 \\ \theta + \frac{\lambda\theta}{4} & -1 < \lambda < 0. \end{cases}
 \end{aligned}$$

آنتروپی:

آنتروپی، اندازه‌ای از اطلاع است که میزان غیرقابل پیش‌بینی بودن نتیجه یک آزمایش را مشخص می‌کند. همچنین آنتروپی نشان دهنده میزان عدم قطعیت است. اگر X

یک متغیر تصادفی با توزیع تجمعی مطلقاً پیوسته $F(x)$ و تابع چگالی $f(x)$ باشد در این صورت معیار عدم قطعیت برای F توسط $H = E[-\log(f(X))]$ تعریف می‌شود. شاخصهای زیادی برای میزان اطلاع و آنتروپی یک توزیع معرفی شده‌اند که از مهمترین آنها می‌توان به آنتروپی رنی و شانون اشاره کرد. برای توزیع چوله‌یکنواخت جدید مقادیر آنتروپی رنی و شانون توسط روابط زیر محاسبه می‌شوند.

الف- آنتروپی رنی: آنتروپی رنی که به صورت

$$g_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \int_{-\theta}^{\theta} f^\gamma(x) dx \right\}, \quad \gamma > 0, \gamma \neq 1$$

تعریف می‌شود، برای توزیع چوله‌یکنواخت جدید عبارت است از:

$$g_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \frac{(1+\lambda)^{\gamma+1} - (1-\lambda)^{\gamma+1}}{\lambda^{2\gamma} \theta^{\gamma-1} (1+\gamma)} \right\}.$$

ب- آنتروپی شانون: آنتروپی شانون به صورت $E[-\log f(X)]$ تعریف می‌شود که مورد خاصی از آنتروپی رنی برای $r \uparrow 1$ است. حاصل آنتروپی شانون برای توزیع چوله‌یکنواخت جدید برابر است با:

$$E[-\log f(X)] = - \int_{-\theta}^{\theta} \frac{1}{2\theta} \left(1 - \frac{\lambda x}{\theta}\right) \log \left(\frac{1}{2\theta} \left(1 - \frac{\lambda x}{\theta}\right)\right) dx \quad (9)$$

$$= \log(2\theta) - \int_{-\theta}^{\theta} \frac{1}{2\theta} \left(1 - \frac{\lambda x}{\theta}\right) \log \left(1 - \frac{\lambda x}{\theta}\right) dx, \quad (10)$$

که با استفاده از تغییر متغیر $u = \frac{1-\lambda x}{\theta}$ داریم:

$$\int_{-\theta}^{\theta} \frac{1}{2\theta} \left(1 - \frac{\lambda x}{\theta}\right) \log \left(1 - \frac{\lambda x}{\theta}\right) dx = \frac{(1-\lambda)^2}{4\lambda} \left\{ \log(1-\lambda) - \frac{1}{2 \ln 10} \right\} \quad (11)$$

$$- \frac{(1+\lambda)^2}{4\lambda} \left\{ \log(1+\lambda) - \frac{1}{2 \ln 10} \right\}, \quad (12)$$

با قرار دادن رابطه (۱۱) در (۹)، آنتروپی شانون به صورت زیر حاصل می شود:

$$E[-\log f(X)] = \frac{1}{\ln 10} \left\{ \ln(2\theta) + \frac{(1-\lambda)^2}{4\lambda} \ln(1-\lambda) - \frac{(1+\lambda)^2}{4\lambda} \ln(1+\lambda) + \lambda \right\}.$$

مقایسه تصادفی و خاصیت تقارن توزیع چوله یکنواخت جدید:

لم ۱: تابع توزیع چوله یکنواخت جدید (۳)، یک تابع توزیع با خاصیت تابع نرخ شکست صعودی^۵ برای $0 \leq \lambda \leq 1$ و دارای خاصیت تابع نرخ شکست نزولی^۶ برای $0 \leq \lambda \leq 1$ است.

لم ۲: فرض کنید $U \sim NSU(\theta, \lambda)$ و $V \sim NSU(\theta, -\lambda)$ ، آنگاه $-U$ و V هم توزیع هستند.

