مجله‌ی بررسی‌های آمار رسمی ایران
سال ۲۱، شماره‌ی ۲، پاییز و زمستان ۱۳۸۹، صص ۱۵۱ – ۱۶۷

مشخصه‌سازی‌هایی برای ترتیب‌های لورنتس در خانواده‌ی توزیع‌های درامد

زهرا بهدانی،[†] غلامرضا محتشمی برزادران[‡,*] و یدالله و اقعی[⋆]

[†] مجتمع آموزش عالی بهبهان
[‡] دانشگاه فردوسی مشهد
[⋆] دانشگاه بیرجند

چکیده: منحنی لورنتس یکی از مهم‌ترین ابزار گرافیکی برای اندازه‌گیری نابرابری اقتصادی است، که برای اولین بار در سال ۱۹۰۵ توسط ماکس اوتو لورنز معرفی شد و از آن پس در زمینه‌های متنوع مورد مطالعه و بررسی قرارگرفت و هم اکنون به‌عنوان یک موضوع کاملاً شناخته‌شده در تحقیقات علمی کاربرد دارد. ترتیب لورنتس یک ترتیب جزئی و یک ابزار ساده برای مقایسه‌ی تغییرپذیری متغیرهای تصادفی نامنفی است که در بسیاری از زمینه‌ها از جمله اقتصاد، مطالعات اجتماعی و قابلیت اعتماد اهمیت دارد. بسط و توسعه این منحنی و ترتیب‌های لورنتس در بسیاری از علوم کاربردهای فراوانی را خواهد داشت که ما بیش‌تر به مشخصه‌هایی از این مفاهیم در آمار با تکیه بر کاربرد آن در اقتصاد توجه خواهیم نمود. در این مقاله در راستای تحقیقات انجام شده ابتدا خانواده‌ی توزیع بتا را معرفی و سپس حالت‌های خاص آن را با جایگذاری مقادیر خاص روی پارامترها و یا حالت‌های حدی به دست می‌آوریم. در ادامه ویژگی‌ها و مشخصه‌سازی‌هایی از ترتیب لورنتس برای خانواده‌ی توزیع درامد را بیان کرده و این ترتیب‌ها را به لحاظ نظری برای توزیع‌های مختلف مقایسه خواهیم کرد. متغیرهای تصادفی نظیر منحنی‌های لورنتس سلسله مراتبی دارای ترتیب‌سازی لورنتس ساده‌تری می‌باشند که در بخشی از مقاله به بیان مشخصه‌های آن‌ها می‌پردازیم. در نهایت منحنی لورنتس و ضریب جینی خانوارهای ایران در سال ۱۳۸۴ را براورد می‌کنیم و داده‌های درامد استان خراسان جنوبی را برحسب داده‌های درامد شهری و روستایی ترتیب‌سازی می‌کنیم.

---

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۸۸/۱۰/۱۵، پذیرش: ۱۳۸۹/۸/۲۵

واژگان کلیدی: توزیع درامد؛ منحنی لورنتس؛ ضریب جینی؛ ترتیب لورنتس؛ ترتیب ستاره؛ ترتیب محدب.

۱- مقدمه

شاید بارها این جمله که ثروتمندان هر روز ثروتمندتر و فقرا فقیرتر می‌شوند را شنیده باشید. بیان این جمله حکایت از وجود نابرابرهای اقتصادی در جامعه دارد. در خصوص توزیع درامد، قرن بیستم شاهد تلاش‌های بی‌سابقه‌ای در جهت کاهش جدی عدم تعادل (ناهمواری‌های اقتصادی) بوده است. توزیع‌های مثبت آماری به‌عنوان ابزاری مهم با نقش بسیار ارزنده در تجزیه و تحلیل داده‌های مثبت به کار می‌روند. با شناسایی منطقی یک توزیع معتبر و مناسب برای داده‌های کمی اقتصادی، استنباط‌ها و تفسیر نتایج به‌راحتی انجام می‌پذیرد. منحنی لورنتس بر اساس داده‌های توزیع درامدها، بین طبقات مختلف- که غالباً به‌طور غیر عادلانه صورت می‌گیرد، بدین معنی که اقلیتی از بین طبقات مختلف جامعه سهم بزرگی از درامد ملی را به خود تخصیص می‌دهند و بقیه‌ی درامد بین گروه‌های متعدد دیگر توزیع می‌شود- حاصل می‌گردد.

در این مقاله ابتدا خانواده‌ی توزیع بتا، که از مهم‌ترین توزیع‌های درامد و یک حالت کلی از دیگر توزیع‌های درامد می‌باشد را معرفی کرده سپس به معرفی منحنی لورنتس به‌عنوان مهم‌ترین ابزار گرافیکی برای مقایسه‌ی نابرابری‌های اقتصادی می‌پردازیم و در ادامه ترتیب لورنتس را تعریف کرده و برخی از مشخصه‌های آن را برای خانواده‌ی توزیع درامد بیان می‌کنیم. در نهایت منحنی لورنتس و ضریب جینی متناظر با داده‌های درامد ایران در سال ۱۳۸۴ را براورد می‌کنیم و داده‌های درامد استان خراسان جنوبی را برحسب داده‌های درامد شهری و روستایی ترتیب‌سازی می‌کنیم. از جمله تحقیقات انجام‌شده در سال‌های اخیر می‌توان به [۲]، [۷]، [۱۵]، [۱۱] و در ایران به [۱] اشاره کرد.

