

مدل سازی تابع توزیع زیان های بیمه ای با بهره گیری از توزیع های ترکیبی و مفهوم کاپیولا

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

چکیده: این تحقیق سعی دارد با بهره گیری همزمان از توزیع های ترکیبی و مفهوم کاپیولا، تابع توزیع توأمان زیان های وارد شده بر اکسپوزرهای مختلف تحت پوشش یک بیمه نامه خاص را نسبت به توزیع های آماری موجود، با دقت بیشتری مدل سازی کند. در این تحقیق از توزیع خاصی که ترکیبی از توزیع تی استودنت چوله هایپربولیک تعمیم یافته و نظریه مقادیر فرین است، برای مدل سازی توابع زیان حاشیه ای و از مفهوم کاپیولا برای مدل سازی ساختار وابستگی میان آنها استفاده شده است. کاپیولا های گوسی، تی، فرانک، گامیل و کلایتون، مهم ترین انواع کاپیولای بررسی شده اند تا از بین آنها بهترین گزینه برای تشریح ساختار وابستگی زیان ها انتخاب شود. داده های مورد استفاده در این تحقیق مقدار خسارت های جانی و مالی بیمه نامه های شخص ثالث وسایل نقلیه موتوری است. نتایج تحقیق نشان می دهد با بهره گیری از توزیع ترکیبی پیشنهادی و کاپیولای کلایتون تابع توزیع توأم، می توان به خوبی زیان های نشست گرفته از بیمه نامه شخص ثالث را مدل سازی کرد.

واژه های کلیدی: توزیع ترکیبی، توزیع توأم، توزیع حاشیه ای، تابع کاپیولا.

۱. دانشجوی دکتری مدیریت مالی، دانشکده مدیریت 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: ۱۳۹۴/۰۹/۱۱

تاریخ پذیرش نهایی مقاله: ۱۳۹۵/۰۴/۰۳

نویسنده مسئول مقاله: سعید باجلان

E-mail: Saeedbajalan@ut.ac.ir

مقدمه

اغلب در انواع بیمه نامه ها چندین اکسپوژر^۱ (مواد در معرض خسارت) تحت پوشش یک بیمه نامه قرار می گیرند که معمولاً خسارت های وارده به این اکسپوژرها مستقل از یکدیگر نیست. برای مثال بیمه های شخص ثالث اتومبیل معمولاً خسارت های بدنی و مالی را پوشش داده و بیمه های آتش سوزی، ساختمان^۲ و محتویات^۳ آن را پوشش می دهند.

از آنجا که میان این زیان ها وابستگی وجود دارد، لازم است هنگام مدل سازی زیان کلی ناشی از یک حوزه خاص فعالیت به نوعی وابستگی میان این زیان ها مد نظر قرار گیرد تا بتوان به درستی حق بیمه متناسب با هر ریسک را تعیین کرد. به بیان دیگر، برای برآورد صحیح تابع زیان، اولاً باید تابع توزیع زیان وارد بر هر اکسپوژر به نحو مطلوبی تخمین زده شود و ثانیاً باید تابع توزیع احتمال توأمان^۴ زیان های وارد بر اکسپوژرها به طور صحیح مدل سازی شود. برای مثال، در صورتی که هدف مدل سازی زیان ناشی از بیمه نامه آتش سوزی باشد، اولاً باید زیان های وارد شده به ساختمان و محتویات آن به صورت جداگانه به دقت مدل سازی شوند، ثانیاً بتوان با استفاده از تابع احتمال توأمان مناسب، خسارت کلی ناشی از یک بیمه نامه آتش سوزی را مدل سازی کرد.

در چنین حالت هایی معمولاً از تحلیل های چندمتغیره استفاده می شود. در خصوص بهره گیری از توزیع نرمال و توزیع های غیرنرمال چندمتغیره در راستای مدل سازی تابع احتمالات مشترک، ادبیات موضوعی گسترده ای وجود دارد (جانسون و ویچرن، ۱۹۸۸؛ کرزانوسکی، ۱۹۸۸؛ اندرسن، ۱۹۵۸؛ جانسون، کوتز و بالا کریشان، ۱۹۹۷) با این حال به این گونه توزیع ها چند ایراد اساسی وارد است: ۱. برای هر تابع توزیع حاشیه ای^۵، خانواده متفاوتی لازم است، ۲. توسعه به حالت های بیش از دو متغیر چندان واضح نیست و ۳. نوع همبستگی متغیرها متأثر از توزیع حاشیه ای آنها است. ساخت توزیع چندمتغیره ای که از این ایرادها مبرا باشد، بر اساس مفهوم تابع کاپیولا^۶ بنا شده است. در واقع کاپیولاها به ما کمک می کنند که تابع احتمال مشترک را مجزا از تابع توزیع حاشیه ای تخمین بزنیم.

بر همین اساس، در این تحقیق سعی شده که تابع توزیع توأم زیان های ناشی از یک نوع بیمه نامه خاص به کمک مفهوم کاپیولا مدل سازی شود.

-
1. Exposure
 2. Buildings
 3. Contents
 4. Joint Probability Distribution Function
 5. Marginal Distribution
 6. Copula Function

پیشینه تجربی

پیشینه مطالعات انجام گرفته را می توان در دو بخش بررسی کرد، مدل سازی توابع زیان حاشیه ای و استفاده از کاپولا برای مدل سازی ساختار وابستگی میان زیان ها. تابع نرمال متداول ترین تابع استفاده شده برای مدل سازی توابع حاشیه ای در حوزه اقتصاد و مالی است. اما به طور کلی توزیع زیان های بیمه ای، چوله^۱ هستند، پدیده ای که باعث می شود توزیع نرمال، برای مدل سازی ریسک های بیمه ای، توزیع مناسبی نباشد (لین، ۲۰۰۰؛ ورنیک، ۲۰۰۶). اگرچه برای مدل سازی ریسک های بیمه ای در خصوص استفاده از سایر توزیع ها، کارهایی انجام شده، اما به طور عمده در این کارها سعی شده که از توزیع های بیضوی^۲ استفاده شود؛ زیرا در عین حال که برخی از خواص توزیع نرمال را دارند، دارای انعطاف بیشتری در خصوص چولگی و کشیدگی هستند (الینگ، ۲۰۱۲). با این حال ایراد اساسی توزیع های بیضوی متقارن بودن آنها است، حال آن که تحقیقات متعدد نشان داده اند توزیع زیان های بیمه ای شکل متقارنی ندارد.

