

مدل سازی تابع زیان بیمه‌ای با استفاده از ترکیب توزیع تی استودنت چوله‌های پیربولیک تعمیم یافته و نظریه مقادیر فرین

سعید باجلان^۱، رضا راعی^۲، شاپور محمدی^۳

چکیده: تحقیق حاضر به بررسی این موضوع می‌پردازد که آیا می‌توان با ترکیب توزیع تی استودنت چوله‌های پیربولیک تعمیم یافته که اخیراً در حوزه مالی و بیمه معرفی شده است و نظریه مقادیر فرین، تابع زیان را به گونه‌ای مدل سازی کرد که هم مقادیر مرکزی را به خوبی تخمین بزند و هم بتواند مقادیر حدی را نیز به شکل مطلوبی مدل سازی کند. داده‌های استفاده شده در این تحقیق، خسارت‌های جانی و مالی بیمه‌نامه‌های شخص ثالث وسایل نقلیه موتوری است. برای کالیبراسیون توزیع تی استودنت چوله‌های پیربولیک تعمیم یافته در این تحقیق از الگوریتم حداکثر سازی انتظارات (EM) و برای مدل سازی اکستریم‌ها بر اساس رویکرد اوج فراتر از آستانه (POT) از روش حداکثر درست‌نمایی (MLE) استفاده شده است. نتایج تحقیق نشان می‌دهد توزیع ترکیبی پیشنهادی، به خوبی می‌تواند زیان‌های ناشی از بیمه شخص ثالث را مدل سازی کند.

واژه‌های کلیدی: الگوریتم حداکثر سازی انتظارات، تابع میانگین مازاد، توزیع تی استودنت چوله‌ها پیربولیک تعمیم یافته، نظریه مقادیر فرین.

۱. دانشجوی دکتری مدیریت مالی، دانشگاه تهران، تهران، ایران

۲. استاد مدیریت مالی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران

۳. دانشیار مدیریت مالی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۲/۰۹/۱۰

تاریخ پذیرش نهایی مقاله: ۱۳۹۳/۰۵/۰۵

نویسنده مسئول مقاله: سعید باجلان

E-mail: saeedbajalan@ut.ac.ir

مقدمه

تابع نرمال متداول‌ترین تابعی است که برای مدل‌سازی در حوزه اقتصاد و مالی استفاده می‌شود، اما به‌طور کلی توزیع زیان‌های بیمه‌ای، چوله^۱ است، پدیده‌ای که سبب می‌شود توزیع نرمال، توزیع مناسبی برای مدل‌سازی ریسک‌های بیمه‌ای نباشد (لین، ۲۰۰۰؛ ورنیک، ۲۰۰۶). اگرچه در خصوص استفاده از سایر توزیع‌ها برای مدل‌سازی ریسک‌های بیمه‌ای کارهایی صورت گرفته، عمدتاً در این کارها تلاش شده از توزیع‌های بیضوی^۲ استفاده شود که هم دارای برخی خواص توزیع نرمال هستند و هم انعطاف بیشتری در خصوص چولگی و کشیدگی دارند (الینگ، ۲۰۱۲). از سویی ایراد اساسی توزیع‌های بیضوی متقارن بودن آن‌هاست، حال آنکه تحقیقات متعدد نشان داده است توزیع زیان‌های بیمه‌ای شکل متقارنی ندارد.

در خصوص استفاده از توزیع‌های نامتقارن برای تخمین تابع زیان، کارهایی صورت گرفته است که از جامع‌ترین آن‌ها می‌توان به کار الینگ اشاره کرد. در تحقیقی که الینگ روی داده‌های مختص به شدت خسارت انجام داد، ۱۸ توزیع متداول در اکچوئری از جمله توزیع‌های نمایی، هندسی، کای دو، کوچی، گاما، گوس، لوجستیک، لوگ نرمال، تی‌استودنت، ویبول، هایپربولیک و... با دو توزیع نرمال چوله و تی‌استودنت چوله از لحاظ برازندگی^۳ مقایسه شدند. نتایج تحقیق نشان داد توزیع تی‌استودنت چوله برازندگی خوبی در مقایسه با رقبا دارد (الینگ، ۲۰۱۲).

با این حال، با توجه به اهمیت وقایع فرین^۴ در امر بیمه‌گری، لازم است که دنباله‌های توزیع زیان با دقت بسیار بیشتری تخمین زده شوند؛ زیرا که اتفاقاتی با احتمال بسیار کم اما شدت زیاد، می‌تواند بقای یک شرکت بیمه را با چالش جدی مواجه کند. بحث بر سر این موضوع که آیا تئوری مقادیر فرین^۵ می‌تواند تخمین‌های بهتری از این‌گونه خسارت‌ها ارائه دهد، موضوع تحقیقات مختلفی بوده است که از آن جمله می‌توان به (مکنیل و سالادین، ۱۹۹۷؛ امبرچتز، رزنیک و سامرودنیتسکی، ۱۹۹۹؛ باسی، امبرچتز و کافتزاک، ۱۹۹۸؛ چاوا، استیفانوسک و ترن بال، ۲۰۰۸؛ لی و فنگ، ۲۰۱۰؛ بیرلانت، جوسنس و سیگرز، ۲۰۰۴) اشاره کرد. بیشتر این تحقیقات نشان داده‌اند این تئوری می‌تواند در تخمین دنباله‌های توزیع زیان بهتر از رقبا عمل کند. با وجود این، در خصوص تخمین تابع زیان با ترکیب‌کردن این دو دسته توزیع برای رسیدن به تخمین‌های صحیح‌تر از شکل تابع زیان تا کنون کار خاصی صورت نگرفته است. بر همین

-
1. Skewed
 2. Elliptical Distribution
 3. Goodness of Fit
 4. Extreme Values
 5. Extreme Value Theory (EVT)

اساس در تحقیق حاضر تلاش شده است با استفاده از ترکیب توزیع تی‌استودنت چوله و تئوری مقادیر فرین، مدل‌سازی بهتری از توزیع زیان انجام شود.

پیشینه پژوهش

ریسک‌های بیمه‌ای معمولاً توزیع‌های چوله دارند که همین مسئله سبب می‌شود توزیع نرمال برای مدل‌سازی ریسک‌های بیمه‌ای و خسارت‌های مرتبط با آن‌ها چندان مناسب نباشد لین، ۲۰۰۰؛ ورنیک، ۲۰۰۶). به جز چولگی، برخی ریسک‌های بیمه‌ای (به‌ویژه آن‌هایی که با حوادث فاجعه‌آمیز مرتبط‌اند) توزیع‌های پهن‌دنباله دارند (امبرچتز، مک‌نیل و استرامان، ۲۰۰۲). توزیع‌های نرمال چوله^۱ و سایر توزیع‌ها از خانواده چوله بیضوی^۲ جایگزین‌های مناسبی برای توزیع نرمال به‌نظر می‌رسند؛ زیرا هم از مزیت‌های توزیع نرمال برخوردارند و هم انعطاف بیشتری در رابطه با چولگی و کشیدگی دارند.

