© Statistical Research and Training Center
J. Statist. Res. Iran 5 (2008): 57–74
مجله‌ی پژوهش‌های آماری ایران
پژوهشکده‌ی آمار
دوره‌ی ۵، شماره‌ی ۱، بهار و تابستان ۱۳۸۷، صص ۵۷–۷۴

روش‌های برآورد پارامترهای فرم‌های تابعی لورنتس و مقایسه‌ی آن‌ها بر اساس داده‌های هزینه‌ی خانوار

آزاده مجیری،[†] غلامرضا محتشمی برزادران[‡,*] و یدالله واقعی[†]

[†] دانشگاه بیرجند

[‡] دانشگاه فردوسی مشهد

چکیده. منحنی لورنتس مهم‌ترین ابزار گرافیکی است که برای توصیف اندازه‌ی تمرکز در جامعه، مانند ثروت استفاده می‌شود، در چنین مواردی بسیاری از اقتصاددانان نابرابری اقتصادی را اندازه‌گیری می‌کنند. شاخص‌های بسیاری بر اساس منحنی لورنتس برای اندازه‌گیری میزان نابرابری تعریف می‌شوند. ضریب جینی یکی از مهم‌ترین این شاخص‌هاست.

در این مقاله، ابتدا به معرفی منحنی لورنتس، ضریب جینی و فرم‌های تابعی لورنتس می‌پردازیم. سپس پارامترهای فرم‌های تابعی لورنتس را به سه روش کم‌ترین توان‌های دوم، ماکسیمم درست‌نمایی و صدک اصلی برآورد می‌کنیم. در نهایت بر اساس داده‌های هزینه‌ی خانوار شهری و روستایی ایران در سال ۱۳۸۴ بهترین فرم تابعی و بهترین روش برآورد را معرفی می‌کنیم.

واژگان کلیدی. منحنی لورنتس؛ ضریب جینی؛ فرم‌های تابعی؛ برآورد ماکسیمم درست‌نمایی؛ برآورد کم‌ترین توان‌های دوم؛ صدک اصلی.

۱ مقدمه

همان‌طور که می‌دانیم آمار یکی از مهم‌ترین علوم کاربردی است که با سایر رشته‌های علمی از جمله اقتصاد مرتبط می‌باشد. در کشورهای کم‌تر توسعه‌یافته، سیستم توزیع درآمدها غیر عادلانه و بیش‌تر به نفع صاحبان

---

*نویسنده‌ی عهده‌دار مکاتبات

زمین‌ها، سرمایه‌دارها و کارفرمایان است. این گروه‌ها، گروه‌های مسلط و حاکم جامعه را تشکیل می‌دهند. در این کشورها کشاورزان اکثراً فاقد زمین، و جمعیت فعال، غالباً بیکار است. در نتیجه‌ی این طبقه‌بندی گروه‌ها، قسمت بیش‌تر درآمدها بین اقلیت محدودی از جمعیت و قسمت کمی از آن، بین اکثریت مردم تقسیم می‌شود. بر اساس این نوع توزیع غیر عادلانه‌ی درآمدها، منحنیی حاصل می‌شود که برای اولین بار توسط ماکس اوتو لورنز در سال ۱۹۰۵ معرفی شد و به منحنی لورنتس معروف است. برای مطالعه‌ی بیش‌تر، خواننده را به کاکوانی (۱۹۸۰)، آرنولد (۱۹۸۷)، موتاتو (۱۹۹۱)، ساربیا و دیگران (۲۰۰۵) و کلیبر (۲۰۰۵) ارجاع می‌دهیم.

ضریب جینی، یک مقیاس پراکندگی آماری است که اغلب شاخصی برای اندازه‌گیری نابرابری توزیع درآمد مورد استفاده قرار می‌گیرد و به‌صورت دو برابر ناحیه‌ی بین منحنی لورنتس و خط قطر تعریف می‌شود که همواره نامنفی است و مقادیری بین صفر و یک را می‌گیرد. اگرچه می‌توان منحنی لورنتس را به طور مستقیم از داده‌های تجربی، محاسبه کرد، اما یکی از راه‌های برآورد منحنی لورنتس این است که منحنی لورنتس داده‌ها را به شکل نقطه به نقطه برآورد کرده و یک خانواده‌ی پارامتری به آن برازش دهیم.

هدف از این تحقیق به دست آوردن یک فرم کلی است که در شرایط تابع لورنتس صدق کند یعنی:

$$\begin{aligned}
&(a)\quad L(0)=0\\
&(b)\quad L(1)=1\\
&(c)\quad L'(u)\geqslant 0,\qquad 0\leqslant u<1\\
&(d)\quad L''(u)\geqslant 0,\qquad 0\leqslant u\leqslant 1\\
&(e)\quad L(u)\leqslant u,\qquad 0<u<1\\
&(f)\quad 0\leqslant \int_0^1 L(u)du\leqslant \frac{1}{2}.
\end{aligned}$$

بنا بر این، در ابتدا منحنی لورنتس و ضریب جینی را معرفی کرده، در بخش بعد برخی از فرم‌های تابعی لورنتس را معرفی می‌کنیم. سپس پارامترهای آن‌ها را به سه روش، برآورد می‌کنیم و در پایان نتایج را بر اساس داده‌های هزینه‌ی خانوار شهری و روستایی بیان می‌کنیم.

۲ منحنی لورنتس و ضریب جینی

فرض کنید درآمد n فرد از جامعه یا نمونه را با $x_۱$، $x_۲$، $\ldots$ ، x_n نشان دهیم. سپس درآمدها را از کوچک به بزرگ مرتب می‌کنیم و مرتب‌شده‌ی آن را از کوچک به بزرگ با $x_{۱:n}$، $x_{۲:n}$ ، $\ldots$ ، $x_{n:n}$ نشان می‌دهیم. بنا بر این، تابع لورنتس با نماد $L(u)$ به‌صورت زیر تعریف می‌شود (کلیبر و کاتز، ۲۰۰۳):

$$L\left(\frac{k}{n}\right)=\frac{\sum_{i=۱}^{k} x_{i:n}}{\sum_{i=۱}^{n} x_{i:n}}, \qquad k=۱,۲...,n \tag{۱}$$

که $L(۰)=۰$ و $x_{i:n}$ درآمد مرتب‌شده‌ی فرد iام است. اگر زوج نقاط $((\frac{k}{n}),L(\frac{k}{n}))$ را برای مقادیر مختلف k در دستگاه محورها رسم کنیم نمودار محدبی به وجود می‌آید که به آن منحنی لورنتس می‌گویند.