۲- توزیع‌های درامد

ویلفرد پاراتو اولین کسی بود که یک مدل از توزیع درامد را به شکل یک تابع چگالی احتمال معرفی کرد [۱۶]. در آمار، توزیع‌های احتمال زیادی وجود دارد که برای مدل‌سازی توزیع متغیرهای تصادفی پیوسته به کار می‌روند ولی از میان تمامی توزیع‌های احتمال،

توزیع‌های احتمال پیوسته با مقادیر نامنفی که عموماً چوله می‌باشند برای مدل‌سازی توزیع درامد مورد استفاده قرار می‌گیرند. توزیع بتای پنج پارامتری و حالت‌های خاص آن را می‌توان یکی از مهم‌ترین اعضای خانواده‌ی توزیع درامد دانست و معرفی کرد.

۱-۲- خانواده توزیع‌های بتا

توزیع بتای تعمیم‌یافته یک توزیع پنج پارامتری و دارای تابع چگالی احتمال زیر است ([۱۱]):

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

که $B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$، $a,b,p,q>0$، $0\le c\le 1$ و $\Gamma(\cdot)$ تابع گاما است.

اگردر رابطه‌ی (۱)، $c=1$ تابع چگالی احتمال بتا تعمیم‌یافته‌ی نوع دوم (فلر پاراتو یا پاراتو تعمیم‌یافته $(GB2(a,b,p,q))$ را نتیجه می‌دهد. در توزیع بتای تعمیم‌یافته‌ی نوع دوم اگر $p=1$ ، $a=1$ و $q=1$ توزیع به ترتیب با توزیع‌های سینگ مادالا $(SM\,(a,b,q))$، بتای نوع دوم $(B2(b,p,q))$ و داگم $(D\,(a,b,p))$ هم‌ارز است. هم چنین با جایگذاری $p=a=1$ توزیع لوماکس (پاراتوی نوع دوم) و $p=q=1$ توزیع لگ لوژستیک (فیسک) نتیجه می‌شود و اگر $a=p=q=b=1$ توزیع لگ لوژستیک تعمیم‌یافته‌ی استاندارد و $a=q=1$ توزیع لوماکس معکوس را نتیجه می‌دهد. توزیع گامای تعمیم‌یافته نیز با قرار دادن $q\to\infty$ و $b=q^{1/a}\beta$ حاصل می‌شود و از آن‌جائی که توزیع‌های گاما و وایبل نیز حالت‌های خاص گامای تعمیم‌یافته‌اند لذا حالت‌های حدی از $GB2$ نیز به‌شمار می‌روند.

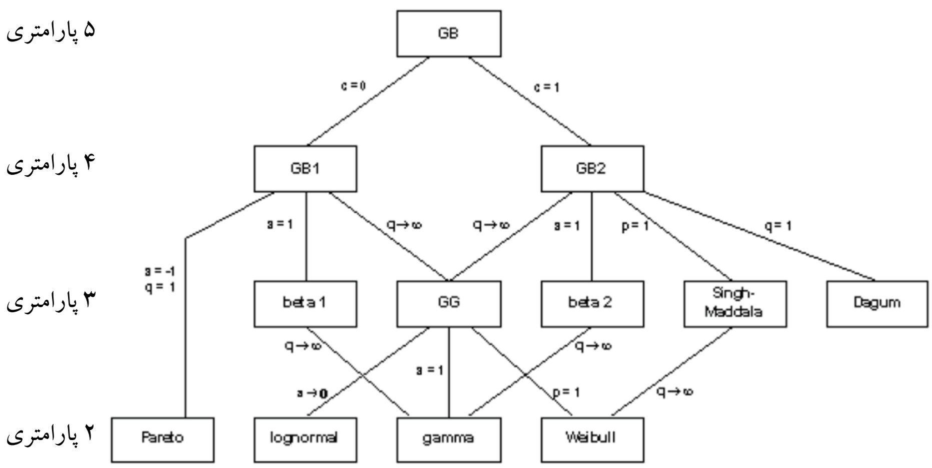
با قرار دادن $c=0$ در رابطه‌ی (۱) توزیع بتای تعمیم‌یافته‌ی نوع اول $(GB1(a,b,p,q))$ را داریم که در این رابطه اگر $q=1$ و $a=-1$ توزیع پاراتو، $a=1$ توزیع بتای نوع اول و $q\to\infty$ توزیع گامای تعمیم‌یافته را نتیجه خواهد

داد. توزیع‌های بنینی، دیویس و چامپرنون نیز از دیگر توزیع‌های درامد می‌باشند که برای مطالعه‌ی آن‌ها [۱۱] می‌تواند مورد استفاده قرار گیرد.

در راستای تحقیقات جانسون و کوتز [۹] یک حالت کلی از همه‌ی توزیع‌های درامد که به توزیع بتای تعمیم‌یافته‌ی نه پارامتری مشهور است را به‌صورت زیر تعریف می‌کنیم:

$$f(x)=cx^{-a}\left\{1-g\left(\frac{x}{b}\right)^{\lambda}\right\}^{q-1}\left\{1+h\left(\frac{x}{c}\right)^{\mu}\right\}^{r}, \quad 0\le x\le b \qquad (۲)$$

که در آن r، μ، h، q، λ، b، g، a و c پارامترهای توزیع می‌باشند. بسیاری از توزیع‌های درامد را می‌توان با جایگذاری مقادیر خاص روی پارامترها و یا حالت‌های حدی به دست آورد. با قرار دادن $e^{-\mu}=a_1$، $c=a_2a_3$، $r=-1-a_3$، $a=1-a_2$، $h=1$ و $g=0$ در رابطه‌ی (۲) یک توزیع سینگ مادالا با پارامترهای a_3، a_2 و a_1 نتیجه می‌شود. ولی کار کردن با توزیع بالا به‌دلیل زیاد بودن تعداد پارامترهای آن مشکل است. شکل ۱ رابطه‌ی بین توزیع بتای تعمیم‌یافته (پنج پارامتری) را با چند توزیع دیگر نشان می‌دهد ([۴]).



