

توزیع پارتو یکنواخت و توزیع پارتو یکنواخت نمایی شده و کاربردهای آنها در داده‌های درآمد

آنیثا عبداللهی نانوایشه^۱، آزیثا عبداللهی نانوایشه^۲

تاریخ دریافت: ۱۳۹۷/۱/۲۵

تاریخ پذیرش: ۱۳۹۷/۸/۲۴

چکیده:

در این مقاله، یک توزیع جدید معرفی می‌گردد که تعمیمی از یک توزیع شناخته شده است. این توزیع انعطاف‌پذیر بوده و در مدل‌بندی داده‌های درآمد کاربرد دارد. در ابتدا برخی از ویژگی‌های ریاضی و توزیعی این الگوی جدید ارائه می‌شود و سپس برای نشان دادن انعطاف‌پذیری توزیع جدید، کاربردهایی از این توزیع در داده‌های واقعی ارائه خواهد شد. نتایج برازش داده‌ها نیز مناسب بودن این الگوی جدید را برای مجموعه داده‌های واقعی در نظر گرفته شده تأیید می‌کنند.

واژه‌های کلیدی: توزیع پارتو، توزیع پارتو یکنواخت نمایی شده، چندک، گشتاورها، آنتروپی، معیار قابلیت اطمینان.

۱ مقدمه

یک توزیع مناسب برای یک مجموعه از داده‌های خاص به ما کمک خواهد کرد تا استدلال‌هایی در باره داده‌ها داشته باشیم. به همین دلیل، استفاده از روش‌هایی برای اصلاح توزیع آماری موجود به منظور ایجاد انعطاف‌پذیری بیشتر در مجموعه داده‌ها در زمینه‌های مختلف صورت گرفته است. توزیع پارتو یکنواخت [۱، ۲] مانند سایر توزیع‌های آماری به تازگی به دلیل کاربرد آن در مدل‌سازی توزیع میزان درآمد شخصی و قابلیت اطمینان و سایر زمینه‌ها توجه زیادی را به خود جلب کرده است. برای بررسی جزئیات انجام شده در این زمینه به منابع [۳، ۴] مراجعه شود. با هدف افزایش انعطاف‌پذیری توزیع پارتو در مدل‌بندی داده‌های درآمد، تغییرات مختلفی در توزیع داده‌اند [۵، ۶]. در این مقاله بسط جدیدی از توزیع پارتو به نام توزیع پارتو یکنواخت نمایی شده، با انگیزه اصلی مدل‌سازی داده‌های درآمد با هر دو نرخ خطر یکنواخت و غیریکنواخت ارائه شده است. در ادامه، مقاله به صورت زیر سازمان‌دهی شده است: در بخش (۳) ابتدا تابع توزیع تجمعی، تابع چگالی احتمال، تابع بقا و تابع خطر توزیع پارتو یکنواخت و توزیع پارتو یکنواخت نمایی شده تعریف

با توجه به این که آمار یکی از مهم‌ترین علوم کاربردی است که با سایر رشته‌های علمی از جمله اقتصاد مرتبط است، از این رو در سال‌های اخیر، علاقه زیادی برای تحقیقات در زمینه مدل‌های پارامتری توزیع درآمد به وجود آمده است. مدل‌های احتمالی مربوط به توزیع درآمد، برای ارزیابی استانداردهای سطح زندگی کل مردم یک کشور و همچنین برای مقایسه استاندارد سطح زندگی طبقات اجتماعی و یا مناطق مختلف یک کشور ارائه شده‌اند. بنا بر این برای ایجاد یک مدل احتمالی ارائه یک تابع توزیع نظری با مشخصه توزیع فراوانی تجربی برای انتخاب روش مناسب تخمین پارامترهای مدل، ضروری است. بنا بر این تحلیل آماری توزیع درآمد جمعیت نشان‌دهنده زمینه لازم برای تصمیم‌گیری در مورد بودجه و سیاست‌های اجتماعی است. در این مقاله توزیع جدیدی ارائه شده است که علاوه بر انعطاف‌پذیری یک توزیع مناسب برای مدل‌سازی داده‌های درآمد است. شناسایی یک توزیع مناسب برای مدل‌بندی مجموعه داده‌ها در تحلیل‌های آماری بسیار با اهمیت است. مشخص کردن

^۱ دانش‌آموخته کارشناسی ارشد آمار، دانشگاه علامه طباطبائی، ایران

^۲ دانشجوی دکتری مکانیک، دانشگاه تهران، ایران

می‌شود. در بخش (۴)، خواص آماری توزیع جدید مورد بحث قرار می‌گیرد و در بخش (۵)، پارامترهای توزیع جدید با استفاده از روش برآورد ماکسیمم درست‌نمایی محاسبه و بیان می‌شوند و در نهایت در بخش (۵) و (۶) کاربردهایی از مدل جدید با استفاده از دو مجموعه از داده‌های واقعی و نتایج شبیه‌سازی ارائه می‌شود.