در خصوص استفاده از توزیع های نامتقارن برای تخمین تابع زیان کارهایی انجام شده که یکی از جامع ترین آنها کار الینگ (۲۰۱۲) است. در تحقیقی که الینگ (۲۰۱۲) روی داده های مربوط به شدت خسارت انجام داد، ۱۶ توزیع متداول در اکچوئری از جمله توزیع های نمایی، هندسی، کای دو، کوچی، گاما، گوس، لوجستیک، لوگ نرمال، تی استودنت، ویبول، هایپربولیک و... با دو توزیع نرمال چوله و تی استودنت چوله از لحاظ برازندگی^۳ مقایسه شدند. نتایج تحقیق وی نشان داد که توزیع تی استودنت چوله در مقایسه با سایر رقبا از برازندگی خوبی برخوردار است.

با این حال به علت اهمیت وقایع فرین^۴ در بیمه گری، لازم است دنباله های توزیع زیان با دقت بسیار بیشتری تخمین زده شوند؛ زیرا اتفاقاتی با احتمال بسیار کم اما شدت زیاد می تواند بقای یک شرکت بیمه را با چالش جدی مواجه کند. بحث سر این موضوع که آیا تئوری مقادیر فرین^۵ می تواند از چنین خسارت هایی، تخمین های بهتری ارائه دهد، موضوع تحقیقات مختلفی بوده است که از آن جمله می توان به (باسی، امبرچتز و کافتزاک، ۱۹۹۸؛ مک نیل و سالادین، ۱۹۹۷؛ امبرچتز، رزنیک و سامورودنیتسکای، ۱۹۹۹؛ بیرلانت، جوسنس و سیگرز، ۲۰۰۴؛ چاوا،

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

۲۰۰۸؛ لی و فانگ، ۲۰۱۰) اشاره کرد. عمده این تحقیقات نشان داده اند که این تئوری می تواند در تخمین دنباله های توزیع زیان بهتر از رقبا عمل کند.

با این حال در خصوص تخمین تابع زیان با ترکیب کردن این دو دسته توزیع برای رسیدن به تخمین های صحیح تر از شکل تابع زیان تاکنون کار زیادی انجام نشده است. در تحقیقی که توسط باجلان و همکارانش انجام گرفت، سعی شد با ترکیب توزیع تی استودنت چوله هایپربولیک تعمیم یافته و تئوری مقادیر فرین مدل سازی بهتری از توزیع زیان انجام شود. نتایج تحقیق نشان داد که توزیع ترکیبی پیشنهادی به خوبی می تواند مقادیر مرکزی و همچنین مقادیر حدی زیان ها را مدل سازی کند (باجلان، راعی و محمدی، ۱۳۹۵).

از سوی دیگر استفاده از تابع کاپیولا به عنوان ابزاری برای مدل سازی ساختار وابستگی متغیرهای تصادفی بسیار فراگیر شده است. فریز، کاریز و والدز (۱۹۹۶) از کاپیولاها برای نشان دادن ساختار وابستگی زمان باقی مانده تا فوت در بیمه های عمر مشترک و آخرین بازمانده استفاده کردند. فریز و والدز (۱۹۹۸) مقدمه ای بر کاپیولاها ارائه کردند که مرور مفیدی بر ادبیات موضوع کاپیولاهاست. آنها در مقاله خود به جستجوی رابطه ای میان خسارت های ناشی از بیمه مسئولیت و هزینه ارزیابی خسارت تخصیص یافته^۱ با استفاده از کاپیولاها پرداختند. بوی (۲۰۰۲) در تحقیق خود روی انتخاب کاپیولا و تخمین آن در حیطه تحلیل ریسک پرتفوی دارایی ها تمرکز کرد. ووتر (۲۰۰۲) مباحثی را در خصوص چندین خانواده پارامتریک کاپیولا به همراه روش هایی برای انتخاب گزینه مناسب از میان آنها برای مدل سازی وابستگی ارائه کرد. وی همچنین در خصوص شبیه سازی خسارت های ناشی از گردبادها در دو ایالت آمریکا، مثالی مطرح کرد. بلقیوس و لوی (۲۰۰۳) با استفاده از کاپیولاها، سعی در مدل سازی وابستگی میان خسارت ناشی از بیمه اتومبیل و آتش سوزی کردند. دیبی و جونز (۲۰۰۶) در مقاله ای نتایج به دست آمده از تحلیل های انجام شده روی مقادیر فرین چندمتغیره را گردآوری کرده و کاربردهای آنها را در اکچوئری تشریح کردند. آنها در مقاله خود به طور خاص روی بحث وابستگی در مقادیر فرین تمرکز کرده و بر اهمیت انتخاب کاپیولای صحیح هنگام کار کردن با دنباله سمت راست تأکید کردند.

از آنجا که به منظور عملکرد صحیح شرکت های بیمه، مهم ترین نیاز داشتن تخمین های دقیق تر از شکل تابع توزیع زیان است (زیرا در عمل بدون داشتن تخمین دقیق از توزیع زیان، تعیین حق بیمه میسر نیست)، در این تحقیق سعی خواهد شد با بهره گیری همزمان از توزیع های ترکیبی و مفهوم کاپیولا، مدل سازی دقیق تری از تابع توزیع زیان نسبت به مدل سازی های موجود انجام شود. در ادامه این دو مفهوم به صورت مختصر تشریح می شوند.

