سیزدهمین کنفرانس
آمار ایران
۲-۴ شهریور ۱۳۹۵

نگاهی بر شاخص جینی در مسائل بیمه

فرناز درخشش[۱،۲]، غلامرضا محتشمی برزادران[۱]، مهدی جباری نوقابی[۱]

[۱] گروه آمار- دانشکده ریاضی- دانشگاه فردوسی مشهد

[۲] دفتر فن‌آوری اطلاعات و ارتباطات- شرکت توزیع نیروی برق استان خراسان رضوی

چکیده: منحنی لورنتس یک ابزار مهم برای بررسی نابرابری اقتصادی است. شاخص‌های بسیاری براساس این منحنی در جهت اندازه‌گیری میزان نابرابری تعریف شده‌اند، که یکی از مهم‌ترین آن‌ها، شاخص جینی می‌باشد. در این مقاله ابتدا منحنی لورنتس و شاخص جینی را ارائه می‌کنیم، سپس به توسیع منحنی لورنتس در مسائل بیمه‌ای از طریق معرفی متغیر نسبیت می‌پردازیم، که آن را منحنی لورنتس مرتب می‌نامند. به کمک این منحنی شاخص جینی متناظر را معرفی خواهیم نمود. سازگاری و نرمال مجانبی این شاخص را ضمن قضایایی بیان و آن را به عنوان معیاری از سود مطرح می‌کنیم. در نهایت روش‌های امتیازدهی فراوانی-شدت و حق‌بیمه خالص را با استفاده از داده‌های اتومبیل ایالت ماساچوست و با کمک منحنی‌های لورنتس مرتب و شاخص‌های جینی در جهت تعیین و پیش‌بینی سود این شرکت‌ها در حوزه بیمه اتومبیل مورد بررسی قرار می‌دهیم.

واژه‌های کلیدی: حق‌بیمه، خسارت، شاخص جینی، متغیر نسبیت و منحنی لورنتس مرتب.

کد موضوع‌بندی ریاضی (۲۰۱۰): 91B30

۱ مقدمه

شرکت‌های بیمه تمایل به مقایسه توزیع ریسک مالی نسبت به قیمت دارند. از آن به عنوان خسارت بیمه یاد می‌کنند و برای قیمت لفظ حق‌بیمه را به کار می‌برند. فرض می‌کنیم، تحلیل‌گران مجموعه‌ای از خصوصیات ریسک را شناخته‌اند، که براساس آن هر دو توزیع ریسک و قیمت وابسته می‌باشند. هدف، توسعه معیاری است که بتوانیم تا حدی حق‌بیمه‌ای را که برای ارزیابی توزیع خسارت استفاده شده است، اندازه‌گیری نماییم. با چنین معیاری می‌توانیم ساختارهای قیمت‌گذاری جایگزین را مقایسه کنیم. امتیازهای بیمه برای تنظیم حق‌بیمه‌ها، مورد استفاده قرار می‌گیرند، که بر پایه‌ی دانشی از خصوصیات بیمه‌گذاران به نام "امتیازها براساس ریسک" شناخته شده‌اند. اغلب آن‌ها براساس مدل‌های خطی تعمیم‌یافته به‌دست می‌آیند. به طور معمول در بیمه از دو نوع مدل فراوانی-شدت و مدل حق‌بیمه خالص استفاده می‌شود. برای جزئیات بیش‌تر در مورد این مدل‌ها و روش‌های امتیازدهی به فرییز و همکاران (۲۰۱۴) مراجعه نمایید. اگر چه آمار کلاسیک مفید است، ولی بنا به دلایلی نمی‌توان آن‌ها را در زمینه‌های مالی به کار برد. از

[۲] فرناز درخشش: derakhshesh.f87@gmail.com

جمله، به علت بهینگی خواص آن‌ها، براساس توزیع‌های متقارن مانند نرمال و لاپلاس نمی‌توان پیچیدگی‌های توزیع ریسک را شرح داد. به عنوان مثال زمانی که خسارات صفر هستند (مربوط به هیچ ادعا) و وقتی که آن‌ها مثبت هستند، توزیع چوله به راست می‌باشد و دنباله پهن دارد. به همین دلیل فرییز و همکاران (۲۰۱۱) معیارهای منحنی لورنتس و شاخص جینی را در مسائل بیمه به کار بردند، که می‌تواند از عهده انتخاب نامطلوب و اندازه‌گیری سود بالقوه برآید.