شکل ۱- چند حالت کلی از خانواده‌ی توزیع بتا و رابطه‌ی بین آن‌ها

۳- منحنی لورنتس

منحنی لورنتس برای اولین بار در سال ۱۹۰۵، توسط ماکس اوتو لورنز معرفی شد و از آن زمان، این منحنی در مطالعه‌ی توزیع‌های درامد در اقتصاد جزء مهم‌ترین ابزارها بوده است اگر تابع چندک (تابع معکوس F^{-1}) برای متغیر تصادفی X با تابع توزیع F به‌صورت زیر تعریف شود:

$$F^{-1}(t)=\sup\{x: F(x)\le t\},\quad t\in[0,1]$$

آن‌گاه منحنی لورنتس آن ($L(u)$) به‌صورت زیر تعریف می‌شود:

$$L(u)=\frac{1}{E(X)}\int_0^u F^{-1}(t)dt,\qquad u\in[0,1] \tag{۳}$$

نتایج زیر برای منحنی لورنتس به‌طور مستقیم از رابطه‌ی (۳) نتیجه می‌شود:

- L روی $[0,1]$ پیوسته با $L(0)=0$ و $L(1)=1$ است.
- L یک تابع صعودی است.
- L یک تابع محدب است.

بر عکس هر تابع با این ویژگی‌ها یک تابع لورنتس از یک توزیع آماری معین است.

شاخص‌های بسیاری بر اساس منحنی لورنتس تعریف می‌شوند ضریب جینی از مهم‌ترین اندازه‌هایی است که برای سنجش میزان نابرابری درامد در جامعه کاربرد دارد. ضریب جینی اولین بار در سال ۱۹۱۲ توسط کورادو جینی [۸] معرفی شد و به‌صورت دو برابر ناحیه‌ی بین منحنی لورنتس و خط قطر، به‌صورت زیرتعریف می شود:

$$G=2\int_0^1\{u-L(u)\}du=1-2\int_0^1 L(u)du$$

که همواره نامنفی و مقداری بین صفر و یک خواهد بود و بهترین حالت آن وقتی است که مقدار آن برابر صفر باشد و این وقتی اتفاق می‌افتد که $L(u)=u$ باشد.

۴- ترتیب‌های لورنتس

روش‌ها و ابزار متفاوتی برای مقایسه‌ی تغییرپذیری و پراکندگی توزیع‌های احتمال وجود دارد مانند واریانس، برد، میانه، واریانس لگاریتم و در این میان ترتیب‌ها و از جمله ترتیب لورنتس یک ابزار ساده برای مقایسه‌ی تغییرپذیری متغیرهای تصادفی نامنفی است. ترتیب لورنتس یکی از ترتیب‌های تصادفی است که با توجه به منحنی لورنتس متغیرهای تصادفی نامنفی $X_۱$ و $X_۲$ با میانگین‌های متناهی به‌صورت زیر تعریف می‌شود:

$$X_۱ \geq_L X_۲ \Leftrightarrow F_۱ \geq_L F_۲ \Leftrightarrow L_۱(u) \leq L_۲(u), \quad \forall u \in [۰,۱]$$

که $X_۱$ بزرگ‌تر از $X_۲$ در حالت لورنتس خوانده می‌شود و یا گوییم $F_۱$ ناهمواری کم‌تری نسبت به $F_۲$ نمایش می‌دهد. ترتیب‌های تصادفی ستاره و محدب ترتیب لورنتس را نتیجه می‌دهند و به‌صورت زیر تعریف می‌شوند:

تعریف ۱. گوییم متغیر تصادفی X از نظر ترتیب محدب کوچک‌تر از متغیر تصادفی Y است و می‌نویسیم $X \leq_{CX} Y$، اگر به‌ازای هر تابع محدب h داشته باشیم:

$$E[h(X)] \leq E[h(Y)], \qquad h: R \to R \ .$$

تعریف۲. برای دو متغیر تصادفی نامنفی X و Y با توابع توزیع F و G و توابع چندک $F^{-۱}$ و $G^{-۱}$، گوییم متغیر تصادفی X از نظر ترتیب ستاره ای کوچک‌تر از متغیر تصادفی Y است و می‌نویسیم $X \leq_* Y$، اگر $\frac{G^{-۱}(u)}{F^{-۱}(u)}$ به‌ازای هر $u \in (۰,۱)$ صعودی باشد.

شیکد و شانتیکومار برای متغیرهای تصادفی نامنفی X و Y با میانگین‌های متناهی تعریف $\frac{X}{E(X)} \leq_{CX} \frac{Y}{E(Y)}$ را برای ترتیب لورنتس ارایه دادند [۱۷].