بولانس، گوپلن، پلیکان و ورنیک (۲۰۰۸) توزیع نرمال چوله و لوگ نرمال چوله‌ای^۳ را روی مجموعه‌ای از ادعاهای دومتغیره^۴ که مختص به صنعت بیمه خودرو اسپانیا بود، برازش کردند. الینگ (۲۰۱۲) کار بولانس را با استفاده از توزیع تی. چوله^۵ و دیتابیس‌های مختلف توسعه داد. ورنیک (۲۰۰۶) در مقاله خود به بحث درباره استفاده از توزیع نرمال چوله چندمتغیره^۶، به‌عنوان جایگزین توزیع نرمال به‌منظور اندازه‌گیری ریسک و تخصیص سرمایه پرداخت. هدف اصلی مقاله او یافتن فرمولی بر پایه معیار TCE^۷ برای تخصیص بهینه سرمایه بود.

ادبیات موضوع در خصوص به‌کارگیری تئوری مقادیر فرین برای مدل‌سازی رویدادهای کرانه^۸ و کاربرد آن در حل مسائل مورد علاقه اکچوئریست‌ها در حال گسترش است. بسیاری از تحقیقات صورت‌گرفته در این حوزه نشان داده‌اند که توزیع پارتو تعمیم‌یافته^۹ برای تخمین دنباله‌های تابع زیان نسبت به سایر توزیع‌ها برتر است (مک‌نیل، ۱۹۹۷؛ مک‌نیل و سالادین، ۱۹۹۷؛ امبرچتز، رزنیگ و سامرودنیتسکی، ۱۹۹۹؛ باسی، امبرچتز و کافتراکی، ۱۹۹۸؛ امبرچتز، کلاپلبرگ و میکوش، ۱۹۹۷؛ رزنیگ، ۱۹۹۷)

1. Skew Normal
2. Skew Elliptical
3. Log Skew Normal
4. Bivariate Claims
5. Skew t
6. Multivariate Skew Normal
7. Tail Conditional Expectation (TCE)
8. External Event
9. Generalized Pareto Distribution (GPD)

بریلانت و تئوگلز (۱۹۹۲) با استفاده از تئوری مقادیر فرین به مدل‌سازی ادعاهای بزرگ در بیمه‌های غیرزندگی پرداختند. روتزن و تجویدی (۲۰۰۰) از تئوری مقادیر فرین برای مدل‌سازی خسارت‌های نشئت‌گرفته از طوفان استفاده کردند. آن‌ها روی مدل‌سازی زیان‌های حدی، مقدار کلی خسارت و شناسایی روند چنین خسارت‌هایی تمرکز کردند. مک‌نیل و سالادین (۲۰۰۰) سناریوهایی را برای پیش‌بینی خسارت‌های فرین آتی با استفاده روش اوج بیش از آستانه (POT) در پرتفویی از خسارت‌های ناشی از طوفان توسعه دادند. بریلانت و جوسنس و سیگرس (۲۰۰۴) در تحقیقی نتیجه گرفتند داده‌های زیان بیمه‌ای، همواره پدیده پهن‌دنباله‌بودن را از خود نشان می‌دهند. پاتریک، جردن و رزنبرگ (۲۰۰۴) قاعده‌های تجربی^۱ در داده‌های زیان عملیاتی را بررسی کردند و دریافتند که داده‌های زیان، برحسب نوع رخداد در میان مؤسسات مختلف مشابه‌اند. چاوا، استیفانسک و ترن بال (۲۰۰۸) بر مدل‌سازی و پیش‌بینی توزیع زیان برای دارایی‌های ریسکی اعتباری، نظیر اوراق قرضه یا وام‌ها تمرکز کردند. داهن و همکارانش (۲۰۱۰) داده‌های مختص به بانک‌های آمریکا را بررسی کردند و نشان دادند بانک‌ها می‌توانند طی یک سال به‌طور متوسط چهار خسارت عمده را تجربه کنند. آن‌ها همچنین از توزیع‌های فرین برای برازش خسارت‌های عملیاتی و تخمین حق بیمه سالانه استفاده کردند. چیانگ لی (۲۰۱۲) در مقاله خود روی مدل‌سازی و تخمین پارامترهای دنباله توزیع زیان برای خسارت‌های به‌وجودآمده از آتش‌سوزی در تایوان تمرکز کرد. وی در مقاله‌اش توزیع پارتو تعمیم‌یافته را با توزیع‌های لوگ‌نرمال، نمایی، گاما و ویبول مقایسه کرد و میزان آستانه را با استفاده از روش‌های نمودارهای هیل^۲، کولموگروف - اسمیرنوف و آزمون نکویی برازش نسبت درست‌نمایی^۳ به‌دست آورد.

در تحقیق حاضر تلاش می‌شود با ترکیب این دو نظریه، مدل‌سازی کامل‌تری از توزیع زیان صورت بگیرد. در ادامه توزیع تی‌استودنت چوله و چگونگی کالیبراسیون آن و همچنین رویکرد اوج بیش از آستانه به‌عنوان دومین قضیه اساسی نظریه مقادیر فرین، تشریح می‌شوند.

توزیع تی‌استودنت چوله^۴

در این تحقیق، توزیع تی‌استودنت چوله به‌صورت حالت خاصی از توزیع هایپربولیک تعمیم‌یافته تعریف می‌شود. گفتنی است توزیع تی‌استودنت چوله به‌صورت‌های مختلفی تعریف شده است که برای مطالعه آنها می‌توان به مقاله آس و هاف مراجعه کرد (۲۰۰۶). به دو دلیل از این تعریف

-
1. Empirical Regularities
 2. Hill Plots
 3. Likelihood ratio goodness-of-fit test
 4. Skewed t distribution

خاص در تحقیق حاضر استفاده شده است؛ نخست اینکه در سایر تعاریف موجود، تمام دنباله‌ها چندجمله‌ای^۱ هستند و به‌خوبی روی توزیع‌های پهن دنباله برازش می‌شوند، اما قادر به مدلسازی چولگی صحیح و شایان توجهی نیستند؛ برخلاف آنچه گفته شد، در این تعریف یکی از دنباله‌ها به‌صورت چندجمله‌ای و دیگری به‌صورت نمایی است و می‌توان به مدلسازی چولگی صحیح و شایان توجهی پرداخت. دوم؛ تخمین حداکثر درست‌نمایی پارامترهای آن به‌کمک الگوریتم حداکثرسازی انتظارات^۲ نیز نسبتاً آسان است.