۱.۱ منحنی لورنتس و ضریب جینی

برای اولین بار ماکس اوتو لورنتس (۱۹۰۵) در رساله خود، منحنی لورنتس را برای نشان دادن نحوه‌ی توزیع درآمد معرفی نمود. فرض کنید $X_۱, X_۲, ..., X_n$ درآمد n فرد از نمونه باشند، که در دسترس هستند. آن‌ها را از کوچک به بزرگ مرتب می‌کنیم و با نماد $X_{۱:n}, X_{۲:n}, ..., X_{n:n}$ نمایش می‌دهیم. تابع لورنتس $L(.)$ برای $n+۱$ نقطه معادل است با:

$$L(\frac{k}{n}) = \frac{\sum_{i=۱}^{k} X_{i:n}}{\sum_{i=۱}^{n} X_{i:n}}, \qquad k = ۰, ۱, ۲, ..., n$$

که در آن $L(\frac{k}{n})$ درصد سهم درآمد $\frac{k}{n}$ از افراد جامعه است (کلیبر و کاتز (۲۰۰۳)). همچنین $L(۱) = ۱$ و $L(۰) = ۰$ می‌باشد، $X_{i:n}$ درآمد iامین فرد مرتب شده است. از ترسیم مجموعه نقاط $(\frac{k}{n}, L(\frac{k}{n}))$ به ازای مقادیر k نمودار محدبی به نام منحنی لورنتس حاصل می‌شود.

کورادو جینی (۱۹۱۲) به کمک منحنی لورنتس، برای سنجش میزان نابرابری درآمدی در مسائل اقتصادی شاخص جینی را معرفی نمود. این شاخص از لحاظ هندسی دو برابر مساحت محصور بین منحنی لورنتس و خط برابری است، که از رابطه:

$$\widehat{G} = \frac{۱}{n}(n + ۱ - ۲\frac{\sum_{i=۱}^{n}(n+۱-i)X_{i:n}}{\sum_{i=۱}^{n} X_{i:n}}),$$

می‌توان محاسبه نمود.

تابع لورنتس را می‌توانیم برای یک متغیر تصادفی با تابع چگالی احتمال متناظر در جامعه به‌دست آوریم. فرض کنید X متغیر تصادفی نامنفی از درآمد با تابع چگالی احتمال $f(x)$، تابع توزیع تجمعی $F(x)$ و امید ریاضی موجود و متناهی $E(X)$ باشد. با در نظر گرفتن تغییر متغیر $t = F(x)$ در نتیجه $\mu = E(X) = \int_۰^۱ F^{-۱}(t)dt$ که در آن $F^{-۱}(t) = \sup\{x|F(x) \leq t\} \forall t \in [۰, ۱]$ تابع چندکی است. لذا تابع لورنتس $L(u)$ به صورت:

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

به‌دست می‌آید. این تابع در بازه $[۰, ۱]$ پیوسته، محدب و صعودی است. همچنین برای متغیر تصادفی X با تابع لورنتس پیوسته $L(u)$ شاخص جینی از رابطه:

$$\begin{aligned} G &= ۲\int_۰^۱ [u - L(u)]du \\ &= ۱ - ۲\int_۰^۱ L(u)du. \end{aligned}$$

محاسبه می‌شود (کلیبر و کاتز (۲۰۰۳)).

۲.۱ ارتباط توزیع‌های حق‌بیمه با خسارت

منحنی لورنتس ابزار گرافیکی مناسبی برای نمایش توزیع‌ها در اقتصاد رفاه است. به ویژه برای تفسیر توزیع‌های چوله مفید می‌باشد. شکلی که تحلیل‌گران بیمه به خوبی با آن آشنا هستند. آیا می‌توان از منحنی‌های لورنتس کلاسیک برای مقایسه حق‌بیمه با توزیع خسارت استفاده نمود؟ خیر، برای مثال محاسبه منحنی لورنتس کلاسیک برای حق‌بیمه‌ها و همچنین خسارات و سپس قرار گرفتن دو منحنی روی شکل مشابه، ساده خواهد بود. اما توزیع جمعیت برای هر منحنی به گونه‌ای متفاوت مرتب می‌شوند (به ترتیب به‌وسیله حق‌بیمه‌ها و خسارات)، طوری که مقایسه حق‌بیمه‌ها و خسارات برای هر گروه بیمه‌گذار معنی‌دار نخواهد بود.

