مجله‌ی بررسی‌های آمار رسمی ایران
سال ۲۶، شماره‌ی ۲، پاییز و زمستان ۱۳۹۴، صص ۱۲۹ – ۱۳۷

خانواده‌ی توزیع درامد بتای تعمیم‌یافته و براوردگر ماکسیمم آنتروپی آن به‌کمک نرم‌افزار متلب

فاطمه عباس‌پور[†]، منیژه صانعی طبس[‡] و غلامرضا محتشمی برزادران[*,†]

[†] دانشگاه فردوسی مشهد
[‡] دانشگاه سیستان و بلوچستان

چکیده: توزیع بتای تعمیم‌یافته یکی از انواع توزیع‌های درامد است. در این مقاله به معرفی خانواده‌ی توزیع تعمیم‌یافته می‌پردازیم و این خانواده‌ی توزیع را از دیدگاه ماکسیمم آنتروپی مورد بررسی قرار می‌دهیم و همچنین براورد ماکسیمم آنتروپی پارامترهای این خانواده‌ی توزیع را تحت شرایط گوناگون، با استفاده از برنامه‌ی نرم‌افزار متلب می‌یابیم.

واژگان کلیدی: براورد ماکسیمم آنتروپی؛ توزیع ماکسیمم آنتروپی؛ توزیع بتای تعمیم‌یافته؛ براورد پارامتر.

۱- مقدمه

بیش از صد سال پیش، تعداد وسیعی از توزیع‌ها برای مدل‌سازی پدیده‌ها به‌خصوص توزیع‌های درامد مورد استفاده قرار می‌گرفت. این توزیع‌ها عبارت‌اند از: پارتو، لگ‌نرمال تعمیم‌یافته، گامای تعمیم‌یافته، بتای تعمیم‌یافته‌ی نوع اول و دوم، داگم و سینگ‌مادالا.

این توزیع‌ها در اقتصاد کاربرد فراوانی دارند. در این مقاله رفتار آنتروپی این توریع‌ها مورد مطالعه قرار می‌گیرد. ماکسیمم آنتروپی، صورت‌های معینی از آنتروپی را انتخاب می‌کند و آن را به‌عنوان یک خانواده‌ی نمایی از توزیع‌ها، تحت شرایط معین قرار

---

* نویسنده‌ی عهده‌دار مکاتبات
دریافت: ۱۳۹۳/۱۲/۱۸، پذیرش۱۳۹۵/۶/۱۶

می‌دهد. یافتن قیدهایی که توزیع‌های درامد، تحت آن ماکسیمم آنتروپی را دارند، نیز مورد بحث است.
در ژوئن ۱۹۰۵، لورنز [۴] مقاله‌ای در مورد توزیع‌های درامد نوشت که در آمار و اقتصاد انقلابی اساسی ایجاد کرد. در واقع آماردان‌های تجربی، توزیع‌های درامد را ازجنبه‌های متفاوت مورد مطالعه قرار می‌دهند و گاهی اوقات آن‌ها را به توابع تحلیلی مناسب برازش می‌دهند.
کاگان و رائو [۳] در سال ۱۹۷۳ کلاسی از توزیع‌های درامد را ارایه دادند که نسبت به قیدهای معین دارای ماکسیمم آنتروپی است.
دراگولسکو و یاکونکو [۲] نیز در سال ۲۰۰۳، در فیزیک اقتصادی در مورد مکانیک آماری، پول، دارایی و درامد بحث کرده‌اند. یاکوونکو [۹] به‌خصوص در این زمینه تلاش های زیادی داشته است.