برهان: از آنجایی که U حول صفر متقارن است داریم:

$$\begin{aligned} F_{-U}(t) = P(-U \leq t) &= 1 - P(U \leq -t) \\ &= 1 - \left(\frac{-t+\theta}{2\theta}\right) \left(1 + \lambda \frac{\theta+t}{2\theta}\right) \\ &= \left(\frac{t+\theta}{2\theta}\right) \left(1 - \lambda \frac{\theta-t}{2\theta}\right) \\ &= P(V \leq t) = F_V(t). \end{aligned}$$

توزیع مجموع دو متغیر تصادفی مستقل و هم توزیع چوله یکنواخت جدید:

اگر X_1 و X_2 دو متغیر تصادفی مستقل و پیوسته با توابع چگالی $f_{X_1}(\cdot)$ و $f_{X_2}(\cdot)$ در بازه $(-\theta, \theta)$ باشند آنگاه تابع چگالی مجموع این دو متغیر تصادفی با استفاده از

^۵ Increasing hazard rate (IHR)

^۶ Decreasing hazard rate (DHR)

رابطه پیچش $f_Z(z) = \int f_{X_1}(x_1)f_{X_2}(z-x_1)dx_1$ برابر است با

$$f_Z(z) = \begin{cases} \int_{-\theta}^{\theta+z} f_{X_1}(x_1)f_{X_2}(z-x_1)dx_1 & -2\theta < z < 0 \\ \int_{z-\theta}^{\theta} f_{X_1}(x_1)f_{X_2}(z-x_1)dx_1 & 0 < z < 2\theta. \end{cases} \quad (13)$$

با فرض آنکه X_1 و X_2 نمونه‌ای تصادفی از توزیع $NSU(\theta, \lambda)$ باشد، با جایگذاری توابع چگالی X_1 و X_2 از رابطه (۴) در (۱۳)، تابع چگالی مجموع $Z = X_1 + X_2$ ، به صورت زیر حاصل می‌شود.

$$f_Z(z) = \begin{cases} \frac{1}{4\theta^4}(\theta^2z + 2\theta^3 - \lambda\theta z^2 - 2\lambda\theta^2z + \frac{\lambda^2z^3}{3} - \lambda^2\theta^2z - \frac{2\lambda^2\theta^3}{3}) & -2\theta < z < 0 \\ \frac{1}{4\theta^4}(-\theta^2z + 2\theta^3 + \lambda\theta z^2 - 2\lambda\theta^2z - \frac{\lambda^2z^3}{3} + \lambda^2\theta^2z - \frac{2\lambda^2\theta^3}{3}) & 0 < z < 2\theta. \end{cases}$$

۳ روش‌های برآورد پارامترهای توزیع

الف- برآورد گشتاوری :

فرض کنید X_1, \dots, X_n نمونه‌ای تصادفی از توزیع چوله‌یکنواخت جدید (۴) باشد. برآورد گشتاوری پارامترهای θ و λ با مساوی قرار دادن گشتاورهای نمونه و جامعه به صورت

$$\begin{cases} \frac{-\theta\lambda}{3} = m_1 \\ \frac{\theta^2(3-\lambda^2)}{9} + \frac{\theta^2\lambda^2}{9} = m_2, \end{cases}$$

به دست می‌آیند، که در آن $m_1 = \frac{1}{n} \sum_{i=1}^n x_i$ و $m_2 = \frac{1}{n} \sum_{i=1}^n x_i^2$ از حل دستگاه معادلات فوق برآورد گشتاوری پارامترهای λ و θ عبارتند از:

$$\tilde{\theta} = \sqrt{3m_2}, \quad \tilde{\lambda} = \frac{-\sqrt{3}m_1}{\sqrt{m_2}}.$$

ب- برآورد ماکسیمم درستمایی :

اگر x_1, \dots, x_n مشاهده از توزیع چوله‌یکنواخت جدید باشند، تابع درستمایی

$$L(\theta, \lambda) = \frac{1}{(2\theta)^n} \prod_{i=1}^n \left(1 - \frac{\lambda x_i}{\theta}\right), \quad -\theta < x_i < \theta, \quad i = 1, \dots, n.$$

۱۰۰ یک توزیع چوله یکنواخت جدید

در $\hat{\theta} = \max |x_i|$ ماکسیمم می شود. با جایگذاری $\hat{\theta}$ در معادله $L(\theta, \lambda)$ ، تابع لگاریتم درستنمایی برابر است با

$$\ell(\lambda) = -n \log(2\hat{\theta}) + \sum_{i=1}^n \log\left(1 - \frac{\lambda x_i}{\hat{\theta}}\right), \quad |\lambda| \leq 1.$$

با مشتق گیری از معادله فوق نسبت به λ داریم:

$$\frac{\partial \ell(\theta, \lambda)}{\partial \lambda} = \sum_{i=1}^n \frac{x_i}{(\lambda x_i - \hat{\theta})} = 0.$$

برآورد ماکسیمم درستنمایی برای λ جواب معادله فوق است که به روشهای عددی حاصل می شود.