به توسعه منحنی لورنتس از طریق معرفی متغیر سومی به نام نسبیت (نسبت امتیازها براساس ریسک بر قیمت)، می‌پردازیم. این متغیر خسارات را به حق‌بیمه‌ها پیوند می‌زند. همچنین طوری که سازگاری بین گروه‌های بیمه‌گذار حفظ شود، مرتب‌سازی می‌نماید. در این راه، می‌توانیم تفاوت‌های بین خسارات و حق‌بیمه‌ها را پی‌گیری کنیم. و از طریق ترتیب‌هایی به نوعی متفاوت، بر اختلافات بین این دو توزیع تأکید نماییم.

نسبیت‌ها همچنین کمک به ارزیابی آسیب‌پذیری ساختار قیمت‌گذاری می‌کنند. در این کار نسبیت‌ها به‌وسیله مقایسه حق‌بیمه‌های فعلی با یک امتیاز بیمه جایگزین پیشنهاد شده، ایجاد می‌شوند. در این راه، می‌توانیم به نارسایی‌های ساختار حق‌بیمه فعلی نسبت به جایگزین پیشنهاد شده، تمرکز کنیم. این، یادآور روش تصمیم‌گیری آزمون فرضیه است. حق‌بیمه فعلی متناظر با فرضیه صفر و رقابت در امتیاز بیمه مطابق با فرضیه جایگزین می‌باشد.

۲ منحنی لورنتس مرتب و ضریب جینی متناظر

منحنی لورنتس مرتب توسط فرییز و همکاران (۲۰۱۱) بیان شده، که نموداری از توزیع خسارت در مقابل حق‌بیمه‌ها است. ابتدا بایستی مقادیر خسارات و حق‌بیمه‌ها به‌وسیله متغیر تصادفی "نسبیت" مرتب شوند. این منحنی در بیمه برای نشان دادن وضعیت سودآوری شرکت‌های بیمه، همچنین به منظور پیش‌گیری از انتخاب نامطلوب پرتفوی[1] (سبد سهام)، مورد استفاده قرار می‌گیرد. محور عمودی در این منحنی نشان‌گر نحوه‌ی توزیع خسارت و محور افقی نشان‌گر نحوه‌ی توزیع حق‌بیمه‌ها است. وقتی توزیع خسارت به طور کامل با توزیع حق بیمه برابر باشد، منحنی روی نیمساز ربع اول قرار می‌گیرد، که معرف "خط برابری" است. به عبارتی در هر نقطه این خط، درصد خسارت با درصد حق‌بیمه برابر می‌باشد. معمولاً منحنی لورنتس مرتب در قسمت پایین خط برابری قرار می‌گیرد. همچنین اگر منحنی از خط برابری بیش‌تر فاصله بگیرد، نشان‌دهنده‌ی سودآوری بیش‌تر بیمه‌گر است.

برای نمادگذاری تعداد n بیمه‌نامه را در نظر می‌گیریم. $i = ۱, ..., n$، Y_i، خسارت بیمه، X_i مجموعه‌ای از خصوصیاتی که در دسترس بیمه‌گر، $P_i =$ $P(X_i)$ حق‌بیمه مربوطه که تابعی از X_i، $S_i = S(X_i)$ امتیاز بیمه با در نظر گرفتن تغییرات نرخ و $R_i = R(X_i) = S(X_i)/P(X_i)$ متغیر تصادفی نسبیت یا نسبت حق‌بیمه است. با داشتن مقادیر (Y_i, P_i, S_i, R_i) می‌توانیم منحنی لورنتس مرتب را برای i بیمه‌نامه به‌دست آوریم.

