تحقیقات مالی
دانشکدهٔ مدیریت دانشگاه تهران
دورهٔ ۱۹، شمارهٔ ۱
بهار ۱۳۹۶
صص. ۴۰- ۲۳

Journal of Financial Research
DOI: 10.22059/jfr.2017.38924.1005629

مدل‌سازی تابع توزیع زیان‌های بیمه‌ای با بهره‌گیری از توزیع‌های ترکیبی و مفهوم کاپیولا

سعید باجلان[۱]، رضا راعی[۲]، شاپور محمدی[۳]

چکیده: این تحقیق سعی دارد با بهره‌گیری همزمان از توزیع‌های ترکیبی و مفهوم کاپیولا، تابع توزیع توأمان زیان‌های واردشده بر اکسپوژرهای مختلف تحت پوشش یک بیمه‌نامهٔ خاص را نسبت به توزیع‌های آماری موجود، با دقت بیشتری مدل‌سازی کند. در این تحقیق از توزیع خاصی که ترکیبی از توزیع تی استودنت چوله هایپربولیک تعمیم‌یافته و نظریهٔ مقادیر فرین است، برای مدل‌سازی توابع زیان حاشیه‌ای و از مفهوم کاپیولا برای مدل‌سازی ساختار وابستگی میان آنها استفاده شده است. کاپیولاهای گوسی، تی، فرانک، گامبل و کلایتون، مهم‌ترین انواع کاپیولای بررسی شده‌اند تا از بین آنها بهترین گزینه برای تشریح ساختار وابستگی زیان‌ها انتخاب شود. داده‌های مورد استفاده در این تحقیق مقدار خسارت‌های جانی و مالی بیمه‌نامه‌های شخص ثالث وسایل نقلیهٔ موتوری است. نتایج تحقیق نشان می‌دهد با بهره‌گیری از توزیع ترکیبی پیشنهادی و کاپیولای کلایتون تابع توزیع توأم، می‌توان به خوبی زیان‌های نشئت گرفته از بیمه‌نامهٔ شخص ثالث را مدل‌سازی کرد.

واژه‌های کلیدی: توزیع ترکیبی، توزیع توأم، توزیع حاشیه‌ای، تابع کاپیولا.

۱. دانشجوی دکتری مدیریت مالی، دانشکدهٔ مدیریت دانشگاه تهران، تهران، ایران
۲. استاد مدیریت مالی، دانشکدهٔ مدیریت دانشگاه تهران، تهران، ایران
۳. دانشیار مدیریت مالی، دانشکدهٔ مدیریت دانشگاه تهران، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹٤/۰۹/۱۱
تاریخ پذیرش نهایی مقاله: ۱۳۹۵/۰٤/۰۳
نویسندهٔ مسئول مقاله: سعید باجلان
E-mail: Saeedbajalan@ut.ac.ir

Archive of SID
www.SID.ir

مقدمه

اغلب در انواع بیمه‌نامه‌ها چندین اکسپوژر[1] (مواد در معرض خسارت) تحت پوشش یک بیمه‌نامه قرار می‌گیرند که معمولاً خسارت‌های وارده به این اکسپوژرها مستقل از یکدیگر نیست. برای مثال بیمه‌های شخص ثالث اتومبیل معمولاً خسارت‌های بدنی و مالی را پوشش داده و بیمه‌های آتش‌سوزی، ساختمان[2] و محتویات[3] آن را پوشش می‌دهند.

از آنجا که میان این زیان‌ها وابستگی وجود دارد، لازم است هنگام مدل‌سازی زیان کلی ناشی از یک حوزهٔ خاص فعالیت به‌نوعی وابستگی میان این زیان‌ها مد نظر قرار گیرد تا بتوان به‌درستی حق بیمهٔ متناسب با هر ریسک را تعیین کرد. به بیان دیگر، برای برآورد صحیح تابع زیان، اولاً باید تابع توزیع زیان وارد بر هر اکسپوژر به‌نحو مطلوبی تخمین زده شود و ثانیاً باید تابع توزیع احتمال توأمان[4] زیان‌های وارد بر اکسپوژرها به‌طور صحیح مدل‌سازی شود. برای مثال، در صورتی که هدف مدل‌سازی زیان ناشی از بیمه‌نامهٔ آتش‌سوزی باشد، اولاً باید زیان‌های واردشده به ساختمان و محتویات آن به‌صورت جداگانه به‌دقت مدل‌سازی شوند، ثانیاً بتوان با استفاده از تابع احتمال توأمان مناسب، خسارت کلی ناشی از یک بیمه‌نامهٔ آتش‌سوزی را مدل‌سازی کرد.