۲- توزیع‌های درامد

پارتو [۶] اولین کسی بود که یک مدل از توزیع‌های درامد را به شکل یک تابع چگالی (تابع چگالی احتمال) معرفی کرد. در آمار، توزیع‌های زیادی وجود دارد که برای مدل‌سازی توزیع متغیرهای تصادفی پیوسته به‌کار می‌روند ولی از میان تمام توزیع‌های احتمال، توزیع‌های احتمال پیوسته با مقادیر نامنفی که عموماً چوله می‌باشند، برای مدل‌سازی توزیع‌های درامد مورد استفاده قرار می‌گیرند. مشخصه‌های اقتصادی کمی مانند درامد، مقادیر نامنفی را اختیار می‌کنند و برای بررسی آن‌ها لازم است از توزیع‌ها برای متغیرهایی که مقادیر مثبت اختیار می‌کنند، استفاده شود. در این رابطه توزیع‌های تعمیم‌یافته به‌عنوان خانواده‌ای از این‌گونه توزیع‌ها با انعطاف‌پذیری بالا برای بررسی این‌گونه مشخصه‌های اقتصادی به‌کار می‌روند. در رأس این توزیع‌ها، توزیع بتای تعمیم‌یافته قرار دارد. توزیع‌های بتای تعمیم‌یافته نوع اول و دوم و توزیع گامای تعمیم‌یافته و همچنین بسیاری از توزیع‌های مثبت آماری از قبیل توزیع بتا، گاما، پارتو حالت‌های خاص و حدی توزیع بتای تعمیم‌یافته هستند. توزیع بتای چندپارامتری و حالت‌های خاص آن یکی از این مدل‌ها است که می‌تواند برای مدل‌سازی توزیع‌های درامد مورد استفاده قرار گیرند.

۱-۲- خانواده‌ی توزیع بتای تعمیم‌یافته‌ی GB

توزیع بتای تعمیم‌یافته، یک توزیع بتای پنج‌پارامتری و دارای تابع چگالی احتمال زیر است:

$$f(x)=\frac{ax^{ap-1}\left[1-(1-c)\left(\frac{x}{b}\right)^{a}\right]^{q-1}}{b^{ap}B(p,q)\left[1+c\left(\frac{x}{b}\right)^{a}\right]^{p+q}},\qquad 0<x^{a}<\frac{b^{a}}{1-c} \tag{1}$$

که در آن $B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ و $a,b,p,q>0$ و $0\le c\le 1$ و $\Gamma(.)$ تابع گاماست. توزیع بتای تعمیم‌یافته، وقتی $E(\ln(x))$ و $E[\ln(1-(1-c)(\frac{x}{b})^{a})]$ و $E\left(\ln\left(1+c\left(\frac{x}{b}\right)^{a}\right)\right)$ مقادیری معلوم باشند، توزیع ماکسیمم آنتروپی است [۳]. اما حالت‌های خاص توزیع بتای تعمیم‌یافته مربوط به زمانی است که در (۱)، c مقدار صفر یا یک بگیرد. اگر در رابطه‌ی (۱)، c را برابر یک قرار دهیم، توزیع بتای نوع دوم (GB2) حاصل می‌شود که دارای چگالی به شکل زیر است:

$$f(x)=\frac{ax^{ap-1}}{b^{ap}B(p,q)\left[1+\left(\frac{x}{b}\right)^{a}\right]^{p+q}},\qquad x>0.$$

توزیع بتای تعمیم‌یافته‌ی نوع دوم، وقتی که

$$E(\ln(X))=\ln b+\frac{1}{a}(\psi(p)-\psi(q))$$

و $E(\ln(1+(\frac{X}{b})^{a}))=\psi(p+q)-\psi(q)$ باشد، توزیع ماکسیمم آنتروپی است که در آن $\psi(\cdot)=\frac{\Gamma'(\cdot)}{\Gamma(\cdot)}$ است.

قضیه‌ی ۱ ([۳]). فرض کنید x، یک بردار تصادفی با تابع چگالی f نسبت به اندازه‌ی μ باشد، چنان‌که به‌ازای هر $x\epsilon(a,b]$, $f(x)>0$ و در سایر نقاط برابر صفر باشد. فرض کنید $h_1,\dots,h_n$ توابعی انتگرال‌پذیر روی $(a,b]$ باشند که در قیدهای

$$\int_{a}^{b}h_i(x)f(x)d\mu(x)=\lambda_i,\qquad i=1,\dots,n$$

صدق کنند در حالی‌که λ_1، λ_2 و ... مقادیر ثابتی باشند. در این صورت ماکسیمم آنتروپی به‌وسیله‌ی توزیعی با چگالی $f(x)$ به‌صورت

$$f(x) = \exp\{c_0 + c_1 h_1(x) + \cdots\}$$

به‌دست می‌آید چنان‌که مقادیر $c_0, c_1, \ldots$ قیدها را نتیجه می‌دهد.