۲ مدل جدید

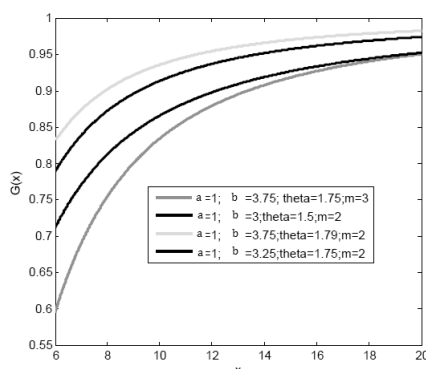
فرض کنید X یک متغیر تصادفی با توزیع پارتو یکنواخت باشد (U-PD) در این صورت تابع توزیع تجمعی آن به صورت زیر است [۷]:

$$G(X) = 1 - \frac{b(\frac{m}{x})^\theta}{b-a}, \quad m(\frac{b-a}{b})^{-\frac{1}{\theta}} < x < \infty \quad (1)$$

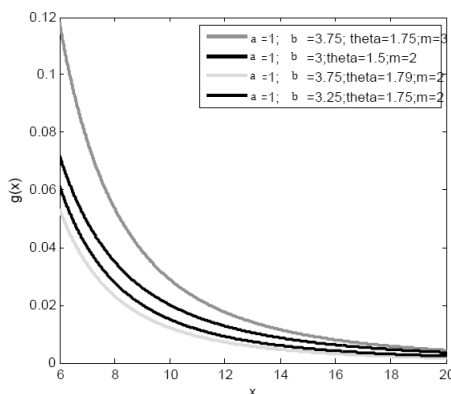
که $-\infty < a \leq b < \infty$ پارامترهای مقیاس و $\theta > 0$ پارامتر شکل است. شکل ۱ نمودار تابع توزیع تجمعی این توزیع

$$S(x) = \frac{b(\frac{m}{x})^\theta}{b-a},$$

$$h(x) = \frac{\frac{b\theta m^\theta}{b-a x^{\theta+1}}}{\frac{b(\frac{m}{x})^\theta}{b-a}}.$$



شکل ۱. تابع توزیع تجمعی پارتو یکنواخت به‌ازای مقادیر مختلف پارامترها



شکل ۲. تابع چگالی احتمال توزیع (U-PD) به‌ازای مقادیر مختلف پارامترها

تابع نرخ خطر این توزیع نیز به صورت زیر است:

$$h(x) = \frac{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}}\right] \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a}\right]^{\gamma-1}}{1 - \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a}\right]}. \quad (5)$$

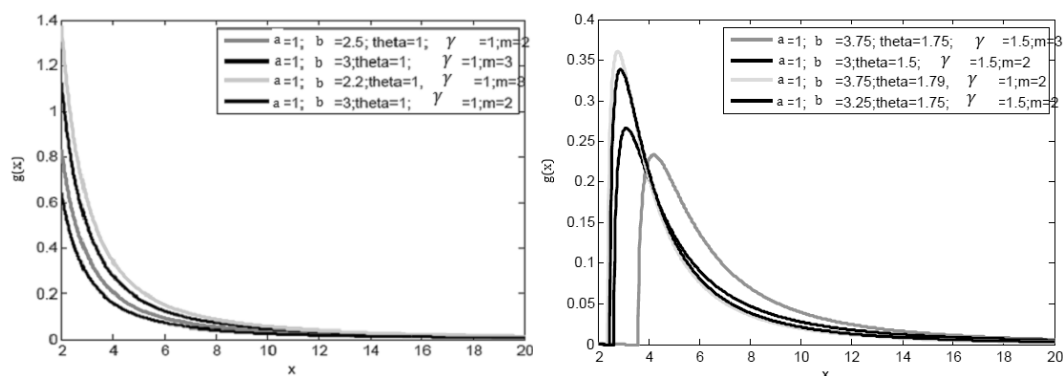
در شکل‌های زیر نمودار تابع چگالی احتمال و تابع نرخ خطر توزیع جدید به‌ازای مقادیر مختلف پارامترها رسم شده است. همان‌طور که مشخص است تابع نرخ خطر این توزیع بر حسب مقادیر پارامترهایش به صورت وانی شکل و وانی شکل وارون است.

شکل ۲ نمودار تابع چگالی احتمال این توزیع را نشان می‌دهد، که تابع چگالی احتمال این توزیع به صورت یکنواخت غیر افزایشی است. حال فرض کنید X دارای تابع توزیع تجمعی به صورت رابطه (۱) باشد، در این صورت تابع توزیع توانی شده آن (EU-PD) به صورت زیر است:

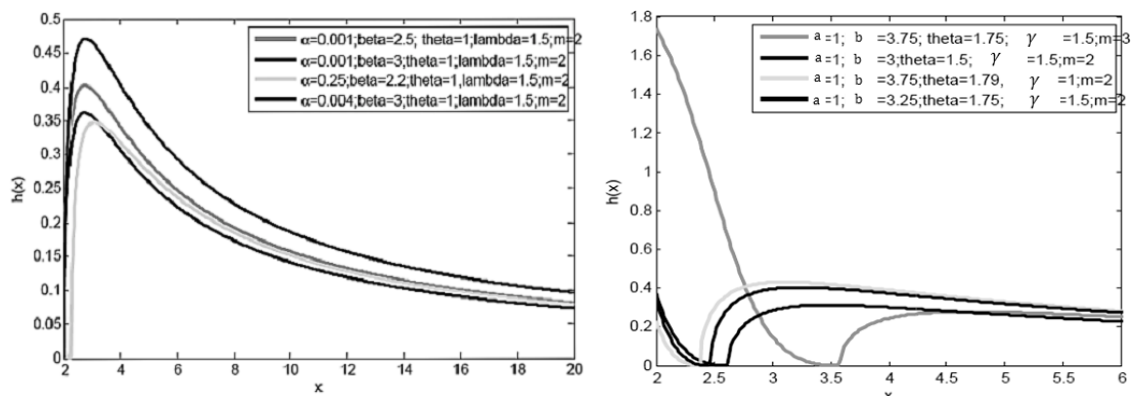
$$G(X) = \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a}\right]^\gamma, \quad m\left(\frac{b-a}{b}\right)^{-\frac{1}{\theta}} < x < \infty \quad (3)$$