۱.۲ منحنی لورنتس مرتب

مجموعه‌ی بیمه‌نامه‌ها را به‌وسیله متغیر نسبیت به صورت غیرنزولی مرتب می‌کنیم. تابع توزیع حق‌بیمه با نماد $F_P(s)$، به صورت:

$$F_P(s) = \frac{E[P(X)I(R \leq s)]}{EP(X)}, \qquad \hat{F}_P(s) = \frac{\sum_{i=1}^{n} P(X_i)I(R_i \leq s)}{\sum_{i=1}^{n} P(X_i)}, \tag{۱.۲}$$

[1] Portfolio

بەدست می‌آید. که مقادیرش در فاصله [۰، ۱] قرار می‌گیرد. که در آن $F_P(\cdot) = \cdot$ حد پایین و $F_P(\infty) = ۱$ حد بالا است. به طور مشابه تابع توزیع خسارت معادل است با:

$$F_L(s) = \frac{E[YI(R \leq s)]}{EY}, \qquad \hat{F}_L(s) = \frac{\sum_{i=۱}^{n} Y_i I(R_i \leq s)}{\sum_{i=۱}^{n} Y_i}. \tag{۲.۲}$$

همچنین $I(.)$ تابع نشان‌گر است، که اگر رویداد درست باشد، مقدار یک و در غیر این صورت صفر، خواهد بود. بنابراین، زوج نقاط $(F_P(s), F_L(s))$ منحنی لورنتس مرتب را تشکیل می‌دهند.

نقش انتخاب نامطلوب به منظور این که بیمه‌گر عملکرد خوبی داشته باشد، باید ریسک کند و یک طرح نرخ‌گذاری دقیق، برنامه‌ریزی نماید. در این حالت به یک رقیب برای بیمه‌گرهای دیگر به حساب می‌آید. لذا بیمه‌گرهای دیگر باید زیان بیش‌تری را متحمل شوند. اگر بیمه‌گر طرح مناسبی برای نرخ‌گذاری تعیین نکند، قربانی انتخاب نامطلوب‌اش خواهد شد. منحنی لورنتس مرتب، به صورت شهودی اندازه پتانسیل این انتخاب نامطلوب را مشخص می‌نماید. به‌وسیله مرتب‌سازی نسبیت، اوراق بهاداری که سودآور است، مشخص می‌شود.

۲.۲ شاخص جینی مرتبط با منحنی لورنتس مرتب

منحنی لورنتس مرتب فقط نشان‌گر وضعیت سودآوری است. به همین سبب فرییز و همکاران (۲۰۱۱) شاخص جینی مرتبط با این منحنی را بیان نمودند، تا مشخص شود بیمه‌گر چه اندازه سود کسب می‌کند. این شاخص به صورت:

$$\begin{aligned} Gini(F_P, F_L) &= ۲\int_{\cdot}^{\infty} [F_P(s) - F_L(s)] dF_P(s) \\ &= ۱ - ۲\int_{\cdot}^{\infty} F_L(s) dF_P(s). \end{aligned}$$

می‌باشد. مقادیر این شاخص در بازه $[-۱, ۱]$ قرار می‌گیرد. همچنین نقش معکوس حق‌بیمه‌ها و خسارات در شاخص جینی معادل، $Gini(F_P, F_L) = -Gini(F_L, F_P)$ است.

۱.۲.۲ برآورد شاخص جینی

فرض کنید $\{(a_{\cdot} = \cdot, b_{\cdot} = \cdot), (a_۱, b_۱), \dots, (a_n = ۱, b_n = ۱)\}$ مجموعه مقادیر تجربی منحنی لورنتس مرتب برای نمونه‌ای به حجم n باشد. با استفاده از $a_j = \hat{F}_P(R_j)$ و $b_j = \hat{F}_L(R_j)$، فرییز و همکاران (۲۰۱۱) شاخص جینی تجربی به صورت:

$$\begin{aligned} \widehat{Gini} &= ۲\sum_{j=\cdot}^{n-۱} (a_{j+۱} - a_j)\{\frac{a_{j+۱} + a_j}{۲} - \frac{b_{j+۱} + b_j}{۲}\} \\ &= ۱ - \sum_{j=\cdot}^{n-۱} (a_{j+۱} - a_j)(b_{j+۱} + b_j), \end{aligned} \tag{۳.۲}$$

تعریف نمودند. شاخص جینی بزرگ‌تر بیان‌گر سودآوری بیش‌تر بیمه‌گر است. فرییز و همکاران (۲۰۱۱) برای این شاخص دو خاصیت سازگاری و نرمال مجانبی که باید از لحاظ معنی‌داری آماری برقرار باشد، را ضمن قضایای زیر بیان نمودند.