با استفاده از تعریف منحنی لورنتس کاملاً مشخص است که ترتیب لورنتس نسبت به مقیاس پایا است:

$$X_۱ \geq_L X_۲ \Leftrightarrow aX_۱ \geq_L bX_۲ \qquad a,b > ۰$$

و می‌توان نشان داد که برای متغیرهای تصادفی نامنفی $X_{۱}$ و $X_{۲}$ با میانگین‌ها و واریانس‌های متناهی به ترتیب $E(X_{۱})$ و $E(X_{۲})$ و $\sigma_{۱}^{۲}$ و $\sigma_{۲}^{۲}$ داریم:

$$F_{۱} \leq_L F_{۲} \Rightarrow \left(\frac{\sigma_{۱}}{E(X_{۱})}\right)^{۲} \leq \left(\frac{\sigma_{۲}}{E(X_{۲})}\right)^{۲}.$$

قضیه‌ی ۱. *اگر $X_{۱}$ و $X_{۲}$ دو متغیر تصادفی مثبت با میانگین متناهی و توابع چندک $F_{۱}^{-۱}(t)$ و $F_{۲}^{-۱}(t)$ باشد و اگر $U(t)=\frac{F_{۱}^{-۱}(t)}{F_{۲}^{-۱}(t)}$، آن‌گاه:*

۱) $X_{۱} \geq_L X_{۲}$ *اگر* $U(t)$ *در* $(۰,۱)$ *صعودی باشد.*

۲) $X_{۱} \leq_L X_{۲}$ *اگر* $U(t)$ *در* $(۰,۱)$ *نزولی باشد.*

۳) $X_{۱} =_L X_{۲}$ *اگر* $U(t)=C$ *در* $(۰,۱)$ *مقداری ثابت باشد.*

قضیه‌ی ۲ رابطه‌ی بین ترتیب لورنتس و ترتیب ستاره را بیان می‌کند.

قضیه‌ی ۲ ([۱۷]). *فرض کنید X و Y متغیرهای تصادفی نامنفی با میانگین متناهی باشد. اگر $X \leq_* Y$ آن‌گاه $X \leq_L Y$.*

این قضیه کاربردی از قسمت ۱ قضیه‌ی ۱ می‌باشد.

کلفسجو رابطه‌ی بین ترتیب ستاره و ترتیب لورنتس را بدون ساختن هیچ محدودیتی روی میانگین‌های دو توزیع اثبات کرد [۱۰]. در حقیقت او ثابت کرد که $X \leq_* Y$ نتیجه می‌دهد $\frac{L_G(u)}{L_F(u)}$ صعودی است در $u \in [۰,۱]$ که در نتیجه $L_F(u) \geq L_G(u)$ برای $u \in [۰,۱]$.

مثال ۱. اگر $P(X=۱)=P(X=۲)=\frac{۱}{۲}$ و $P(Y=۲)=P(Y=۵)=\frac{۱}{۲}$ آنگاه $\frac{G^{-۱}(u)}{F^{-۱}(u)}=\begin{cases} ۲ & ۰<u \leq ۱/۲, \\ ۵/۲ & ۱/۲<u<۱. \end{cases}$ تابعی صعودی است یعنی $X \leq_* Y$ و در نتیجه $X \leq_L Y$ ([۱۷]).

قضیه‌ی ۳ ([۷]). *فرض کنید $g(x)$ یک تابع پیوسته و اکیداً صعودی باشد. اگر $\mu^* = E\{g(X)\}$ موجود باشد آن‌گاه $L^*(u)$ (منحنی لورنتس $Y = g(X)$) موجود و نتایج زیر برقرار است:*

۱) $L^*(u) \geq L(u)$ *اگر* $\frac{g(x)}{x}$ *اکیداً نزولی باشد.*

۲) $L^*(u) = L(u)$ *اگر* $\frac{g(x)}{x}$ *ثابت باشد.*

۳) $L^*(u) \leq L(u)$ *اگر* $\frac{g(x)}{x}$ *اکیداً صعودی باشد.*

خواننده را برای مشاهده‌ی اثبات این قضیه به [۷] ارجاع می‌دهیم.

جدول ۱ برخی از توزیع‌های خاص را با مشخص کردن $Y = g(X)$ و استفاده از قضیه‌ی ۳ ترتیب‌سازی می‌کند که توسط نویسندگان انجام شده است.

۵- توزیع‌های پارامتری از منحنی‌های لورنتس

با توجه به اهمیت منحنی لورنتس [۱۴] در تحلیل‌های اقتصادی و آماری ناهمواری‌های درامد، مدل‌های پارامتری زیادی برای تقریب منحنی‌های لورنتس تجربی پیشنهاد شده‌اند. اخیراً یک خانواده از منحنی‌های لورنتس را که ترکیبی از دیگر منحنی‌های لورنتس می‌باشد (منحنی‌های لورنتس سلسله مراتبی) معرفی شده است که در این دیدگاه از هر منحنی لورنتس $L_{\circ}$ دنباله منحنی‌های بعدی، تعمیمی از مدل اولیه‌ی $L_{\circ}$ خواهد بود.

- $L_{۱}(u;\alpha) = u^{\alpha} L_{\circ}(u)$، $\circ < u < ۱$ که یا $\alpha \geq ۱$ و یا $\circ \leq \alpha < ۱$ و $L_{\circ}'''(u) \geq \circ$.
- $L_{۲}(u;\gamma) = \{L_{\circ}(u)\}^{\gamma}$، $\circ < u < ۱$ که $\gamma \geq ۱$.
- $L_{۳}(u;\alpha,\gamma) = u^{\alpha}\{L_{\circ}(u)\}^{\gamma}$، $\circ < u < ۱$ و $\alpha, \gamma \geq ۱$.