تابع چگالی احتمال این توزیع با مشتق‌گیری از رابطه بالا به صورت زیر به دست می‌آید:

$$g(x) = \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}}\right] \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a}\right]^{\gamma-1}, \quad m\left(\frac{b-a}{b}\right)^{-\frac{1}{\theta}} < x < \infty \quad (4)$$



(آ) تابع چگالی احتمال توزیع (EU-PD) به‌ازای مقادیر مختلف پارامترها (ب) تابع چگالی احتمال توزیع (EU-PD) به‌ازای مقادیر مختلف پارامترها



(ج) تابع خطر توزیع (EU-PD) به‌ازای مقادیر مختلف پارامترها (د) تابع خطر توزیع (EU-PD) به‌ازای مقادیر مختلف پارامترها

شکل ۳. نمودار تابع چگالی احتمال و تابع نرخ خطر توزیع جدید

۳ ویژگی‌های آماری توزیع (U-PD) و توزیع (EU-PD)

گشتاورها مورد نیاز است، پرداخته می‌شود. اگر

$$g(X) = \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]^{\gamma-1},$$

در این صورت بسط آن به صورت:

$$g(X) = \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(\frac{m}{x})^\theta)^j}{(b-a)^j} \quad (۶)$$

و به طور مشابه بسط تابع توزیع آن نیز به صورت زیر است:

$$G(X) = \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]^\gamma$$

$$G(X) = \sum_{j=0}^{\infty} \binom{\gamma}{j} \frac{(-1)^j (b(\frac{m}{x})^\theta)^j}{(b-a)^j}. \quad (۷)$$

در این صورت گشتاور مرتبه r -ام توزیع (EU-PD) به صورت زیر محاسبه می‌شود:

$$E[X^r] = \int_w^\infty x^r \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(\frac{m}{x})^\theta)^j}{(b-a)^j} dx$$

$$= \int_w^\infty x^{-\theta j+r-\theta-1} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right] dx$$

$$= \frac{x^{-\theta j+r-\theta}}{-\theta j+r-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right]_x^w,$$

که $w = m(\frac{b-a}{b})^{-\frac{1}{\theta}}$ بنا بر این داریم:

$$E[X^r] = - \frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+r-\theta}}{-\theta j+r-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \times \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right]. \quad (۸)$$

بنا بر این گشتاور اول توزیع جدید به صورت زیر است:

$$E[X] = - \frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+1-\theta}}{-\theta j+1-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \times \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right].$$

به همین ترتیب سایر گشتاورها نیز با استفاده از رابطه (۸) و با دادن مقدار به r به دست می‌آیند. در این بخش به محاسبه گشتاورهای ناقص رابطه (۲) پرداخته می‌شود، که در محاسبه منحنی لورنز و معادلات منحنی بونفرونی، همچنین در محاسبه معیارهای میانگین انحراف از میانگین و میانگین انحراف از میانه

در این بخش برخی از ویژگی‌های توزیع پارتو یکنواخت و توزیع پارتو یکنواخت نمایی شده ارائه می‌شود. اگر X دارای توزیع پارتو یکنواخت باشد، آن گاه میانگین، واریانس، ضریب تغییرات، چولگی و کشیدگی و گشتاور مرتبه r -ام آن به ترتیب به صورت زیر است:

$$E(X) = \frac{b^{\frac{1}{\theta}}}{(b-a)^{\frac{1}{\theta}}} \frac{\theta m}{\theta-1},$$

$$Var(X) = \frac{b^{\frac{1}{\theta}}}{(b-a)^{\frac{1}{\theta}}} \frac{(\theta m^2)}{(\theta-2)} - \left(\frac{b^{\frac{1}{\theta}}}{(b-a)^{\frac{1}{\theta}}} \frac{\theta m}{\theta-1} \right)^2,$$

$$CV = \frac{\sqrt{\frac{b^{\frac{1}{\theta}} \theta m}{(b-a)^{\frac{1}{\theta}}} \left(\frac{(\theta-1)-2\theta-2\mu}{(\theta-2)(\theta-1)} \right) + \mu}}{\frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}} (\theta-1)}},$$

$$CS = \frac{\frac{b^{\frac{1}{\theta}} (\theta m^3)}{(b-a)^{\frac{1}{\theta}} (\theta-3)} - 3\mu \frac{b^{\frac{1}{\theta}} (\theta m^2)}{(b-a)^{\frac{1}{\theta}} (\theta-2)}}{\left(\frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}}} \left(\frac{(\theta-1)-2\theta-2\mu}{(\theta-2)(\theta-1)} \right) + \mu^2 \right)^2} + \frac{2\mu^2 \frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}} (\theta-1)} - \mu^2}{\left(\frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}}} \left(\frac{(\theta-1)-2\theta-2\mu}{(\theta-2)(\theta-1)} \right) + \mu^2 \right)^2},$$

$$CK = \frac{\frac{b^{\frac{1}{\theta}} (\theta m^4)}{(b-a)^{\frac{1}{\theta}} (\theta-4)} - 4\mu \frac{b^{\frac{1}{\theta}} (\theta m^3)}{(b-a)^{\frac{1}{\theta}} (\theta-3)}}{\left(\frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}}} \left(\frac{(\theta-1)-2\theta-2\mu}{(\theta-2)(\theta-1)} \right) + \mu^2 \right)^2} + \frac{6\mu^2 \frac{b^{\frac{1}{\theta}} (\theta m^2)}{(b-a)^{\frac{1}{\theta}} (\theta-2)} - 4\mu^3 \frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}} (\theta-1)} + \mu^4}{\left(\frac{b^{\frac{1}{\theta}} (\theta m)}{(b-a)^{\frac{1}{\theta}}} \left(\frac{(\theta-1)-2\theta-2\mu}{(\theta-2)(\theta-1)} \right) + \mu^2 \right)^2}$$