با توجه به قضیه‌ی بالا، توزیع GB2، تحت چند قید مشخص‌شده دارای ماکسیمم آنتروپی است. در صورتی که فرض شود

$$\mu_1 = E(\ln(x)), \quad \mu_2 = E(\ln(1 + \left(\frac{X}{b}\right)^a))$$

مقادیری معلوم باشند، تابع چگالی ماکسیمم آنتروپی با توجه به قضیه‌ی بالا به‌صورت زیر خواهدبود:

$$f(x) = e^{-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln\left[1 + (\frac{x}{b})^a\right]}, \qquad x > 0 \quad (2)$$

با توجه به مقادیر معلوم a, b, q, p می‌توان براورد ماکسیمم آنتروپی (در جدول‌های این مقاله، این براورد را با ME نشان داده‌ایم) پارامترها را به‌دست آورد. با استفاده از برنامه‌ی نرم‌افزار متلب به‌ازای چندین مقدار معلوم a, b, q, p، براورد ماکسیمم آنتروپی پارامترهای توزیع دارای ماکسیمم آنتروپی را به‌صورت جدول ۱ ارایه می‌دهیم.

جدول ۱- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی نمایش ۲

λ_2	λ_1	λ_0	q	p	b	a
۱٫۶۶۴۱	-۱٫۶۶۵۸	-۰٫۱۹۴۳	۱	۲	۱	۲
-۱۵٫۵۵۵۴	۳٫۵۳۵۸	۱۳٫۷۸۲	۲	۳	۳	۲
۲٫۴۷۹۳	۱٫۳۱۴۳	۴٫۷۴۸۳	۱	۴	۲	۳
۳٫۹۷۷	-۱۳٫۹۲۵۶	-۲٫۶۸۹۱	۱	۳	۱	۵
-۵٫۱۳۰۱	۱٫۶۳۲۵	۵٫۵۹۹۲	۲	۴	۳	۳

حالت خاص توزیع GB2 هنگامی‌است که $p = 1$ یا $q = 1$ باشد. اگر فرض شود که $q = 1$ باشد، توزیع سینگ‌مادالا حاصل می‌شود. در واقع اگر فرض شود که

$$E(\ln(X)) = \ln\ b + \frac{1}{a}\left(\psi\left(1\right) - \psi(q)\right), \qquad E(\ln\left(1 + \left(\frac{x}{b}\right)^a\right)) = \frac{1}{q}$$

باشند، توزیع سینگ‌مادالا، توزیع دارای ماکسیمم آنتروپی خواهد بود. توزیع سینگ مادالا یک توزیع سه‌پارامتری با پارامترهای a, b, p است. به‌ازای چندین مقدار معلوم a, b, p، براورد ماکسیمم آنتروپی پارامترهای λ_0 و λ_1 و λ_2 در نمایش تابع چگالی دارای ماکسیمم آنتروپی به‌صورت جدول ۲ ارایه می‌شود.

جدول ۲- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی (۲)

λ_2	λ_1	λ_0	q	b	a
-۲۴٫۵۶۹۴	۳٫۰۲۵	۱۰٫۹۹	۲	۵	۲
۰٫۹۹۷۳	-۰٫۰۰۳۲	۰٫۸۱۵۶	۲	۲	۳
۴٫۰۰۹۶	-۳٫۰۰۷۵	-۲٫۴۹۰۹	۳	۱	۴

همچنین حالت خاص توزیع سینگ‌مادالا، وقتی است که $a = 1$ باشد در این صورت، توزیع پاراتوی نوع دوم نتیجه می‌شود. اگر داشته‌باشیم

$$E(\ln(X)) = \ln\ b + \psi\left(1\right) - \psi(q), \qquad E(\ln\left(1 + \frac{x}{b}\right)) = \frac{1}{q}$$

توزیع لوماکس یا پارتوی نوع دوم، توزیع ماکسیمم آنتروپی خواهد بود. به‌ازای سه مقدار متفاوت b, p براورد ماکسیمم آنتروپی پارامترهای چگالی نمایش (۲) را به‌دست آورده‌ایم که در جدول ۳ آمده‌است.