بیان توزیع تی‌استودنت چوله به‌صورت حالت خاص هایپربولیک تعمیم‌یافته^۳

توزیع تی‌استودنت چوله^۴ هایپربولیک تعمیم‌یافته، حالت خاصی از توزیع هایپربولیک تعمیم‌یافته است. چگالی توزیع هایپربولیک تعمیم‌یافته تک‌متغیره را می‌توان به‌صورت زیر تعریف کرد (پراس، ۱۹۹۹):

$$f_x(x) = \frac{(\alpha^2 - \beta^2)^{\lambda/2} K_{\lambda-1/2}(\alpha\sqrt{\delta^2 + (x - \mu)^2}) \exp(\beta(x - \mu))}{\sqrt{2\pi} \alpha^{\lambda-1/2} \delta^\lambda K_\lambda(\delta\sqrt{\alpha^2 - \beta^2}) (\sqrt{\delta^2 + (x - \mu)^2})^{1/2-\lambda}} \quad \text{رابطه ۱}$$

که در آن پارامترهای μ و δ به‌ترتیب تعیین‌کننده موقعیت^۴ و مقیاس^۵ هستند و α و β شکل^۶ توزیع را مشخص می‌کنند. $K_\lambda(x)$ که $x > 0$ معرف تابع بسل تعدیل‌یافته^۷ نوع دوم با رتبه λ است^۸. بیان انتگرالی از این تابع به‌صورت زیر است (بارندورف - نیلسن و بلسید، ۱۹۸۱):

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} e^{-\frac{x}{2}(y+y^{-1})} dy, x > 0 \quad \text{رابطه ۲}$$

و پارامترها باید شرایط زیر را به‌دست آورند:

$$\begin{cases} \delta \geq 0, |\beta| < \alpha & \text{if } \lambda > 0 \\ \delta > 0, |\beta| < \alpha & \text{if } \lambda = 0 \\ \delta > 0, |\beta| \leq \alpha & \text{if } \lambda < 0 \end{cases} \quad \text{رابطه ۳}$$

1. Polynomial
2. Expectation Maximization (EM)
3. Generalized Hyperbolic Skew Student's t-distribution
4. Location
5. Scale
6. Shape
7. Modified Bessel Function

۸. یادآوری می‌شود که در متون قدیمی‌تر به آن نوع ۳ گفته می‌شود.

توزیع تی استودنت چوله هایپربولیک تعمیم یافته را می توان به صورت آمیخته نرمال میانگین- واریانس^۱ با توزیع گوسی معکوس تعمیم یافته (GIG)^۲ نیز تعریف کرد که در آن توزیع GIG دارای تابع چگالی زیر است (بارندورف-نیلسن و بلسید، ۱۹۸۱):

$$f(z; \lambda, \delta, \gamma) = \left(\frac{\gamma}{\delta}\right)^\lambda \frac{z^{\lambda-1}}{2K_\lambda(\gamma\delta)} \exp\left\{-\frac{1}{2}(\delta^2 z^{-1} + \gamma^2 z)\right\} \quad \text{رابطه ۴}$$

بنابراین، می توان متغیر X که توزیع هایپربولیک تعمیم یافته دارد را به صورت زیر تعریف کرد:

$$X = \mu + \beta Z + \sqrt{Z}Y \quad \text{رابطه ۵}$$

که در آن؛ $Y \sim N(0, 1)$ ، $Z \sim GIG(\lambda, \delta, \gamma)$ و Y مستقل از Z هستند و همچنین $\gamma = \sqrt{\alpha^2 - \beta^2}$

اگر در رابطه ۱ مقدار $\lambda = -\frac{v}{2}$ و $|\beta| \rightarrow \alpha$ قرار داده شود، توزیع تی استودنت چوله هایپربولیک تعمیم یافته به دست می آید. تابع چگالی این توزیع به صورت رابطه ۶ است:

$$f_x(x) = \begin{cases} \frac{2^{\frac{1-v}{2}} \delta^v |\beta|^{\frac{v+1}{2}} K_{\frac{v+1}{2}}(\sqrt{\beta^2(\delta^2 + (x-\mu)^2)}) \exp(\beta(x-\mu))}{\Gamma\left(\frac{v}{2}\right) \sqrt{\pi} (\sqrt{\delta^2 + (x-\mu)^2})^{\frac{v+1}{2}}}, & \beta \neq 0 \\ \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi} \delta \Gamma\left(\frac{v}{2}\right)} \left[1 + \frac{(x-\mu)^2}{\delta^2}\right]^{-(v+1)/2}, & \beta = 0 \end{cases}$$

که در آن Γ معرف تابع گاما است.^۳

کالیبراسیون توزیع تی. چوله با استفاده از الگوریتم حداکثرسازی انتظارات^۴

بیان توزیع تی. چوله به صورت آمیخته نرمال میانگین - واریانس، این مزیت مهم را دارد که اجازه می دهد الگوریتم حداکثرسازی انتظارات (EM) برای کالیبراسیون آن استفاده شود. این الگوریتم را که دمپستر، لیرد و روبین در سال ۱۹۷۷ ارائه کردند، ابزار قدرتمندی برای تخمین حداکثر درست نمایی داده هایی است که برخی مقادیر آن مفقود شده است. این الگوریتم به طور خاص برای تخمین توزیع های آمیخته بسیار مناسب است؛ زیرا عملیات آمیخته کردن به طور ذاتی به

1. Normal Mean-Variance
 2. Generalized Inverse Gaussian
 3. $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}. dt$
 4. Expectation Maximization Algorithm (EM)

مفقودشدن برخی داده‌ها منجر می‌شود. در ادامه چگونگی به‌کارگیری این الگوریتم برای تخمین پارامترهای توزیع تی‌استودنت چوله‌هایپربولیک تعمیم‌یافته براساس مقاله آس و هاف (۲۰۰۶) تشریح می‌شود:

فرض می‌کنیم داده‌های صحیح شامل یک بخش مشاهده پذیر X و یک بخش مشاهده ناپذیر مشاهده Z هستند. الگوریتم حداکثرسازی انتظارات، تکرار متوالی دو مرحله است: مرحله انتظارات (E-Step) و مرحله حداکثرسازی انتظارات (M-Step). در مرحله E، فرد انتظارات بخش مشاهده ناپذیر را با در نظر گرفتن مقادیر فعلی پارامترها محاسبه می‌کند و در مرحله M درست‌نمایی عبارت $f_x(x, z) = f_{x|z}(x|z)f_z(z)$ با استفاده از انتظارات محاسبه‌شده در مرحله E به‌دست می‌آید.

مرحله E: مرحله E شامل محاسبه انتظارات شرطی آماره‌های Z ، Z^{-1} و $\log Z$ است. برای توزیع هایپربولیک تعمیم‌یافته داریم: $Z|X \sim GIG\left(\lambda - \frac{1}{\gamma}, \sqrt{\delta^2 + (x - \mu)^2}, \alpha\right)$. بنابراین برای توزیع تی‌استودنت چوله‌هایپربولیک تعمیم‌یافته نیز خواهیم داشت:

$$Z|X \sim GIG\left(-\frac{v+1}{2}, \sqrt{\delta^2 + (x - \mu)^2}, |\beta|\right) \quad \text{رابطه ۷}$$

گشتاورهای توزیع $GIG(\lambda, \delta, \gamma)$ به‌صورت رابطه ۸ هستند (کارلیس، ۲۰۰۲)

$$E(Z^r) = \left(\frac{\delta}{\gamma}\right)^r \frac{K_{\lambda+r}(\delta\gamma)}{K_{\lambda}(\delta\gamma)} \quad \text{رابطه ۸}$$