و گشتاور مرتبه r -ام این توزیع به صورت زیر است:

$$E(X^r) = \frac{b^{r/\theta} (\theta m^r)}{(b-a)^{r/\theta} (\theta-r)}.$$

حال در این بخش برخی از خواص آماری توزیع (EU-PD) مانند گشتاورها و گشتاورهای ناقص، میانگین انحرافها، منحنی لورنز و معادلات منحنی بونفرونی ارائه می‌شوند.

۱.۳ گشتاورها و گشتاورهای ناقص توزیع

در این بخش گشتاورهای مدل معرفی شده (EU-PD) ارائه می‌شوند ولی قبل از آن به ارائه بسط $g(x)$ که در محاسبه

کاربرد دارند. اگر X دارای توزیع رابطه (۲) باشد، آن گاه داریم:

$$\begin{aligned}\int_0^w x^r g(x) dx &= \int_0^w x^r \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] \\ &\times \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(\frac{m}{x})^\theta)^j}{(b-a)^j} dx \\ &= \int_0^w x^{-\theta j+r-\theta-1} \sum_{j=0}^{\infty} \\ &\times \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right] dx \\ &= \frac{x^{-\theta j+r-\theta}}{-\theta j+r-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \\ &\times \left[\frac{b\theta\gamma}{b-a} m^\theta \right]_0^w\end{aligned}$$

که $w = m(\frac{b-a}{b})^{-\frac{1}{\theta}}$ بنا بر این داریم:

$$\begin{aligned}\int_0^w x^r g(x) dx &= \frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+r-\theta}}{-\theta j+r-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \\ &\times \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right]\end{aligned}\quad (9)$$

۲.۳ میانگین انحرافها، منحنی لورنز و معادلات منحنی بونفرونی

اگر X دارای توزیع رابطه (۲) باشد، آن گاه انحراف از میانگین از رابطه زیر به دست می آید:

$$\begin{aligned}\delta_1(x) &= \int_0^\infty |x-M|f(x)dx = 2MF(M) - M + \mu - 2I(M) \\ &= 2M \sum_{j=0}^{\infty} \binom{\gamma}{j} \frac{(-1)^j (b(\frac{m}{M})^\theta)^j}{(b-a)^j} - M + \mu \\ &= -2 \left[\frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+1-\theta}}{-\theta j+1-\theta} \right. \\ &\times \left. \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right] \right]\end{aligned}\quad (10)$$

که $I(b) = \int_0^b xf(x)dx$ و انحراف از میانه نیز به صورت زیر است:

$$\begin{aligned}\delta_2(x) &= \int_0^\infty |x-\mu|f(x)dx = 2\mu F(\mu) - 2I(\mu) \\ &= 2\mu \sum_{j=0}^{\infty} \binom{\gamma}{j} \frac{(-1)^j (b(\frac{m}{\mu})^\theta)^j}{(b-a)^j} \\ &= -2 \left[\frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+1-\theta}}{-\theta j+1-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \right. \\ &\times \left. \left[\frac{b\theta\gamma}{b-a} m^\theta \right] \right]\end{aligned}\quad (11)$$

در این قسمت معادلات منحنی لورنز و معادلات منحنی بونفرونی ارائه خواهد شد که کاربردهای زیادی در اقتصاد برای بررسی درآمد و فقر، قابلیت اطمینان، پزشکی و بیمه دارند. منحنی لورنز، ارتباط بین نسبت تجمعی دارندگان درآمد و نسبت تجمعی درآمد دریافت شده توسط آنان را نشان می دهد، به شرط آن که دارندگان درآمد بر حسب میزان درآمدشان به ترتیب صعودی مرتب شده باشند. به این ترتیب، هر نقطه از منحنی لورنز مبین سهمی از کل درآمد جامعه است که توسط نسبتی از افراد جامعه کسب شده است که دارای مقدار معینی درآمد و یا کمتر از آن هستند. اگر X دارای توزیع ارائه شده در رابطه (۲) باشد، آن گاه معادلات منحنی بونفرونی به صورت زیر است:

$$\begin{aligned}B_F[F(x)] &= \frac{1}{\mu F(x)} \int_0^x uf(u)du \\ &= \frac{1}{\mu \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]^\gamma} \left[\frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+1-\theta}}{-\theta j+1-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right] \right]\end{aligned}\quad (12)$$

و منحنی لورنز نیز به صورت زیر خواهد بود:

$$\begin{aligned}L_F[F(x)] &= B_F[F(x)] = \frac{1}{\mu} \int_0^x uf(u)du \\ &= \frac{1}{\mu \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]^\gamma} \left(\frac{(m(\frac{b-a}{b})^{-\frac{1}{\theta}})^{-\theta j+1-\theta}}{-\theta j+1-\theta} \sum_{j=0}^{\infty} \binom{\gamma-1}{j} \right. \\ &\times \left. \frac{(-1)^j (b(m)^\theta)^j}{(b-a)^j} \left[\frac{b\theta\gamma}{b-a} m^\theta \right] \right)\end{aligned}\quad (13)$$