یکی از مهم‌ترین شاخص‌هایی که توسط کوارادو جینی در سال ۱۹۱۲ برای اندازه‌گیری میزان نابرابری معرفی شد ضریب جینی است. برای یک جامعه با مقادیر $i=۱,۲,...,n$ عبارت زیر معرف ضریب جینی است:

$$G=\frac{۱}{n}\left\{n+۱-۲\frac{\sum_{i=۱}^{n}(n+۱-i)x_{i:n}}{\sum_{i=۱}^{n} x_{i:n}}\right\}. \tag{۲}$$

مقدار ضریب جینی تجربی با استفاده از نرم‌افزار S-PLUS محاسبه می‌شود و از نرم‌افزارهای دیگری مانند STATA هم می‌توان استفاده کرد.

اگر برای درآمد، یک متغیر تصادفی X با تابع چگالی احتمال $f(x)$ در نظر بگیریم باز هم می‌توانیم یک تابع لورنتس برای آن تعریف کنیم. در این حالت اگر $F(x)=Pr(X\leqslant x)=\int_۰^x f(y)dy$ تابع توزیع تجمعی، $F^{-۱}(t)=\sup\{x:F(x)\leqslant t\}$ تابع چندکی و $E(X)=\int_۰^\infty yf(y)dy$ موجود، متناهی و نامنفی باشد، آنگاه $L(u)$ به‌صورت زیر تعریف می‌شود:

$$L(u)=\frac{۱}{E(X)}\int_۰^u F^{-۱}(t)dt, \qquad u\in[۰,۱]. \tag{۳}$$

تابع لورنتس روی بازه‌ی $[۰,۱]$ پیوسته، صعودی و محدب با $L(۰)=۰$ و $L(۱)=۱$ است. اگر متغیر تصادفی X دارای تابع لورنتس $L(u)$ باشد ضریب جینی آن به‌صورت زیر تعریف می‌شود:

$$G=۲\int_۰^۱[u-L(u)]du=۱-۲\int_۰^۱ L(u)du. \tag{۴}$$

۳ معرفی فرم‌های تابعی لورنتس و برآورد آن‌ها

با توجه به ویژگی‌های منحنی لورنتس، هر تابع که از نقاط $(\circ,\circ)$ و $(۱,۱)$ عبور کرده، به‌طور یکنواخت صعودی و محدب باشد، یک منحنی لورنتس را مشخص می‌کند. فرم‌های تابعی زیادی برای برآورد تابع لورنتس توسط کاکوانی و پودر (۱۹۷۳)، راسچ و دیگران (۱۹۸۰)، گوپتا (۱۹۸۴)، کاکوانی (۱۹۸۰)، ویلاسنور و آرنولد (۱۹۸۹)، ارتگا و دیگران (۱۹۹۱)، چوتیکاپانیچ (۱۹۹۳)، ساربیا و دیگران (۱۹۹۹) و ... معرفی گردید. همچنین برخی از فرم‌های تابعی از توزیع‌های مشهوری چون لگ‌نرمال، سینگ–مادالا و داگم گرفته می‌شوند (اسچادر و اسمید، ۱۹۸۸).

در این‌جا به معرفی مدل‌های پیشنهادی چوتیکاپانیچ (تابع $L_۱$)، ارتگا و دیگران (تابع $L_۲$)، راسچ و دیگران (تابع $L_۳$)، ساربیا و دیگران (تابع $L_۴$) و کاکوانی (تابع $L_۵$) و ضرایب جینی آن‌ها می‌پردازیم (ضرایب جینی بر اساس رابطه‌ی (۴) محاسبه شده است).

$$
\begin{aligned}
L_۱(u;k) &= \frac{e^{ku}-۱}{e^k-۱}, \qquad k \geqslant \circ, \qquad G_۱ = \frac{(k-۲)e^k+(k+۲)}{k(e^k-۱)},\\
L_۲(u;\alpha,k) &= u^\alpha[۱-(۱-u)^k], \quad \alpha \geqslant \circ,\ \circ < k \leqslant ۱,\\
G_۲ &= ۱-۲[B(\alpha+۱,۱)-B(\alpha+۱,k+۱)],\\
L_۳(u;k,\gamma) &= [۱-(۱-u)^k]^\gamma, \quad \circ < k \leqslant ۱,\ \gamma \geqslant ۱,\\
G_۳ &= ۱-\frac{۲}{k}B\left(\frac{۱}{k},\gamma+۱\right)\\
L_۴(u;\alpha,k,\gamma) &= u^\alpha[۱-(۱-u)^k]^\gamma, \quad \alpha \geqslant \circ,\ \gamma \geqslant ۱,\ \circ < k \leqslant ۱,\\
G_۴ &= ۱-۲\sum_{i=\circ}^{\infty}\left[\frac{\Gamma(i-\gamma)}{\Gamma(i+۱)\Gamma(-\gamma)}B(\alpha+۱,ki+۱)\right],\\
L_۵(u;\alpha,\beta,\delta) &= u-\alpha u^\delta(۱-u)^\beta, \quad \alpha > \circ,\ \circ < \beta \leqslant ۱,\ \circ < \delta \leqslant ۱,\\
G_۵ &= ۲\alpha B(\delta+۱,\beta+۱).
\end{aligned}
$$

توجه کنید که هر گاه $\gamma = ۱$ باشد تابع $L_۴$، $L_۲$ را نتیجه می‌دهد و هرگاه $\alpha = \circ$ باشد تابع $L_۴$، $L_۳$ را نتیجه می‌دهد. همچنین $B(.,.)$ تابع بتا است.

۳٫۱ برآورد پارامترهای فرم‌های تابعی لورنتس به روش کم‌ترین توان‌های دوم

اگر برای منحنی لورنتس یک فرم پارامتری در نظر بگیریم، می‌توانیم آن را با استفاده از روش کم‌ترین توان‌های دوم برآورد کنیم، به این مفهوم که مجموعه مقادیر $(u_۱, L(u_۱))$، ...، $(u_n, L(u_n))$ را محاسبه کرده و یک تابع (فرم تابعی لورنتس پارامتری) به داده‌ها برازش می‌دهیم، حال باید مقادیر تابع $\sum_{i=۱}^{n}(L(u_i) - L(u_i;\theta))^۲$ بر حسب پارامترهای θ مینیمم شود. اگر فرم پارامتری، یک تابع خطی باشد، پارامترهای آن به روش‌های کلاسیک و در غیر این صورت به روش‌های عددی با کامپیوتر برآورد می‌گردد. از آن‌جا که اغلب فرم‌های تابعی لورنتس پارامتری توابع غیر خطی از پارامترها است، برآوردهای کم‌ترین توان‌های دوم پارامترها با روش‌های عددی محاسبه می‌گردد. در این‌جا از فرمان nls در نرم افزار R استفاده می‌کنیم.