۴ مطالعه شبیه سازی

برای شبیه سازی از توزیع مورد نظر می توان از دو روش استفاده کرد:

روش اول: ابتدا در هر مرحله نمونه تصادفی دوتایی x_1 و x_2 از توزیع $U(-\theta, \theta)$ (θ معلوم) را تولید می شود، سپس مینیمم و ماکسیمم دو نمونه تولید شده با احتمال $\frac{1+\lambda}{2}$ به صورت زیر ترکیب می شود:

$$x = \frac{1+\lambda}{2}y_1 + \frac{1-\lambda}{2}y_2,$$

که در آن $y_1 = \min(x_1, x_2)$ و $y_2 = \max(x_1, x_2)$.

روش دوم: با توجه به پیوسته بودن توزیع و استفاده از قضیه تبدیل انتگرال احتمال، کافی است از حل معادله $F(x) = u$ ، که در آن u عددی تصادفی بین صفر و یک است، استفاده شود. در این صورت خواهیم داشت

$$\left(\frac{x+\theta}{2\theta}\right)\left(1 + \lambda \frac{\theta-x}{2\theta}\right) = u.$$

پس از حل معادله فوق، x به صورت زیر حاصل می‌شود:

$$x = \frac{-\theta \pm \theta \sqrt{1 + \lambda(2 + \lambda - 4u)}}{-\lambda},$$

از آن جایی که ریشه $x = \frac{\theta(1 + \sqrt{1 + \lambda(2 + \lambda - 4u)})}{\lambda}$ به ازای $\lambda > 0$ ، بزرگ‌تر از θ و برای $\lambda < 0$ کوچک‌تر از θ است و در بازه $(-\theta, \theta)$ قرار نمی‌گیرد، غیر قابل قبول می‌باشد. در نهایت جواب مورد قبول x از معادله فوق برابر است با

$$x = \frac{\theta(1 - \sqrt{1 + \lambda(2 + \lambda - 4u)})}{\lambda}.$$

در این بخش با استفاده از شبیه‌سازی یک نمونه تصادفی به حجم معلوم از روش دوم به مقایسه برآورد پارامترها پرداخته می‌شود. در این جا برای شبیه‌سازی از حجم نمونه‌های $n = 20, 50, 70, 100$ و مقادیر مختلف پارامترها استفاده می‌شود. برنامه شبیه‌سازی با نرم‌افزار R تهیه و به ازای مقادیر $\theta = 2, 5$ و $\lambda = -0/9, -0/5, -0/1, 0/1, 0/5, 0/9$ با ۱۰۰۰۰ بار تکرار اجرا شده‌اند. به ازای هر یک از پارامترها و در هر یک از حجم‌های نمونه برآورد پارامترها به دو روش گشتاوری و ماکسیمم درست‌نمایی محاسبه شده‌اند و میانگین توان‌های دوم خطا (MSE)، برای برآوردگرها محاسبه و در جداول ۱ و ۲ ارائه شده است.

برای نشان دادن برتری توزیع چوله‌یکنواخت جدید، نتایج حاصل از برآوردن توزیع چوله‌یکنواخت جدید به یک سری از داده‌ها با زیر مدل آن یعنی توزیع یکنواخت مقایسه می‌شود. برای این منظور نمونه‌ای تصادفی به حجم ۱۰۰ از توزیع چوله‌یکنواخت جدید با پارامترهای $\lambda = 0/1$ و $\theta = 2$ تولید کرده، سپس برآوردگرهای ماکسیمم درست‌نمایی پارامترهای دو مدل محاسبه شده‌اند. جدول ۳ مقادیر ماکسیمم درست‌نمایی پارامترهای دو مدل، مقادیر آماره کولموگروف-اسمیرنوف^۷ به همراه p -مقدارهای مربوطه، $2 \log L$ و مقادیر آماره‌های کرامرون-میسز^۸ و واتسون^۹ را نشان می‌دهد. این مقادیر نشان می‌دهند

^۷ Kolmogorov-Smirnov (K-S)