در چنین حالت‌هایی معمولاً از تحلیل‌های چندمتغیره استفاده می‌شود. در خصوص بهره‌گیری از توزیع نرمال و توزیع‌های غیرنرمال چندمتغیره در راستای مدل‌سازی تابع احتمالات مشترک، ادبیات موضوعی گسترده‌ای وجود دارد (جانسون و ویچرن، ۱۹۸۸؛ کرزانوسکی، ۱۹۸۸؛ اندرسن، ۱۹۵۸؛ جانسون، کوتز و بالاکریشنان، ۱۹۹۷) با این حال به این گونه توزیع‌ها چند ایراد اساسی وارد است: ۱. برای هر تابع توزیع حاشیه‌ای[5]، خانوادهٔ متفاوتی لازم است، ۲. توسعه به حالت‌های بیش از دو متغیر چندان واضح نیست و ۳. نوع همبستگی متغیرها متأثر از توزیع حاشیه‌ای آنها است. ساخت توزیع چندمتغیره‌ای که از این ایرادها مبرا باشد، بر اساس مفهوم تابع کاپیولا[6] بنا شده است. در واقع کاپیولاها به ما کمک می‌کنند که تابع احتمال مشترک را مجزا از تابع توزیع حاشیه‌ای تخمین بزنیم.

بر همین اساس، در این تحقیق سعی شده که تابع توزیع توأم زیان‌های ناشی از یک نوع بیمه‌نامهٔ خاص به کمک مفهوم کاپیولا مدل‌سازی شود.

---

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

Archive of SID

www.SID.ir

پیشینهٔ تجربی

پیشینهٔ مطالعات انجام‌گرفته را می‌توان در دو بخش بررسی کرد، مدل‌سازی توابع زیان حاشیه‌ای و استفاده از کاپیولا برای مدل‌سازی ساختار وابستگی میان زیان‌ها.

تابع نرمال متداول‌ترین تابع استفاده‌شده برای مدل‌سازی توابع حاشیه‌ای در حوزهٔ اقتصاد و مالی است. اما به‌طور کلی توزیع زیان‌های بیمه‌ای، چوله[1] هستند، پدیده‌ای که باعث می‌شود توزیع نرمال، برای مدل‌سازی ریسک‌های بیمه‌ای، توزیع مناسبی نباشد (لین، ۲۰۰۰؛ ورنیک، ۲۰۰۶). اگرچه برای مدل‌سازی ریسک‌های بیمه‌ای در خصوص استفاده از سایر توزیع‌ها، کارهایی انجام شده، اما به‌طور عمده در این کارها سعی شده که از توزیع‌های بیضوی[2] 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 توزیع‌های بیضوی متقارن بودن آنها است، حال آن که تحقیقات متعدد نشان داده‌اند توزیع زیان‌های بیمه‌ای شکل متقارنی ندارد.

در خصوص استفاده از توزیع‌های نامتقارن برای تخمین تابع زیان کارهایی انجام شده که یکی از جامع‌ترین آنها کار الینگ (۲۰۱۲) است. در تحقیقی که الینگ (۲۰۱۲) روی داده‌های مربوط به شدت خسارت انجام داد، ۱۶ توزیع متداول در اکچوئری از جمله توزیع‌های 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[3] 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 به‌علت اهمیت وقایع فرین[4] در بیمه‌گری، لازم است دنباله‌های توزیع زیان با دقت بسیار بیشتری تخمین زده شوند؛ زیرا اتفاقاتی با احتمال بسیار کم اما شدت زیاد می‌تواند بقای یک شرکت بیمه را با چالش جدی مواجه کند. بحث سر این موضوع که آیا تئوری مقادیر فرین[5] می‌تواند از چنین خسارت‌هایی، تخمین‌های بهتری ارائه دهد، موضوع تحقیقات مختلفی بوده است که از آن جمله می‌توان به (باسی، امبرچتز و کافتزاکی، ۱۹۹۸؛ مک‌نیل 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)

Archive of SID

www.SID.ir

۲۰۰۸؛ لی و فانگ، ۲۰۱۰) اشاره کرد. عمدهٔ این تحقیقات نشان داده‌اند که این تئوری می‌تواند در تخمین دنباله‌های 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 تخمین‌های صحیح‌تر 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 تعیم‌یافته و تئوری مقادیر فرین مدل‌سازی بهتری از توزیع زیان انجام شود. نتایج تحقیق نشان داد که توزیع ترکیبی پیشنهادی به‌خوبی می‌تواند مقادیر مرکزی و همچنین مقادیر حدی زیان‌ها را مدل‌سازی کند (باجلان، راعی و محمدی، ۱۳۹۵).