جدول ۳- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی نمایش ۲

λ_2	λ_1	λ_0	q	b
۰٫۲۳۲۵	۰٫۸۰۶۲	۱٫۵۶۱۳	۳	۲
-۶٫۷۴۶۵	۱٫۶۶۵۳	۴٫۹۱۵۱	۲	۳
-۱۷٫۶۳۰۵	۲٫۴۷۳۳	۸٫۱۲۱۴	۳	۶

در توزیع بتای تعمیم‌یافته‌ی نوع دوم اگر $q = 1$ باشد، توزیع داگم حاصل می‌شود. اگر

$$E(\ln(X)) = \ln\ b + \frac{1}{a}(\psi(p) - \psi(1)),$$

$$E(\ln\left(1+(\frac{x}{b})^{a}\right))=\psi(p+1)-\psi(1)$$

باشند، توزیع ماکسیمم آنتروپی، توزیع داگم خواهد بود. در ادامه در جدول ۴ براورد پارامترهای λ_0 وλ_1 و λ_2 توزیع دارای ماکسیمم آنتروپی به ازای چند مقدار معلوم a, b, p آمده است.

جدول ۴- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی نمایش ۲

λ_2	λ_1	λ_0	p	b	a
-۰٫۹۲۰۲۵	۰٫۶۰۹۷	۲٫۴۵۲۶	۲	۲	۳
۱٫۴۵۴۲	-۱٫۹۵۷۷	۰٫۳۳۹۴	۳	۱	۲
-۰٫۱۹۲۶	۰٫۲۳۹۵	۱٫۷۵۱۷	۲	۲	۴

در توزیع داگم اگر $a=1$ باشد، حالت خاص این توزیع حاصل می‌شود که توزیعی به نام توزیع لوماکس معکوس را نتیجه می‌دهد.

اگر $E(\ln(X))=\ln b+\psi(p)-\psi(1)$ و $E(\ln(1+\left(\frac{X}{b}\right)^{a}))$ باشند، توزیع لوماکس معکوس توزیع ماکسیمم آنتروپی خواهد بود. براورد ماکسیمم آنتروپی پارامترهای λ_0 وλ_1 و λ_2 به‌ازای چند مقدار b و p در جدول ۵ آمده است.

جدول ۵- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی نمایش (۲)

λ_2	λ_1	λ_0	p	b
-۲۵٫۷۰۴۶	۶٫۹۸۸۱	۲۹٫۷۰۳۱	۲	۲
-۷٫۷۲۱۵	۲٫۴۸۴۶	۹٫۹۸۲۱	۲	۱

اگر در توزیع بتای تعمیم‌یافته، $c=0$ باشد، توزیعی به نام توزیع بتای تعمیم‌یافته‌ی نوع اول نتیجه می‌شود که به‌صورت زیر است:

$$f(x)=\frac{ax^{ap-1}\left[1-\left(\frac{x}{b}\right)^{a}\right]^{q-1}}{b^{ap}B(p,q)},\qquad 0<x<b$$

اگر $\mu = E(\ln(X))$ و $\mu = E(\ln(1-\left(\frac{X}{b}\right)^a))$ مقادیری معلوم باشند با توجه به قضیه‌ی ۱، تابع چگالی دارای ماکسیمم آنتروپی به‌صورت زیر خواهدبود:

$$f(x) = e^{-\lambda-\lambda\ln x-\lambda\ln\left[1-\frac{x^a}{b}\right]}, \qquad 0 < x < b. \tag{۳}$$

حال اگر

$$E(\ln(X)) = \ln b + \frac{1}{a}(\psi(p) - \psi(q))$$

و

$$E(\ln(1-\left(\frac{X}{b}\right)^a)) = \psi(p+q) - \psi(q)$$

باشند، توزیع بتای نوع اول، توزیع ماکسیمم آنتروپی خواهد بود. با معلوم بودن مقادیر a, b, p, q، براورد ماکسیمم آنتروپی پارامترهای λ_0 و λ_1 و λ_2 در نمایش (۳) را می‌توان یافت.

جدول ۶- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی نمایش (۳)

λ_2	λ_1	λ_0	q	p	b	a
-۰/۰۰۱	-۰/۰۰۰۱	-۰/۰۰۱	۳	۴	۱	۲
-۰/۰۰۰۹	-۰/۰۰۰۵	۱/۰۹۷۸	۱	۲	۳	۲
۰/۰۰۵۷	۰/۰۱	۱/۳۸۵۷	۱	۳	۴	۲
۰/۱۳۳۶	۰/۰۵۴۷	۰/۱۱۶۳	۲	۳	۱	۳
۰/۴۵۱۲	۰/۱۵۷۵	۰/۳۵۹۸	۱	۴	۱	۴

در توزیع بتای تعمیم‌یافته‌ی نوع اول، اگر a را برابر یک قرار دهیم، توزیع بتای نوع اول به دست می‌آید. اگر داشته باشیم:

$$E(\ln(X)) = \ln b + \psi(p) - \psi(q)$$

و

$$E(\ln(1+\frac{X}{b})) = \psi(p+q) - \psi(q),$$

توزیع بتای نوع اول، توزیع ماکسیمم آنتروپی است. در جدول ۷، براورد پارامترها در نمایش (۳) را به‌ازای چندین مقدارمعلوم b, p, q آورده‌ایم.