۴ تابع ماکسیم درست نمایی (MLE)

اگر $X_1, \dots, X_n$ نمونه ای تصادفی از توزیع (EU-PD) با اندازه n باشند، آن گاه تابع ماکسیم درست نمایی این توزیع به صورت زیر است:

$$\begin{aligned}g(X) &= \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]^{\gamma-1} \\ L &= \prod_{i=1}^n \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x_i^{\theta+1}} \right] \prod_{i=1}^n \left[1 - \frac{b(\frac{m}{x_i})^\theta}{b-a} \right]^{\gamma-1}.\end{aligned}$$

در نتیجه لگاریتم تابع درست نمایی به صورت زیر است:

$$l(a, b, \theta, m, \gamma) = \sum_{i=1}^n \log \left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right] + (\gamma - 1) \sum_{i=1}^n \log \left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right] \quad (14)$$

با مشتق گیری از رابطه (۱۴) نسبت به پارامترها داریم:

$$\begin{aligned} \frac{\partial l_n}{\partial a} &= \sum_{i=1}^n \frac{\frac{b\theta\gamma}{(b-a)^\gamma} \frac{m^\theta}{x^{\theta+1}}}{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right]} + (\gamma - 1) \sum_{i=1}^n \frac{-\frac{b(\frac{m}{x})^\theta}{(b-a)^\gamma}}{\left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]}, \\ \frac{\partial l_n}{\partial b} &= \sum_{i=1}^n \frac{\frac{\theta\gamma(b-a) - b\theta\gamma}{(b-a)^\gamma} \frac{m^\theta}{x^{\theta+1}}}{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right]} + (\gamma - 1) \sum_{i=1}^n \frac{\frac{-(\frac{m}{x})^\theta(b-a) + b(\frac{m}{x})^\theta}{(b-a)^\gamma}}{\left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]}, \\ \frac{\partial l_n}{\partial \theta} &= \sum_{i=1}^n \frac{\frac{(b\gamma)}{(b-a)} \frac{m^\theta \ln m - (x^{\theta+1})(\ln x)(m^\theta)}{x^{\theta+1}}}{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right]} + (\gamma - 1) \sum_{i=1}^n \frac{\frac{b(\frac{m}{x})^\theta(b-a) + (1 \ln \frac{m}{x})}{(b-a)^\gamma}}{\left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]}, \\ \frac{\partial l_n}{\partial m} &= \sum_{i=1}^n \frac{\frac{b\theta\gamma}{(b-a)^\gamma} \frac{\theta m^{\theta-1}}{x^{\theta+1}}}{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right]} + (\gamma - 1) \sum_{i=1}^n \frac{\frac{-(\frac{b}{(b-a)}) \frac{\theta m^{\theta-1}}{x^{\theta+1}}}{(b-a)^\gamma}}{\left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]}, \\ \frac{\partial l_n}{\partial \gamma} &= \sum_{i=1}^n \frac{\frac{b\theta}{(b-a)^\gamma} \frac{m^\theta}{x^{\theta+1}}}{\left[\frac{b\theta\gamma}{b-a} \frac{m^\theta}{x^{\theta+1}} \right]} + \sum_{i=1}^n \frac{\frac{b(\frac{m}{x})^\theta}{(b-a)^\gamma}}{\left[1 - \frac{b(\frac{m}{x})^\theta}{b-a} \right]}. \end{aligned}$$

بر آورد ماکسیم درست نمایی پارامترها با برابر صفر قرار دادن عبارات بالا و با استفاده از حل تابع غیرخطی به دست می آیند که دارای شکل صریحی نیستند. برای ارائه برآورد بازه‌ای، فاصله اطمینان و آزمون فرضیه آماری روی پارامترها، نیاز به ماتریس اطلاع است. ماتریس اطلاع مشاهده شده به صورت زیر است:

$$J(\vartheta) = \begin{bmatrix} I_{aa} & I_{ab} & I_{a\theta} & I_{am} & I_{a\gamma} \\ I_{ab} & I_{bb} & I_{b\theta} & I_{bm} & I_{b\gamma} \\ I_{a\theta} & I_{a\theta} & I_{\theta\theta} & I_{\theta m} & I_{\theta\gamma} \\ I_{am} & I_{bm} & I_{\theta m} & I_{mm} & I_{m\gamma} \\ I_{a\gamma} & I_{b\gamma} & I_{\theta\gamma} & I_{m\gamma} & I_{\gamma\gamma} \end{bmatrix}$$

که $\vartheta = (a, b, \theta, m, \gamma)'$ عبارات واضحی برای عناصر ماتریس اطلاع مشاهده شده وجود دارد که از مشتق دوم لگاریتم تابع درست نمایی نسبت به پارامترها به دست می آیند. تحت شرایطی که برای پارامترها در خارج از فضای پارامتری انجام می شود اما نه روی مرزها، توزیع مجانبی $\sqrt{n}(\hat{\Theta} - \Theta)$ همگرا به توزیع $N_0(0, I^{-1}(\vartheta))$ است که $I(\vartheta)$ ماتریس اطلاع مورد انتظار است. استنباط مجانبی برای بردار پارامترها تحت شرایط نظم [۱۴]، بر