از سوی دیگر استفاده از تابع کاپیولا به‌عنوان ابزاری برای مدل‌سازی ساختار وابستگی متغیرهای تصادفی بسیار فراگیر شده است. فریز،کاریر و والدز (۱۹۹۶) از کاپیولاها برای نشان دادن ساختار وابستگی زمان باقی‌مانده تا فوت در بیمه‌های عمر مشترک و آخرین بازمانده استفاده کردند. فریز و والدز (۱۹۹۸) مقدمه‌ای بر کاپیولاها ارائه کردند که مرور مفیدی بر ادبیات موضوع کاپیولاهاست. آنها در مقالهٔ خود به جست‌وجوی رابطه‌ای میان خسارت‌های ناشی از بیمهٔ مسئولیت و هزینهٔ ارزیابی خسارت تخصیص‌یافته[1] با استفاده از کاپیولاها پرداختند. بوی (۲۰۰۲) در تحقیق خود روی انتخاب کاپیولا و تخمین آن در حیطهٔ تحلیل ریسک پرتفوی دارایی‌ها تمرکز کرد. ونتر (۲۰۰۲) مباحثی را در خصوص چندین خانوادهٔ پارامتریک کاپیولا به‌همراه روش‌هایی برای انتخاب گزینهٔ مناسب از میان آنها برای مدل‌سازی وابستگی ارائه کرد. وی همچنین در خصوص شبیه‌سازی خسارت‌های ناشی از گردبادها در دو ایالت آمریکا، مثالی مطرح کرد. بلقیوس و لوی (۲۰۰۳) با استفاده از کاپیولاها، سعی در مدل‌سازی وابستگی میان خسارت ناشی از بیمهٔ اتومبیل و آتش‌سوزی کردند. دیبی و جونز (۲۰۰۶) در مقاله‌ای نتایج به‌دست آمده از تحلیل‌های انجام‌شده روی مقادیر فرین چندمتغیره را گردآوری کرده و کاربردهای آنها را در اکچوئری تشریح کردند. آنها در مقالهٔ خود به‌طور خاص روی بحث وابستگی در مقادیر فرین تمرکز کرده و بر اهمیت انتخاب کاپیولای صحیح هنگام کار کردن با دنبالهٔ سمت راست تأکید کردند.

از آنجا که به‌منظور عملکرد صحیح شرکت‌های بیمه، مهم‌ترین نیاز داشتن تخمین‌های دقیق‌تر 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 بهره‌گیری همزمان از توزیع‌های ترکیبی و مفهوم کاپیولا، مدل‌سازی دقیق‌تری از تابع توزیع زیان نسبت به مدل‌سازی‌های موجود انجام شود. در ادامه این دو مفهوم به‌صورت مختصر تشریح می‌شوند.

---

1. Allocated Loss Adjustment Expenses (ALAE)

Archive of SID

www.SID.ir

توزیع ترکیبی[1]

در پژوهشی که باجلان و همکارانش (۱۳۹۵) انجام دادند، به‌منظور مدل‌سازی زیان‌های بیمه‌ای از یک توزیع ترکیبی که با استفاده از توزیع تی استودنت چوله هایپربولیک تعمیم‌یافته و نظریهٔ مقادیر فرین ساخته شده بود، استفاده شد. علت استفاده از این توزیع ترکیبی نیاز به ساخت توزیعی بود که به‌خوبی بتواند مقادیر مرکزی توزیع و مقادیر حدی را مدل‌سازی کند، مقادیری که علی‌رغم احتمال پایین رخ دادن در صورت وقوع می‌توانند نتایج ویران‌کننده‌ای به‌دنبال داشته باشند و حتی بقای یک شرکت بیمه را به مخاطره بیندازند.

آنها در تحقیق خود بر پایهٔ مقالهٔ آس و هاف (۲۰۰۶) توزیع تی استودنت چوله را به‌صورت حالت خاصی از توزیع هایپربولیک تعمیم‌یافته، تعریف کردند. علت بهره‌گیری از این تعریف خاص داشتن یک دنباله با رفتار چندجمله‌ای و یک دنباله به‌صورت نمایی در این توزیع است که برای مدل‌سازی داده‌هایی با دنبالهٔ پهن و چولگی زیاد منحصربه‌فرد محسوب می‌شود.

تابع چگالی این توزیع به‌صورت زیر است:

رابطهٔ ۱)

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

که در آن پارامترهای μ و δ به‌ترتیب موقعیت[2] و مقیاس[3] را تعیین کرده و α و β شکل[4] توزیع را مشخص می‌کنند. $K_\lambda(x)$ که $x>0$ معرف تابع بسل تعدیل‌یافتهٔ[5] نوع دوم با رتبهٔ λ است[6]. Γ معرف تابع گاماست و υ تعداد درجات آزادی را نشان می‌دهد. این توزیع برای مدل‌سازی مقادیر مرکزی و دنبالهٔ سمت چپ توزیع زیان‌ها به‌کار گرفته شد.

به‌منظور مدل‌سازی مقادیر حدی نیز بر اساس رویکرد اوج فراتر از آستانه (POT) به‌عنوان یکی از رویکردهای اساسی تئوری مقادیر فرین (EVT) از توزیع پارتوی تعمیم‌یافته (GPD) استفاده شد. بر اساس این رویکرد بر پایهٔ قضیه بالکما، دی‌هان (۱۹۷۴) و پیکاندز (۱۹۷۵)

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

۶. شایان ذکر است که در متون قدیمی‌تر به آن نوع ۳ گفته می‌شود.