اگر $q(x_i) = \sqrt{\delta^2 + (x - \mu)^2}$ باشد، برای توزیع تی‌استودنت چوله‌هایپربولیک تعمیم‌یافته رابطه ۹ و ۱۰ به‌دست می‌آید:

$$\xi_i = E(Z_i | X_i = x_i) = \frac{q(x_i) K_{\frac{1-v}{2}}(|\beta|q(x_i))}{|\beta| K_{\frac{v+1}{2}}(|\beta|q(x_i))} \quad \text{رابطه ۹}$$

$$\rho_i = E(Z_i^{-1} | X_i = x_i) = \frac{|\beta| K_{\frac{v+3}{2}}(|\beta|q(x_i))}{q(x_i) K_{\frac{v+1}{2}}(|\beta|q(x_i))} \quad \text{رابطه ۱۰}$$

به‌علاوه، خواهیم داشت:

$$\begin{aligned} \chi_i &= E(\log Z_i^{-1} | X_i = x_i) \\ &= \log\left(\frac{q(x_i)}{|\beta|}\right) + \frac{1}{K_{\frac{v+1}{2}}(|\beta|q(x_i))} \frac{\partial K_{\frac{v+1}{2}}(|\beta|q(x_i))}{\partial \left(\frac{v+1}{2}\right)} \end{aligned} \quad \text{رابطه (۱۱)}$$

که از رابطه زیر استنتاج می‌شود (مینکا و سنتانا، ۲۰۰۴):

$$E(\log Z) = \left. \frac{\partial E(Z^r)}{\partial r} \right|_{r=0} = \log\left(\frac{\delta}{\gamma}\right) + \frac{1}{K_\lambda(\delta\gamma)} \frac{\partial}{\partial \lambda} K_\lambda(\delta\gamma) \quad \text{رابطه (۱۲)}$$

مشتق‌های تابع بسل تعمیم‌یافته نوع سوم نسبت به مرتبه λ را می‌توان با استفاده از فرمول‌های موجود محاسبه کرد. با این حال، آن قدر این فرمول‌ها پیچیده‌اند که تخمین‌های عددی ترجیح داده می‌شوند.

مرحله M: در مرحله M فرد تخمین‌های پارامترهایی را محاسبه می‌کند که از حداکثرسازی درست‌نمایی $f_x(x, z) = f_{x|z}(x|z)f_z(z)$ با استفاده از مقادیر ساختگی ξ_i ، ρ_i و χ_i نتیجه می‌شود. اگر $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ و $\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$ باشد، در تکرار k ام الگوریتم، تخمین‌های β و μ براساس رابطه‌های زیر به‌روز می‌شوند:

$$\beta^{(k+1)} = \frac{\sum_{i=1}^n x_i \rho_i - \bar{x} \sum_{i=1}^n \rho_i}{\bar{n} - \bar{\xi} \sum_{i=1}^n \rho_i} \quad \text{رابطه (۱۳)}$$

$$\mu^{(k+1)} = \bar{x} - \beta^{(k+1)} \bar{\xi} \quad \text{رابطه (۱۴)}$$

پارامتر v به‌عنوان جواب معادله زیر محاسبه می‌شود:

$$\log \frac{n}{2} - \log\left(\sum_{i=1}^n \rho_i\right) - \frac{1}{n} \sum_{i=1}^n \chi_i = \psi\left(\frac{v^{k+1}}{2}\right) - \log v^{(k+1)} \quad \text{رابطه (۱۵)}$$

که در آن $\psi(\cdot)$ تابع دی گاما است که به‌صورت زیر تعریف می‌شود:

$$\psi(x) = \frac{d(\log(\Gamma(x)))}{dx} = \frac{d(\Gamma(x))/dx}{\Gamma(x)} \quad \text{رابطه (۱۶)}$$

سرانجام:

$$\delta^{(k+1)} = \frac{n v^{(k+1)}}{\sum_{i=1}^n \rho_i} \quad \text{رابطه ۱۷}$$

همگرایی الگوریتم EM به سمت تخمین‌زنده‌های حداکثر درست‌نمایی تضمین شده است؛ زیرا این الگوریتم استاندارد است.

تئوری مقادیر فرین^۱

وقایعی که طی سه دهه اخیر رخ داده است، نشان می‌دهد نباید زیاد به داده‌های تاریخی برای پیش‌بینی احتمال و شدت خسارت‌ها در آینده اتکا کرد. طوفان اندرو^۲، زلزله نورث‌ریج^۳، زلزله کوب^۴ و حمله‌های تروریستی ۱۱ سپتامبر، همگی نمونه‌هایی از خسارت‌های فاجعه‌آمیز طبیعی یا ساخته دست بشر هستند که شاید برمبنای شواهد تاریخی نتوان آنها را مدل‌سازی کرد. براساس قضیه بالکما - دی‌هان و پیکاندز (۱۹۷۴) به‌عنوان دومین قضیه اصلی تئوری مقادیر فرین، تحت شرط $F \in MDA(H)$ ، توزیع پارتو تعمیم‌یافته توزیع حدی برای توزیع مقادیر بیش از یک آستانه به‌اندازه کافی بزرگ u است؛ به این معنا که ما می‌توانیم تابع مثبت $\sigma(u)$ را به‌گونه‌ای بیابیم که رابطه زیر برقرار شود:

$$\lim_{u \rightarrow X_F} \sup_{0 \leq y \leq X_F - u} |F_u(y) - G_{\xi, \sigma(u)}(y)| = 0 \quad \text{رابطه ۱۸}$$

$$u \rightarrow X_F \quad 0 \leq y \leq X_F - u$$

$$F \in MDA(H_\xi) \quad \text{اگر و تنها اگر}$$

به بیان دیگر:

$$F_{u(y) \approx G_{\xi, \sigma(n)}(y)} \quad \text{رابطه ۱۹}$$

که در آن:

1. Extreme value Theory
2. Hurricane Andrew
3. Northridge Earth quake
4. Kobe Earthquake

$$G_{\xi, \sigma}(y) = \begin{cases} 1 - (1 + \frac{\xi}{\sigma} y)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - e^{-y/\sigma} & \text{if } \xi = 0 \end{cases} \quad \text{رابطه ۲۰}$$

برای $y \in [0, x_F - u]$ اگر $\xi = 0$ و اگر $y \in [0, \frac{\sigma}{\xi}]$ اگر $\xi < 0$. می‌توان اثبات کرد که تابع میانگین مازاد خسارت^۱ برای توزیع GPD با پارامتر $\xi < 1$ نوعی تابع خطی به صورت زیر است (امبرجتز، کلاپلبرگ و میکوش، ۱۹۹۷):

$$e(u) = E(X - u | X > u) = \frac{\sigma + \xi u}{1 + \xi} \quad \text{رابطه ۲۱}$$

که در آن X دارای توزیع پارتو تعمیم‌یافته با پارامترهای σ ، ξ است. این نتیجه در تعیین آستانه u بسیار کاربرد دارد که در قسمت بعد به آن پرداخته خواهد شد.

چگونگی انتخاب حد آستانه

تابع میانگین مازاد تجربی برای نمونه مفروض X_1, X_2, \dots, X_n به صورت زیر تعریف می‌شود:

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u)^+}{\sum_{i=1}^n 1_{\{X_i > u\}}} \quad \text{رابطه ۲۲}$$

با توجه به این تعریف و این نکته که تابع میانگین مازاد برای توزیع پارتو تعمیم‌یافته یک تابع خطی با شیب مثبت است، می‌توان رویکرد نموداری زیر را برای تعیین آستانه u به دست آورد: «آستانه $u > 0$ را به گونه‌ای انتخاب کنید که نمودار $e_n(x)$ تقریباً برای $x \geq u$ خطی باشد.»