۵ کاربردها

در این بخش کاربردهایی از توزیع ارائه شده در بخش های قبل بیان می شود و نشان داده می شود که توزیع (EU-PD) برازش بهتری نسبت به سایر مدل ها برای مجموعه داده های درآمد فراهم می کند. یکی از معیارها و شاخص های تعیین رفاه جامعه در اقتصاد متکی به بازار آزاد، درآمد ملی است. اگرچه این فاکتور بیان کننده توزیع مناسب و عادلانه درآمد در جامعه نیست، اما خود، به نوعی بیان کننده واقعیت درآمد جامعه است. درآمد ملی عبارتست از مجموع درآمدهای کسب شده اتباع یک کشور، در چارچوب یک اقتصاد ملی در یک دوره معین اقتصادی پس از کسر سرمایه های مصرفی. (نظریه u وارون که توسط کوزنتس (۱۹۵۵) معرفی شد اساساً سطح فعالیت اقتصادی را به عنوان عامل تعیین کننده اصلی نابرابری درآمدی مشخص می کند. به طور دقیق تر این امر تأکید دارد که در مرحله اولیه رشد اقتصادی، نابرابری درآمدی بدتر می شود و فقط در مراحل بعدی بهبود می یابد. کوزنتس سطح فعالیت درآمد یا فعالیت اقتصادی را به عنوان عامل اصلی نابرابری درآمدی شناسایی کرد. وی تأکید کرد در مراحل اولیه توسعه نابرابری درآمد بدتر می شود و زمانی که نیروی کار از مناطق روستایی به مناطق شهری مهاجرت می کنند، بهتر می شود. از آنجا که الگوی حرکت نابرابری در طول زمان شبیه به شکل u وارون است، به عنوان الگوی u وارون شناخته می شود) [۸]. مجموعه داده های ارائه شده، داده های درآمد ملی سال های (۱۳۹۳ - ۱۳۳۸) (بر حسب میلیارد ریال) مربوط به کشور ایران است که از سایت بانک مرکزی جمهوری اسلامی ایران گرفته شده است.

۱۶۶۰۸۴/۹، ۱۵۳۷۸۹/۴، ۱۳۶۹۸۱/۴، ۱۲۴۲۹۸/۴، ۱۱۱۲۵۷/۸،

کمترین مقدار) مدل ارائه شده بهترین برازش برای مجموعه داده‌ها فراهم می‌کند. شایان ذکر است که توزیع لیندلی ارائه شده در جدول، یک توزیع تک پارامتری است که برای داده‌های طول عمر مناسب بوده و نسبت به بعضی از توزیع‌ها در بسیاری از خواص نظیر تابع چگالی، تابع نرخ شکست و میانگین طول عمر باقی‌مانده از انعطاف بیشتری برخوردار است. تابع چگالی این توزیع به صورت زیر است:

$$f_L(x) = \frac{\beta^2}{1+\beta} (1+x)e^{-\beta x}, \quad x \geq 0, \beta > 0$$

به منظور ارزیابی عملکرد روش ماکسیم درست‌نمایی، مطالعه شبیه‌سازی با تکرار ۱۰۰۰۰ بار با استفاده از نرم‌افزار R انجام می‌شود. ارزیابی برآوردها بر مبنای موارد زیر انجام می‌شود: میانگین توان دوم خطاهای تجربی و برآورد ماکسیم درست‌نمایی به ازای هر داده شبیه‌سازی برای

$$i = 1, 2, \dots, 10000$$

به صورت $(\hat{a}, \hat{b}, \hat{\theta}, \hat{m}, \hat{\gamma})$ تعیین می‌شوند. مقدار اریبی و میانگین توان دوم خطاها از فرمول‌های زیر به دست می‌آیند:

$$bias_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h}_i - h);$$

$$MSE_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h}_i - h)^2$$

۲۳۴۱۹۷/۳، ۲۰۷۱۷۴/۴، ۲۰۲۵۶۴/۷، ۱۸۸۸۲۱/۲، ۱۶۷۴۲۵/۳،
 ۶۸۹۸۷۴/۲، ۴۵۵۲۸۵/۸، ۳۸۸۹۸۵/۷، ۲۸۴۸۳۰/۵، ۲۵۸۹۸۲/۰،
 ۹۹۸۶۲۸/۷، ۴۲۷۲۳۵/۸، ۶۸۵۶۲۶/۳، ۵۴۵۷۰۰/۵، ۷۱۱۳۵۷/۳،
 ۷۱۸۲۲۴/۰، ۸۸۶۰۹۲/۳، ۶۵۱۴۰۹/۶، ۶۳۸۴۲۷/۲، ۵۸۹۲۵۷/۶،
 ۴۸۴۰۵۰/۳، ۸۶۳۲۸۴/۱، ۹۰۰۱۵۱/۸، ۸۵۵۵۶۶/۲، ۸۷۸۱۹۸/۷،
 ۱۰۷۷۶۰۳/۷، ۸۰۰۰۷۷/۵، ۷۵۹۹۸۱/۳، ۶۲۳۹۸۴/۰، ۶۷۴۶۵۶/۱،
 ۱۲۴۳۸۰/۱۰، ۱۱۶۳۰۲۹/۱، ۹۸۹۶۳۷/۳، ۹۵۵۵۲۱/۶، ۹۲۴۲۵۰/۱،
 ۱۵۸۰۳۴۸/۰، ۹۷۷۴۹۲/۵، ۱۱۹۵۴۹۸/۹، ۱۲۵۹۴۶۵/۵، ۱۰۶۹۰۵۱/۵،
 ۱۷۷۸۱۷۷/۴، ۱۷۸۹۹۸۴/۵، ۱۶۰۰۳۴۴/۸، ۱۵۰۷۳۵۹/۸، ۱۳۳۴۴۲۴/۷،
 ۱۶۰۶۳۱۶/۰، ۱۶۴۹۲۴۸/۱، ۱۹۹۰۲۶۱/۳، ۱۸۷۱۶۳۷/۶، ۱۷۰۴۵۶۵/۱.