Archive of SID

www.SID.ir

به‌عنوان دومین قضیهٔ اصلی تئوری مقادیر فرین، تحت شرط $F \in MDA(H)$، توزیع پارتوی تعمیم‌یافتهٔ توزیع حدی برای توزیع مقادیر بیش از یک آستانه به اندازهٔ کافی بزرگ u است.

آنها در تحقیق خود برای تخمین آستانهٔ u از تابع میانگین مازاد و برای کالیبراسیون توزیع تی استونت چوله هایپربولیک تعمیم‌یافته از الگوریتم حداکثرسازی انتظارات (EM) استفاده کردند. پارامترهای توزیع GPD نیز با بهره‌گیری از روش حداکثر درست‌نمایی (MLL) تخمین زده شدند. به‌منظور مطالعهٔ بیشتر در این خصوص می‌توان به تحقیق اصلی مراجعه کرد.

کاپیولا

هر تابع توزیع توأم برداری تصادفی از عوامل ریسک، به‌طور ضمنی شامل شرحی از رفتار حاشیه‌ای عوامل ریسک به‌طور منفرد و شرحی از ساختار وابستگی بین آنها است. رویکرد کاپیولا راهی برای تفکیک ساختار وابستگی از رفتار حاشیه‌ای عوامل فراهم می‌آورد.

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

$$C: [0\,,1]^d \to [0,1] \qquad \text{رابطهٔ ۲)}$$

قضیهٔ اسکلار (۱۹۵۹)[1]: اگر F یک تابع توزیع توأمان با توابع حاشیه‌ای $F_1, \dots, F_d$ باشد، یک کاپیولا $C: [0\,,1]^d \to [0,1]$ وجود دارد به‌گونه‌ای که برای همهٔ $x_1, \dots, x_d$ عضو $\mathcal{R} = [-\infty, \infty]$:

$$F(x_1, \dots, x_d) = C\big(F_1(x_1), \dots, Fd(x_d)\big) \qquad \text{رابطهٔ ۳)}$$

اگر توابع حاشیه‌ای پیوسته باشند، C منحصربه‌فرد است[2] و برعکس اگر C یک کاپیولا و $F_1, \dots, F_d$ توابع توزیع تک‌متغیره باشد، تابع F که در رابطهٔ ۳ تعریف شده است، یک تابع توزیع توأمان با توابع حاشیه‌ای $F_1, \dots, F_d$ است. اگر داشته باشیم $x_i = F_i^{\leftarrow}(u_i)$ و $0 \le u_i \le 1$ و $i = 1,\dots,d$ می‌توان رابطهٔ ۳ را به‌صورت زیر نوشت:

---

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

Archive of SID

www.SID.ir

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

رابطهٔ ۴، ارائهٔ صریحی از C بر حسب تابع توزیع توأم F و توابع حاشیه‌ای آن می‌دهد، بنابراین بیان‌کنندهٔ منحصربه‌فرد 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 رابطهٔ 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[1] که معرف ساختارهای وابستگی مهم خاص (نظیر استقلال، وابستگی کامل مثبت و وابستگی کامل منفی) هستند. کاپیولاهای ضمنی[2] که از توزیع‌های چندمتغیرهٔ معروف با استفاده از قضیهٔ اسکلار استخراج شده‌اند، اما لزوماً دارای تصریح ساده‌ای[3] نیستند، نظیر خانوادهٔ کاپیولاهای گوسی. کاپیولاهای صریح[4] که دارای تصریح ساده‌ای هستند و با استفاده از اصول عمومی ریاضی ساخته شده‌اند، نظیر کاپیولاهای ارشمیدسی (مک‌نیل، فری و امبرچتز، ۲۰۰۵).

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

روش‌شناسی پژوهش

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

۱. مدل‌سازی تابع توزیع هر یک از ریسک‌های تحت پوشش بیمه‌نامه توسط توزیع ترکیبی. بر اساس این توزیع مقادیر مرکزی و دنبالهٔ سمت چپ توزیع با استفاده از توزیع تی استودنت چوله و مقادیر دنبالهٔ 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

Archive of SID

www.SID.ir

۲. مدل‌سازی همبستگی میان ریسک‌های تحت پوشش بیمه‌نامه: بعد از اینکه توابع زیان تک تک خسارت‌های تحت پوشش یک بیمه‌نامه مدل‌سازی شد، لازم است ساختار همبستگی میان آنها مدل‌سازی شود. همان‌گونه که در قسمت مربوط به کاپیولاها توضیح داده شد، کاپیولاها اجازه می‌دهند که ساختار وابستگی مستقل از توزیع‌های حاشیه‌ای مدل‌سازی شود و همچنین بسیاری از معایب ضرایب همبستگی را ندارند. بر این اساس بعد از تخمین پارامترهای توزیع‌های حاشیه‌ای، با استفاده از روش استنباط برای حاشیه‌ها (IFM) پنج خانواده از کاپیولاها شامل گوسی، تی، گامبل، کلایتون و فرانک، که بیشترین کاربرد را در حوزهٔ اکچوئری دارند، تخمین زده می‌شوند و از میان آنها کاپیولایی که دارای بیشترین مقدار لگاریتم درستنمایی باشد، به‌عنوان کاپیولای برتر انتخاب می‌شود.