بر اساس این رویکرد، برای تعیین آستانه u مناسب، باید مقدار $e_n(u)$ را به ازای مقادیر مختلف u محاسبه کرد و پس از آن به ترسیم نمودار میانگین مازاد پرداخت. در واقع، مجموع نقاط $(u, e_n(u)), x_{n:n} < u < x_{1:n}$ را باید روی محورهای مختصات ترسیم کرد که در آن $x_{n:n}$ و $x_{1:n}$ به ترتیب بزرگ‌ترین و کوچک‌ترین داده‌های موجود در نمونه هستند. از آنجا که تابع میانگین مازاد توزیع پارتو تعمیم‌یافته، نوعی تابع خطی با شیب مثبت است، از یک آستانه

مشخص به بعد شکل منحنی تقریباً به صورت یک خط راست با شیب مثبت در خواهد آمد. آن آستانه مشخص مقدار مناسب u خواهد بود (امبرجتز، کلاپلبرگ و میکوش، ۱۹۹۷).

تخمین پارامترهای توزیع‌های مقادیر فرین

روش‌های گوناگونی برای برازش توزیع‌های مقادیر فرین تعمیم‌یافته (GEV) و پارتو تعمیم‌یافته (GPD) روی داده‌های واقعی ارائه شده است که از جمله آن‌ها می‌توان به روش حداکثر درست‌نمایی (از این پس ML) و گشتاورهای موزون شده برحسب احتمالات (از این پس PWM) اشاره کرد.

در خصوص متدلوژی‌های تخمین پارامترهای توزیع GPD باید گفت که هر دو روش ML و PWM در مطالعات شبیه‌سازی که توسط هاسکینگ، والیس و وود (۱۹۸۷) و همچنین روتزن و تجویدی (۱۹۹۶) انجام شده است، مقایسه شده‌اند. هاسکینگ و والیس دریافتند که برای داده‌هایی با توزیع پارتو تعمیم‌یافته با پارامتر شکل در دامنه $0 \leq \xi \leq 0.4$ ، به‌ویژه برای نمونه‌هایی با اندازه کوچک، روش گشتاورهای موزون شده برحسب احتمالات نسبت به روش حداکثر درست‌نمایی مزیت بیشتری دارد؛ زیرا تخمین‌های این روش انحراف کمتری از مقدار واقعی دارند؛ اما به موازات افزایش اندازه نمونه این مزیت ناپدید می‌شود. از سوی دیگر روتزن و تجویدی نشان دادند که برای داده‌هایی با دنباله پهن با $\xi \geq 0.5$ روش گشتاورهای موزون شده برحسب احتمالات در تخمین پارامترها تورش‌های جدی دارد؛ در حالیکه تخمین‌های روش حداکثر درست‌نمایی سازگارند. از آنجا که در این تحقیق با داده‌های خسارت بیمه‌ای که توزیع پهن دنباله دارند سروکار داریم، روش حداکثر درست‌نمایی را انتخاب می‌کنیم.

روش حداکثر درست‌نمایی

فرض کنید توزیع H_θ تابع چگالی تابع h_θ دارد. تابع درست‌نمایی براساس داده‌های $y = \{y_1, y_2, \dots, y_n\}$ که در آنها y ها i.i.d هستند، توسط رابطه ۲۳ بیان می‌شود (گیلی و کلزی، ۲۰۰۶):

$$l(\theta; y) = \prod_{i=1}^n h_\theta(y_i) \quad \text{رابطه ۲۳}$$

اگر $L(\theta; y) = \ln(l(\theta; y))$ باشد، تخمین زننده حداکثر درست‌نمایی (MLE) برای θ برابر است با:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta; y) \quad \text{رابطه ۲۴}$$

که در آن Θ مجموعه مقادیری است که θ می‌تواند اتخاذ کند. درباره نمونه i.i.d از متغیر $y = \{y_1, y_2, \dots, y_n\}$ با تابع توزیع GPD، تابع لگاریتم درست‌نمایی به صورت زیر است:

$$L(\xi, \sigma | y) = \begin{cases} -n \log - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\sigma} y_i\right) & \text{if } \xi \neq 0 \\ -n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^n y_i & \text{if } \xi = 0 \end{cases} \quad \text{رابطه ۲۵}$$

جامعه آماری و نمونه تحقیق

جامعه آماری تحقیق کلیه داده‌های خسارتی مختص به بیمه‌نامه شخص ثالث صادرشده توسط شرکت‌های بیمه ایرانی است. به منظور گردآوری و پایش داده‌ها، ابتدا داده‌های خسارتی بیمه‌نامه‌های شخص ثالث صادرشده از کلیه شرکت‌های بیمه که اطلاعات آنها در سامانه سنهاب موجود است، گردآوری شد. این اطلاعات ۱۴۹۳۰۵ مشاهده را دربرمی‌گیرد که هر یک شامل شماره بیمه‌نامه، تاریخ صدور بیمه‌نامه، مبلغ خسارت و نوع خسارت (جانی، مالی) در بازه زمانی فروردین ۱۳۸۴ تا مرداد ۱۳۹۲ می‌شود.

در گام بعد، از آنجا که در برخی موارد به‌ازای یک شماره بیمه‌نامه مشخص بیش از یک خسارت رخ داده بود، کلیه خسارت‌ها جمع شدند تا مشخص شود به‌ازای هر بیمه‌نامه صادرشده، خسارت جانی پرداختی و خسارت مالی پرداختی چه میزان بوده است. بعد از اعمال این اصلاحات، در مجموع ۳۷۸۰۴ زوج مشاهده از خسارت‌های جانی و مالی به‌ازای هر بیمه‌نامه به‌دست آمد. با توجه به تغییرات نرخ دیه و همچنین تغییرات اجرت تعمیر خودرو و قیمت خودرو، لازم بود روی میزان خسارت جانی و مالی سال‌های قبل از ۱۳۹۲ تعدیل صورت گیرد. به‌منظور اعمال این تعدیلات، ابتدا مبلغ دیه و شاخص هزینه حمل و نقل استخراج شدند، سپس نرخ رشد آنها

مشخص شد. از نرخ رشد محاسبه‌شده به‌عنوان عامل تعدیل خسارت جانی و مالی در سال‌های قبل از ۱۳۹۲ استفاده شد.

