توزیع پارتو یکنواخت و توزیع پارتو یکنواخت نمایی‌شده و کاربردهای آنها در داده‌های درآمد

آنیتا عبداللهی نانواپیشه،[۱] آزیتا عبداللهی نانواپیشه[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 مجموعه‌داده‌های واقعی در نظر گرفته شده تأیید می‌کنند.

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

۱ مقدمه

با توجه به این‌که آمار یکی از مهم‌ترین علوم کاربردی است که با سایر رشته‌های علمی از جمله اقتصاد مرتبط است، از این رو در سال‌های اخیر، علاقهٔ زیادی برای تحقیقات در زمینهٔ مدل‌های پارامتری توزیع درآمد به وجود آمده است. مدل‌های احتمالی مربوط به توزیع درآمد، برای ارزیابی استانداردهای سطح زندگی کلّ مردم یک کشور و همچنین برای مقایسهٔ استاندارد سطح زندگی طبقات اجتماعی و یا مناطق مختلف یک کشور ارائه شده‌اند. بنا بر این برای ایجاد یک مدل احتمالی ارائهٔ یک تابع توزیع نظری با مشخصهٔ توزیع فراوانی تجربی برای انتخاب روش مناسب تخمین پارامترهای مدل، ضروری است. بنا بر این تحلیل آماری توزیع درآمد جمعیت نشان‌دهندهٔ 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 مجموعه‌داده‌ها در تحلیل‌های آماری بسیار با اهمیت است. مشخص کردن یک توزیع مناسب برای یک مجموعه از داده‌های خاص به ما کمک خواهد کرد تا استدلال‌هایی در بارهٔ داده‌ها داشته باشیم. به همین دلیل، استفاده از روش‌هایی برای اصلاح توزیع آماری موجود به‌منظور ایجاد انعطاف‌پذیری بیشتر در مجموعه‌داده‌ها 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 نمایی‌شده، با انگیزهٔ اصلی مدل‌سازی داده‌های درآمد با هر دو نرخ خطر یکنواخت و غیریکنواخت ارائه شده است. در ادامه، مقاله به‌صورت 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 نمایی‌شده تعریف

[۱] دانش‌آموختهٔ کارشناسی ارشد آمار، دانشگاه علامه طباطبایی، ایران
[۲] دانشجوی دکترای مکانیک، دانشگاه تهران، ایران

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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 به‌صورت زیر است [۷]:

$$G(X) = ۱ - \frac{b(\frac{m}{x})^{\theta}}{b-a}, \qquad m(\frac{b-a}{b})^{-\frac{۱}{\theta}} < x < \infty \quad (۱)$$

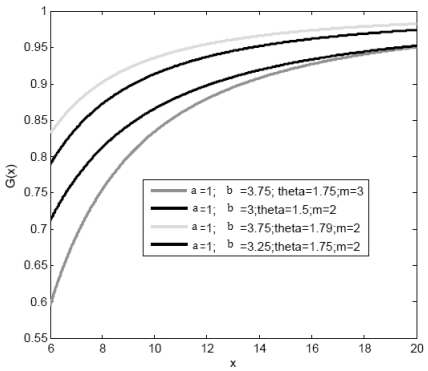
که $-\infty < a \leq b < \infty$ پارامترهای مقیاس و $\theta > ۰$ پارامتر شکل است. شکل ۱ نمودار تابع توزیع تجمعی این توزیع را به‌ازای مقادیر مختلف پارامترها نشان می‌دهد که به‌صورت یکنواخت غیرکاهشی است و در نقطهٔ یک پایدار می‌شود.