۳. در گام سوم با ترکیب کردن نتایج مراحل ۱ و ۲ مدل‌سازی نهایی از تابع توزیع زیان کلی تحت پوشش یک بیمه‌نامه انجام می‌گیرد.

جامعهٔ آماری

جامعهٔ آماری تحقیق داده‌های خسارتی کلیهٔ شرکت‌های بیمهٔ ایرانی است. از آنجا که در مباحث اکچوئری هر مشاهده‌ای حاوی اطلاعات بسیار ارزشمندی در خصوص شکل توزیع است که در صورت نمونه‌گیری، احتمال از دست رفتن آن وجود دارد (به‌دلیل حذف شدن مشاهده)، در تحقیقات این حوزه عملاً بحث نمونه‌گیری مطرح نیست و باید سعی شود مدل‌ها با استفاد از حداکثر اطلاعات موجود تخمین زده شوند. از آنجا که در این تحقیق ساختار وابستگی میان زیان‌های تحت پوشش یک بیمه‌نامه نیز مدل‌سازی خواهد شد، لازم است مشاهدات گردآوری شده به‌صورت زوجی باشند. بررسی‌های انجام‌شده نشان داد با توجه به ضعف سیستم‌های اطلاعاتی صنعت بیمهٔ کشور، گردآوری داده‌های خسارتی برای اکثر رشته‌ها به تفکیک امکان‌پذیر نیست و صرفاً می‌توان داده‌های مربوط به خسارت جانی و مالی ناشی از رشتهٔ شخص ثالث را گردآوری و به‌نوعی تفکیک کرد. به همین منظور در این تحقیق روی مدل‌سازی زیان‌های ناشی از رشتهٔ شخص ثالث تمرکز شده است.

به‌منظور گردآوری و پایش داده‌ها، در ابتدا داده‌های خسارتی مربوط به بیمه‌نامه‌های شخص ثالث صادرشده توسط کلیهٔ شرکت‌های بیمه که اطلاعات آنها در سامانهٔ سنهاب موجود است، گردآوری شد. این اطلاعات شامل ۱۴۹۳۰۵ مشاهده است که هر یک شامل شمارهٔ بیمه‌نامه، تاریخ صدور بیمه‌نامه، مبلغ خسارت و نوع خسارت (جانی و مالی) در بازهٔ زمانی فرودین ۱۳۸۴ تا مرداد ۱۳۹۲ است.

Archive of SID

www.SID.ir

در گام بعد، از آنجا که در مواردی به‌ازای یک شماره بیمه‌نامهٔ مشخص، بیش از یک خسارت رخ داده بود، کلیهٔ این خسارت‌ها تجمیع شدند تا مشخص شود به‌ازای هر بیمه‌نامهٔ صادرشده، چه مقدار خسارت جانی و خسارت مالی پرداخت شده است. بعد از انجام 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 و قیمت خودرو، لازم بود که میزان خسارت‌های جانی و مالی مربوط به سال‌های قبل از ۱۳۹۲ تعدیل شوند. به‌منظور انجام این تعدیلات ابتدا مبلغ دیه و شاخص هزینهٔ حمل‌ونقل استخراج شد؛ سپس نرخ رشد آنها به‌دست آمد. از نرخ رشد محاسبه‌شده به‌عنوان عامل تعدیل خسارت‌های جانی و مالی در سال‌های قبل از ۱۳۹۲ استفاده شد.

یافته‌های پژوهش

مدل‌سازی خسارت‌های مالی و جانی با استفاده از توزیع ترکیبی

تعیین آستانهٔ u: برای تعیین آستانهٔ u به‌منظور مشخص‌کردن مشاهدات حدی که دارای توزیع GPD هستند (بر اساس دومین قضیهٔ بنیادی تئوری مقادیر فرین) با توجه به توضیحات ارائه‌شده در تحقیق باجلان و همکاران (۱۳۹۵)، از تابع میانگین مازاد استفاده شده است. بر این اساس آستانهٔ u برای لگاریتم خسارت‌های مالی برابر ۱۷/۲ و برای لگاریتم خسارت‌های جانی ۲۱/۴ انتخاب شد. مشاهدات فراتر از این آستانه‌ها به‌عنوان مشاهدات حدی در نظر گرفته خواهد شد که با استفاده از توزیع GPD مدل‌سازی می‌شوند و مقادیر کمتر از آنها به‌عنوان مقادیر مرکزی و دنبالهٔ سمت چپ در نظر گرفته می‌شود که با استفاده از توزیع تی استودنت چوله مدل‌سازی خواهند شد.