تجزیه و تحلیل داده‌های تحقیق

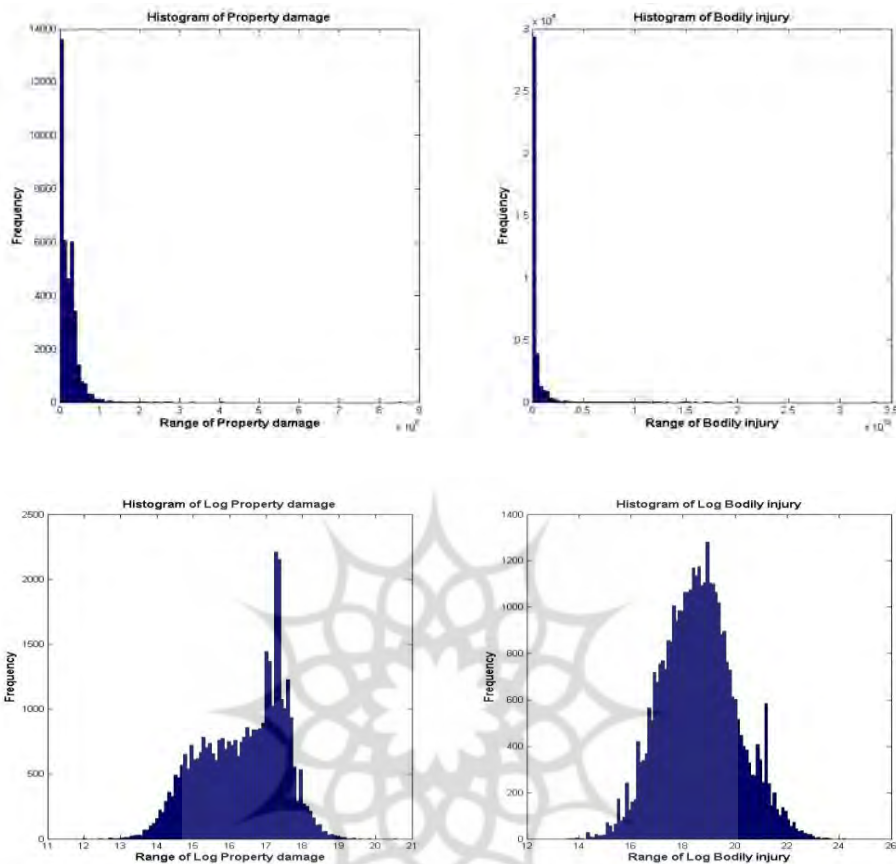
توصیف آماری نمونه تحقیق

جدول ۱ آماره‌های توصیفی خسارت‌های مالی و جانی را به تفکیک نشان می‌دهد. در این جدول علاوه بر تعداد مشاهدات و سنجه‌های هر یک از گشتاورهای مرتبه اول تا چهارم (میانگین، انحراف معیار، چولگی و کشیدگی) کوانتایل ۹۵ درصد مشاهدات و همچنین میانگین مشاهدات بالای کوانتایل ۹۵ درصد نمایش داده شده‌اند:

جدول ۱. آماره‌های توصیفی خسارت‌های جانی و مالی

| سنجه | خسارت‌های مالی | خسارت‌های جانی | لگاریتم خسارت‌های جانی | لگاریتم خسارت‌های مالی |
|-----------------------------------|----------------|----------------|------------------------|------------------------|
| تعداد مشاهدات | ۳۷,۸۰۴ | ۳۷,۸۰۴ | ۳۷,۸۰۴ | ۳۷,۸۰۴ |
| میانگین | ۲۱,۸۳۳,۸۵۷ | ۳۵۱,۳۱۰,۳۲۴ | ۱۶/۳۸ | ۱۸/۶۱ |
| انحراف معیار | ۲۲,۰۴۳,۴۲۶ | ۸۰۷,۴۶۰,۲۴۴ | ۱/۱۳ | ۱/۴۳ |
| چولگی | ۳/۸۷ | ۸/۴۵ | -۰/۳۹ | -۰/۱۷ |
| کشیدگی | ۷۰/۹۴ | ۱۵۲/۰۴ | ۲/۲۹ | ۲/۹۴ |
| کوانتایل ۹۵٪ | ۵۹,۹۵۸,۲۹۴ | ۱,۵۲۰,۰۰۰,۰۰۰ | ۱۷/۹۱ | ۲۱/۱۴ |
| میانگین مشاهدات فراتر از کوانتایل | ۸۵,۶۵۸,۸۲۵ | ۲,۸۷۳,۲۹۹,۰۹۹ | ۱۸/۲۲ | ۲۱/۶۳ |

بررسی‌های اولیه نشان داد برآزش توزیع تی استودنت چوله روی مقادیر اصلی خسارت‌های مالی مشکلات خاصی دارد؛ زیرا مقدار درجه آزادی کمتر از چهار است. پایین‌ترین مقدار درجه آزادی از چهار، به این معناست که گشتاورهای توزیع وجود ندارند. در واقع، حداقل چهار درجه آزادی لازم است تا چهار گشتاور اول موجود باشند. برای حل این مشکل براساس پیشنهاد الینگ (۲۰۱۲) به جای مدل‌سازی مقادیر اصلی، لگاریتم مشاهدات مدل‌سازی شده است. هیستوگرام مشاهدات اصلی و لگاریتم آنها در شکل ۱ ارائه شده است.



شکل ۱. هیستوگرام خسارت‌های جانی (بالا سمت راست) و مالی (بالا سمت چپ)، لگاریتم خسارت‌های جانی (پایین سمت راست) و لگاریتم خسارت مالی (پایین سمت چپ)

مدل‌سازی خسارت‌های مالی و جانی

تعیین آستانه u

بررسی‌ها نشان می‌دهد به‌ازای $u \geq 17/2$ شکل تابع میانگین مازاد خسارت‌های مالی، کمابیش به‌صورت یک خط راست است. بر همین اساس آستانه u برابر $17/2$ در نظر گرفته می‌شود. تعداد مشاهدات فراتر از این آستانه ۱۱۶۳۴ مشاهده است. این مشاهدات برای تخمین تابع GPD به‌منظور مدل‌سازی دنباله سمت راست توزیع خسارت‌های مالی به‌کار گرفته خواهند شد. مشاهدات کمتر از آستانه یادشده ۲۶۱۷۰ مشاهده است که به‌منظور مدل‌سازی مقادیر مرکزی و

دنباله سمت چپ توزیع خسارت‌های مالی استفاده می‌شود. برای خسارت‌های جانی، مشاهده می‌شود به‌ازای $u \geq 21/4$ شکل تابع به‌صورت یک خط دیده می‌شود، بنابراین آستانه $21/4$ به‌عنوان مرز مشخص‌کننده مشاهدات حدی انتخاب می‌شود. 1773 مشاهده فراتر از این آستانه قرار دارند که به‌منظور تخمین توزیع GPD به‌کار گرفته خواهند شد. 36031 مشاهده باقی‌مانده نیز برای تخمین مقادیر مرکزی و دنباله سمت چپ توزیع مقادیر لگاریتم خسارت‌های جانی با استفاده از توزیع تی‌استودنت چوله استفاده می‌شوند.

مدلسازی مقادیر حدی

نتیجه برازش توزیع GPD روی مشاهدات حدی با روش حداکثر درست‌نمایی در جدول ۲ ارائه شده است.