^۸ Cramer-von Mises (CM)

^۹ Watson (WA)

جدول ۱: برآورد گشتاوری پارامترها به ازای مقادیر مختلف λ و θ

$MSE(\tilde{\theta})$	$MSE(\tilde{\lambda})$	$\tilde{\theta}$	$\tilde{\lambda}$	θ	λ	n
۰/۰۴۲	۰/۰۹۰	۱/۹۹۲	-۰/۸۹۶	۲	-۰/۹	۲۰
۰/۰۱۵	۰/۰۳۴	۱/۹۹۴	-۰/۹۰۰	۲	-۰/۹	۵۰
۰/۰۱۲	۰/۰۲۵	۱/۹۹۴	-۰/۸۹۸	۲	-۰/۹	۷۰
۰/۰۰۸	۰/۰۱۹	۲/۰۰۱	-۰/۹۰۰	۲	-۰/۹	۱۰۰
۰/۰۴۵	۰/۱۲۸	۱/۹۸۸	-۰/۴۹۹	۲	-۰/۵	۲۰
۰/۰۱۷	۰/۰۵۶	۲/۰۰۴	-۰/۵۰۲	۲	-۰/۵	۵۰
۰/۰۱۲	۰/۰۳۷	۱/۹۹۷	-۰/۵۰۳	۲	-۰/۵	۷۰
۰/۰۰۸	۰/۰۲۷	۱/۹۹۸	-۰/۴۹۹	۲	-۰/۵	۱۰۰
۰/۰۳۹	۰/۱۴۶	۱/۹۸۵	-۰/۰۸۶	۲	-۰/۱	۲۰
۰/۰۱۷	۰/۰۶۰	۱/۹۹۰	-۰/۰۹۹	۲	-۰/۱	۵۰
۰/۰۱۱	۰/۰۴۲	۱/۹۹۸	-۰/۱۰۲	۲	-۰/۱	۷۰
۰/۰۰۸	۰/۰۳۰	۱/۹۹۴	-۰/۱۰۲	۲	-۰/۱	۱۰۰
۰/۰۴۰	۰/۱۵۵	۱/۹۹۴	۰/۱۱۴	۲	۰/۱	۲۰
۰/۰۱۵	۰/۰۶۵	۱/۹۹۴	۰/۱۰۹	۲	۰/۱	۵۰
۰/۰۱۱	۰/۰۴۳	۱/۹۹۷	۰/۱۰۷	۲	۰/۱	۷۰
۰/۰۰۸	۰/۰۲۹	۱/۹۹۷	۰/۱۰۳	۲	۰/۱	۱۰۰
۰/۰۳۷۵	۰/۱۳۶	۱/۹۸۵	۰/۴۸۳	۲	۰/۵	۲۰
۰/۰۱۶	۰/۰۵۱	۱/۹۹۶	۰/۴۸۸	۲	۰/۵	۵۰
۰/۰۱۱	۰/۰۳۹	۱/۹۹۴	۰/۴۹۳	۲	۰/۵	۷۰
۰/۰۰۸	۰/۰۲۹	۱/۹۹۶	۰/۴۹۶	۲	۰/۵	۱۰۰
۰/۰۳۷	۰/۰۸۲	۱/۹۹۴	۰/۸۹۹	۲	۰/۹	۲۰