یک مزیت این دیدگاه محاسبه‌ی آسان ترتیب‌های لورنتس است در حقیقت:

جدول ۱- ترتیب‌سازی لورنتس برای برخی از توزیع‌ها با استفاده از قضیه‌ی ۳

ترتیب لورنتس	وضعیت تابع $\frac{g(x)}{x}$ در کل دامنه	تابع $y = g(x)$	X و Y
$L_Y(u) \le L_X(u)$	یک تابع صعودی	$g(x) = \frac{\beta}{\alpha} \cdot \frac{x}{1-x}$	$X \sim B(\alpha, \beta)$ $Y \sim F(2\alpha, 2\beta)$
برای $\alpha > 1$ ، $L_Y(u) \le L_X(u)$ برای $\alpha < 1$ ، $L_Y(u) \ge L_X(u)$ برای $\alpha = 1$ ، $L_Y(u) = L_X(u)$	برای $\alpha > 1$ یک تابع صعودی و برای $0 < \alpha < 1$ یک تابع نزولی و برای $\alpha = 1$ ثابت	$g(x) = x^{1/\alpha}$	$X \sim \exp(\lambda)$ $Y \sim \text{wei}(\alpha, \lambda)$
برای $x > e$ $L_Y(u) \ge L_X(u)$ $0 < x < e$ ، $L_Y(u) \ge L_X(u)$ برای $x = e$ ، $L_Y(u) = L_X(u)$	برای $x > e$ یک تابع نزولی برای $0 < x < e$ یک تابع صعودی و برای $x = e$ ثابت	$g(x) = \log(x)$	$X \sim \text{Par}(\alpha, x_0)$ $Y \sim \exp(\alpha)$ $X \sim \ln(\mu, \sigma^2)$ $Y \sim N(\mu, \sigma^2)$

- $L_1(u;\alpha_1) \ge_L L_1(u;\alpha_2)$ اگر و فقط اگر $\alpha_1 \ge \alpha_2 > 0$.
- $L_2(u;\gamma_1) \ge_L L_2(u;\gamma_2)$ اگر و فقط اگر $\gamma_1 \ge \gamma_2 > 0$.
- یک ترکیب از حالت‌های قبل نتایج را برای L_3 نتیجه می‌دهد.

۶- ترتیب لورنتس برای خانواده‌ی توزیع‌های درامد

بعد از اتکینسون، [۳]، در زمینه‌ی ترتیب‌های تصادفی و از جمله ترتیب لورنتس برای مقایسه‌ی توزیع‌های درامد تلاش‌های زیادی انجام گرفته است ولی تلاش‌های اخیر فقط ویژگی‌هایی از ترتیب لورنتس در خانواده‌ی توزیع‌های درامد پارامتری را مشخصه‌سازی نموده است. ترتیب لورنتس برای تابع‌های چگالی احتمال یک و دو پارامتری ساده مانند لگ‌نرمال و پاراتو دارای فرم خطی است. برای مثال برای توزیع پاراتو داریم:

$$L(u)=۱-(۱-u)^{۱-\frac{۱}{\alpha}}, \qquad ۰<u<۱.$$

می‌توان نشان داد:

$$X_۱ \geq_L X_۲ \Leftrightarrow \alpha_۱ \leq \alpha_۲.$$

به‌عبارت دیگر در توزیع پاراتو با افزایش پارامتر α منحنی لورنتس متناظر با آن توزیع نیز افزایش می‌یابد و از منحنی لورنتسی با پارامتر α کوچک‌تر، بالاتر است و برای توزیع لگ‌نرمال منحنی لورنتس آن به‌صورت زیر است:

$$L(u)=\Phi\left\{\Phi^{-۱}(u)-\sigma^۲\right\}, \qquad ۰<u<۱$$

و این نتیجه می‌دهد که $X_۱ \geq_L X_۲ \Leftrightarrow \sigma_۱^۲ \geq \sigma_۲^۲$ ([۱۱]). به‌عبارت دیگر در توزیع لگ‌نرمال هر چه پارامتر $\sigma^۲$ بیش‌تر باشد تابع لورنتس پایین‌تر خواهد بود. در خانواده‌های سه و چهار پارامتری ترتیب لورنتس دیگر دارای فرم خطی نیست.

قضیه‌ی ۴ ([۱۲]). *فرض کنید $X_۱$ و $X_۲$ متغیرهای تصادفی نامنفی و $X_۱ \sim GB۲(a_۱,b_۱,p_۱,q_۱)$ و $X_۲ \sim GB۲(a_۲,b_۲,p_۲,q_۲)$ در این صورت:*

۱) *اگر $a_۱ p_۱ \leq a_۲ p_۲$ ، $a_۱q_۱ \leq a_۲q_۲$ و $a_۱ \leq a_۲$ آنگاه $X_۱ \geq_L X_۲$*

۲) *اگر $X_۱ \geq_L X_۲$ آنگاه $a_۱p_۱ \leq a_۲p_۲$ و $a_۱q_۱ \leq a_۲q_۲$.*

از آنجا که بیش‌تر توزیع‌های درامد حالت خاصی از توزیع $GB۲$ اند لذا قضیه‌ی بالا برای همه‌ی زیرخانواده‌های این توزیع نیز نتایج زیر را ارایه می‌دهد [۱۲].

برای توزیع سینگ مادلا $SM(a,b,q) \equiv GB۲(a,b,۱,q)$ داریم:

$$X_۱ \geq_L X_۲ \Leftrightarrow a_۱ \leq a_۲ \text{ و } a_۱q_۱ \leq a_۲q_۲.$$

برای توزیع داگم با $D(a,b,p) \equiv GB۲(a,b,p,۱)$ داریم:

$$X_۱ \geq_L X_۲ \Leftrightarrow a_۱ \leq a_۲ \text{ و } a_۱p_۱ \leq a_۲p_۲.$$

برای توزیع بتای نوع دوم، $B۲(b,p,q) \equiv GB۲(۱,b,p,q)$

$$X_1 \ge_L X_2 \Leftrightarrow p_1 \le p_2 \text{ و } q_1 \le q_2,$$

و برای توزیع لگ لوژستیک یا پاراتو $LL(a,b) \equiv GB2(a,b,1,1)$ داریم:

$$X_1 \ge_L X_2 \Leftrightarrow a_1 \le a_2$$

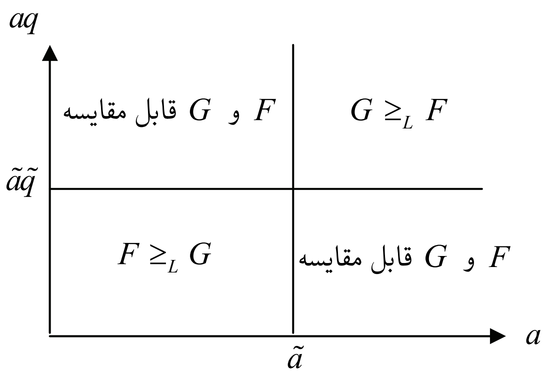
شکل ۲ تجزیه‌ی فضای (a,aq) را برای ترتیب لورنتس بین توزیع‌های $F \sim SM(a,q)$ و $G \sim SM(\tilde{a},\tilde{q})$ نشان می‌دهد [۲].

قضیه‌ی بعد مشخصه‌هایی از ترتیب لورنتس را برای دو توزیع متفاوت بیان می‌کند.

قضیه‌ی ۵ ([۱۲]). *فرض کنید یکی از شرط‌های زیر برقرار باشد:*

۱- $Y \sim GB2(a,p,q)$، $X \sim GG(\tilde{a},\tilde{p})$ با $\tilde{a}\tilde{p} \ge a$ و $aq > 1$، $\tilde{a} \ge a$
۲- $X \sim GB1(a,p,q)$، $Y \sim GB2(\tilde{a},\tilde{p},\tilde{q})$ با $\tilde{a}\tilde{q} > 1$، $\tilde{a} \ge a$، $ap \ge \tilde{a}\tilde{p}$ و $aq > \tilde{a}\tilde{q}$
۳- $X \sim GB1(a,p,q)$، $Y \sim GG(\tilde{a},\tilde{p})$ با $ap \ge \tilde{a}\tilde{p}$ و $\tilde{a} \ge a$

آن‌گاه $X \le_L Y$.



شکل ۲- تجزیه‌ی فضای (a,aq) برای ترتیب لورنتس توزیع‌های $F \sim SM(a,q)$ و $G \sim SM(\tilde{a},\tilde{q})$

در نتیجه می‌توان حالت خاص این قضیه را برای زیرخانواده‌هایی از این توزیع‌ها بیان کرد.
نتیجه‌ی ۱ ([۱۲]). اگر یکی از شرایط زیر برقرار باشد:

$X \sim G(\tilde{p})$، $Y \sim B2(p,q)$ با $\tilde{p} \ge p$ و $q > 1$؛
$X \sim W(\tilde{a})$، $Y \sim SM(a,q)$ با $aq > 1$ و $\tilde{a} \ge a$؛
$X \sim E(1)$، $Y \sim L(q)$ با $q > 1$؛
$X \sim W(\tilde{a})$، $Y \sim LL(a)$ با $\tilde{a} \ge a > 1$؛
$X \sim GG(\tilde{a},\tilde{p})$، $Y \sim D(a,p)$ با $a > 1$، $\tilde{a}\tilde{p} \ge ap$ و $\tilde{a} \ge a$؛
آن‌گاه $X \le_L Y$.

قضیه‌ی بعد به بیان ویژگی‌های ترتیب لورنتس برای میانگین نمونه می‌پردازد.

قضیه‌ی ۶ ([۲]). *برای هر* n، $\bar{X}_n \le_L \bar{X}_{n-1}$.

جدول ۲ که توسط نویسندگان تهیه شده است شرایط لازم برای برقراری ترتیب لورنتس بین برخی از توزیع‌های درامد را نشان می‌دهد. رابطه‌ی واقع در سطر i ام و ستون j ام جدول شرط لازم و کافی برای برقراری ترتیب لورنتس برای توزیع واقع در خانه‌ی j ام سطر اول جدول و خانه‌ی i ام ستون اول می‌باشد.

۷- مثال‌های کاربردی

شکل ۳ منحنی لورنتس خانوار ایران در سال ۱۳۸۴ را نشان می‌دهد، این شکل منحنی لورنتس متناظر با داده‌های هزینه‌ی خانوار نمونه‌گیری‌شده از طرح «آمارگیری از هزینه و درامد خانوارهای شهری و روستایی» سال ۱۳۸۴ می‌باشد. جدول ۳ مقدار ضریب جینی متناظر با این منحنی‌ها را نشان می‌دهد.

شکل ۴ منحنی لورنتس داده‌های درامد حاصل از مشاغل مزد و حقوق‌بگیری شهری و روستایی خانوار استان خراسان جنوبی (سمت راست) و ضریب جینی هر یک را نشان می‌دهد. همان‌طور که از شکل مشخص است در این سال در کشور، داده‌های درامد خانوار شهری ناهمواری کم‌تری نسبت به افراد روستایی نشان می‌دهد به بیان دیگر خانوار شهری استان از نظر ترتیب لورنتس کوچک‌تر از خانوار روستایی آن می‌باشند در حالی که در

جدول ۲- ترتیب‌سازی برای متغیرهای تصادفی X_1 و X_2 (توزیع‌های واقع در سطر اول مربوط به متغیرهای تصادفی X_1 و توزیع‌های واقع در ستون اول مربوط به متغیر تصادفی X_2 می‌باشد) داریم: $X_1 \geq_L X_2$