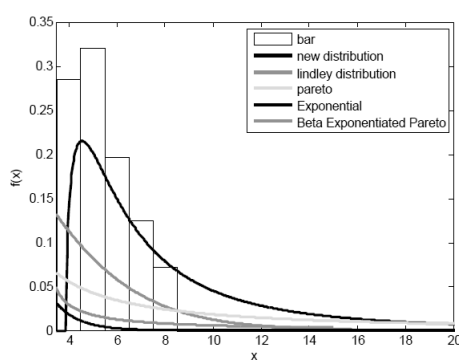
در شکل ۴ نمودار تابع چگالی احتمال چندین توزیع و بافت‌نگاشت داده‌های درآمد، جهت انتخاب بهترین توزیع برای برازش به داده‌ها رسم شده است. همان‌طور که از شکل‌ها مشخص است توزیع (EU-PD)، نسبت به سایر مدل‌ها، برازش بهتری روی این مجموعه داده‌ها فراهم می‌کند. جهت بررسی نیکویی برازش توزیع‌های آماری ارائه شده با توزیع داده‌های واقعی، از آزمون $k-s$ (کلموگروف-اسمیرنوف) و AIC استفاده گردید. در جدول ۱، MLE داده‌ها و آزمون کلموگروف-اسمیرنوف، مقدار AIC و BIC برای مجموعه داده‌های واقعی، ارائه شده‌اند. همان‌طور که مشخص است با توجه به معیارهای AIC (بر اساس

جدول ۱. برآورد ماکسیم درست‌نمایی پارامترها و آزمون نیکویی برازش برای مجموعه داده‌های واقعی

BIC	AIC	(K-S)	برآوردها					مدل
			γ	m	θ	b	a	
۱۶۷/۳۷	۱۵۷/۳۱	۰/۴۹۱۵	۰/۹۶	۰/۳۴	۱/۷۵	۲/۷۵	۱	مدل جدید
۱۷۴/۱	۱۷۳/۲	۱/۳۹۳۷	—	۱	۱	۵۹۴۳/۵۹	۱۸/۷۹	پارتو یکنواخت
۱۰۰۷۲	۵۰۳۶۵	۰/۸۵۷۴	—	—	—	۰/۱۰۶۰	—	لیندلی
۱۵۳۳۰	۱۵۳۱۰	۰/۶۳۱۹	—	—	$۱/۱۷۸ \times ۱۰^{-۵}$	—	—	نمایی
۱۹۶/۶۴	۱۸۹/۵۶	۱/۴۸۳۲	۰/۷۱۵	۰/۲۷۱۹	۱/۰۱۳۳۷	۰/۰۰۰۰۰۰۱	۰/۷۴۹۲۳	بتا پارتو نمایی شده

جدول ۲. مقادیر اریبی و میانگین توان دوم خطاهای برآوردها تحت روش ماکسیمم درست‌نمایی

میانگین توان دوم خطاها					اریبی					مقادیر واقعی					اندازه نمونه
$\hat{\gamma}$	$\hat{m}$	$\hat{\theta}$	$\hat{b}$	$\hat{a}$	$\hat{\gamma}$	$\hat{m}$	$\hat{\theta}$	$\hat{b}$	$\hat{a}$	γ	m	θ	b	a	n
۰/۰۱	۰/۰۲	۰/۰۴	۰/۰۵	۰/۰۶	۰/۰۱	۰/۰۲	۰/۰۳	-۰/۰۸	-۰/۱۲	۱	۰/۵	۲	۰/۵	۰/۵	۱۰۰
۰/۰۹	۰/۰۸	۰/۰۶	۰/۰۵	۰/۰۲	-۰/۲۴	۰/۳۱	۰/۱۱	-۰/۰۹	-۰/۱۳	۱	۰/۳	۳	۰/۵	۰/۵	
۰/۱۴	۰/۰۳	۰/۰۱	۰/۱۲	۰/۰۲	-۰/۲۱	۰/۱۶	-۰/۱۸	۰/۱۲	۰/۱۴	۱	۰/۲	۴	۰/۸	۰/۷	
۰/۰۹	۰/۰۸	۰/۰۶	۰/۰۵	۰/۰۴	۰/۰۲	۰/۰۳	۰/۰۶	۰/۰۷	۰/۰۸	۱	۰/۵	۲	۰/۵	۰/۵	۳۰۰
۰/۰۲	۰/۰۳	۰/۰۴	۰/۰۵	۰/۰۰۷	۰/۰۳	۰/۰۵	۰/۰۷	۰/۰۹	۰/۱۲	۱	۰/۳	۳	۰/۵	۰/۵	
۰/۰۱	۰/۰۲	۰/۰۴	۰/۰۶	۰/۰۰۸	۰/۰۴	۰/۰۵	۰/۰۶	۰/۱۱	۰/۰۹	۱	۰/۲	۴	۰/۸	۰/۷	
۰/۰۱	۰/۰۳	۰/۰۶	۰/۰۵	۰/۰۰۸	-۰/۰۱	-۰/۰۲	-۰/۰۳	-۰/۰۴	-۰/۰۴	۱	۰/۵	۲	۰/۵	۰/۵	۵۰۰
۰/۰۸	۰/۰۵	۰/۰۴	۰/۰۳	۰/۰۰۵	-۰/۰۲	-۰/۰۳	-۰/۰۵	-۰/۱۴	-۰/۰۶	۱	۰/۳	۳	۰/۵	۰/۵	
۰/۰۱	۰/۰۲	۰/۰۳	۰/۰۴	۰/۰۰۶	-۰/۰۱	-۰/۰۳	۰/۰۵	-۰/۰۶	-۰/۰۸	۱	۰/۲	۴	۰/۸	۰/۷	