قضیه ۱.۲. *تحت شرایط نظم ضعیف، در صورتی که $n \to \infty$، با احتمال یک داریم:* $\widehat{Gini} \to Gini$

برای بیان خاصیت نرمال مجانبی، $h_{۱}(X,Y)$ را در نظر می‌گیریم.

$$h_{۱}(X,Y) = \frac{۱}{۲}(\mu_Y P(X) F_L(R) + Y\mu_P[۱ - F_P(R)]).$$

قضیه ۲.۲. *با برقراری شرایط نظم ضعیف، شاخص جینی به صورت مجانبی دارای توزیع نرمال به این صورت است:*

$$\sqrt{n}(\widehat{Gini} - Gini) \longrightarrow_D N(\cdot, \Sigma_{Gini}),$$

که درآن Σ_{Gini} برابر خواهد بود، با:

$$\Sigma_{Gini} = \frac{۴}{\mu_Y^{۲}\mu_P^{۲}}(۴\Sigma_h + \frac{\mu_h^{۲}}{\mu_Y^{۲}}\Sigma_Y + \frac{\mu_h^{۲}}{\mu_P^{۲}}\Sigma_P - \frac{۴\mu_h}{\mu_Y}\Sigma_{hY} - \frac{۴\mu_h}{\mu_P}\Sigma_{hP} + \frac{۲\mu_h^{۲}}{\mu_Y\mu_P}\Sigma_{YP}),$$

که $\mu_h = \frac{۱}{۲}\mu_Y\mu_P(۱ - Gini)$ می‌باشد.

جدول ۱: برآوردگرهای گشتاوری در جهت برآورد Σ_{Gini}

$\hat{h}_1(X,Y) = \frac{1}{2}(\bar{Y}P(X)\hat{F}_L(R) + Y\bar{P}[1 - \hat{F}_P(R)]),$	$\bar{h}_1 = n^{-1}\sum_{i=1}^{n}\hat{h}_1(X_i,Y_i),$
$\hat{\Sigma}_h = n^{-1}\sum_{i=1}^{n}\hat{h}_1(X_i,Y_i)^2 - \bar{h}_1^2,$	$\hat{\Sigma}_{hY} = n^{-1}\sum_{i=1}^{n}\hat{h}_1(X_i,Y_i)Y_i - \bar{h}_1\bar{Y},$
$\hat{\Sigma}_{hP} = n^{-1}\sum_{i=1}^{n}\hat{h}_1(X_i,Y_i)P(X_i) - \bar{h}_1\bar{P},$	$\hat{\Sigma}_Y = n^{-1}\sum_{i=1}^{n}Y_i^2 - \bar{Y}^2,$
$\hat{\Sigma}_P = n^{-1}\sum_{i=1}^{n}P(X_i)^2 - \bar{P}^2,$	$\hat{\Sigma}_{YP} = n^{-1}\sum_{i=1}^{n}Y_iP(X_i) - \bar{Y}\bar{P}.$

قضیه ۳.۲. *تحت شرایط نظم ضعیف، فریز و همکاران (۲۰۱۱) نشان دادند که یک برآورد سازگار برای Σ_{Gini} برابر است با:*

$$\hat{\Sigma}_{Gini} = \frac{۴}{\bar{Y}^{۲}\bar{P}^{۲}}(۴\hat{\Sigma}_h + \frac{\bar{h}_{۱}^{۲}}{\bar{Y}^{۲}}\hat{\Sigma}_Y + \frac{\bar{h}_{۱}^{۲}}{\bar{P}^{۲}}\hat{\Sigma}_P - \frac{۴\bar{h}_{۱}}{\bar{Y}}\hat{\Sigma}_{hY} - \frac{۴\bar{h}_{۱}}{\bar{P}}\hat{\Sigma}_{hP} + \frac{۲\bar{h}_{۱}^{۲}}{\bar{Y}\bar{P}}\hat{\Sigma}_{YP}).$$