بعد از برآورد پارامترهای فرم‌های تابعی لورنتس نیاز به معیارهایی است که از روی آن‌ها بتوان بهترین فرم تابعی را تشخیص داد. برای این منظور، از میانگین توان‌های دوم خطا (MSE)، میانگین قدر مطلق خطا (MAE) و ماکسیمم قدر مطلق خطا (MAX) با روابط زیر استفاده می‌کنیم:

$$\mathrm{MSE} = \frac{۱}{n}\sum_{i=۱}^{n}\{L(u_i) - L(u_i;\hat{\theta})\}^۲,$$

$$\mathrm{MAE} = \frac{۱}{n}\sum_{i=۱}^{n}\left|L(u_i) - L(u_i;\hat{\theta})\right|,$$

$$\mathrm{MAX} = \max_{i=۱,۲,\ldots,n}\left|L(u_i) - L(u_i;\hat{\theta})\right|,$$

که $L(u_i)$ منحنی لورنتس تجربی و $L(u_i;\hat{\theta})$ منحنی لورنتس برازش‌شده (فرم تابعی لورنتس) است. هر چه مقادیر MSE، MAE و MAX کم‌تر باشد نشان‌دهنده‌ی بهتر بودن فرم تابعی است.

جداول ۱ و ۲ برآورد پارامترهای فرم‌های تابعی لورنتس را به همراه میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا، ماکسیمم قدر مطلق خطا و ضریب جینی را برای داده‌های هزینه‌ی خانوار شهری و روستایی، به روش کم‌ترین توان‌های دوم نشان می‌دهند.

با توجه به جداول ۱ و ۲، فرم تابعی پیشنهادی راسچ (ردیف سوم جدول، تابع $L_۳$) با داشتن کم‌ترین مقدار میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا و ماکسیمم قدر مطلق خطا بهترین فرم تابعی است.

جدول ۱. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار شهری به روش کم‌ترین توان‌های دوم

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۲٫۷۸۳$	۰٫۰۰۰۸۵۴	۰٫۰۲۲۳	۰٫۰۷۱۱	۰٫۴۱۳۲
۲*	$\alpha = ۰٫۴۶۱,\ k = ۰٫۵۳۷$	۰٫۰۰۰۰۲۲	۰٫۰۰۳۸	۰٫۰۱۶۴	۰٫۴۱۹
۳*	$k = ۰٫۶۱۶,\ \gamma = ۱٫۴۰۸$	۰٫۰۰۰۰۰۸	۰٫۰۰۲۳	۰٫۰۰۹۴	۰٫۴۱۷۸
۴*	$\alpha = ۰,\ k = ۰٫۶۱۶,$ $\gamma = ۱٫۴۰۸$	۰٫۰۰۰۰۰۸	۰٫۰۰۲۳	۰٫۰۰۹۴	۰٫۴۱۷۸
۵	$\alpha = ۰٫۷۸۶,\ \beta = ۰٫۵۰۷,$ $\delta = ۱$	۰٫۰۰۰۰۰۹	۰٫۰۰۲۳	۰٫۰۱۳۵	۰٫۴۱۶۱

جدول ۲. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار روستایی به روش کم‌ترین توان‌های دوم

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۲٫۶۳۲$	۰٫۰۰۰۶۷۸	۰٫۰۱۹۶	۰٫۰۶۴۲	۰٫۳۹۵۱
۲*	$\alpha = ۰٫۴۵۷,\ k = ۰٫۵۶۲$	۰٫۰۰۰۰۱۴	۰٫۰۰۳	۰٫۰۱۳۷	۰٫۴۰۱۶
۳*	$k = ۰٫۶۴,\ \gamma = ۱٫۴۱$	۰٫۰۰۰۰۰۵	۰٫۰۰۱۸	۰٫۰۰۷۵	۰٫۴۰۰۲
۴*	$\alpha = ۰,\ k = ۰٫۶۴,$ $\gamma = ۱٫۴۱$	۰٫۰۰۰۰۰۵	۰٫۰۰۱۸	۰٫۰۰۷۵	۰٫۴۰۰۲
۵	$\alpha = ۰٫۷۶۶,\ \beta = ۰٫۵۲۸,$ $\delta = ۰٫۹۹۱$	۰٫۰۰۰۰۰۶	۰٫۰۰۱۸	۰٫۰۱۱۱	۰٫۳۹۹۱

گاستویریث (۱۹۷۲) کران پایین زیر را برای بررسی درستی برآورد ضریب جینی معرفی کرد:

$$G \geqslant ۱ - \sum_{i=۱}^{n} (u_i - u_{i-۱})[Lu_i + Lu_{i-۱}], \qquad (۵)$$

که $u_۰ = Lu_۰ = ۰$ است.

مقدار این کران برای داده‌های هزینه‌ی خانوار شهری و روستایی به ترتیب برابر ۰٫۴۱۶۲ و ۰٫۳۹۹۲ است، که ضریب جینی برآورد شده‌ی فرم‌های تابعی ارتگا، راسچ و ساربیا در این کران صدق می‌کنند (این فرم‌های تابعی در جداول ۱ و ۲ با علامت $*$ مشخص شده‌اند).

۳٫۲ برآورد پارامترهای فرم‌های تابعی لورنتس به روش ماکسیمم درست‌نمایی

برای برآورد پارامترهای فرم‌های تابعی لورنتس به روش دیگر، ابتدا داده‌ها را به T رده، رده‌بندی می‌کنیم. در نظر می‌گیریم بردار سهم درآمد برابر $Q_i = L(u_i;\theta) - L(u_{i-1};\theta)$ متغیرهای تصادفی با میانگین $E(Q_i) = L(u_i;\theta) - L(u_{i-1};\theta)$ باشند، که $\theta = (\theta_1, ..., \theta_T)'$ است.