شکل ۴. نمودارهای تابع‌های چگالی احتمال برآزش داده شده توزیع‌های در نظر گرفته به همراه بافت‌نگاشت برای مجموعه داده‌ها

۶ مطالعات شبیه‌سازی

به‌عنوان نتیجه کلی، می‌توان گفت که توزیع ارائه‌شده با توجه به تابع چگالی و تابع نرخ شکست خود، مکانیسم نسبتاً انعطاف‌پذیری برای برآزش طیف وسیعی از مجموعه داده‌های واقعی مثبت (داده‌های درآمد)، فراهم می‌کند و بنا بر این می‌تواند جایگزین مناسبی برای الگوهای موجود دیگر در منابع، در الگوبندی داده‌های واقعی در زمینه‌های اقتصاد، مهندسی، تحلیل بقا و هیدرولوژی باشد.

چندین ویژگی از این توزیع در این مقاله ارائه نشده است که در تحقیقات آینده مورد بررسی قرار خواهند گرفت. برخی از این ویژگی‌ها عبارت‌اند از: آماره‌های ترتیبی، گشتاورهای آماره ترتیبی، توزیع مجانبی آماره‌های اکسترم، ترتیب‌های تصادفی، آنتروپی و آنتروپی تجمعی باقیمانده، توزیع نسبت دو متغیر تصادفی. همچنین موضوع‌های استنباطی نیز می‌تواند زمینه‌ای برای تحقیق در آینده را فراهم آورد. برخی از این موضوع‌ها عبارت‌اند از: برآوردیابی بیزی پارامترها، برآوردیابی بر اساس

برای مقادیر مختلف پارامترها $(a, b, \theta, m, \gamma)$ اندازه نمونه برابر $n = 100, 300, 500$ در نظر گرفته می‌شود. نتایج در جدول ۲ ارائه شده‌اند، که نشان‌دهنده این است که نتایج کاملاً پایدار و مهم‌تر این که برای این اندازه از نمونه، به مقادیر واقعی نزدیک هستند. علاوه بر این با افزایش اندازه نمونه مقدار MSE کاهش می‌یابد.

نتیجه‌گیری

در این مقاله مدل جدیدی از توزیع (EU-PD) و ویژگی‌های آن ارائه شد. این مدل نسبت به سایر مدل‌ها دارای فرم‌های ساده توابع توزیع و چگالی به لحاظ ریاضی بوده و با توجه به مقادیر پارامترهای دارای تابع نرخ خطر وانی‌شکل و وانی‌شکل وارون است.

انواع سانسورها، برآوردیابی به کمک روش‌های دیگر مانند روش کمترین توان دوم قطری، پیش‌بینی مشاهده‌های آینده از این توزیع و غیره.

مراجع

- [1] Akinsete, A., Famoye, F. and Lee, C. (2008). The Beta-Pareto distribution, *Statistics*, **42**(4), 547–563.
- [2] Bahmani-Oskooee, M. and Motavallizadeh-Ardakani, A. (2018). On the effects of income volatility on income distribution: asymmetric evidence from state level data in the U.S, *Journal of Research in Economics*, **72**(2), 224-239.
- [3] Clementi, F., Gallegati, M. and Kaniadakis, G. (2007). The Kk-generalised statistics in personal income distribution, *The European Physical Journal*. **B 57**, 187–193.
- [4] Clementi, F., Matteo, T. Di., Gallegati, M. and Kaniadakis, G. (2008). The k-generalised distribution: A new descriptive model for the size distribution of incomes, *Physica*, **387**, 3201–3208.
- [5] Clementi, F., Gallegati, M. and Kaniadakis, G. (2009). A k-generalized statistical mechanics approach to income analysis, *Journal of Statistical Mechanics Theory and Experiment*, arXiv:0902.0075v2.
- [6] Dragulescu, A. and Yakovenko, V. M. (2001). Evidence for the exponential distribution of income in the USA, *The European Physical Journal*, **B20**, 585–599.
- [7] Figueira, F. C., Moura, Jr., N. J. and Ribeiro, M. B. (2010). The Gompertz Pareto income distribution, *Physica A: statistical mechanics and it's applications*, **A 390**, 689–698.
- [8] Kareema, A. and Ameer, M. (2017). Uniform-Pareto Distribution, *Journal of University of Babylon, Pure and Applied Sciences*, **25**(5).
- [9] Moura, N. J. and Ribeiro, Jr., M. B. (2009). Evidence for the Gompertz curve in the income distribution of Brazil 1978–2005, *The European Physical Journal*, **B 67**, 101–120.
- [10] Pareto, V. (1987). *Cours Dconomie Politique*. Lausanne.
- [11] Silva, C. A. (2001). *Applications of Physics to Finance and Economics: Returns, Trading activity and income* (Ph.D. Thesis), University of Maryland.
- [12] Yakovenko, V. M. and Rosser, J. B. (2009). Colloquium: Statistical mechanics of money, wealth and income, *Rev. Modern Physics*, **81**, 1703–1725.
- [13] Willis, G. and Mimkes, J. (2004). Evidence for the independence of waged and unwaged income, evidence for Boltzmann distributions in waged income and the outlines of a coherent theory of income distribution, e-print. arXiv:cond-mat/0406694v1.

- [14] Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, Second Edition, Springer-Verlag, New York.