1. Allocated Loss Adjustment Expenses (ALAE)

توزیع ترکیبی^۱

در پژوهشی که باجلان و همکارانش (۱۳۹۵) انجام دادند، به منظور مدل سازی زبان های بیمه ای از یک توزیع ترکیبی که با استفاده از توزیع تی استودنت چوله هایپربولیک تعمیم یافته و نظریه مقادیر فرین ساخته شده بود، استفاده شد. علت استفاده از این توزیع ترکیبی نیاز به ساخت توزیعی بود که به خوبی بتواند مقادیر مرکزی توزیع و مقادیر حدی را مدل سازی کند، مقادیری که علی رغم احتمال پایین رخ دادن در صورت وقوع می توانند نتایج ویران کننده ای به دنبال داشته باشند و حتی بقای یک شرکت بیمه را به مخاطره بیندازند.

آنها در تحقیق خود بر پایه مقاله آس و هاف (۲۰۰۶) توزیع تی استودنت چوله را به صورت حالت خاصی از توزیع هایپربولیک تعمیم یافته، تعریف کردند. علت بهره گیری از این تعریف خاص داشتن یک دنباله با رفتار چند جمله ای و یک دنباله به صورت نمایی در این توزیع است که برای مدل سازی داده هایی با دنباله پهن و چولگی زیاد منحصر به فرد محسوب می شود. تابع چگالی این توزیع به صورت زیر است:

$$f_x(x) = \begin{cases} \frac{2^{\frac{1-v}{2}} \delta^v |\beta|^{\frac{v+1}{2}} K_{\frac{v+1}{2}}(\sqrt{\beta^2(\delta^2 + (x-\mu)^2)}) \exp(\beta(x-\mu))}{\Gamma(\frac{v}{2}) \sqrt{\pi} (\sqrt{\delta^2 + (x-\mu)^2})^{\frac{v+1}{2}}}, & \beta \neq 0 \\ \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi} \delta \Gamma(\frac{v}{2})} \left[1 + \frac{(x-\mu)^2}{\delta^2} \right]^{-(v+1)/2}, & \beta = 0 \end{cases} \quad \text{رابطه ۱}$$

که در آن پارامترهای μ و δ به ترتیب موقعیت^۲ و مقیاس^۳ را تعیین کرده و β و α شکل^۴ توزیع را مشخص می کنند. $K_\lambda(x)$ که $x > 0$ معرف تابع بسل تعدیل یافته^۵ نوع دوم با رتبه λ است. Γ معرف تابع گاماست و v تعداد درجات آزادی را نشان می دهد. این توزیع برای مدل سازی مقادیر مرکزی و دنباله سمت چپ توزیع زبان ها به کار گرفته شد.

به منظور مدل سازی مقادیر حدی نیز بر اساس رویکرد اوج فراتر از آستانه (POT) به عنوان یکی از رویکردهای اساسی تئوری مقادیر فرین (EVT) از توزیع پارتوی تعمیم یافته (GPD) استفاده شد. بر اساس این رویکرد بر پایه قضیه بالکما، دی هان (۱۹۷۴) و پیکاندز (۱۹۷۵)

1. Compound distribution
2. Location
3. Scale
4. Shape
5. Modified Bessel Function

ع شایان ذکر است که در متون قدیمی تر به آن نوع ۳ گفته می شود.

به عنوان دومین قضیه اصلی تئوری مقادیر فرین، تحت شرط $F \in MDA(H)$ ، توزیع پارتوی تعمیم یافته توزیع حدی برای توزیع مقادیر بیش از یک آستانه به اندازه کافی بزرگ u است. آنها در تحقیق خود برای تخمین آستانه u از تابع میانگین مازاد و برای کالیبراسیون توزیع تی استونت چوله هایپر بولیک تعمیم یافته از الگوریتم حداکثر سازی انتظارات (EM) استفاده کردند. پارامترهای توزیع GPD نیز با بهره گیری از روش حداکثر درست نمایی (MLL) تخمین زده شدند. به منظور مطالعه بیشتر در این خصوص می توان به تحقیق اصلی مراجعه کرد.

کاپیولا

هر تابع توزیع توأم برداری تصادفی از عوامل ریسک، به طور ضمنی شامل شرحی از رفتار حاشیه ای عوامل ریسک به طور منفرد و شرحی از ساختار وابستگی بین آنها است. رویکرد کاپیولا راهی برای تفکیک ساختار وابستگی از رفتار حاشیه ای عوامل فراهم می آورد.

تعریف کاپیولا: یک کاپیولای d بعدی یک تابع توزیع روی $[0, 1]^d$ با توزیع های حاشیه ای یکنواخت استاندارد است. از نماد $C(u) = C(u_1, u_2, \dots, u_d)$ در این متن برای نمایش توابع توزیع چند متغیره ای که کاپیولا هستند، استفاده خواهد شد. بنابراین C نگاشتی به شکل زیر است:

$$C: [0, 1]^d \rightarrow [0, 1] \quad \text{رابطه ۲}$$

قضیه اسکالر (۱۹۵۹):^۱ اگر F یک تابع توزیع توأم با توابع حاشیه ای $F_1, \dots, F_d$ باشد، یک کاپیولا $C: [0, 1]^d \rightarrow [0, 1]$ وجود دارد به گونه ای که برای همه $x_1, \dots, x_d$ عضو $\mathcal{R} = [-\infty, \infty]$

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad \text{رابطه ۳}$$

اگر توابع حاشیه ای پیوسته باشند، C منحصر به فرد است^۲ و برعکس اگر C یک کاپیولا و $F_1, \dots, F_d$ توابع توزیع تک متغیره باشد، تابع F که در رابطه ۳ تعریف شده است، یک تابع توزیع توأم با توابع حاشیه ای $F_1, \dots, F_d$ است. اگر داشته باشیم $x_i = F_i^{\leftarrow}(u_i)$ و $0 \leq u_i \leq 1$ و $i = 1, \dots, d$ می توان رابطه ۳ را به صورت زیر نوشت:

1. $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \times dt$
2. Range

$$C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)) \quad \text{رابطه ۴}$$

رابطه ۴، ارائه صریحی از C بر حسب تابع توزیع توأم F و توابع حاشیه‌ای آن می‌دهد، بنابراین بیان‌کننده منحصر به فرد بودن آن نیز هست.