۰/۰۱۵	۰/۰۳۴	۱/۹۹۶	۰/۹۰۲	۲	۰/۹	۵۰
۰/۰۱۰	۰/۰۲۴	۱/۹۹۵	۰/۸۹۹	۲	۰/۹	۷۰
۰/۰۰۸	۰/۰۱۶	۲/۰۰۰	۰/۹۰۰	۲	۰/۹	۱۰۰
۰/۲۴۴	۰/۰۹۰	۴/۹۸۹	-۰/۸۹۸	۵	-۰/۹	۲۰
۰/۰۱۰۲	۰/۰۳۵	۴/۹۸۴	-۰/۸۹۲	۵	-۰/۹	۵۰
۰/۰۷۱	۰/۰۲۶	۴/۹۷۹	-۰/۸۹۰	۵	-۰/۹	۷۰
۰/۰۵۴	۰/۰۱۷	۵/۰۰۳	-۰/۹۰۴	۵	-۰/۹	۱۰۰
۰/۲۳۴	۰/۱۲۶	۴/۹۷۴	-۰/۴۹۸	۵	-۰/۵	۲۰
۰/۰۹۹	۰/۰۵۰	۴/۹۸۸	-۰/۵۰۵	۵	-۰/۵	۵۰
۰/۰۷۵	۰/۰۳۹	۵/۰۰۲	-۰/۴۹۳	۵	-۰/۵	۷۰
۰/۰۵۰	۰/۰۲۶	۴/۹۹۷	-۰/۵۰۰	۵	-۰/۵	۱۰۰
۰/۲۶۶	۰/۱۵۲	۴/۹۷۵	-۰/۰۸۸	۵	-۰/۱	۲۰
۰/۰۱۰۳	۰/۰۶۳	۴/۹۹۲	-۰/۰۸۹	۵	-۰/۱	۵۰
۰/۰۷۲	۰/۰۴۰	۴/۹۸۸	-۰/۱۰۵	۵	-۰/۱	۷۰
۰/۰۵۲	۰/۰۲۷	۴/۹۹۱	-۰/۰۹۹	۵	-۰/۱	۱۰۰
۰/۲۵۹	۰/۱۶۱	۴/۹۹۸	۰/۰۹۷	۵	۰/۱	۲۰
۰/۰۱۰۹	۰/۰۶۲	۴/۹۹۴	۰/۰۹۷	۵	۰/۱	۵۰
۰/۰۷۰	۰/۰۳۹	۴/۹۸۵	۰/۰۹۸	۵	۰/۱	۷۰
۰/۰۴۹	۰/۰۲۹	۵/۰۱۱	۰/۱۰۲	۵	۰/۱	۱۰۰
۰/۲۴۴	۰/۱۲۹	۴/۹۸۶	۰/۴۹۸	۵	۰/۵	۲۰
۰/۰۹۵	۰/۰۵۱	۴/۹۸۹	۰/۴۹۰	۵	۰/۵	۵۰
۰/۰۷۶	۰/۰۳۹	۴/۹۹۰	۰/۴۹۴	۵	۰/۵	۷۰
۰/۰۵۰	۰/۰۲۵	۴/۹۹۹	۰/۵۰۶	۵	۰/۵	۱۰۰
۰/۲۴۳	۰/۰۷۶	۴/۹۸۷	۰/۸۹۶	۵	۰/۹	۲۰
۰/۰۹۲	۰/۰۳۵	۴/۹۸۲	۰/۹۰۳	۵	۰/۹	۵۰
۰/۰۷۱	۰/۰۲۴	۴/۹۷۱	۰/۸۸۹	۵	۰/۹	۷۰
۰/۰۵۳	۰/۰۱۷	۴/۹۷۹	۰/۸۹۴	۵	۰/۹	۱۰۰