X_2 \ X_1	$GB2(a_1,b_1,p_1,q_1)$	$GG(a_1,b_1,p_1)$	$D(a_1,b_1,p_1)$	$SM(a_1,b_1,q_1)$	$B2(b_1,p_1,q_1)$
$GB2(a_2,b_2,p_2,q_2)$	$a_1 \leq a_2$ $a_1p_1 \leq a_2p_2$ $a_1q_1 \leq a_2q_2$	$a_1 \leq a_2$ $a_1p_1 \leq a_2$ $1 \leq a_2q_2$	$a_1 \leq a_2$ $a_1p_1 \leq a_2$ $1 \leq a_2q_2$	$a_1 \leq a_2$ $a_1 \leq a_2p_2$ $a_1q_1 \leq a_2q_2$	$q_1 \geq 1$ $p_1 \leq a_2p_2$ $q_1 \leq a_2q_2$
$GG(a_2,b_2,p_2)$	$a_1 \leq a_2$ $a_1 \leq a_2p_2$ $1 \leq a_1q_1$	$a_1 \leq a_2$ $a_1p_1 \leq a_2p_2$ $1 \leq a_2$	$a_1 \leq a_2$ $a_1p_1 \leq a_2p_2$ $1 \leq a_1$	$a_1 \leq a_2$ $a_1 \leq a_2p_2$ $1 \leq a_1q_1$	$p_1 \leq p_2$ $q_1 \geq 1$
$D(a_2,b_2,p_2)$	$a_1 \leq a_2$ $a_1 \leq a_2p_2$ $1 \leq a_1q_1$	$a_1 \leq a_2$ $a_1p_1 \leq a_2p_2$ $1 \leq a_2$	$a_1 \leq a_2$ $a_1p_1 \leq a_2p_2$	$a_1 \leq a_2$ $1 \leq a_1q_1$	$p_1 \leq p_2$ $q_1 \geq 1$
$SM(a_2,b_2,q_2)$	$a_1 \leq a_2$ $a_1p_1 \leq a_2$ $a_1q_1 \leq a_2q_2$	$a_1 \leq a_2$ $1 \leq a_2q_2$	$a_1 \leq a_2$ $1 \leq a_2q_2$	$a_1 \leq a_2$ $a_1q_1 \leq a_2q_2$	$q_1 \leq q_2$ $a_2q_2 \geq 1$
$B2(b_2,p_2,q_2)$	$a_1q_1 > 1$ $a_1p_1 \leq p_2$ $a_1q_1 \leq q_2$	$p_1 \leq p_2$ $q_2 \geq 1$	$p_1 \leq p_2$ $q_2 \geq 1$	$q_1 \leq q_2$ $a_1q_1 \geq 1$	$p_1 \leq p_2$ $q_1 \leq q_2$

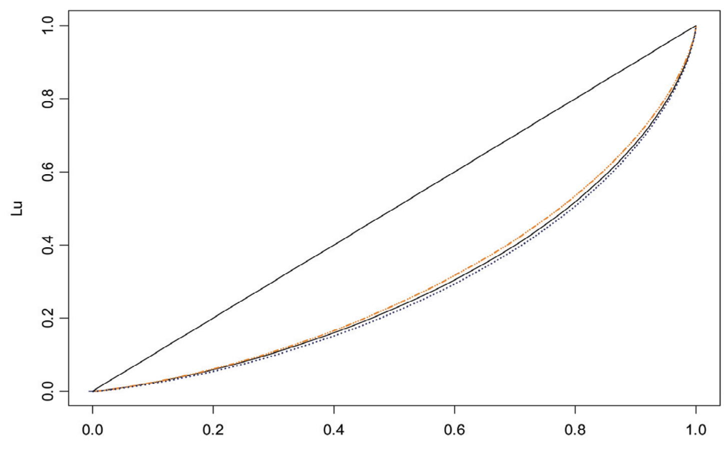
جدول ۳- ضریب جینی داده‌های خانوار شهری و روستایی در سال ۱۳۸۴

	هزینه‌ی کل	هزینه‌ی شهری	هزینه‌ی روستایی
ضریب جینی	۰/۴۳۲	۰/۴۱۶	۰/۳۹۹

استان، منحنی لورنتس درامد خانوار شهری و روستایی متقاطع و نسبتاً مشابه است و با توجه به ضریب جینی آن‌ها می‌توان گفت خانوار روستایی در ردیف درامدهای حاصل از مشاغل مزد و حقوق‌بگیری دارای تغییرات و ناهمواری کم‌تری نسبت به خانوار شهری می‌باشند.

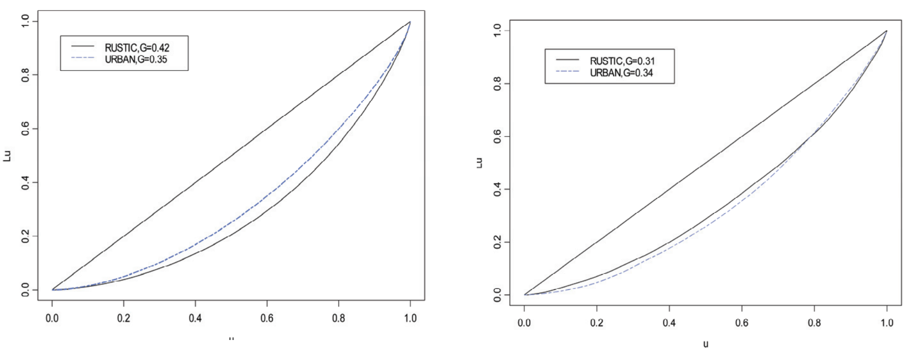
۸- نتیجه

در این تحقیق به بررسی مشخصه‌هایی از ترتیب‌های لورنتس برای برخی از خانواده‌ی توزیع‌های درامد پرداختیم و برخی از شرایط لازم و کافی برای برقراری ترتیب لورنتس بین متغیرهای تصادفی این توزیع‌ها را ذکر نمودیم و مشاهده کردیم که این ترتیب‌ها در خانواده‌های چند پارامتری غیر خطی، ولی در خانواده‌ی یک و دو پارامتری دارای فرمی خطی می‌باشند. نتایج تحقیقات روی هزینه‌های شهری و روستایی کشور در