بردار $q = (q_1, ..., q_T)'$ از توزیع دیریکله با تابع چگالی احتمال زیر تولید شده است:

$$Q \sim f(q|\alpha) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_T)}{\Gamma(\alpha_1)\ldots\Gamma(\alpha_T)} q_1^{\alpha_1 - 1} \ldots q_T^{\alpha_T - 1}, \tag{۶}$$

که $\alpha = (\alpha_1, \ldots, \alpha_T)'$ پارامترهای تابع چگالی دیریکله و $\Gamma(\cdot)$ تابع گاما است. همچنین در نظر می‌گیریم α تابعی از پارامترهای منحنی با رابطه‌ی زیر باشد:

$$\alpha_i = \lambda[L(u_i;\theta) - L(u_{i-1};\theta)], \tag{۷}$$

که λ پارامتر اضافی مجهول است. با استفاده از روابط (۶) و (۷) تابع چگالی احتمال برای q به‌صورت زیر است:

$$f(q|\phi) = \Gamma(\lambda) \prod_{i=1}^{T} \frac{q_i^{\lambda[L(u_i;\theta) - L(u_{i-1};\theta)] - 1}}{\Gamma[\lambda\{L(u_i;\theta) - L(u_{i-1};\theta)\}]}.$$

برآورد ماکسیمم درست‌نمایی برای $\phi = (\theta', \lambda)'$، بر مبنای ماکسیمم ساختن تابع $\ln$ به‌صورت زیر است (چوتیکاپانیچ و گریفیتث، ۲۰۰۲):

$$\begin{aligned} \ln[f(q|\phi)] = \ln\Gamma(\lambda) &+ \sum_{i=1}^{T} [\lambda\{L(u_i;\theta) - L(u_{i-1};\theta\} - 1] \\ &\times \ln q_i - \sum_{i=1}^{T} \ln\Gamma[\lambda\{L(u_i;\theta) - L(u_{i-1};\theta)\}]. \end{aligned} \tag{۸}$$

با قرار دادن فرم‌های تابعی لورنتس در رابطه‌ی (۸) و استفاده از روش‌های عددی برآورد ماکسیمم درست‌نمایی پارامترهای فرم‌های تابعی به دست می‌آید.

با توجه به این‌که روش‌های برآورد را روی داده‌های هزینه به کار می‌بریم منظور ما در همه جا از بردار سهم درآمد، بردار سهم هزینه است. برای داده‌های شهری، ۱۸ رده و برای داده‌های روستایی، ۲۱ رده در نظر

جدول ۳. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار شهری به روش ماکسیمم درست‌نمایی

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۰/۴۹۱$	۰/۰۳۳۲۴	۰/۱۶۷۴	۰/۲۴۸۳	۰/۰۸۱۵
۲	$\alpha = ۰/۳۸۸,\ k = ۰/۵۵۷$	۰/۰۰۰۲۸	۰/۰۱۳۸	۰/۰۲۶۶	۰/۳۸۹۵
۳	$k = ۰/۶۱۸,\ \gamma = ۱/۳۴۹$	۰/۰۰۰۱۶۷	۰/۰۱۱	۰/۰۲	۰/۳۹۴۸
۴	$\alpha = ۰/۰۰۷,\ k = ۰/۶۲۲,$ $\gamma = ۱/۳۴۹$	۰/۰۰۰۱۹	۰/۰۱۱۶	۰/۰۲۱۶	۰/۳۹۳۳
۵	$\alpha = ۰/۷۶۴,\ \beta = ۰/۵۳۲,$ $\delta = ۰/۹۹۹$	۰/۰۰۰۱۶۸	۰/۰۱۱۱	۰/۰۲۲۳	۰/۳۹۴۲

جدول ۴. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار روستایی به روش ماکسیمم درست‌نمایی

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۰/۰۶۲$	۰/۰۴۴۲۲	۰/۱۹۴۴	۰/۲۸۰۴	۰/۰۱۰۳
۲	$\alpha = ۰/۳۶, k = ۰/۵۷۹$	۰/۰۰۰۳۴	۰/۰۱۵۸	۰/۰۲۷۵	۰/۳۶۸۲
۳	$k = ۰/۶۳۸, \gamma = ۱/۳۲۹$	۰/۰۰۰۲۵	۰/۰۱۳۸	۰/۰۲۲۸	۰/۳۷۱۹
۴	$\alpha = ۰/۰۳۵,\ k = ۰/۶۳۷,$ $\gamma = ۱/۳۰۱$	۰/۰۰۰۲۸	۰/۰۱۴۶	۰/۰۲۴۳	۰/۳۷۰۳
۵	$\alpha = ۰/۷۱۹,\ \beta = ۰/۵۴۶,$ $\delta = ۰/۹۹$	۰/۰۰۰۳	۰/۰۱۵۷	۰/۰۲۵۹	۰/۳۶۷۹

گرفته شده است.

جداول ۳ و ۴، برآورد پارامترهای فرم‌های تابعی لورنتس را به همراه میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا، ماکسیمم قدر مطلق خطا و ضریب جینی را برای داده‌های هزینه‌ی خانوار شهری و روستایی به روش ماکسیمم درست‌نمایی نشان می‌دهند.

با توجه به جداول ۳ و ۴ فرم تابعی پیشنهادی راسچ (ردیف سوم جدول، تابع $L_۳$) با داشتن کم‌ترین مقدار میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا و ماکسیمم قدر مطلق خطا بهترین فرم تابعی است.

ضریب جینی برآورد شده‌ی هیچ‌کدام از فرم‌های تابعی در کران پایین گاستویریث، رابطه‌ی (۵)، صدق نمی‌کنند.

۳٫۳ برآورد پارامترهای فرم‌های تابعی لورنتس به روش صدک اصلی (EPM)

در این بخش یک روش کلی برای برآورد پارامترهای فرم‌های تابعی لورنتس که توسط کاستیلو و هادی (۱۹۹۵) پیشنهاد شده است را معرفی می‌کنیم که قابل استفاده برای همه‌ی خانواده‌ی توابع لورنتس است. همچنین این روش برای هر تابع لورنتس یا تابع توزیع مربوط به تابع لورنتس می‌تواند به کار گرفته شود. مزیت این روش این است که برای تعیین متغیرهایی که تابعی از دیگر فرم‌های تابعی لورنتس هستند پایا است، همچنین این روش برای داده‌های خام و گروه‌بندی شده نیز مورد استفاده قرار می‌گیرد.

روش صدک اصلی شامل دو مرحله است، در مرحله‌ی اول برآوردهای اولیه‌ی پارامترها به دست می‌آید و در مرحله‌ی دوم برآوردهای اولیه به روش‌های مناسب ترکیب می‌شوند و برآوردهای نهایی به دست می‌آید. در این‌جا به برآورد پارامترها با استفاده از تابع لورنتس می‌پردازیم (کاستیلو و دیگران، ۱۹۹۸).

فرض کنید $v = L(u)$ تابع لورنتس داده‌ها و مشاهدات شامل n زوج، $(u_۱, v_۱)$، ...، (u_n, v_n) باشند، که $u_i = \frac{i}{n}$، $v_i = \frac{s_i}{s_n}$ و $i = ۱, ۲, \ldots, n$، $s_i = x_{۱:n} + x_{۲:n} + \cdots + x_{i:n}$ است. مجموعه‌ای rتایی از $I = \{i_۱, \ldots, i_r\}$ انتخاب می‌کنیم که در آن $i_j \in \{۱, ۲, \ldots, n\}$ به معنای انتخاب $j \in \{۱, ۲, \ldots, r\}$ در دفعه‌ی iام و r تعداد پارامترها در θ است. با حل معادله‌ی زیر برآوردهای اولیه به دست می‌آید:

$$v_i = L(u_i; \theta), \qquad i \in I. \tag{۹}$$

حال با قرار دادن فرم‌های تابعی لورنتس معرفی‌شده به جای $L(u; \theta)$ برآوردهای اولیه را به دست می‌آوریم.