جدول ۲. پارامترهای توزیع GPD برازش‌شده روی لگاریتم مشاهدات حدی خسارت‌های جانی و مالی

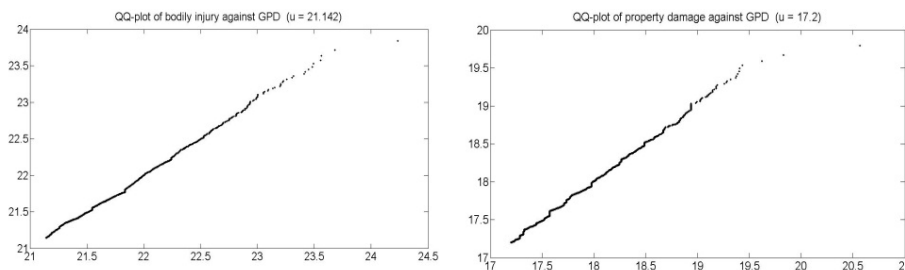
| پارامتر تخمین زده‌شده | لگاریتم خسارت‌های مالی | لگاریتم خسارت‌های جانی |
|-----------------------|------------------------|------------------------|
| | ۱۷/۲ | ۲۱/۴۲ |
| | -۰/۰۹۲۵ | -۰/۱۸۵۹ |
| | -۰/۴۱۴۲ | -۰/۶۶۷۳ |

به‌منظور بررسی نیکویی برازش‌های صورت‌گرفته، نمودار کوانتایل-کوانتایل^۱ مقادیر واقعی در مقابل مقادیر تخمین زده‌شده توسط GPD در شکل ۲ ترسیم شده است. گفتنی است که این نمودار مکان هندسی نقاط زیر است:

$$\left\{ \left(X_{k,n}, F^{\leftarrow} \left(\frac{n-k+1}{n+1} \right) \right) : k = 1, \dots, n \right\} \quad \text{رابطه ۲۶}$$

که در آن $X_{k,n}$ مشاهده K ام از تعداد کل n مشاهده است (مشاهدات به‌صورت صعودی مرتب شده‌اند) و F^{\leftarrow} تابع معکوس توزیع تخمین‌زده‌شده است. طبق قضیه گلیونکو-کانتلی^۲ هرچه این نمودار با خط $y = x$ شباهت بیشتری داشته باشد، معرف بهتر بودن برازش صورت گرفته است؛ از این رو، شباهت بسیار زیاد نمودارهای کوانتایل-کوانتایل ترسیم‌شده با خط $y = x$ گویای مطلوب بودن برازش‌های صورت‌گرفته روی مشاهدات حدی است.

1. Q-Q Plot
2. Gelivenkov-Cantelli Theorem



شکل ۲. نمودار کوانتایل-کوانتایل مقادیر واقعی در مقابل توزیع GPD تخمینی (راست: لگاریتم خسارت‌های مالی؛ چپ: لگاریتم خسارت‌های جانی)

مدل‌سازی مقادیر مرکزی و دنباله سمت چپ

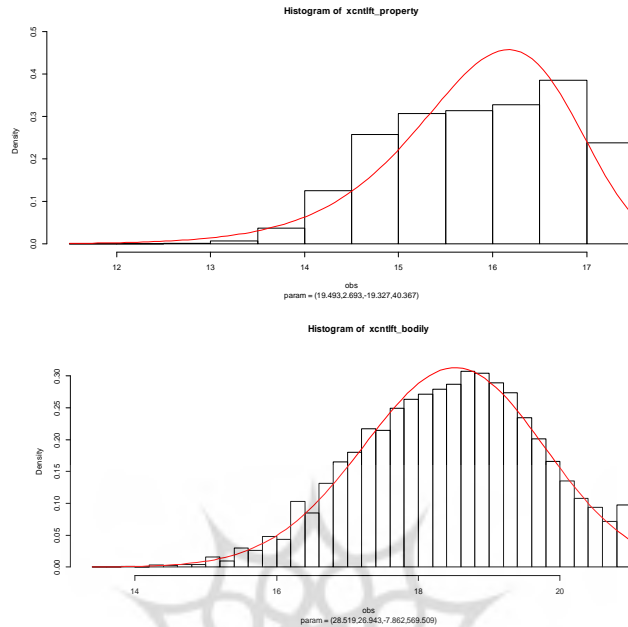
نتایج کالیبراسیون توزیع تی‌استودنت چوله روی مقادیر واقع در ناحیه مرکزی و دنباله سمت چپ توزیع خسارت‌ها با استفاده از الگوریتم EM در جدول ۳ ارائه شده است.

جدول ۳. نتایج کالیبراسیون توزیع تی‌استودنت چوله روی لگاریتم خسارت‌ها

| پارامتر تخمین زده شده | لگاریتم خسارت‌های مالی | لگاریتم خسارت‌های جانی |
|-----------------------|------------------------|------------------------|
| μ | ۱۹/۴۹۳ | ۲۸/۵۱۹ |
| δ | ۲/۶۹۳ | ۲۶/۹۴۳ |
| β | -۱۹/۳۲۷ | -۷/۸۶۲ |
| ϑ | ۴۰/۳۶۷ | ۵۶۹/۵۰۹ |
| Likelihood | -۳۴۴۲/۱۴ | -۵۹۹۳۵/۸۹ |
| Iterations | ۱۰۰۱ | ۶۵۹ |

شکل ۳ هیستوگرام مقادیر واقع در قسمت مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت‌های مالی و خسارت‌های جانی را نشان می‌دهد، همچنین توزیع تی چوله برازش شده نیز با آن مقایسه شده است.

همان‌گونه که در شکل ۳ مشاهده می‌شود، توزیع تی‌استودنت چوله به خوبی توانسته است مقادیر مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت‌های جانی و مالی را مدل‌سازی کند. با این حال، عملکرد آن در خصوص لگاریتم خسارت‌های جانی به مراتب بهتر از لگاریتم خسارت‌های مالی بوده است. با توجه به شواهد موجود یکی از دلایل این نتیجه را می‌توان به چوله بودن بیش از حد توزیع تجربی لگاریتم خسارت‌های مالی نسبت داد.



شکل ۳. هیستوگرام مقادیر واقع در قسمت مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت در مقایسه با تابع تی استودنت چوله برازش شده روی آنها (بالا: خسارت مالی؛ پایین: خسارت جانی)

نتیجه‌گیری و پیشنهادها

در این تحقیق به منظور مدل‌سازی زیان‌های ناشی از ریسک‌های مختلف تحت پوشش بیمه‌نامه‌ای خاص، تابع توزیع زیان‌های هر ریسک به کمک توزیع ترکیبی که با استفاده از توزیع تی استودنت چوله هایپربولیک تعمیم یافته و نظریه مقادیر فرین ساخته شده بود، مدل‌سازی شد. علت استفاده از این توزیع ترکیبی، نیاز به ساخت توزیعی بود که به خوبی بتواند هم مقادیر مرکزی توزیع و هم مقادیر حدی را مدل‌سازی کند؛ مقادیری که به رغم احتمال پایین رخ دادن، در صورت وقوع می‌توانند نتایج ویران‌کننده‌ای را به دنبال داشته باشند و حتی بقای شرکت بیمه‌ای را به مخاطره بیندازند.

نتایج بررسی‌های صورت گرفته نشان داد که به خوبی می‌توان مقادیر مرکزی و دنباله سمت چپ توزیع مقادیر لگاریتم زیان‌ها را به کمک توزیع تی استودنت چوله و مقادیر واقع در دنباله سمت راست آن را به کمک توزیع پارتو تعمیم یافته مدل‌سازی کرد.