۳.۲ شاخص جینی به عنوان معیاری از سود

شاخص جینی به طور مستقیم تفسیر اقتصادی دارد. این به یکی از دلایلی که شاخص جینی با اهمیت است، اشاره دارد. به طور مثال فریز و همکاران (۲۰۱۳) نشان دادند که میانگین سود معادل است با:

$$\frac{۱}{n}\sum_{i=۱}^{n}(\hat{F}_P(R_i) - \hat{F}_L(R_i)) \approx \frac{\widehat{Gini}}{۲}. \qquad (۴.۲)$$

تحلیل‌گران از میانگین گرفتن در تصمیم‌گیری‌ها بیش‌تر از هر استراتژی دیگری، استفاده می‌کنند. هر استراتژی بیمه‌نامه‌هایی با نسبیت‌های کم‌تر یا مساوی R_i را حفظ می‌کند. به این معنا که بیمه‌گرها ساختار نرخ‌گذاری با یک شاخص جینی بزرگ اتخاذ می‌کنند، تا از سود بیش‌تری لذت ببرند. بنابراین شاخص جینی را به عنوان میانگین سود مورد انتظار در نظر می‌گیریم، که به‌وسیله استفاده از نسبیت‌ها (به شکل پرتفوی) به‌دست خواهد آمد. شاخص جینی عدم تطابق بالقوه بین خسارات و حق‌بیمه‌ها را می‌سنجد، که می‌توانیم به‌وسیله نسبیت‌های مختلف آن را تغییر دهیم. این متغیر جایی که عدم تطابق بالقوه رخ دهد، ایده می‌دهد. به ویژه یک نسبیت پایین به این معنی است، که بیمه‌نامه‌ای بسیار سود آور و کاندیدای خوبی برای حفظ است.

۴.۲ مقایسه شاخص جینی به‌واسطه دو امتیاز

استفاده از شاخص جینی کلاسیک در جهت درک نابرابری‌های اقتصادی میان زیرگروه‌های جامعه امری عادی محسوب می‌شود. در زمینه بیمه، تمایل داریم، به مقایسه تأثیر دو امتیاز S_A و S_B مربوط به حق‌بیمه پایه $P(x)$ بر روی شاخص جینی بپردازیم. با عنایت بر برآورد تجربی از اختلاف شاخص‌های جینی، نسبیت‌های متناظر R_A و R_B به کار رفته است.

قضیه ۴.۲. *تحت شرایط نظم ضعیف، اختلاف شاخص‌های جینی دارای توزیع نرمال مجانبی به فرم،*

$$\sqrt{n}\left[\begin{pmatrix}\widehat{Gini}_A\\ \widehat{Gini}_B\end{pmatrix}-\begin{pmatrix}Gini_A\\ Gini_B\end{pmatrix}\right]\longrightarrow_D N\left[\cdot,\begin{pmatrix}\Sigma Gini_A & \Sigma Gini_{AB}\\ \Sigma Gini_{AB} & \Sigma Gini_B\end{pmatrix}\right]$$

می‌باشد، که درآن $\Sigma Gini_A$ و $\Sigma Gini_B$ به ترتیب برای امتیازات S_A و S_B از (۲.۲) و همچنین،

$$\Sigma_{Gini_{AB}}=\frac{۴}{\mu_Y^۲\mu_P^۲}\Big(۴\Sigma_{hA,hB}-\frac{۲\mu_{hB}}{\mu_Y}\Sigma_{hA,Y}-\frac{۲\mu_{hB}}{\mu_P}\Sigma_{hA,P}-\frac{۲\mu_{hA}}{\mu_Y}\Sigma_{hB,Y}-\frac{۲\mu_{hA}}{\mu_P}\Sigma_{hB,P}$$
$$+\frac{\mu_{hA}\mu_{hB}}{\mu_Y^۲}\Sigma_Y+\frac{۲\mu_{hA}\mu_{hB}}{\mu_Y\mu_P}\Sigma_{YP}+\frac{\mu_{hA}\mu_{hB}}{\mu_P^۲}\Sigma_P\Big)$$

با $\mu_{hA}=\frac{۱}{۲}\mu_Y\mu_P(۱-Gini_A)$ و $\mu_{hB}=\frac{۱}{۲}\mu_Y\mu_P(۱-Gini_B)$ به‌دست می‌آید. برای اثبات به فرییز و همکاران (۲۰۱۱) مراجعه نمایید.