فرم پارامتری چوتیکاپانیچ

فرم تابعی پیشنهادی چوتیکاپانیچ (تابع $L_۱$) شامل پارامتر k است. بنا بر این، $r = ۱$ می‌باشد. با قرار دادن این فرم در رابطه‌ی (۹) و حل معادله‌ی زیر برآوردهای اولیه‌ی k به دست می‌آیند.

$$h(k) = \exp(ku_i) - v_i \exp(k) + v_i - ۱ = ۰, \qquad k \geqslant ۰,\ i = ۱, ۲, \ldots, n. \tag{۱۰}$$

برای حل چنین معادلاتی از روش دوبخشی استفاده می‌کنیم. اگر $\beta = \exp(k)$ قرار دهیم، رابطه‌ی (۱۰) به‌صورت زیر تبدیل می‌شود:

$$h(\beta) = \beta^{u_i} - v_i\beta + v_i - ۱ = ۰, \qquad \beta \geqslant ۰,\ i = ۱, ۲, \ldots, n. \tag{۱۱}$$

با انجام روش دوبخشی روی بازه‌ی $\left(\beta_\circ, ۲\beta_\circ - \frac{h(۲\beta_\circ)}{h'(۲\beta_\circ)}\right)$ که $\beta_\circ = (\frac{u_i}{v_i})^{\frac{۱}{۱-u_i}}$ است $\hat{\beta}$ را به دست می‌آوریم. بنا بر این، $\hat{k} = \log \hat{\beta}$ می‌باشد.

فرم پارامتری ارتگا و دیگران

فرم تابعی پیشنهادی ارتگا و دیگران (تابع $L_۲$) شامل پارامترهای α و k است. بنا بر این، $r = ۲$ می‌باشد. با انتخاب نقاط مجزا (u_i, v_i)، (u_j, v_j) و قرار دادن این فرم در رابطه‌ی (۹) و حل معادله‌ی غیر خطی زیر برآوردهای اولیه‌ی k به دست می‌آیند.

$$h(k) = \frac{\log v_i - \log\{۱ - (۱ - u_i)^k\}}{\log v_j - \log\{۱ - (۱ - u_j)^k\}} - \frac{\log u_i}{\log u_j} = \circ, \qquad i, j \in \{۱, ۲, \ldots, n\},\ i \neq j. \tag{۱۲}$$

با به کار بردن روش دوبخشی روی بازه‌های $(\circ, k_\circ)$ و $(k_\circ, ۱)$ که در آن $k_\circ = \frac{\log(۱-v_j)}{\log(۱-u_j)}$، $\hat{k}_{ij}$ به دست می‌آید. بعد از به دست آوردن $\hat{k}_{ij}$ و قرار دادن آن در رابطه‌ی زیر برآوردهای اولیه‌ی α به دست می‌آید (ساربیا و دیگران، ۱۹۹۹):

$$\hat{\alpha}_{ij} = \frac{\log v_i - \log\{۱ - (۱ - u_i)^{\hat{k}_{ij}}\}}{\log u_i}. \tag{۱۳}$$

فرم پارامتری راسچ و دیگران

فرم تابعی پیشنهادی راسچ و دیگران (تابع $L_۳$) شامل پارامترهای k و γ است. بنا بر این، $r = ۲$ می‌باشد. با انتخاب نقاط مجزا (u_i, v_i)، (u_j, v_j) و قرار دادن این فرم در رابطه‌ی (۹) و حل معادله‌ی غیر خطی زیر برآوردهای اولیه‌ی k به دست می‌آیند.

$$h(k) = \frac{\log\{۱ - (۱ - u_i)^k\}}{\log\{۱ - (۱ - u_j)^k\}} - \frac{\log v_i}{\log v_j} = \circ, \qquad i, j \in \{۱, ۲, ..., n\},\ i \neq j. \tag{۱۴}$$

اگر $u_i < u_j$، آنگاه $v_i < v_j$. با قرار دادن این شرط در رابطه‌ی (۱۴) و به کاربردن روش دوبخشی روی بازه‌ی $(\circ, ۱)$، $\hat{k}_{ij}$ به دست می‌آید. بعد از به دست آوردن $\hat{k}_{ij}$ و قرار دادن آن در رابطه‌ی زیر برآوردهای اولیه‌ی γ به دست می‌آید.

$$\hat{\gamma}_{ij} = \frac{\log v_i}{\log\{۱ - (۱ - u_i)^{\hat{k}_{ij}}\}}. \qquad (۱۵)$$

فرم پارامتری ساربیا و دیگران

فرم تابعی پیشنهادی ساربیا و دیگران (تابع $L_۴$) شامل پارامترهای k، α و γ است. بنا بر این، $r = ۳$ می‌باشد. با انتخاب نقاط مجزا (u_i, v_i)، (u_j, v_j)، (u_s, v_s) و قرار دادن این فرم در رابطه‌ی (۹) و حل معادله‌ی غیر خطی زیر برآوردهای اولیه‌ی k به دست می‌آیند.

$$h(k) = \frac{U(i,j,s) + U(j,s,i) + U(s,i,j)}{\log\{۱ - (۱ - u_i)^k\}\log u_j - \log\{۱ - (۱ - u_j)^k\}\log u_i} = ۰,$$
$$i, j, s \in \{۱, ۲, \ldots, n\},\ i \neq j \neq s, \qquad (۱۶)$$

که در آن $U(i,j,s) = \log\{۱ - (۱ - u_s)^k\}(\log u_j \log v_i - \log u_i \log v_j)$ است.