شکل ۳- منحنی لورنتس برای داده‌های هزینه‌ی کشور در سال ۱۳۸۴ (شهری ممتد و روستایی خط‌چین)

سال ۱۳۸۴ بیان‌گر ناهمواری بیش‌تر هزینه‌ی شهرنشینان نسبت به روستاییان بود. از طرفی مقایسه‌ی درامد مزد و حقوق‌بگیری کشور و استان خراسان جنوبی در این سال ناهمواری کم‌تر در استان نسبت به کشور را نشان می‌دهد. نتایج نشان می‌دهد در میان این قشر از جامعه در استان، روستاییان دارای ضریب جینی کم‌تری (۰/۳۱) نسبت به شهرنشینان (۰/۳۴) هستند. در حالی که در کشور روستاییان دارای ضریب جینی (۰/۴۲) بالاتر از ضریب جینی شهرنشینان (۰/۳۵) می‌باشند. ترتیب‌های تصادفی یکی از مهم‌ترین ترتیب‌های تصادفی است که در بسیاری از زمینه‌ها و شاخه‌های علوم، کاربرد دارد و در ادامه‌ی این تحقیق کارهای زیادی از جمله ترتیب‌های لورنتس برای توزیع آماره‌های مرتب بر اساس یک نمونه‌ی تصادفی $X_۱,X_۲,...,X_n$ یا توزیع رکوردها نیز قابل بررسی و محاسبه می‌باشد که در بعضی شرایط نکات جالبی را نتیجه می‌دهد. در راستای تحقیقات کوچار [۱۳]، بارتزویچ [۵] و بارتزویچ و اسکولیموسکا [۶] ترتیب‌های لورنتس برای آماره‌های مرتب، رکوردها و توزیع‌های وزنی از موضوعات جالب توجه می‌باشد که به‌عنوان ادامه‌ی تحقیق مد نظر قرار گرفته است.



شکل ۴- منحنی لورنتس داده‌های درامد حاصل از مشاغل مزد و حقوق‌بگیری افراد شهری (خط ممتد) و افراد روستایی (خط‌چین) استان خراسان جنوبی (نمودار سمت راست) و کشور (نمودار سمت چپ)

سپاسگزاری

نویسنده‌ی اول این مقاله از حمایت پژوهشکده‌ی آمار و نویسنده‌ی دوم از حمایت قطب داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد تشکر می‌نمایند.

مرجع‌ها

[۱] پروین، سهیلا (۱۳۸۴). نقش انواع درامدها در نابرابری توزیع درامد در ایران. مجله تحقیقات اقتصادی، شماره‌ی ۷۵، صص ۱۱۱-۱۲۸.

[2] Arnold, B.C, Villasenor, J.A. (1985). Lorenz ordering of mean and medians. *Statistics & Probability Letters*, **4**, 47–49.

[3] Atkinson, A.B. (1970). On the measurement of inequality. *Journal of Economic Theory*. **2**, 244–263.

[4] Bandourian, R., McDonald, J. (2002). A comparison of parametric models of income distribution across countries and over time. Luxembourg income study working paper. No, 305.

[5] Bartoszewicz, J. (2005). Dispersive ordering between order statistics and spacings from an IRFR distribution. University of Wroc Law.

[6] Bartoszewicz, J., Skolimowska, M. (2006). Preservation of classes of life distributions under weighting. *Statistics & Probability Letters*, **76**, 587–596.

[7] Fellman, J. (1976). The effect of transformations of Lorenz curves. *Econometrica*. **44**, 823–824.

[8] Gini, C. (1912). Variabilita' e mutabilita, studio Economicogiuridici, universita di Cagliari Anno III, Parte 2a, reprinted in C, 211–382.

[9] Johnson, N.Lm., Kotz, S. (1970). *Continuous Univariate Probability Distributions*. Houghton Mifflin, Boston.

[10] Klefsjo, B. (1984). Reliability interpretations of some concepts from economics. *Naval Res. Logistics Quart*. **31**, 301–308.

[11] Kleiber, C., Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, wiley, New York.

[12] Kleiber, Ch. (1999). On the Lorenz order within parametric familes of income distributions. *The Indian Journal of Statistics*. **61**, 514–517.

[13] Kochar, S. (2006). Lorenz ordering of order statistics. *Statistics & Probability Letters*, **76**, 1855–1860.

[14] Lorenz, M.O. (1905). Method of measuring the concentration of wealth. *Journal of the American Statistical Association*. **9**, 209–219.

[15] Moothathu, T.S.K., (1991). On the sufficient condition for two non-intersecting Lorenz curves .*The Indian Journal of statistics*. **53**, series, B. pt, 2, 268–274.

[16] Pareto, V. (1895). La Legge della Domanda, Giornale degli Economisti, Gennaio. 59–68.

[17] Shaked, M., Shanthikumar, J. (1994). *Stochastic Orders and Their Applications*. Academic Press, New York.

زهرا بهدانی
کارشناس ارشد آمار
بهبهان، مجتمع آموزش عالی بهبهان، گروه ریاضی.
رایانشانی: zbehdani@yahoo.com

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

یدالله واقعی
دکتری آمار
بیرجند، دانشگاه بیرجند، دانشکده‌ی علوم، گروه آمار.
رایانشانی: ywaghei@yahoo.com