اگر تابع توزیع تجمعی توزیع (U-PD) به‌صورت رابطهٔ (۱) باشد، در این صورت تابع چگالی احتمال این توزیع به‌صورت زیر است:

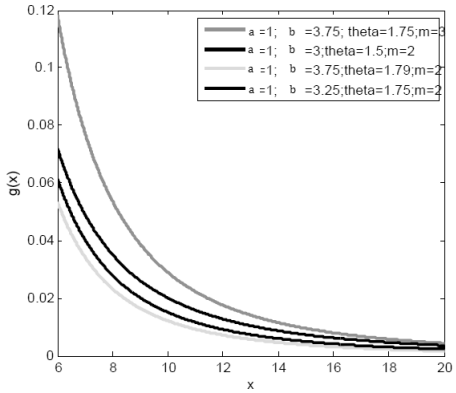
$$g(X) = \frac{b}{b-a}\frac{\theta m^{\theta}}{x^{\theta+۱}}, \qquad m(\frac{b-a}{b})^{-\frac{۱}{\theta}} < x < \infty \quad (۲)$$

تابع بقا و تابع نرخ خطر (نرخ خطر معیاری است برای سنجیدن شکست یک سیستم در یک بازهٔ زمانی مشخص) این توزیع نیز به‌ترتیب به‌صورت زیر هستند:

$$S(x) = \frac{b(\frac{m}{x})^{\theta}}{b-a},$$
$$h(x) = \frac{\frac{b\theta m^{\theta}}{b-ax^{\theta+۱}}}{\frac{b(\frac{m}{x})^{\theta}}{b-a}}.$$



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



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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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

$$G(X) = [1 - \frac{b(\frac{m}{x})^{\theta}}{b-a}]^{\gamma}, \qquad m(\frac{b-a}{b})^{-\frac{1}{\theta}} < x < \infty \quad (۳)$$

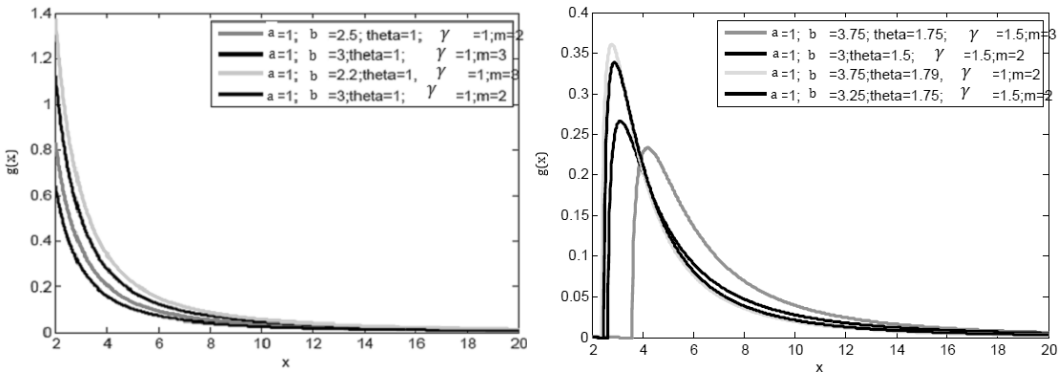
تابع چگالی احتمال این توزیع با مشتق‌گیری از رابطهٔ بالا به‌صورت زیر به دست می‌آید:

$$g(x) = [\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}][1 - \frac{b(\frac{m}{x})^{\theta}}{b-a}]^{\gamma-1}, \quad m(\frac{b-a}{b})^{-\frac{1}{\theta}} < x < \infty \quad (۴)$$

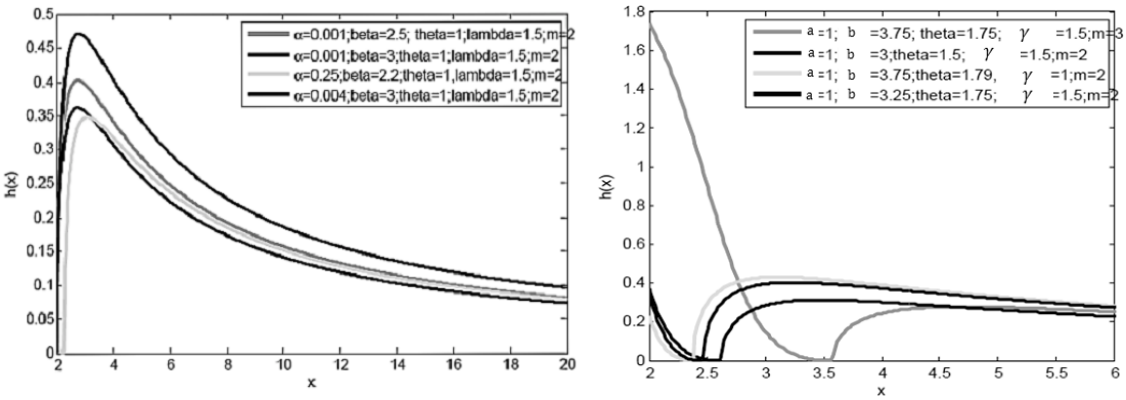
تابع نرخ خطر این توزیع نیز به‌صورت زیر است:

$$h(x) = \frac{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}][1 - \frac{b(\frac{m}{x})^{\theta}}{b-a}]^{\gamma-1}}{1 - [1 - \frac{b(\frac{m}{x})^{\theta}}{b-a}]}. \quad (۵)$$

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 به‌صورت وانی‌شکل و وانی‌شکل وارون است.



(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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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

در این بخش برخی از ویژگی‌های توزیع پارتو یکنواخت و توزیع پارتو یکنواخت نمایی‌شده ارائه می‌شود. اگر X دارای توزیع پارتو یکنواخت باشد، آن‌گاه میانگین، واریانس، ضریب تغییرات، چولگی و کشیدگی و گشتاور مرتبه r-ام آن به‌ترتیب به‌صورت زیر است:

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

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

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

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

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

و گشتاور مرتبه r-ام این توزیع به‌صورت 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 معرفی‌شده (EU-PD) ارائه می‌شوند ولی قبل از آن به ارائه بسط $g(x)$ که در محاسبهٔ گشتاورها مورد نیاز است، پرداخته می‌شود. اگر

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

در این صورت بسط آن به‌صورت:

$$g(X) = \Big[\frac{b\theta\gamma}{b-a}\frac{m^\theta}{x^{\theta+1}}\Big]\sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(\frac{m}{x})^\theta)^j}{(b-a)^j} \quad (6)$$

و به‌طور مشابه بسط تابع توزیع آن نیز به‌صورت زیر است:

$$G(X) = \Big[1-\frac{b(\frac{m}{x})^\theta}{b-a}\Big]^{\gamma}$$
$$G(X) = \sum_{j=0}^{\infty}\binom{\gamma}{j}\frac{(-1)^j(b(\frac{m}{x})^\theta)^j}{(b-a)^j}. \quad (7)$$

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

$$\begin{aligned}E[X^r] &= \int_w^\infty x^r\Big[\frac{b\theta\gamma}{b-a}\frac{m^\theta}{x^{\theta+1}}\Big]\sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(\frac{m}{x})^\theta)^j}{(b-a)^j}dx\\ &= \int_w^\infty x^{-\theta j+r-\theta-1}\sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]dx\\ &= \frac{x^{-\theta j+r-\theta}}{-\theta j+r-\theta}\sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]_x^w,\end{aligned}$$

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

$$E[X^r] = -\frac{\big(m(\frac{b-a}{b})^{-\frac{1}{\theta}}\big)^{-\theta j+r-\theta}}{-\theta j+r-\theta}\sum_{j=0}^{\infty}\binom{\gamma-1}{j} \times \frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]. \quad (8)$$

بنا بر این گشتاور اول توزیع جدید به‌صورت زیر است:

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

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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

کاربرد دارند. اگر X دارای توزیع رابطهٔ (۲) باشد، آن‌گاه داریم:

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

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

$$\int_0^w x^r g(x)dx = \frac{\big(m(\frac{b-a}{b})^{-\frac{1}{\theta}}\big)^{-\theta j+r-\theta}}{-\theta j+r-\theta}\sum_{j=0}^{\infty}\binom{\gamma-1}{j} \times \frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big] \qquad (۹)$$