به‌علت این‌که فرم ساربیا ترکیب دو فرم تابعی ارتگا و راسچ است با به کار بردن روش دوبخشی روی بازه‌ی $(۰, ۱)$، $\hat{k}_{ijs}$ به دست می‌آید. بعد از به دست آوردن $\hat{k}_{ijs}$ و قرار دادن آن در روابط زیر برآوردهای اولیه‌ی γ و α به دست می‌آیند.

$$\hat{\gamma}_{ijs} = \frac{\log u_i \log v_j - \log u_j \log v_i}{\log\{۱ - (۱ - u_j)^{\hat{k}_{ijs}}\}\log u_i - \log\{۱ - (۱ - u_i)^{\hat{k}_{ijs}}\}\log u_j}, \qquad (۱۷)$$

$$\hat{\alpha}_{ij} = \frac{\log v_i - \hat{\gamma}_{ijs}\log\{۱ - (۱ - u_i)^{\hat{k}_{ijs}}\}}{\log u_i}. \qquad (۱۸)$$

فرم پارامتری کاکوانی

فرم تابعی پیشنهادی کاکوانی (تابع $L_۵$) شامل پارامترهای α، β و δ است. بنا بر این، $r = ۳$ می‌باشد. با انتخاب نقاط مجزا (u_i, v_i)، (u_j, v_j)، (u_s, v_s) و قرار دادن این فرم در رابطه‌ی (۹) و حل معادله‌ی غیر خطی زیر برآوردهای اولیه‌ی α به دست می‌آیند.

$$\log(\alpha) = \frac{M(i,j,s) + M(j,s,i) + M(s,i,j)}{N(i,j) + N(j,s) + N(s,i)} = \circ, \qquad i,j,s \in \{۱,۲,\ldots,n\},\ i \neq j \neq s, \tag{۱۹}$$

که $M(i,j,s)$ و $N(i,j)$ به‌صورت زیر تعریف می‌شوند:

$$M(i,j,s) = \log(u_s - v_s)\big\{\log(u_i)\log(۱-u_j) - \log(u_j)\log(۱-u_i)\big\},$$
$$N(i,j) = \log(u_i)\log(۱-u_j) - \log(u_j)\log(۱-u_i).$$

با استفاده از روابط زیر برآوردهای اولیه‌ی پارامترهای β و δ به دست می‌آیند.

$$\hat{\beta} = \frac{\log(u_j - v_j)\log(u_i) - \log(u_i - v_i)\log(u_j)}{\log(۱-u_j)\log(u_i) - \log(۱-u_i)\log(u_j)} + \log(\hat{\alpha})\frac{\log(u_j) - \log(u_i)}{\log(۱-u_j)\log(u_i) - \log(۱-u_i)\log(u_j)} \tag{۲۰}$$

$$\hat{\delta} = \frac{\log(u_i - v_i) - \log(\hat{\alpha}) - \hat{\beta}\{\log(۱-u_i)\}}{\log u_i}. \tag{۲۱}$$

برآوردهای به دست آمده از $L(u;\theta)$ به تعداد مشاهدات r بستگی دارد. برای مقادیر بزرگ n و r تعداد برآوردهای اولیه بسیار زیاد است. بنا بر این، در این حالت به جای در نظر گرفتن تمام برآوردهای اولیه، تعداد از پیش تعیین‌شده‌ای از مجموعه‌ی مقادیر اولیه‌ی (N) را انتخاب می‌کنیم که $N < n$ است. مجموعه‌ی مقادیر اولیه‌ی $\hat{\theta}_{j۱}$، ...، $\hat{\theta}_{jN}$ که در آن $j = ۱,۲,\ldots,r$ است را با روابط زیر ترکیب می‌کنیم و برآورد نهایی به دست می‌آید.

$$\hat{\theta}(\text{MED}) = \text{med}(\hat{\theta}_{j۱}, \hat{\theta}_{j۲}, \ldots, \hat{\theta}_{jN}), \tag{۲۲}$$

$$\hat{\theta}(\text{LMS}) = \text{LMS}(\hat{\theta}_{j۱}, \hat{\theta}_{j۲}, \ldots, \hat{\theta}_{jN}), \tag{۲۳}$$

جدول ۵. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار شهری به روش صدک اصلی (رابطه‌ی (۲۰))

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۲/۶۰۸$	۰/۰۰۰۹۸۹	۰/۰۲۰۶	۰/۰۷۸۹	۰/۳۹۲۲
۲*	$\alpha = ۰/۳۶۵,\ k = ۰/۵۰۱$	۰/۰۰۰۰۹۴	۰/۰۰۵۲	۰/۰۳۱۷	۰/۴۲۳۹
۳*	$k = ۰/۵۹۲,\ \gamma = ۱/۳۶$	۰/۰۰۰۰۲۳	۰/۰۰۲۷	۰/۰۱۶۴	۰/۴۱۹۲
۴*	$\alpha = ۰/۷۴۵,\ \beta = ۰/۴۷۵,\ \delta = ۰/۹۶۶$	۰/۰۰۰۰۱۹	۰/۰۰۲۲	۰/۰۲۰۷	۰/۴۱۷۷

که $\text{med}(\alpha_1, \ldots, \alpha_n)$ میانه‌ی بردار $(\alpha_1, \ldots, \alpha_n)$ و $\text{LMS}(\alpha_1, \ldots, \alpha_n)$ برآورد به دست آمده به روش کم‌ترین میانه‌ی توان‌های دوم (LMS) می‌باشد [۲] $\min \text{med}(L(u_i) - L(u_i; \theta))$ که برابر نقطه‌ی میانی کوچک‌ترین بازه شامل نصف تعداد $\alpha_1, \ldots, \alpha_n$ است (روسیو و لروی، ۱۹۸۷).

برای داده‌های هزینه‌ی خانوار شهری و روستایی، به‌علت حجم زیاد داده‌ها یک نمونه‌ی سیستماتیک با حجم ۲۰۰ $(N = ۲۰۰)$ گرفته شده و زیرفاصله‌ها در نمونه‌ی خانوار شهری و روستایی به ترتیب ۶۵ و ۷۰ می‌باشد.

لازم به ذکر است که در فرم تابعی ارتگا جواب‌های به دست آمده در بازه‌ی $(۰, k_۰)$ نتایج غیر قابل قبولی داشت (مقادیر α منفی است). بنا بر این، محاسبات بر اساس بازه‌ی $(k_۰, ۱)$ انجام شده است. همچنین محاسبات مربوط به فرم تابعی ساربیا به‌علت داشتن نتایج غیر قابل قبول (مقادیر منفی α) انجام نشد.

جداول ۵ و ۶ برآورد پارامترهای فرم‌های تابعی لورنتس که با استفاده از رابطه‌ی (۲۰) به دست آمده است را به همراه میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا، ماکسیمم قدر مطلق خطا و ضریب جینی را برای داده‌های هزینه‌ی خانوار شهری و روستایی به روش صدک اصلی نشان می‌دهند.

با توجه به جداول ۵ و ۶، برای داده‌های هزینه‌ی خانوار شهری فرم تابعی پیشنهادی کاکوانی (ردیف چهارم جدول، تابع L_5) و برای داده‌های هزینه‌ی خانوار روستایی فرم تابعی پیشنهادی راسچ (ردیف سوم جدول، تابع L_3) با داشتن کم‌ترین مقدار میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا و ماکسیمم قدر مطلق خطا بهترین فرم تابعی است.