نتیجه ۵.۲. *تحت شرایط قضیه ۴.۲ اختلاف شاخص‌های جینی از توزیع نرمال مجانبی پیروی می‌کند:*

$$\sqrt{n}\big[(\widehat{Gini_A}-\widehat{Gini_B})-(Gini_A-Gini_B)\big]\longrightarrow_D N(\cdot,\Sigma Gini_A+\Sigma Gini_B-۲\times\Sigma Gini_{AB}).$$

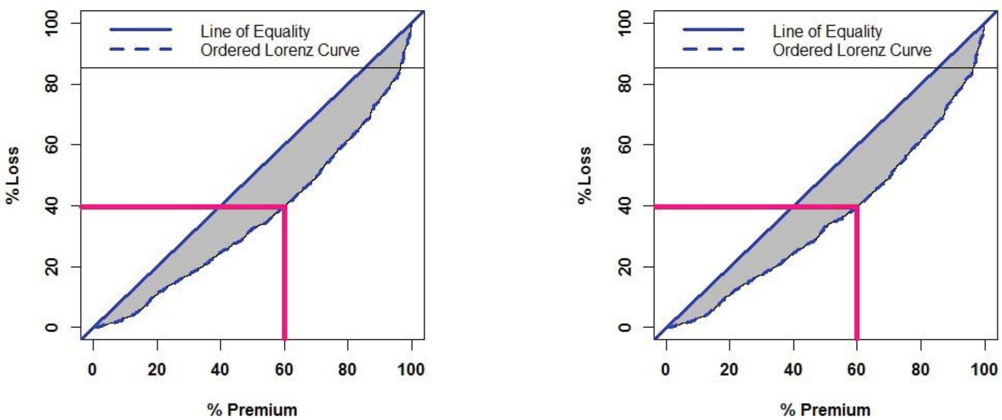
۳ مثال کاربردی: خسارات اتومبیل ماساچوست

مجموعه داده‌های این موضوع توسط فریرا و مینکل (۲۰۱۰) مورد مطالعه قرار گرفته است. این داده‌های عمومی از طریق اداره اجرایی امور انرژی و محیط زیست[۲] $(EOEEA)$ ماساچوست جمع‌آوری شده و متشکل از حدود ۳ میلیون بیمه‌نامه است، که بیان‌گر تجارب چند شرکت بیمه می‌باشد. با این حال، اطلاعات پایه نرخ‌گذاری موجود است، که معمولاً برای تمام شرکت‌ها شامل خصوصیات اصلی راننده و گروه‌بندی‌های منطقه می‌باشند. از مجموعه داده‌های فریرا و مینکل (۲۰۱۰)، یک نمونه تصادفی از ۱۰۰۰۰۰ بیمه‌گذار برای تحلیل انتخاب شده است. دو عامل منطقه و گروه نرخ‌گذاری را به عنوان متغیرهای توضیحی در مدل‌های فراوانی-شدت و حق‌بیمه خالص مورد بررسی قرار می‌دهیم، که از مقادیر برازش شده این مدل‌ها در محاسبه امتیازات بیمه استفاده می‌شود. سپس منحنی لورنتس مرتب را از روابط (۱.۲) و (۲.۲) و برآورد شاخص جینی را از رابطه (۳.۲) به‌دست آورده و تحلیل می‌نماییم.