جدول ۷- براورد ME پارامترهای λ_0 و λ_1 و λ_2 در چگالی (۲)

λ_2	λ_1	λ_0	q	p	b
-۰٫۰۰۵۹	-۰٫۰۰۹	۰٫۶۸۴	۱	۲	۲
۰٫۰۰۲۲	-۰٫۰۰۱۴	۱٫۱۰۰۳	۲	۳	۳
-۴٫۳۵۳۷	-۶٫۱۰۴۷	-۳٫۲۰۷۸	۲	۴	۲

همچنین در صورتی‌که در توزیع بتای تعمیم‌یافته‌ی نوع اول $a = -1$ و $q = 1$ باشد، توزیع پارتو، توزیع ماکسیمم آنتروپی خواهد بود.

۳- بحث و نتیجه‌گیری

در این مقاله، خانواده‌ی توزیع‌های بتای تعمیم‌یافته را تعریف و حالت‌های خاص این توزیع را معرفی کردیم و این خانواده از توزیع‌ها را از دیدگاه آنتروپی و ماکسیمم آنتروپی مورد بررسی قرار دادیم. همچنین براورد ماکسیمم آنتروپی پارامترهای این خانواده را با استفاده از نرم‌افزار متلب به‌دست آوردیم و در قالب هفت جدول ارایه کردیم.

مرجع‌ها

[۱] تارمست، قاسم؛ فخارزاده جهرمی، علیرضا و نیاپرست، مهرداد (۱۳۸۱). توزیع بتای تعمیم‌یافته‌ی بهینه برای داده‌های درامد و محاسبه‌ی ضریب جینی در استان خوزستان. مجله‌ی پژوهشی علوم پایه‌ی دانشگاه اصفهان، دوره‌ی ۱۶، شماره‌های ۲-۱، صص ۹۶-۷۹.

[2] Dragulescu, A.A. and Yakovenko, V.M. (2003). Statistical Mechanics of Money, Income, and Wealth: A Short Survey,

Modeling of Complex Systems: Seventh Granada Lectures, AIP Conference Proceedings 661, New York, 180-183.

[3] Kagan, A.M. and Rao, C.R. (1973). *Characterization Problems in Mathematical Statistics*, Wiley, New York.

[4] Lorenz, M.O. (1905). Method of Measuring the Concentration of Wealth, *Journal of American Statistical Association*, **9**, 209-219.

[5] Mohtashami Borzadaran, G.R. and Behdani, Z. (2009). Maximum Entropy and Entropy of Mixing for Income Distributions. *Journal of Income Distribution*, New Series, **18**: 2.

[6] Pareto, V. (1895). La Legga Della Domanda, Girnale Degli Economisti,/gennaio. 59-68.

[7] Shannon, C.E. (1948). A Mathematical Theory of Communication, *Bell System and Information Dynamics*, **13**, 75-90.

[8] Shorrocks, A.F. (1983). Ranking Income Distributions, *Economica*, **50**, 3-17.

[9] Yakovenko, V.M. (2003). Research in Econophysics, Cond-math. Stat-mech., published: *The Photon*, Issue **24**, January- February.

فاطمه عباس‌پور
فوق لیسانس آمار
مشهد، میدان آزادی، دانشگاه فردوسی مشهد، دانشکده‌ی علوم ریاضی، گروه آمار.

منیژه صانعی طبس
فوق لیسانس آمار
زاهدان، بلوار دانشگاه، دانشگاه سیستان و بلوچستان، دانشکده‌ی ریاضی، گروه آمار.
رایانشانی: manijesanei@gmail.com

غلامرضا محتشمی برزادران
دکتری آمار
مشهد، میدان آزادی، دانشگاه فردوسی مشهد، دانشکده‌ی علوم ریاضی، گروه آمار.
رایانشانی: grmohtashami@um.ac.ir