ضریب جینی برآورد شده‌ی همه‌ی فرم‌های تابعی به جز فرم تابعی چوتیکاپانیچ درکران پایین گاستویریث (رابطه‌ی (۵)) صدق می‌کنند (این فرم‌های تابعی در جداول، با علامت $*$ مشخص شده‌اند).

جدول ۶. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار روستایی به روش صدک اصلی (رابطه‌ی (۲۰))

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k = ۲/۵۰۹$	۰/۰۰۰۷۴۹	۰/۰۱۸	۰/۰۶۹۴	۰/۳۸
۲*	$\alpha = ۰/۳۸۵,\ k = ۰/۵۳۶$	۰/۰۰۰۰۴۶	۰/۰۰۴	۰/۰۲۳	۰/۴۰۳۲
۳*	$k = ۰/۶۲۵,\ \gamma = ۱/۳۸$	۰/۰۰۰۰۱	۰/۰۰۲	۰/۰۰۱۱	۰/۴۰۰۸
۴*	$\alpha = ۰/۷۲۸,\ \beta = ۰/۴۹۸,\ \delta = ۰/۹۵۸$	۰/۰۰۰۰۱۲	۰/۰۰۲	۰/۰۱۷	۰/۴۰۰۵

جداول ۷ و ۸، برآورد پارامترهای فرم‌های تابعی لورنتس که با استفاده از رابطه‌ی (۲۱) به دست آمده است را به همراه میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا، ماکسیمم قدر مطلق خطا و ضریب جینی را برای داده‌های هزینه‌ی خانوار شهری و روستایی به روش صدک اصلی نشان می‌دهند.

با توجه به جداول ۷ و ۸ برای داده‌های هزینه‌ی خانوار شهری فرم تابعی پیشنهادی راسچ (ردیف سوم جدول، تابع $L_۳$) و برای داده‌های هزینه‌ی خانوار روستایی فرم تابعی پیشنهادی کاکوانی (ردیف چهارم جدول، تابع $L_۵$) با داشتن کم‌ترین مقدار میانگین توان‌های دوم خطا، میانگین قدر مطلق خطا و ماکسیمم قدر مطلق خطا بهترین فرم تابعی است.

برای داده‌های هزینه‌ی خانوار شهری ضریب جینی برآورد شده‌ی همه‌ی فرم‌های تابعی به جز فرم تابعی چوتیکاپانیچ و برای داده‌های هزینه‌ی خانوار روستایی فقط ضریب جینی برآورد شده‌ی فرم تابعی کاکوانی در کران پایین گاستویریث (رابطه‌ی (۵)) صدق می‌کنند (این فرم‌های تابعی در جداول با علامت $*$ مشخص شده‌اند).

۴ نتیجه‌گیری

برای مقایسه‌ی سه روش برآورد ذکر شده برای فرم‌های تابعی از معیار اندازه‌ی اطلاع عدم صحت که توسط تهیل (۱۹۶۷) معرفی شد استفاده می‌کنیم که به‌صورت زیر تعریف می‌شود:

$$I = \sum_{i=۱}^{T} q_i \ln\left(\frac{q_i}{\hat{q}_i}\right), \qquad (۲۴)$$

جدول ۷. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار شهری به روش صدک اصلی (رابطه‌ی (۲۱))

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k=۲/۵۰۸$	۰/۰۰۱۱۹۳	۰/۰۲۱۳	۰/۰۸۳۶	۰/۳۷۹۹
۲*	$\alpha=۰/۳۵۴,\ k=۰/۵$	۰/۰۰۰۰۹۶	۰/۰۰۵۷	۰/۰۳۲۱	۰/۴۲۲۳
۳*	$k=۰/۵۹۲,\ \gamma=۱/۳۵۵$	۰/۰۰۰۰۲۲	۰/۰۰۳	۰/۰۱۵۹	۰/۴۱۷۳
۴*	$\alpha=۰/۷۲۹,\ \beta=۰/۴۵۴,$ $\delta=۰/۹۵۶$	۰/۰۰۰۰۴۱	۰/۰۰۳	۰/۰۲۷۱	۰/۴۲۰۸

جدول ۸. برآورد پارامترهای فرم‌های تابعی لورنتس برای داده‌های هزینه‌ی خانوار روستایی به روش صدک اصلی (رابطه‌ی (۲۱))

ردیف	برآورد پارامترها	میانگین توان‌های دوم خطا	میانگین قدر مطلق خطا	ماکسیمم قدر مطلق خطا	ضریب جینی
۱	$k=۲/۴۲۲$	۰/۰۰۰۸۸۷	۰/۰۱۸۳	۰/۰۷۳۳	۰/۳۶۹
۲	$\alpha=۰/۳۷۹,\ k=۰/۵۴۳$	۰/۰۰۰۰۴	۰/۰۰۴۸	۰/۰۲	۰/۳۹۷
۳	$k=۰/۶۳۱,\ \gamma=۱/۳۷۳$	۰/۰۰۰۰۲	۰/۰۰۳۸	۰/۰۰۸۶	۰/۳۹۳۷
۴*	$\alpha=۰/۷۲,\ \beta=۰/۴۸۷,$ $\delta=۰/۹۵۳$	۰/۰۰۰۰۱۹	۰/۰۰۱۷	۰/۰۲	۰/۴۰۲۱

که $q_i=L(u_i)-L(u_{i-۱})$ و $\hat{q}_i=L(u_i;\hat{\theta})-L(u_{i-۱};\hat{\theta})$ است. هر چه مقادیر q_i و $\hat{q}_i$ نزدیک به هم باشند مقدار I کاهش می‌یابد و نشان‌دهنده‌ی بهتر بودن روش برآورد است.

برای هر فرم تابعی لورنتس مقدار I با خی‌دو $(\chi^۲)$ با $T-K$ درجه‌ی آزادی قابل مقایسه است که T تعداد رده‌ها و K تعداد پارامترهای مدل است. هرگاه $I>\chi^۲_{T-K}$ باشد نشان‌دهنده‌ی عدم برازش فرم تابعی لورنتس به داده‌ها است.

جداول ۹ و ۱۰ مقدار را برای هر سه روش برآورد برای داده‌های هزینه‌ی خانوار شهری و روستایی نشان می‌دهند.