جدول ۲: برآورد ماکسیمم درست‌نمایی پارامترها به ازای مقادیر مختلف λ و θ

$MSE(\hat{\theta})$	$MSE(\hat{\lambda})$	$\hat{\theta}$	$\hat{\lambda}$	θ	λ	n
۰/۰۱۶	۰/۰۵۸	۱/۹۰۶	-۰/۸۲۰	۲	-۰/۹	۲۰
۰/۰۰۳	۰/۰۲۲	۱/۹۵۹	-۰/۸۶۷	۲	-۰/۹	۵۰
۰/۰۰۲	۰/۰۱۶	۱/۹۷۲	-۰/۸۷۱	۲	-۰/۹	۷۰
۸e-۰۴	۰/۰۱۲	۱/۹۸۰	-۰/۸۸۶	۲	-۰/۹	۱۰۰
۰/۰۱۸	۰/۱۱۶	۱/۹۰۴	-۰/۴۷۹	۲	-۰/۵	۲۰
۰/۰۰۳	۰/۰۵۴	۱/۹۶۰	-۰/۴۹۴	۲	-۰/۵	۵۰
۰/۰۰۱	۰/۰۳۵	۱/۹۷۳	-۰/۴۹۳	۲	-۰/۵	۷۰
۹e-۰۴	۰/۰۲۵	۱/۹۷۹	-۰/۴۹۷	۲	-۰/۵	۱۰۰
۰/۰۱۶	۰/۱۴۰	۱/۹۰۹	-۰/۰۸۳	۲	-۰/۱	۲۰
۰/۰۰۳	۰/۰۶۰	۱/۹۶۱	-۰/۰۹۷	۲	-۰/۱	۵۰
۰/۰۰۱	۰/۰۴۲	۱/۹۷۴	-۰/۱۰۰	۲	-۰/۱	۷۰
۸e-۰۴	۰/۰۳۰	۱/۹۸۰	-۰/۱۰۰	۲	-۰/۱	۱۰۰
۰/۰۱۷	۰/۱۴۹	۱/۹۰۷	۰/۱۱۱	۲	۰/۱	۲۰
۰/۰۰۳	۰/۰۶۴	۱/۹۵۹	۰/۱۰۷	۲	۰/۱	۵۰
۰/۰۰۲	۰/۰۴۳	۱/۹۷۲	۰/۱۰۶	۲	۰/۱	۷۰
۸e-۰۴	۰/۰۲۸	۱/۹۸۰	۰/۱۰۳	۲	۰/۱	۱۰۰
۰/۰۱۶	۰/۱۲۴	۱/۹۰۷	۰/۴۶۲	۲	۰/۵	۲۰
۰/۰۰۳	۰/۰۵۰	۱/۹۶۳	۰/۴۸۲	۲	۰/۵	۵۰
۰/۰۰۲	۰/۰۳۸	۱/۹۷۲	۰/۴۸۷	۲	۰/۵	۷۰
۹e-۰۴	۰/۰۲۸	۱/۹۷۸	۰/۴۹۲	۲	۰/۵	۱۰۰
۰/۰۱۵۶	۰/۰۵۱۵	۱/۹۰۹۴	-۰/۸۲۹۴	۲	۰/۹	۲۰
۰/۰۰۲۹	۰/۰۲۱۸	۱/۹۶۱۴	-۰/۸۷۴۵	۲	۰/۹	۵۰
۰/۰۰۱۷	۰/۰۱۵۲	۱/۹۷۰۴	-۰/۸۷۷۵	۲	۰/۹	۷۰
۷e-۰۴	۰/۰۱۱	۱/۹۸۰	-۰/۸۸۴	۲	۰/۹	۱۰۰
۰/۱۰۴	۰/۰۵۶	۴/۷۶۹	-۰/۸۲۹	۵	-۰/۹	۲۰
۰/۰۲۱	۰/۰۲۲	۴/۹۰۰	-۰/۸۶۵	۵	-۰/۹	۵۰
۰/۰۱۰	۰/۰۱۷	۴/۹۳۰	-۰/۸۷۱	۵	-۰/۹	۷۰
۰/۰۰۴	۰/۰۱۱	۴/۹۵۴	-۰/۸۸۶	۵	-۰/۹	۱۰۰
۰/۱۰۰	۰/۱۱۴	۴/۷۶۷	-۰/۴۷۷	۵	-۰/۵	۲۰
۰/۰۱۹	۰/۰۴۸	۴/۹۰۲	-۰/۴۷۶	۵	-۰/۵	۵۰
۰/۰۱۰	۰/۰۳۷	۴/۹۳۰	-۰/۴۸۶	۵	-۰/۵	۷۰
۰/۰۰۴	۰/۰۲۵	۴/۹۵۲	-۰/۴۹۶	۵	-۰/۵	۱۰۰
۰/۱۲۰	۰/۱۴۴	۴/۷۵۵	-۰/۰۸۴	۵	-۰/۱	۲۰
۰/۰۱۶	۰/۰۶۲	۴/۹۰۸	-۰/۰۸۷	۵	-۰/۱	۵۰
۰/۰۱۰	۰/۰۴۰	۴/۹۲۹	-۰/۱۰۴	۵	-۰/۱	۷۰
۰/۰۰۵	۰/۰۲۷	۴/۹۴۹	-۰/۰۹۸	۵	-۰/۱	۱۰۰
۰/۱۰۶	۰/۱۵۲	۴/۷۶۷	۰/۰۹۴	۵	۰/۱	۲۰
۰/۰۱۸	۰/۰۶۰	۴/۹۰۳	۰/۰۹۵	۵	۰/۱	۵۰
۰/۰۱۰	۰/۰۳۹	۴/۹۲۹	۰/۰۹۵	۵	۰/۱	۷۰
۰/۰۰۵	۰/۰۲۸	۴/۹۵۲	۰/۱۰۱	۵	۰/۱	۱۰۰
۰/۰۹۵	۰/۱۱۵	۴/۷۷۲	۰/۴۷۶	۵	۰/۵	۲۰
۰/۰۱۹	۰/۰۴۹	۴/۹۰۳	۰/۴۸۱	۵	۰/۵	۵۰
۰/۰۱۰	۰/۰۳۷	۴/۹۳۱	۰/۴۸۹	۵	۰/۵	۷۰
۰/۰۰۴	۰/۰۲۴	۴/۹۵۲	۰/۵۰۲	۵	۰/۵	۱۰۰
۰/۱۰۶	۰/۰۴۹	۴/۷۷۰	۰/۸۳۱	۵	۰/۹	۲۰
۰/۰۱۸	۰/۰۲۱	۴/۹۰۳	۰/۸۶۸	۵	۰/۹	۵۰
۰/۰۱۱	۰/۰۱۷	۴/۹۲۶	۰/۸۷۴	۵	۰/۹	۷۰
۰/۰۰۵	۰/۰۱۲	۴/۹۴۹	۰/۸۸۴	۵	۰/۹	۱۰۰