۲.۳ میانگین انحراف‌ها، منحنی لورنز و معادلات منحنی بونفرونی

اگر X دارای توزیع رابطهٔ (۲) باشد، آن‌گاه انحراف از میانگین از رابطهٔ زیر به دست می‌آید:

$$\begin{aligned}\delta_1(x) &= \int_0^\infty |x-M|f(x)dx = 2MF(M)-M+\mu-2I(M) \\ &= 2M\sum_{j=0}^{\infty}\binom{\gamma}{j}\frac{(-1)^j(b(\frac{m}{M})^\theta)^j}{(b-a)^j} - M+\mu \\ &= -2\Bigg[\frac{\big(m(\frac{b-a}{b})^{-\frac{1}{\theta}}\big)^{-\theta j+1-\theta}}{-\theta j+1-\theta} \\ &\times \sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]\Bigg]\end{aligned} \qquad (۱۰)$$

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

$$\begin{aligned}\delta_2(x) &= \int_0^\infty |x-\mu|f(x)dx = 2\mu F(\mu)-2I(\mu) \\ &= 2\mu\sum_{j=0}^{\infty}\binom{\gamma}{j}\frac{(-1)^j(b(\frac{m}{\mu})^\theta)^j}{(b-a)^j} \\ &= -2\Bigg[\frac{\big(m(\frac{b-a}{b})^{-\frac{1}{\theta}}\big)^{-\theta j+1-\theta}}{-\theta j+1-\theta}\sum_{j=0}^{\infty}\binom{\gamma-1}{j}\frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j} \\ &\times \Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]\Bigg]\end{aligned} \qquad (۱۱)$$

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 می‌دهد، به شرط آن‌که 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 ارائه‌شده در رابطهٔ (۲) باشد، آن‌گاه معادلات منحنی بونفرونی به‌صورت زیر است:

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

(۱۲)

و منحنی لورنز نیز به‌صورت زیر خواهد بود:

$$\begin{aligned}L_F[F(x)] = B_F[F(x)] &= \frac{1}{\mu}\int_0^x uf(u)du \\ &= \frac{1}{\mu\big[1-\frac{b(\frac{m}{x})^\theta}{b-a}\big]^\gamma}\Big(\frac{\big(m(\frac{b-a}{b})^{-\frac{1}{\theta}}\big)^{-\theta j+1-\theta}}{-\theta j+1-\theta}\sum_{j=0}^{\infty}\binom{\gamma-1}{j} \\ &\times \frac{(-1)^j(b(m)^\theta)^j}{(b-a)^j}\Big[\frac{b\theta\gamma}{b-a}m^\theta\Big]\Big)\end{aligned} \qquad (۱۳)$$

۴ تابع ماکسیمم درست‌نمایی (MLE)

اگر $X_1, \ldots, X_n$ نمونه‌ای تصادفی از توزیع (EU-PD) با اندازهٔ n باشند، آن‌گاه تابع ماکسیمم درست‌نمایی این توزیع به‌صورت زیر است:

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

$$L = \prod_{i=1}^{n}\Big[\frac{b\theta\gamma}{b-a}\frac{m^\theta}{x^{\theta+1}}\Big]\prod_{i=1}^{n}\Big[1-\frac{b(\frac{m}{x})^\theta}{b-a}\Big]^{\gamma-1}.$$

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

در نتیجه لگاریتم تابع درست‌نمایی به‌صورت 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 مشتق‌گیری از رابطهٔ (۱۴) نسبت به پارامترها داریم:

$$\frac{\partial l_n}{\partial a}=\sum_{i=1}^{n}\frac{\frac{b\theta\gamma}{(b-a)^2}\frac{m^{\theta}}{x^{\theta+1}}}{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}]}+(\gamma-1)\sum_{i=1}^{n}\frac{-\frac{b(\frac{m}{x})^{\theta}}{(b-a)^2}}{[1-\frac{b(\frac{m}{x})^{\theta}}{b-a}]},$$

$$\frac{\partial l_n}{\partial b}=\sum_{i=1}^{n}\frac{\frac{\theta\gamma(b-a)-b\theta\gamma}{(b-a)^2}\frac{m^{\theta}}{x^{\theta+1}}}{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}]}+(\gamma-1)\sum_{i=1}^{n}\frac{\frac{-(\frac{m}{x})^{\theta}(b-a)+b(\frac{m}{x})^{\theta}}{(b-a)^2}}{[1-\frac{b(\frac{m}{x})^{\theta}}{b-a}]},$$