رابطه‌های ۳ و ۴ در برخورد با کاپیولاها بسیار اساسی هستند. نخستین رابطه نشان می‌دهد که چگونه تابع توزیع توأم F از طریق زوج کردن توزیع‌های حاشیه‌ای با کاپیولای C شکل می‌گیرد و رابطه دوم نشان می‌دهد که چگونه کاپیولاها از یک تابع توزیع چندمتغیره و توابع حاشیه‌ای پیوسته استخراج می‌شوند.

در دیدگاهی کلی می‌توان کاپیولاها را در سه طبقه جا داد: کاپیولاهای بنیادی^۱ که معرف ساختارهای وابستگی مهم خاص (نظیر استقلال، وابستگی کامل مثبت و وابستگی کامل منفی) هستند. کاپیولاهای ضمنی^۲ که از توزیع‌های چندمتغیره معروف با استفاده از قضیه اسکالر استخراج شده‌اند، اما لزوماً دارای تصریح ساده‌ای^۳ نیستند، نظیر خانواده کاپیولاهای گوسی. کاپیولاهای صریح^۴ که دارای تصریح ساده‌ای هستند و با استفاده از اصول عمومی ریاضی ساخته شده‌اند، نظیر کاپیولاهای ارشمیدسی (مکنیل، فری و امبرچتر، ۲۰۰۵).

از دیدگاهی دیگر می‌توان کاپیولاها را به دو دسته پارامتریک و ناپارامتریک دسته‌بندی کرد (چروبینی، لویانکو و وچیاتو، ۲۰۰۴). کاپیولاهای پارامتریک دارای پارامترهایی برای کنترل شدت وابستگی هستند. برخی محققان نیز کاپیولاها را در سه خانواده کاپیولاهای بیضوی، ارشمیدسی و کاپیولاهای مقادیر فرین جا داده‌اند (نلسن، ۱۹۹۹).

روش‌شناسی پژوهش

با توجه به موضوع تحقیق روش انجام آن را می‌توان به صورت زیر تشریح کرد:

۱. مدل‌سازی تابع توزیع هر یک از ریسک‌های تحت پوشش بیمه‌نامه توسط توزیع ترکیبی. بر اساس این توزیع مقادیر مرکزی و دنباله سمت چپ توزیع با استفاده از توزیع تی استودنت چوله و مقادیر دنباله سمت راست توزیع با استفاده از تئوری مقادیر فرین مدل‌سازی می‌شود. مراحل انجام این مدل‌سازی مطابق تحقیق باجلان و همکاران (۱۳۹۵) انجام خواهد گرفت.

1. Fundamental Copulas
2. Implicit Copulas
3. Simple Closed Form Expression
4. Explicit Copulas

۲. مدل سازی همبستگی میان ریسک های تحت پوشش بیمه نامه: بعد از اینکه توابع زیان تک تک خسارت های تحت پوشش یک بیمه نامه مدل سازی شد، لازم است ساختار همبستگی میان آنها مدل سازی شود. همان گونه که در قسمت مربوط به کاپیولاها توضیح داده شد، کاپیولاها اجازه می دهند که ساختار وابستگی مستقل از توزیع های حاشیه ای مدل سازی شود و همچنین بسیاری از معایب ضرایب همبستگی را ندارند. بر این اساس بعد از تخمین پارامترهای توزیع های حاشیه ای، با استفاده از روش استنباط برای حاشیه ها (IFM) پنج خانواده از کاپیولاها شامل گوسی، تی، گامبل، کلایتون و فرانک، که بیشترین کاربرد را در حوزه اکچوئری دارند، تخمین زده می شوند و از میان آنها کاپیولایی که دارای بیشترین مقدار لگاریتم درست نمایی باشد، به عنوان کاپیولای برتر انتخاب می شود.
۳. در گام سوم با ترکیب کردن نتایج مراحل ۱ و ۲ مدل سازی نهایی از تابع توزیع زیان کلی تحت پوشش یک بیمه نامه انجام می گیرد.

جامعه آماری

جامعه آماری تحقیق داده های خسارتی کلیه شرکت های بیمه ایرانی است. از آنجا که در مباحث اکچوئری هر مشاهده ای حاوی اطلاعات بسیار ارزشمندی در خصوص شکل توزیع است که در صورت نمونه گیری، احتمال از دست رفتن آن وجود دارد (به دلیل حذف شدن مشاهده)، در تحقیقات این حوزه عملاً بحث نمونه گیری مطرح نیست و باید سعی شود مدل ها با استفاد از حداکثر اطلاعات موجود تخمین زده شوند. از آنجا که در این تحقیق ساختار وابستگی میان زیان های تحت پوشش یک بیمه نامه نیز مدل سازی خواهد شد، لازم است مشاهدات گردآوری شده به صورت زوجی باشند. بررسی های انجام شده نشان داد با توجه به ضعف سیستم های اطلاعاتی صنعت بیمه کشور، گردآوری داده های خسارتی برای اکثر رشته ها به تفکیک امکان پذیر نیست و صرفاً می توان داده های مربوط به خسارت جانی و مالی ناشی از رشته شخص ثالث را گردآوری و به نوعی تفکیک کرد. به همین منظور در این تحقیق روی مدل سازی زیان های ناشی از رشته شخص ثالث تمرکز شده است.

به منظور گردآوری و پایش داده ها، در ابتدا داده های خسارتی مربوط به بیمه نامه های شخص ثالث صادر شده توسط کلیه شرکت های بیمه که اطلاعات آنها در سامانه سنهاب موجود است، گردآوری شد. این اطلاعات شامل ۱۴۹۳۰۵ مشاهده است که هر یک شامل شماره بیمه نامه، تاریخ صدور بیمه نامه، مبلغ خسارت و نوع خسارت (جانی و مالی) در بازه زمانی فرودین ۱۳۸۴ تا مرداد ۱۳۹۲ است.