جدول ۳: مقایسه توزیع چوله یکنواخت جدید و توزیع یکنواخت

Model	MLEs	$-2 \log L$	K-S	CM	WA	p-value
NSU	$\hat{\theta} = 1/992$ $\hat{\lambda} = 0/049$	276/388	0/049	0/131	0/131	0/969
Uniform	$\hat{\theta} = 1/992$	276/473	0/054	0/148	0/141	0/937

که توزیع چوله یکنواخت جدید نسبت به توزیع یکنواخت از برازش بهتری برای این سری از داده‌ها برخوردار است.

۵ بحث و نتیجه‌گیری

در این مقاله توزیع چوله یکنواخت جدید در بازه $(-\theta, \theta)$ به عنوان تعمیمی از توزیع یکنواخت در بازه $(-\theta, \theta)$ معرفی شده است. این توزیع می‌تواند کاربردهای بسیاری را در زمینه‌های آزمون فرض، آزمون‌های ناپارامتری، فیزیک، تولید اعداد تصادفی و نمونه‌گیری طبقه‌بندی به خود اختصاص دهد. در این مقاله خواص مهم توزیع از جمله فرم تابع چگالی، تابع توزیع و تابع نرخ شکست، گشتاورهای مرتبه k ، تابع مولد گشتاور، تابع مشخصه، واریانس، ضرایب چولگی و کشیدگی، میانگین انحرافات از میانگین، میانگین و مد به دست آمده و روش‌های مختلف برآوردیابی بررسی شده است. همچنین برای بررسی سازگاری برآوردهای حاصل از روش گشتاوری و روش ماکسیمم درست‌نمایی، شبیه‌سازی‌هایی انجام شده است. در پایان برای نشان دادن برتری توزیع چوله یکنواخت جدید نسبت به توزیع یکنواخت با استفاده از شبیه‌سازی مقایسه‌هایی صورت گرفته است.

تشکر و قدردانی

نویسندگان مقاله نهایت تشکر و قدردانی را از داوران و سردبیر محترم نشریه، که با رهنمودهای ارزنده خود باعث بهتر شدن مقاله گردیده‌اند، دارند. همچنین

نویسندگان از دانشگاه یزد به خاطر حمایت‌های مالی تقدیر و تشکر می‌کنند.

مراجع

- Azzalini, A. (1985), A Class of Distributions which Includes the Normal Ones, *Scandinavian Journal of Statistics*, **12**, 171-178.
- Drouet-Mari, D. and Kotz, S. (2001), *Correlation and Dependence*, Imperial College Press: London.
- Gupta, A. K., Chang, F. C. and Huang, W. J. (2002), Some Skew-Symmetric Models, *Random Operators and Stochastic Equations*, **10**, 133-140.
- Jones, M. C. (2004), Families of Distributions Arising from Distributions of Order Statistics (with Discussion), *Test*, **13**, 1-43.
- Marshall, A. W. and Olkin, I. (1997), Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families, *Biometrika*, **84**, 641-652.
- Mirhoseini, S. M., Dolati, A. and Amini, M. (2011), On a Class of Distributions Generated by Stochastic Mixture of the Extreme Order Statistics of a Sample of Size Two, *Journal of Statistical Theory and Applications*, **10**, 455-468.
- Nadarajah, S. and Aryal, G. (2004), On the Skew Uniform Distribution, *Random Operators and Stochastic Equations*, **12**, 319-330.
- Nadarajah, S. and Kotz, S. (2007), Skew Models I, *Acta Applicandae Mathematica*, **98**, 1-28.