$$\frac{\partial l_n}{\partial \theta}=\sum_{i=1}^{n}\frac{\frac{(b\gamma)}{(b-a)}\frac{m^{\theta}\ln m-(x^{\theta+1})(\ln x)(m^{\theta})}{x^{\theta+1}}}{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}]}+(\gamma-1)\sum_{i=1}^{n}\frac{-\frac{b(\frac{m}{x})^{\theta}(b-a)+(\ln\frac{m}{x})}{(b-a)^2}}{[1-\frac{b(\frac{m}{x})^{\theta}}{b-a}]},$$

$$\frac{\partial l_n}{\partial m}=\sum_{i=1}^{n}\frac{[\frac{b\theta\gamma}{(b-a)}\frac{\theta m^{\theta-1}}{x^{\theta+1}}]}{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}]}+(\gamma-1)\sum_{i=1}^{n}\frac{[-\frac{b}{(b-a)}\frac{\theta m^{\theta-1}}{x^{\theta+1}}]}{[1-\frac{b(\frac{m}{x})^{\theta}}{b-a}]},$$

$$\frac{\partial l_n}{\partial \gamma}=\sum_{i=1}^{n}\frac{\frac{b\theta}{(b-a)^2}\frac{m^{\theta}}{x^{\theta+1}}}{[\frac{b\theta\gamma}{b-a}\frac{m^{\theta}}{x^{\theta+1}}]}+\sum_{i=1}^{n}\frac{-\frac{b(\frac{m}{x})^{\theta}}{(b-a)^2}}{[1-\frac{b(\frac{m}{x})^{\theta}}{b-a}]}.$$

برآورد ماکسیمم درست‌نمایی پارامترها با برابر صفر قرار دادن عبارات بالا و با استفاده از حل تابع غیرخطی به دست می‌آیند که دارای شکل صریحی نیستند. برای ارائهٔ برآورد بازه‌ای، فاصلهٔ اطمینان و آزمون فرضیهٔ آماری روی پارامترها، نیاز به ماتریس اطلاع است. ماتریس اطلاع مشاهده‌شده به‌صورت 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 مشاهده‌شده وجود دارد که از مشتق دوم لگاریتم تابع درست‌نمایی نسبت به پارامترها به دست می‌آیند. تحت شرایطی که برای پارامترها در خارج از فضای پارامتری انجام می‌شود امّا 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 اساس تقریب نرمال امکان‌پذیر است. توزیع نرمال چندمتغیره $N_5(0,I^{-1}(\vartheta))$ می‌تواند برای ایجاد بازه‌های اطمینان دوطرفه $100(1-\eta)$ درصدی برای پارامترهای مدل به کار رود که η سطح معناداری است.

۵ کاربردها

در این بخش کاربردهایی از توزیع ارائه‌شده در بخش‌های قبل بیان می‌شود و نشان داده می‌شود که توزیع (EU-PD) برازش بهتری نسبت به سایر مدل‌ها برای مجموعه‌داده‌های درآمد فراهم می‌کند. یکی از معیارها و شاخص‌های تعیین رفاه جامعه در اقتصاد متکی به بازار آزاد، درآمد ملی است. اگرچه این فاکتور بیان‌کنندهٔ توزیع مناسب و عادلانه درآمد در جامعه نیست، اما خود، به نوعی بیان‌کنندهٔ واقعیت درآمد جامعه است. درآمد ملّی عبارتست از مجموع درآمدهای کسب‌شدهٔ اتباع یک کشور، در چارچوب یک اقتصاد ملّی در یک دورهٔ معین اقتصادی پس از کسر سرمایه‌های مصرفی. (نظریه u وارون که توسط کوزنتس (۱۹۵۵) معرفی شد اساساً سطح فعالیت اقتصادی را به‌عنوان عامل تعیین‌کنندهٔ اصلی نابرابری درآمدی مشخص می‌کند. به‌طور دقیق‌تر این امر تأکید دارد که در مرحلهٔ اوّلیهٔ رشد اقتصادی، نابرابری درآمدی بدتر می‌شود و فقط در مراحل بعدی بهبود می‌یابد. کوزنتس سطح فعالیت درآمد یا فعالیت اقتصادی را به‌عنوان عامل اصلی نابرابری درآمدی شناسایی کرد. وی تأکید کرد در مراحل اوّلیه توسعهٔ نابرابری درآمد بدتر می‌شود و زمانی که نیروی کار از مناطق روستایی به مناطق شهری مهاجرت می‌کنند، بهتر می‌شود. از آن‌جا که الگوی حرکت نابرابری در طول زمان شبیه به شکل u وارون است، به‌عنوان الگوی u وارون شناخته می‌شود) [۸]. مجموعه‌داده‌های ارائه‌شده، داده‌های درآمد ملّی سال‌های (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,