مدل‌سازی مقادیر حدی

پارامترهای توزیع GPD برازش‌شده روی مشاهدات حدی با استفاده از روش حداکثر درستنمایی در جدول ۱ ارائه شده است.

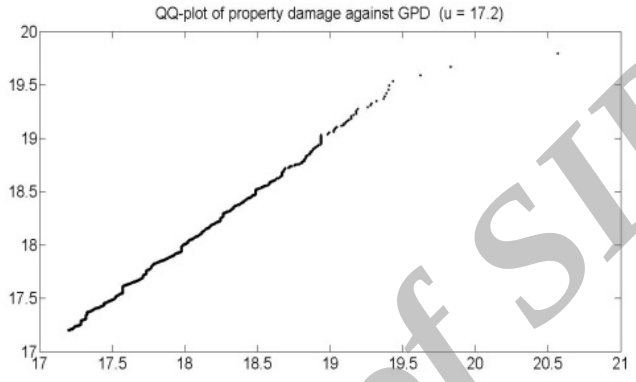
جدول ۱. پارامترهای توزیع GPD برازش‌شده روی لگاریتم مشاهدات حدی خسارت‌های جانی و مالی

پارامتر تخمین‌زده‌شده	لگاریتم خسارت‌های مالی	لگاریتم خسارت‌های جانی
μ	۱۷/۲	۲۱/۴۲
ξ	-۰/۰۹۲۵	-۰/۱۸۵۹
σ	۰/۴۱۴۲	۰/۶۶۷۳

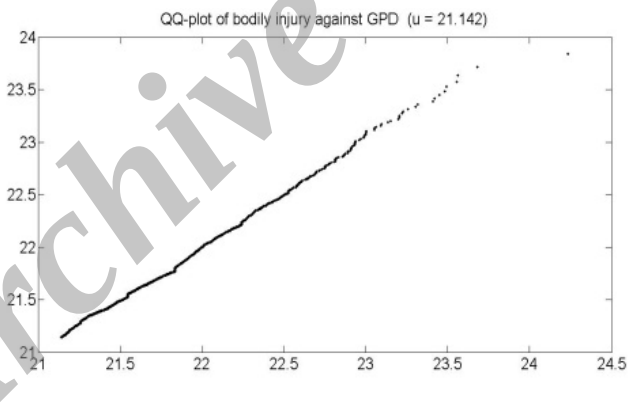
Archive of SID

www.SID.ir

به‌منظور بررسی نیکویی برازش‌های انجام‌شدهٔ نمودار کوانتایل ـ کوانتایل[1]، مقادیر واقعی در مقابل مقادیر تخمین‌زده‌شده توسط GPD در شکل ۱ ترسیم شده است. طبق قضیهٔ گلیونکو ـ کانتلی[2]، شباهت هر چه بیشتر این شکل با یک خط راست با زاویهٔ ۴۵ درجه که از مبدأ عبور می‌کند، معرف بهتر بودن برازش انجام‌شده است.



لگاریتم خسارت‌های مالی



لگاریتم خسارت‌های جانی

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

شباهت بسیار زیاد نمودارهای کوانتایل ـ کوانتایل ترسیم‌شده با خط $y = x$ نشان‌دهندهٔ مطلوب بودن برازش‌های انجام‌شده روی مشاهدات حدی است.

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

Archive of SID

www.SID.ir

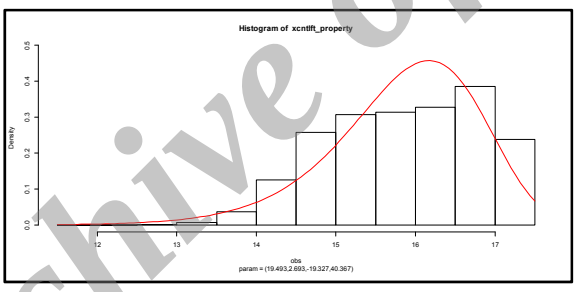
مدل‌سازی مقادیر مرکزی و دنبالهٔ سمت چپ

نتایج کالیبراسیون توزیع تی استودنت چوله هایپربولیک تعمیم‌یافته روی مقادیر مرکزی و دنبالهٔ سمت چپ با استفاده از الگوریتم حداکثرسازی انتظارات در جدول ۲ ارائه شده است.

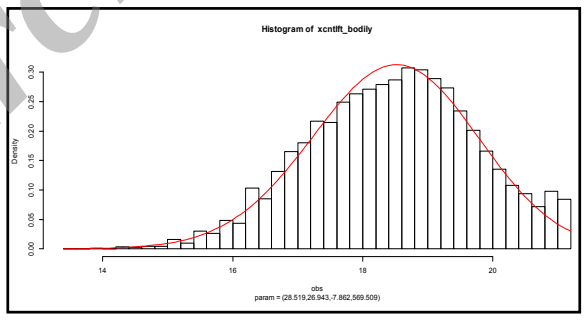
جدول۲. نتایج کالیبراسیون توزیع تی استودنت چوله روی لگاریتم خسارت‌ها

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	−۳۴۴۴۲/۱۴	−۵۹۹۳۵/۸۹
Iterations	۱۰۰۱	۶۵۹

هیستوگرام مقادیر واقع در قسمت مرکزی و دنبالهٔ سمت چپ توزیع لگاریتم خسارت‌های مالی و جانی در شکل ۲ ترسیم شده و توزیع تی چولهٔ برازش‌شده نیز با آن مقایسه شده است.