در گام بعد، از آنجا که در مواردی به ازای یک شماره بیمه نامه مشخص، بیش از یک خسارت رخ داده بود، کلیه این خسارت ها تجمیع شدند تا مشخص شود به ازای هر بیمه نامه صادر شده، چه مقدار خسارت جانی و خسارت مالی پرداخت شده است. بعد از انجام این اصلاحات در مجموع ۳۷۸۰۴ زوج مشاهده از خسارت های جانی و مالی به ازای هر بیمه نامه به دست آمد.

با توجه به تغییرات نرخ دیه و همچنین تغییرات اجرت تعمیر و قیمت خودرو، لازم بود که میزان خسارت های جانی و مالی مربوط به سال های قبل از ۱۳۹۲ تعدیل شوند. به منظور انجام این تعدیلات ابتدا مبلغ دیه و شاخص هزینه حمل و نقل استخراج شد؛ سپس نرخ رشد آنها به دست آمد. از نرخ رشد محاسبه شده به عنوان عامل تعدیل خسارت های جانی و مالی در سال های قبل از ۱۳۹۲ استفاده شد.

یافته های پژوهش

مدل سازی خسارت های مالی و جانی با استفاده از توزیع ترکیبی

تعیین آستانه u : برای تعیین آستانه u به منظور مشخص کردن مشاهدات حدی که دارای توزیع GPD هستند (بر اساس دومین قضیه بنیادی تئوری مقادیر فرین) با توجه به توضیحات ارائه شده در تحقیق باجلان و همکاران (۱۳۹۵)، از تابع میانگین مازاد استفاده شده است. بر این اساس آستانه u برای لگاریتم خسارت های مالی برابر $۱۷/۲$ و برای لگاریتم خسارت های جانی $۲۱/۴$ انتخاب شد. مشاهدات فراتر از این آستانه ها به عنوان مشاهدات حدی در نظر گرفته خواهد شد که با استفاده از توزیع GPD مدل سازی می شوند و مقادیر کمتر از آنها به عنوان مقادیر مرکزی و دنباله سمت چپ در نظر گرفته می شود که با استفاده از توزیع تی استودنت چوله مدل سازی خواهند شد.

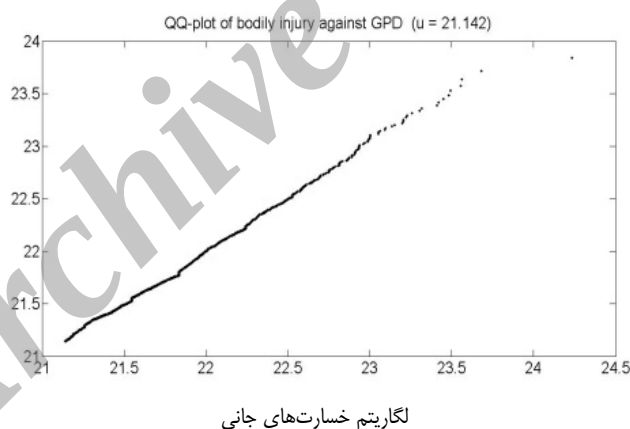
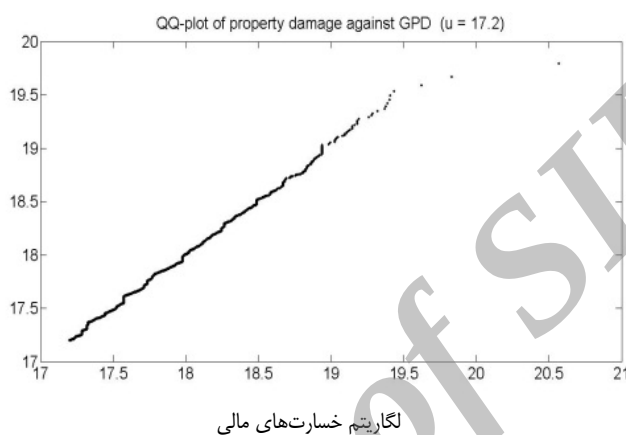
مدل سازی مقادیر حدی

پارامترهای توزیع GPD برازش شده روی مشاهدات حدی با استفاده از روش حداکثر درست نمایی در جدول ۱ ارائه شده است.

جدول ۱. پارامترهای توزیع GPD برازش شده روی لگاریتم مشاهدات حدی خسارت های جانی و مالی

پارامتر تخمین زده شده	لگاریتم خسارت های مالی	لگاریتم خسارت های جانی
μ	۱۷/۲	۲۱/۴۲
ξ	-۰/۰۹۲۵	-۰/۱۸۵۹
σ	۰/۴۱۴۲	۰/۶۶۷۳

به منظور بررسی نیکویی برازش های انجام شده نمودار کوانتایل - کوانتایل^۱، مقادیر واقعی در مقابل مقادیر تخمین زده شده توسط GPD در شکل ۱ ترسیم شده است. طبق قضیه گلیونکو - کانتلی^۲، شباهت هر چه بیشتر این شکل با یک خط راست با زاویه ۴۵ درجه که از مبدأ عبور می کند، معرف بهتر بودن برازش انجام شده است.



شکل ۱. نمودار کوانتایل - کوانتایل مقادیر واقعی در مقابل توزیع GPD تخمینی

شباهت بسیار زیاد نمودارهای کوانتایل - کوانتایل ترسیم شده با خط $y = x$ نشان دهنده مطلوب بودن برازش های انجام شده روی مشاهدات حدی است.

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

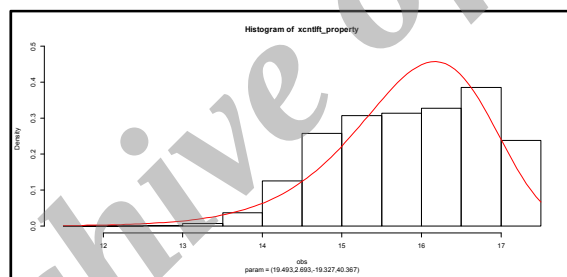
مدل سازی مقادیر مرکزی و دنباله سمت چپ

نتایج کالیبراسیون توزیع تی استودنت چوله هایپربولیک تعمیم یافته روی مقادیر مرکزی و دنباله سمت چپ با استفاده از الگوریتم حداکثر سازی انتظارات در جدول ۲ ارائه شده است.