۱.۳ نتایج

برای بررسی میزان سود شرکت‌های بیمه و مقایسه بین روش‌های امتیازدهی، منحنی لورنتس مرتب و شاخص جینی متناظر با آن را به کار می‌بریم. شکل ۱ منحنی لورنتس مرتب برای هر دو روش امتیازدهی فراوانی-شدت و حق‌بیمه خالص را به کمک نرم‌افزار آماری R نمایش می‌دهد. نمودار

[۲] Executive Office of Energy and Environmental Affairs



شکل ۱: منحنی لورنتس مرتب برای امتیازات فراوانی-شدت و حق‌بیمه خالص

سمت راست، اختلاف بین خسارات و حق‌بیمه‌ها را براساس امتیاز فراوانی-شدت، از لحاظ گرافیکی مشخص نموده است. همچنین، براساس این منحنی به ازای ۶۰ درصد حق‌بیمه پرداخت شده، ۴۰ درصد خسارت وارد شده، که به صورت کلی بیان‌گر سود خالص شرکت‌های بیمه برای داده‌های بیمه اتومبیل ماساچوست است. نمودار سمت چپ، منحنی لورنتس مرتب برای امتیاز حق‌بیمه خالص است، که نتایج مشابه با منحنی سمت راست است. به‌علاوه، این منحنی‌ها به معنی‌دار نبودن تفاوت بین دو روش امتیازدهی، برای داده‌های بیمه اتومبیل ماساچوست اشاره دارند.

در جدول ۲ شاخص‌های جینی، خطاهای استاندارد و همچنین اختلاف آن‌ها آورده شده است. بر طبق این نتایج، درصد مقادیر شاخص‌های

جدول ۲: خلاصه نتایج به‌دست آمده توسط شاخص جینی برای داده‌های بیمه اتومبیل

$t-Diff$	$StdErrDiff$	$StdErrG_{Pure-P}$	G_{Pure-P}	$StdErrG_{F-S}$	G_{F-S}
۰٫۱۸۰۳	۰٫۱۰۵۴	۱٫۶۴۴۹	۲۸٫۲۷۰۷	۱٫۶۴۹۰	۲۸٫۲۵۱۶

جینی و خطاهای استاندارد این دو امتیاز بسیار نزدیک به هم هستند. آماره t برای اختلاف بین درصد شاخص‌ها از دو روش، کوچک می‌باشد. لذا دو روش امتیازدهی فراوانی-شدت و حق‌بیمه خالص از لحاظ آماری تفاوت معنی‌داری ندارند. در نتیجه، امتیاز حق‌بیمه خالص در سطح معنای ۵ درصد عملکرد خوبی داشته است. گرچه استفاده از مدل فراوانی-شدت جزئیات بیش‌تری را در اختیار تحلیل‌گر قرار می‌دهد.

بحث و نتیجه‌گیری

به منظور این که بیمه‌گر عملکرد خوبی داشته باشد، باید ریسک کند و یک طرح نرخ‌گذاری دقیق، برنامه‌ریزی نماید. همان‌طور که بیان شد، منحنی لورنتس مرتب به صورت شهودی سودآوری شرکت بیمه را مشخص می‌کند. به‌وسیله مرتب‌سازی متغیر نسبیت، اوراق بهاداری که سودآور است، تعیین می‌گردد. در این جهت از شاخص جینی به عنوان میانگین سود مورد انتظار استفاده می‌شود. همچنین می‌توان این شاخص را در انتخاب مدل‌های بیمه‌ای به کار برد.

مراجع

Ferreira, J. and Minikel, E. (2010). Pay-as-you-drive auto insurance in Massachusetts: A risk assessment and report on consumer, industry and environmental benefits. *In Conservation Law Foundation, Boston*, http://www.clf.org/wp-content/uploads/2010/12/CLF-PAYD-Study November-2010.pdf.

Frees, E. W., Meyers, G. and Cummings, D. A. (2011). Summarizing insurance scores using a Gini index. *Journal of the American Statistical Association*, 106:495, 1085-1098.

Frees, E. W., Meyers, G. and Cummings, D. A. (2013). Insurance ratemaking and a Gini index. *Journal of Risk and Insurance*, 81, 2, 335:366.

Frees, E. W., Derrig, R. A. Meyers, G. (2014). *Predictive Modeling Applications in Actuarial* . Cambridge University Press, New York.

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

Kleiber, C. and Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley Series in Probability and Statistics, John Wiley.

Lorenz, Max O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*. 9, 70, 209-219.