با توجه به نتایج تحقیق به شرکت‌های بیمه توصیه می‌شود که حتماً برای برآورد خسارت‌های ناشی از بیمه‌نامه‌های مختلف و به تبع آن، تعیین حق بیمه مناسب برای آن‌ها از

این گونه مدل‌ها استفاده کنند. البته، در این راستا باید بیمه مرکزی جمهوری اسلامی ایران، به‌عنوان متولی امر بیمه در کشور، اولاً هرچه سریع‌تر نظام تعرفه‌ای را تغییر دهد و از سوی دیگر شرکت‌های بیمه را مکلف به استفاده از اکچوئرها برای تخمین توابع شدت و خسارت زیان کند. در این تحقیق توزیع تی‌استودنت چوله به‌صورت حالت خاصی از توزیع هایپربولیک تعمیم‌یافته در نظر گرفته شده است، سایر محققانی که علاقه‌مندند در این زمینه کار کنند، می‌توانند مدل‌سازی تابع زیان را با استفاده از سایر اعضای این خانواده انجام دهند، یا اینکه ابتدا مدل‌سازی را براساس فرم کلی تابع هایپربولیک تعمیم‌یافته انجام دهند و پس از آن با توجه به اینکه تابع توزیع زیان به کدام یک از اعضای این خانواده نزدیک‌تر است، اظهار نظر کنند. همچنین در این تحقیق، توابع زیان خسارت‌های جانی و مالی به‌صورت مجزا مدل‌سازی شد، در حالیکه موضوع بسیار مهم، تعیین تابع توزیع توأم این متغیرهاست که بر پایه آن می‌توان درباره خسارت کلی ناشی از یک بیمه‌نامه اظهار نظر کرد، بر این اساس مدل‌سازی این تابع توزیع توأم می‌تواند موضوع تحقیقات آتی باشد.

References

- Aas, K. & Haff, I. (2006). The Generalized Hyperbolic Skew Student's t-Distribution. *Journal of Financial Econometrics* 4(2): 275-309.
- Balkema, A. A. & de Haan, L. (1974). Residual Life Time at Great Age. *Annals of Probability*, 2(5): 792-804.
- Barndorff-Nielsen, O. E. & Blæsild, P. (1981). Hyperbolic distributions and ramifications: Contributions to theory and application. *Statistical Distributions in Scientific Work*, 4: 19-44.
- Bassi, F., Embrechts, P. & Kafetzaki, M. (1998). Risk management and Quantile Estimation. In *A Practical Guide to Heavy Tails*, Adler, R. J., Feldman, F., and Taqqu, M. (eds), 111-130. Birkhäuser.
- Beirlant, J., Joossens, E. & Segers, J. (2004). Generalized Pareto Fit to the Society of Actuaries' Large Claims Database. *North American Actuarial Journal* 8(2): 108-111.
- Beirlant, J. & Teugels, J. (1992). Modeling large claims in non-life insurance. *Insurance: Mathematics and Economics*, 11 (1): 17-29.
- Bolancé, C., Guillen, M., Pelican, E. & Vernic, R. (2008). Skewed Bivariate Models and Nonparametric Estimation for the CTE Risk Measure. *Insurance: Mathematics and Economics*, 43 (3): 386-393.
- Chang Lee, W. (2012). Fitting Generalized Pareto Distribution to Commercial Fire Loss Severity: Evidence From Taiwan. *The Journal of Risk*, 14(3): 63-80.

- Chava, S., Stefanescu, C. & Turnbull, S. (2008). Modeling the Loss Distribution. *Working Paper*. Available in: http://faculty.london.edu/cstefanescu/Chava_Stefanescu_Turnbull.pdf.
- Dahen, H., Dionne, G. & Zajdenweber, D. (2010). A Practical Application of Extreme Value Theory to Operational Risk in Banks. *The Journal of Operational Risk*, 5(2): 1–16.
- Dempster, A. P., Laird, N. M. & Rubin, D. (1977). Maximum likelihood from incomplete data using the EM algorithm. *Journal of the royal statistical society, Series B*, 39 (1): 1–38.
- Eling, E. (2012). Fitting Insurance Claims to Skewed Distributions: Are the Skew-Normal and Skew-Student Good Models? *Insurance: Mathematics and Economics*, 51(2): 239-248.
- Embrechts, P., Kluppelberg, S. & Mikosch, T. (1997). *Extremal Events in Finance and Insurance*. Berlin: Springer.
- Embrechts, P., McNeil, A. & Straumann, D. (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls. In: *Dempster, M.A.H. (Ed.), Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, 176–223.
- Embrechts, P., Resnick, S. I. & Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal*, 3(2): 30-41.
- Gilli, M. & Kellezi, E. (2006). An Application of Extreme Value Theory for Measuring Financial Risk. *Computational Economic*, 27(1): 1-23.
- Hosking, J.R.M., Wallis, J.R. & Wood, E.F. (1985). Estimation of the generalized extreme value distribution by the method of probability weighted moments. *Technometrics*, 27(3): 251-261.
- Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation of the normal inverse Gaussian distribution. *Statistics & Probability Letters*, 57(1): 43–52.
- Lane, M.N. (2000). Pricing Risk Transfer Transactions. *ASTIN Bulletin*, 30 (2): 259-293.
- Lee, W. C. & Fang, C. J. (2010). The Measurement of Capital for Operational Risk of Taiwanese Commercial Banks. *The Journal of Operational Risk*, 5(2): 79-102.
- McNeil, A.J. (1997). Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, 27(1):117–137.
- McNeil, A. J. & Saladin, T. (1997). *The Peaks Over Thresholds Method for Estimating High Quantiles of Loss Distributions*. Preprint, Department Mathematik, ETH Zentrum, Zurich.

- McNeil, A.J. & Saladin, T. (2000). *Developing Scenarios for Future Extreme Losses Using the Peaks-over-Thresholds Method*. *Extremes and Integrated Risk Management*, edited by P. Embrechts, London, England, RISK Books, Chapter 18, 253-267.
- Mencia, F. J. & Sentana, E. (2004). Estimation and testing of dynamic models with generalised hyperbolic innovations. *CMFI Working Paper 0411*, Madrid, Spain.
- Patrick, D. F., Jordan, J. S. & Rosengren, E. S. (2004). Implications of Alternative operational Risk Modeling Techniques. *National Bureau of Economic Research Working Paper*, W11103.
- Pickands, J. I. (1975). Statistical Inference Using Extreme Value Order Statistics. *Annals of Statistics*, 3(1): 119-131.
- Prause, K. (1997). *Modelling financial data using generalized hyperbolic distributions*. FDM preprint 48, University of Freiburg.
- Resnick, S. I. (1997). Discussion of the Danish data on large fire insurance losses. *ASTIN Bulletin*, 27(1): 139-151.
- Rootzen, H. & Tajvidi, N. (1996). Extreme value statistics and wind storm losses: a case study. *Scandinavian Actuarial Journal*, 1997(1): 70-94.
- Rootzen, H. & Tajvidi, N. (2000). Extreme value statistics and wind storm losses: a case study. *In Extremes and Integrated Risk Management*, Embrechts, P. (ed). Risk Books, London.
- Vernic, R. (2006). Multivariate Skew-Normal Distributions with Applications in Insurance. *Insurance: Mathematics and Economics*, 38(2): 413-426.