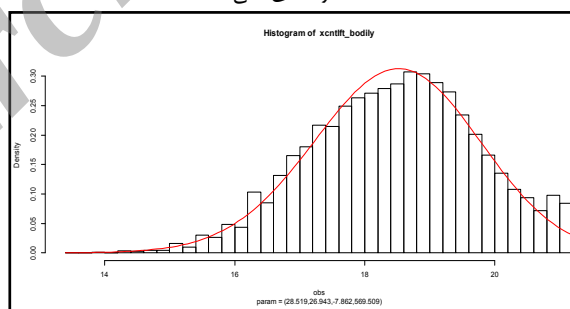
جدول ۲. نتایج کالیبراسیون توزیع تی استودنت چوله روی لگاریتم خسارت ها

پارامتر تخمین زده شده	لگاریتم خسارت های مالی	لگاریتم خسارت های جانی
μ	۱۹/۴۹۳	۲۸/۵۱۹
δ	۲/۶۹۳	۲۶/۹۴۳
β	-۱۹/۳۲۷	-۷/۸۶۲
θ	۴۰/۳۶۷	۵۶۹/۵۰۹
Likelihood	-۳۴۴۴۲/۱۴	-۵۹۹۳۵/۸۹
Iterations	۱۰۰۱	۶۵۹

هیستوگرام مقادیر واقع در قسمت مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت های مالی و جانی در شکل ۲ ترسیم شده و توزیع تی چوله برازش شده نیز با آن مقایسه شده است.



خسارت های مالی



خسارت های جانی

شکل ۲. هیستوگرام مقادیر واقع در قسمت مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت در مقایسه با تابع تی استودنت چوله برازش شده روی آنها

همان گونه که در شکل ۲ مشاهده می شود، توزیع تی استودنت چوله به خوبی توانسته مقادیر مرکزی و دنباله سمت چپ توزیع لگاریتم خسارت های جانی و مالی را مدل سازی کند، هر چند عملکرد آن در خصوص لگاریتم خسارت های جانی به مراتب بهتر بوده است.

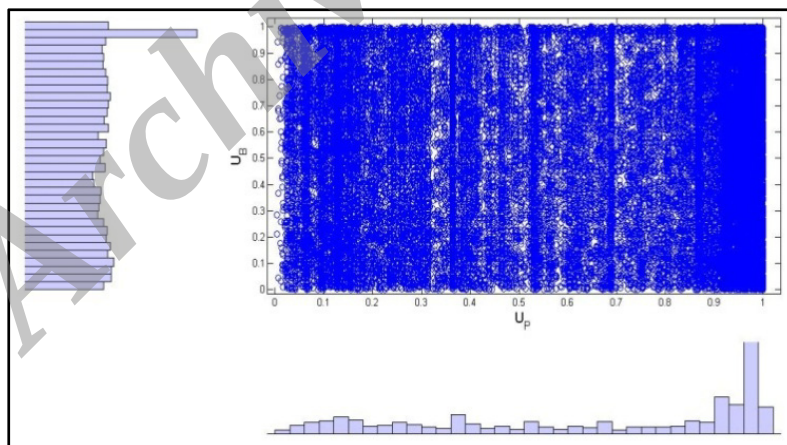
مدل سازی ساختار وابستگی خسارت های مالی و جانی

بعد از مدل سازی توزیع لگاریتم خسارت های مالی و جانی به طور مجزا، باید ساختار وابستگی میان آنها مدل سازی شود.

به منظور محاسبه $F_1()$ و $F_2()$ (تابع توزیع لگاریتم خسارت های مالی و لگاریتم خسارت های جانی) به ازای مقادیر مختلف x_{1t} و x_{2t} به این صورت عمل شده که برای مشاهدات واقع در دنباله سمت چپ و ناحیه مرکزی توزیع از تابع چگالی احتمال توزیع تی استودنت چوله با استفاده از روش های عددی انتگرال گیری شده و برای مشاهداتی که بالاتر از آستانه u قرار دارند، از رابطه زیر استفاده شده است:

$$F(x) = F_{Skew t}(u) + (1 - F_{Skew t}(u)) \times F_{GPD}(y) \quad \text{رابطه ۵}$$

شکل ۳ ساختار پراکنش میان مقادیر تابع توزیع لگاریتم خسارت های مالی و خسارت های جانی را نشان می دهد.



شکل ۳. پراکنش مقادیر تابع توزیع لگاریتم خسارت های مالی و خسارت های جانی

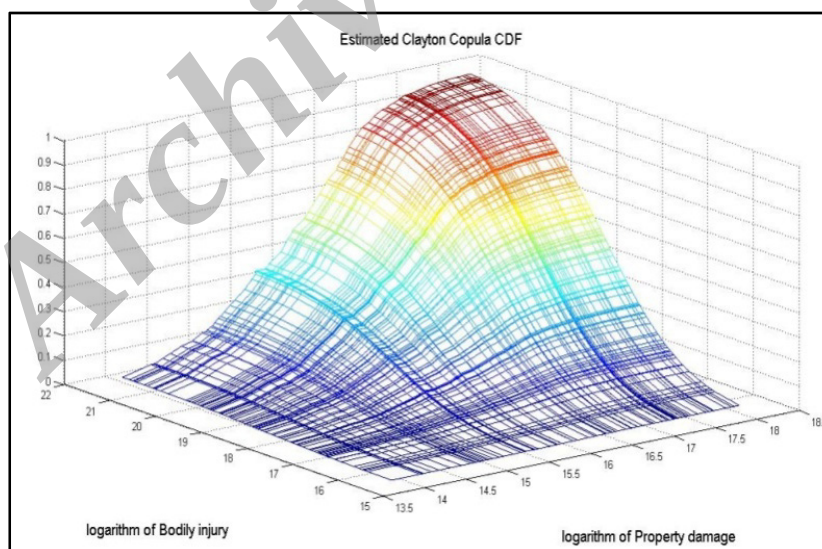
بعد از تخمین کاپیولا‌های مختلف، بیشینه مقدار تابع درستنمایی آنها محاسبه شده است که نتیجه آن را می‌توان در جدول ۳ مشاهده کرد.