خسارت‌های مالی



خسارت‌های جانی

شکل ۲. هیستوگرام مقادیر واقع در قسمت مرکزی و دنبالهٔ سمت چپ توزیع لگاریتم خسارت در مقایسه با تابع تی استودنت چولهٔ برازش‌شده روی آنها

Archive of SID

www.SID.ir

همان‌گونه که در شکل ۲ مشاهده می‌شود، توزیع تی استودنت چوله به‌خوبی توانسته مقادیر مرکزی و دنبالهٔ سمت چپ توزیع لگاریتم خسارت‌های جانی و مالی را مدل‌سازی کند، هر چند عملکرد آن در خصوص لگاریتم خسارت‌های جانی به‌مراتب بهتر بوده است.

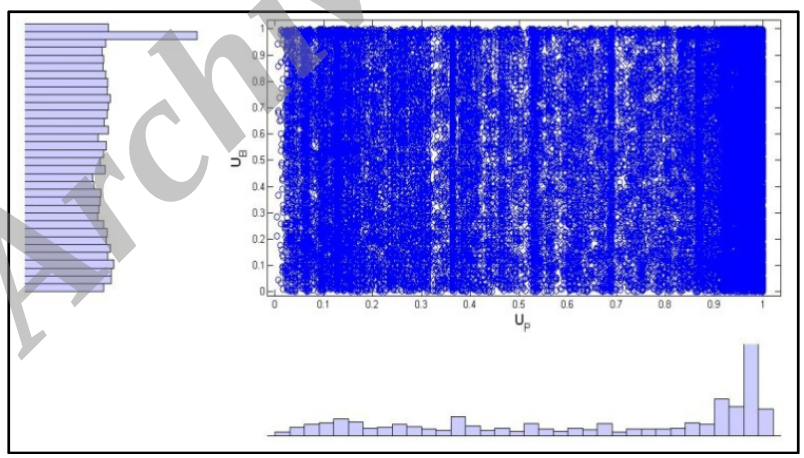
مدل‌سازی ساختار وابستگی خسارت‌های مالی و جانی

بعد از مدل‌سازی توزیع لگاریتم خسارت‌های مالی و جانی به‌طور مجزا، باید ساختار وابستگی میان آنها مدل‌سازی شود.

به‌منظور محاسبهٔ $F_1(\,)$ و $F_2(\,)$ (تابع توزیع لگاریتم خسارت‌های مالی و لگاریتم خسارت‌های جانی) به‌ازای مقادیر مختلف x_{1t} و x_{2t} به این صورت عمل شده که برای مشاهدات واقع در دنبالهٔ سمت چپ و ناحیهٔ مرکزی توزیع از تابع چگالی احتمال توزیع تی استودنت چوله با استفاده از روش‌های عددی انتگرال‌گیری‌شده و برای مشاهداتی که بالاتر از آستانهٔ u قرار دارند، از رابطهٔ زیر استفاده شده است:

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

شکل ۳ ساختار پراکنش میان مقادیر تابع توزیع لگاریتم خسارت‌های مالی و خسارت‌های جانی را نشان می‌دهد.



شکل ۳. پراکنش مقادیر تابع توزیع لگاریتم خسارت‌های مالی و خسارت‌های جانی

Archive of SID

www.SID.ir

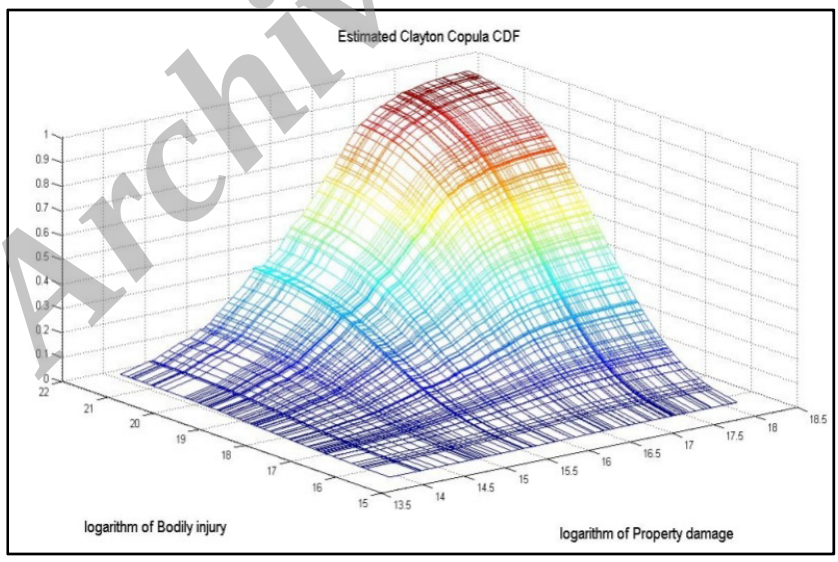
بعد از تخمین کاپیولاهای مختلف، بیشینهٔ مقدار تابع درستنمایی آنها محاسبه شده است که نتیجهٔ آن را می‌توان در جدول ۳ مشاهده کرد.