با توجه به جداول ۹ و ۱۰ روش برآورد کم‌ترین توان‌های دوم با داشتن کم‌ترین مقدار اندازه‌ی اطلاع عدم صحت (I) در همه‌ی فرم‌های تابعی لورنتس بهترین روش برآورد برای داده‌های هزینه‌ی خانوار شهری و روستایی است. همچنین از این معیار برای بهتر بودن فرم تابعی لورنتس می‌توان استفاده کرد که فرم

جدول ۹. مقادیر اندازه‌ی اطلاع عدم صحت برای روش‌های برآورد کم‌ترین توان‌های دوم، ماکسیمم درست‌نمایی و صدک اصلی برای داده‌های هزینه‌ی خانوار شهری

روش‌های برآورد پارامترهای فرم‌های تابعی		تابع L_1	تابع L_2	تابع L_3	تابع L_4	تابع L_5
روش کم‌ترین توان‌های دوم		۰/۰۳۵۳	۰/۰۰۲۲	۰/۰۰۰۸	۰/۰۰۸	۰/۰۰۱۴
روش ماکسیمم درست‌نمایی		۰/۲۱۶۸	۰/۰۰۲۵	۰/۰۰۱۴	۰/۰۰۱۴	۰/۰۰۱۹
روش	برآور پارامترها با استفاده از رابطه‌ی (۲۰)	۰/۰۳۶۴	۰/۰۰۵	۰/۰۰۱۵	–	۰/۰۰۲۲
صدک اصلی	برآورد پارامترها با استفاده از رابطه‌ی (۲۱)	۰/۰۳۷۹	۰/۰۰۵۲	۰/۰۰۱۵	–	۰/۰۰۳۶

جدول ۱۰. مقادیر اندازه‌ی اطلاع عدم صحت برای روش‌های برآورد کم‌ترین توان‌های دوم، ماکسیمم درست‌نمایی و صدک اصلی برای داده‌های هزینه‌ی خانوار روستایی

روش‌های برآورد پارامترهای فرم‌های تابعی		تابع L_1	تابع L_2	تابع L_3	تابع L_4	تابع L_5
روش کم‌ترین توان‌های دوم		۰/۰۲۷۹	۰/۰۰۱۴	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۸
روش ماکسیمم درست‌نمایی		۰/۲۶۵۷	۰/۰۰۲۴	۰/۰۰۱۶	۰/۰۰۱۷	۰/۰۰۲۵
روش	برآورد پارامترها با استفاده از رابطه‌ی (۲۰)	۰/۰۲۸۶	۰/۰۰۲۶	۰/۰۰۰۷	–	۰/۰۰۱۲
صدک اصلی	برآورد پارامترها با استفاده از رابطه‌ی (۲۱)	۰/۰۲۹۶	۰/۰۰۲۲	۰/۰۰۰۶	–	۰/۰۰۱۷

تابعی راسچ (تابع L_3) دارای کم‌ترین مقدار I در هر سه روش برآورد است که تقریباً نتایج همانند نتایجی است که با معیار MSE به دست آمد. با مقایسه‌ی مقادیر I با $\chi^2_{\alpha,T-k}$ که مقدار $\alpha = ۰/۰۵$ و T برای داده‌های هزینه‌ی خانوار شهری و روستایی به ترتیب برابر ۱۸ و ۲۱ است همه‌ی فرم‌های تابعی لورنتس به داده‌ها می‌برازد.

سپاسگزاری

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

مرجع‌ها

Arnold, B.C. (1987). *Majorization and the Lorenz order*. Lecture Notes in Statistics, **43**, Springer, Berlin and New York.

Castillo, E. and Hadi, A.S. (1995). A method for estimating parameters and quantiles of continuous distribution of random variables. *Comput. Statist. Data Anal.* **2**0, 421-439.

Castillo, E., Hadi, A.S. and Sarabia, J.M. (1998). A method for estimating Lorenz curves. *Comm. Statist. Theory Methods* **27**, 2037-2063.

Chotikapanich, D. (1993). A comparsion of alternative functional forms for the Lorenz curve. *Econom. Lett.* **41**, 129-138.

Chotikapanich, D. and Griffiths, W.E. (2002). Estimating Lorenz curves a Dirichlet distribution. *J. Bus. Econom. Statist.* **20**, 290-295.

Gastwirth, J.L. (1972). The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, 306-316.

Gini, C. (1912). *Variailita' e mutabilita, studio Economicogiuridici*. Universita di Cagliari Anno III, Parte 2a, reprinted in C, 211-382.

Gupta, M.R. (1984). Functional form for estimating the Lorenz curve. *Econometrica* **52**, 1313-1314.

Kakwani, N.C. (1980). On a class of poverty measures. *Econometrica* **48**, 437-446.

Kakwani, N.C. and Podder, N. (1973). On the estimation of Lorenz curve from grouped observations. *International Economic Review* **14**, 278-292.

Kleiber, C. (2005). The Lorenz curve in economics and econometrics. Invited Paper, *Gini-Lorenz Centennial Conference*. Siena.

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

Lorenz, M.O. (1905). Method of measuring the concentration of wealth. *J. Amer. Statist. Assoc.* **9**, 209-219.

Moothathu, T.S.K. (1991). On the sufficient condition for two non-intersecting Lorenz curves. *The Indian Journal of Statistics* **53**, 268-274.

Ortega, P., Fernandez, M.A., Lodoux, M. and Garcia, A. (1991). A new functional form for estimating Lorenz Curve. *Review of Income and Wealth* **37**, 447-452.

Rasche, R.H., Gaffeny, J., Koo, A. and Obst, N. (1980). Functional forms for estimating the Lorenz curve. *Econometrica* **48**, 1061-1062.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.

Sarabia, J.M., Castillo, E., Pascual, M. and Sarabia, M. (2005). Mixture Lorenz curves. *Econom. Lett.* **89**, 89-94.

Sarabia, J.M.; Castillo, E. and Slottje, D. (1999). An ordered family of Lorenz curves. *J. Econometrics* **91**, 43-60.

Schader, M. and Schmid, F. (1988). Zur messung der relativen konzentration aus gruppierten daten. *Jahrbucher Fur Nationalokonomie und Statistik* **204**, 437-455.

Theil, H. (1967). *Economics and Information Theory.* Amesterdam, North Holland.

Villasenor, J. and Arnold, B. (1989). Elliptical Lorenz curves. *J. Econometrics* **40**, 327-338.

آزاده مجیری
گروه آمار، دانشکده‌ی علوم،
دانشگاه بیرجند،
بیرجند، ایران.
رایانشانی: *a.mojiri3230@yahoo.com*

غلامرضا محتشمی برزادران
گروه آمار، دانشکده‌ی علوم ریاضی،
عضو قطب علمی داده‌های ترتیبی و فضایی،
دانشگاه فردوسی مشهد،
مشهد، ایران.
رایانشانی: *gmb1334@yahoo.com*
grmohtashami@um.ac.ir

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