جدول ۳. مقادیر حداکثر لگاریتم درستنمایی کاپیولا‌های بررسی شده

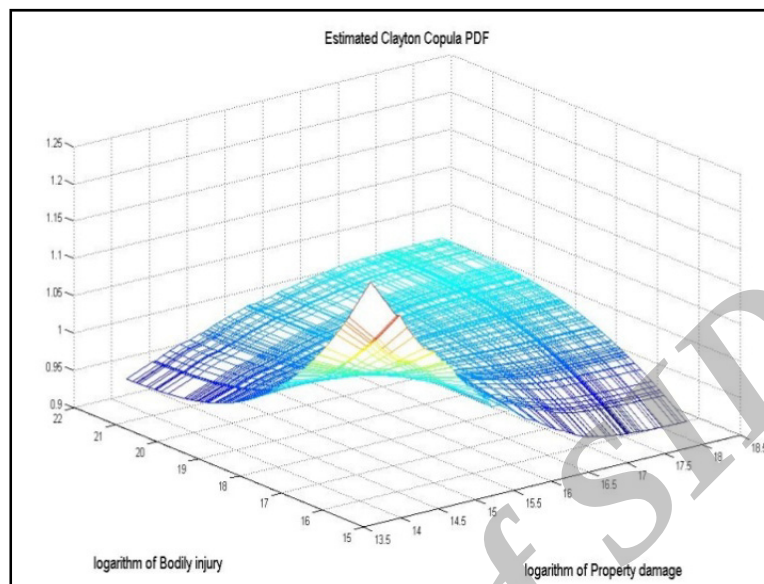
نام کاپیولا	مقدار حداکثر لگاریتم درستنمایی
گوسی	-۳۸/۴۹۹۹
تی	-۱۴۸/۵۹۶۲
کلایتون	-۱۰/۱۵۱۵
فرانک	-۴۲/۰۷۳۰
گامبل	-۱۵۶/۳۲۰۲

با توجه به مقادیر جدول ۳، بهترین کاپیولا برای مدل‌ساز ساختار وابستگی میان لگاریتم خسارت‌ها کاپیولای کلایتون است. پارامتر α تخمین زده شده برای این کاپیولا برابر ۰/۰۲۶۷ است.

در شکل ۴ تابع توزیع احتمال توأمان لگاریتم خسارت‌های مالی و جانی به همراه تابع چگالی توأمان آنها با توجه به کاپیولای تخمین زده شده محاسبه و نمایش داده شده است.



شکل ۴- الف. تابع توزیع احتمال توأمان لگاریتم خسارت‌های مالی و جانی



شکل ۴- ب. تابع چگالی احتمال توأمان خسارت های مالی و جانی
 $Clayton\ Copula, \alpha = 0.0267$

مقایسه مدل تحقیق با مدل رقیب

به منظور مقایسه عملکرد مدل ارائه شده در تحقیق با رقبای موجود، مراحل تحقیق با استفاده از تابع توزیع تی استودنت چوله هایپربولیک تعمیم یافته چندمتغیره نسخه مک نیل و همکاران (۲۰۰۵) به جای توزیع ترکیبی پیشنهادی مجدداً اجرا شده و نتایج آن با مدل تحقیق مقایسه شد. استفاده از این توزیع خاص برای مقایسه به دو دلیل انجام شده است، اولاً بر اساس نتایج تحقیق جامع الینگ، توزیع تی استودنت چوله در مقایسه با ۱۶ توزیع پرکاربرد دیگر در حوزه اکچوئری دارای مزیت رقابتی است، از این رو این توزیع یکی از برترین رقبای موجود برای آزمون مزیت مدل ارائه شده در تحقیق حاضر است. ثانیاً از آنجا که در این تحقیق نیز توزیع ارائه شده ترکیبی از توزیع تی استودنت چوله و پارتوی تعمیم یافته است، در عمل استفاده از این توزیع باعث خواهد شد که تنها تفاوت موجود میان مدل سازی های انجام شده، استفاده از توزیع پارتوی تعمیم یافته برای مدل سازی مقادیر فرین و تابع کاپیولا برای مدل سازی ساختار وابستگی باشد که در واقع همان نوآوری های تحقیق نسبت به کارهای موجود است.

نتایج شبیه سازی های انجام شده بر اساس هر دو مدل و مقایسه آنها با مقادیر خسارت واقعی نشان داد مدل تحقیق به مراتب دارای معیار میانگین مجذور خطای (RMSE) کمتری نسبت به مدل رقیب است. همچنین نکته بسیار مهم دیگر زمان لازم برای کالیبراسیون توزیع ها است.

زمان لازم برای انجام محاسبات با یک کامپیوتر با پردازنده Core i5 2.5GH و رم ۴ گیگابایت برای مدل ارائه شده در تحقیق ۲۳ دقیقه و ۱۸ ثانیه است، در حالی که این زمان برای مدل رقیب ۴ ساعت و ۱۶ دقیقه و ۲۵ بود. این اختلاف در سرعت انجام محاسبات به خصوص زمانی که به بازنگری در مفروضات و داده‌ها نیاز باشد، بسیار تأثیرگذار خواهد بود.

در مجموع از مقایسه بالا می‌توان نتیجه‌گیری کرد که بر اساس دو معیار دقت مدل‌سازی و سرعت انجام محاسبات مدل تحقیق نسبت به رقبای موجود دارای مزیت رقابتی است و نوآوری‌های انجام شده در آن می‌تواند به بهبود مدل‌سازی داده‌های خسارتی کمک کند.