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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,
۱۲۴۳۸۰۱/۰, 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 مجموعه‌داده‌ها فراهم می‌کند. جهت بررسی نیکوئی 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 مجموعه‌داده‌های واقعی، ارائه شده‌اند. همان‌طور که مشخص است با توجه به معیارهای AIC (بر اساس کمترین مقدار) مدل ارائه‌شده بهترین برازش برای مجموعه‌داده‌ها فراهم می‌کند. شایان ذکر است که توزیع لیندلی ارائه‌شده در جدول، یک توزیع تک‌پارامتری 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 به‌صورت زیر است:

$$f_L(x) = \frac{\beta^{۲}}{۱+\beta}(۱+x)e^{-\beta x}, \quad x \geq ۰, \beta > ۰$$

به‌منظور ارزیابی عملکرد روش ماکسیمم درست‌نمایی، مطالعهٔ 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 ماکسیمم درست‌نمایی به‌ازای هر داده شبیه‌سازی برای

$$i = ۱, ۲, \ldots, ۱۰۰۰۰$$

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

$$bias_h(n) = \frac{۱}{۱۰۰۰۰}\sum_{i=۱}^{۱۰۰۰۰}(\hat{h}_l - h);$$

$$MSE_h(n) = \frac{۱}{۱۰۰۰۰}\sum_{i=۱}^{۱۰۰۰۰}(\hat{h}_l - h)^{۲}$$

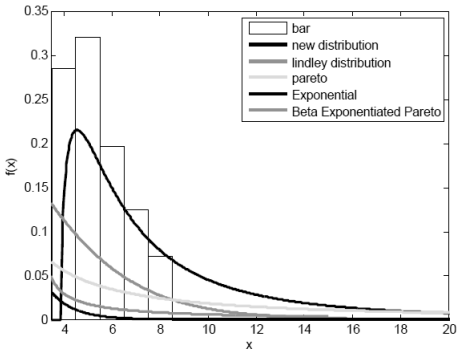
جدول ۱. برآورد ماکسیمم درست‌نمایی پارامترها و آزمون نیکویی برازش برای مجموعه‌داده‌های واقعی

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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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

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



شکل ۴. نمودارهای تابع‌های چگالی احتمال برازش داده شدهٔ توزیع‌های در نظر گرفته به همراه بافت‌نگاشت برای مجموعه‌داده‌ها

۶ مطالعات شبیه‌سازی

برای مقادیر مختلف پارامترها $(a, b, \theta, m, \gamma)$ اندازهٔ نمونه برابر $n = ۱۰۰, ۳۰۰, ۵۰۰$ در نظر گرفته می‌شود. نتایج در جدول ۲ ارائه شده‌اند، که نشان‌دهندهٔ این است که نتایج کاملاً پایدار و مهم‌تر این‌که برای این اندازه از نمونه، به مقادیر واقعی نزدیک هستند. علاوه بر این با افزایش اندازهٔ 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.

به‌عنوان نتیجهٔ کلّی، می‌توان گفت که توزیع ارائه‌شده با توجه به تابع چگالی و تابع نرخ شکست خود، مکانیسم نسبتاً انعطاف‌پذیری برای برازش طیف وسیعی از مجموعه‌داده‌های 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 آمارهٔ ترتیبی، توزیع مجانبی آماره‌های اکسترمم، 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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

انواع سانسورها، برآوردیابی به‌کمک 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.

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]

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

[Downloaded from andisheyeamari.irstat.ir on 2024-05-08]