نتیجه‌گیری و پیشنهادها

در این تحقیق به منظور مدل‌سازی زیان‌های وارد شده بر اکسپوژرهای مختلف تحت پوشش یک بیمه‌نامه خاص، تابع توزیع زیان‌های وارد شده بر هر ریسک به کمک یک توزیع ترکیبی مدل‌سازی شد. از آنجا که به منظور مدل‌سازی زیان کلی ناشی از یک بیمه‌نامه خاص، به مدل‌سازی تابع توزیع توأمان زیان‌ها نیاز است، در ادامه تحقیق سعی شد که این تابع توزیع به کمک مفهوم کاپیولا مدل‌سازی شود.

نتایج بررسی‌های انجام شده روی نمونه تحقیق نشان داد که اولاً به خوبی می‌توان زیان‌های وارد بر هر اکسپوژر را با استفاده از توزیع ترکیبی پیشنهادی مدل‌سازی کرد، ثانیاً کاپیولای کلایتون به خوبی می‌تواند تابع توزیع توأمان زیان‌ها را مدل‌سازی کند.

با توجه به اینکه نتایج تحقیق نشان می‌دهد مدل‌سازی انجام شده از تابع توزیع زیان از دقت بیشتری در مقایسه با رقبای موجود برخوردار است و همچنین زمان لازم برای این مدل‌سازی نسبت به رقبا بسیار کمتر است، از این رو به شرکت‌های بیمه توصیه می‌شود در راستای برآورد خسارت‌های ناشی از بیمه‌نامه‌های مختلف و به تبع آن تعیین حق بیمه مناسب برای آنها از این مدل استفاده کنند.

البته در این رابطه لازم است که اقدام‌های اساسی از سوی بیمه مرکزی جمهوری اسلامی ایران، به عنوان متولی امر بیمه در کشور انجام شود. در واقع نظام تعرفه‌ای موجود، که مشکلات عدیده آن یکی از عمده دلایل عدم رشد صنعت بیمه در کشور است، باعث می‌شود انگیزه چندانی برای استفاده از مدل‌های اکچوئری در شرکت‌های بیمه ایرانی وجود نداشته باشد. از این رو لازم است که اولاً بیمه مرکزی و شورای عالی بیمه هر چه سریع‌تر در نظام تعرفه‌ای بیمه بازنگری اساسی کنند و از سوی دیگر شرکت‌های بیمه را به استفاده از مدل‌های اکچوئری برای تخمین تابع زیان مکلف کنند.

از بعد پیشنهاد برای تحقیقاتی آتی لازم به توضیح است که در تحقیق حاضر ساختار وابستگی در سراسر توزیع زیان ها یکسان فرض شد، حال آنکه بر پایه برخی مطالعات، لزوماً ساختار وابستگی در سراسر توزیع یکسان نیست، از این رو محققان علاقه مند می توانند ساختار وابستگی را در مقادیر مرکزی توزیع و مقادیر حدی آن به صورت مجزا مدل سازی کنند و به تخمین های بهتری از تابع توزیع توأمان دست یابند.

فهرست منابع

- باجلان، س.، راعی، ر.، محمدی، ش. (۱۳۹۵). مدل سازی تابع زیان بیمه ای با استفاده از ترکیب توزیع تی استودنت چوله هایپربولیک تعمیم یافته و نظریه مقادیر فرین. *تحقیقات مالی*. ۱۸ (۱)، ۸۵-۳۹.
- Aas, K., Haff, I. (2006). The Generalized Hyperbolic Skew Student's t-Distribution. *Journal of Financial Econometrics* 4(2), 275-309.
- Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley & Sons, New York.
- Balkema, A. A., de Haan, L. (1974). Residual Life Time at Great Age. *Annals of Probability*, 2(5), 792-804.
- 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.
- Belguise, O., Levy, C. (2003). Tempêtes : Etude des dépendances entre les branches Automobile et Incendie à l'aide de la théorie des copulas Topic 1 Risk evaluation. 2003 *ASTIN Colloquium*, Berlin, Germany. Available in: http://www.actuaries.org/ASTIN/Colloquia/Berlin/Belguise_Levi.pdf.
- Bouyé, E. (2002). Multivariate Extremes at Work for Portfolio Risk Measurement. *Warwick Business School Working Paper Series*, WP01-10.
- Chava, S., Stefanescu, C., Turnbull, S. (2008). Modeling the Loss Distribution. *Working Paper*. Available in: http://faculty.london.edu/cstefanescu/Chava_Stefanescu_Turnbull.pdf.
- Cherubini, U., Luciano, E., Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons, Sussex, England.

- Debbie, J. D., Jones, B. L. (2006). Multivariate Extreme Value Theory and Its Usefulness in Understanding Risk. *North American Actuarial Journal*, 10(4), 1-27.
- 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., Resnick, S. I., Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal* 3(2), 30-41.
- Frees, E.W., Carrière, J.F., Valdez, E. (1996). Annuity Valuation with Dependent Mortality. *Journal of Risk and Insurance*, 63(2), 229-261.
- Frees, E.W., Valdez, E. A. (1998). Understanding Relationships Using Copulas. *North American Actuarial Journal* 2(1), 1-25.
- Johnson, N. L., Kotz, S., Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. John Wiley & Sons, New York.
- Johnson, R. A., Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis* (2 ed). Prentice-Hall Inc, New Jersey.
- Krzanowski, W. J. (1988). *Principles of Multivariate Analysis: a user's perspective*. Oxford University Press, Oxford.
- 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., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton Series in Finance, Princeton University Press. New Jersey.
- McNeil, A. J., Saladin, T. (1997). The Peaks Over Thresholds Method for Estimating High Quantiles of Loss Distributions. Preprint, Department Mathematik, ETH Zentrum, Zurich.
- Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer, New York.
- Pickands, J. I. (1975). Statistical Inference Using Extreme Value Order Statistics. *Annals of Statistics*, 3, 119-131.

۴۰ _____ مدل سازی تابع توزیع زیان های بیمه ای با بهره گیری از توزیع های

Venter, G.G. (2002). Tails of Copulas. *Proceedings of the Casualty Actuarial Society*, 89(1), 68–113.

Vernic, R. (2006). Multivariate Skew-Normal Distributions with Applications in Insurance. *Insurance: Mathematics and Economics*, 38(2), 413-426.

Archive of SID