جدول۳. مقادیر حداکثر لگاریتم درست‌نمایی کاپیولاهای بررسی‌شده

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 تخمین‌زده‌شده برای این کاپیولا برابر ۰/۰۲۶۷ است.

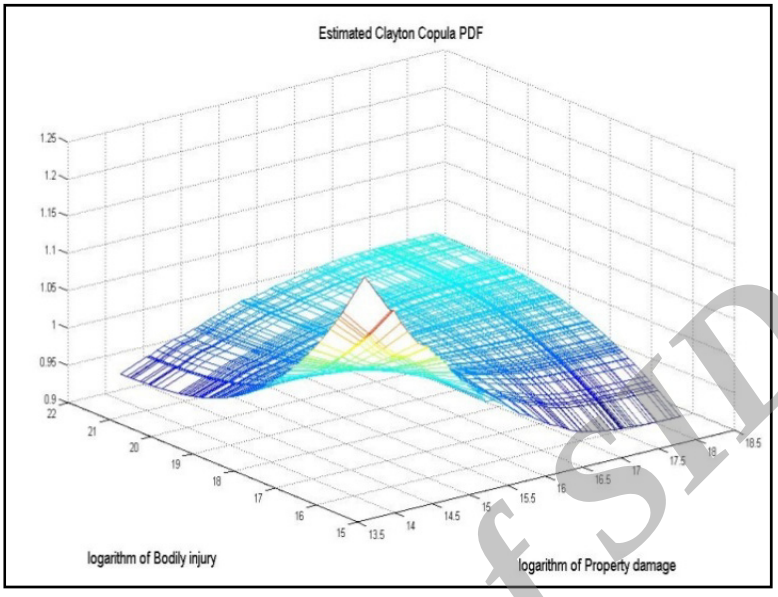
در شکل ۴ تابع توزیع احتمال توأمان لگاریتم خسارت‌های مالی و جانی به‌همراه تابع چگالی توأمان آنها با توجه به کاپیولای تخمین‌زده‌شده محاسبه و نمایش داده شده است.



شکل ٤- الف. تابع توزیع احتمال توأمان لگاریتم خسارت‌های مالی و جانی

Archive of SID

www.SID.ir



شکل ٤- ب. تابع چگالی احتمال توأمان خسارت‌های مالی و جانی
$Clayton\ Copula, \alpha = 0.0267$

مقایسهٔ مدل تحقیق با مدل رقیب

به‌منظور مقایسهٔ عملکرد مدل ارائه‌شده در تحقیق با رقبای موجود، مراحل تحقیق با استفاده از تابع توزیع تی استودنت چوله هایپربولیک تعمیم‌یافتهٔ چندمتغیرهٔ نسخهٔ مک نیل و همکاران (۲۰۰۵) به‌جای 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 استونت چوله در مقایسه با ۱۶ توزیع پرکاربرد دیگر در حوزهٔ 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 ارائه‌شدهٔ ترکیبی از توزیع تی استودنت چوله و پارتوی تعمیم‌یافته است، در عمل استفاده از این توزیع باعث خواهد شد که تنها تفاوت موجود میان مدل‌سازی‌های انجام‌شده، استفاده از توزیع پارتوی تعمیم‌یافته برای مدل‌سازی مقادیر فرین و تابع کاپیولا برای مدل‌سازی ساختار وابستگی باشد که در واقع همان نوآوری‌های تحقیق نسبت به کارهای موجود است.

نتایج شبیه‌سازی‌های انجام‌شده بر اساس هر دو مدل و مقایسهٔ آنها با مقادیر خسارت واقعی نشان داد مدل تحقیق به‌مراتب دارای معیار میانگین مجذور خطای (RMSE) کمتری نسبت به مدل رقیب است. همچنین نکتهٔ بسیار مهم دیگر زمان لازم برای کالیبراسیون توزیع‌ها است.

Archive of SID

www.SID.ir

زمان لازم برای انجام محاسبات با یک کامپیوتر با پردازندهٔ Core i5 2.5GH و رم ۴ گیگابایت برای مدل ارائه‌شده 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 به‌خصوص زمانی که به بازنگری در مفروضات و داده‌ها نیاز باشد، بسیار تأثیرگذار خواهد بود.

در مجموع از مقایسهٔ 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 انجام‌شده در آن می‌تواند به بهبود مدل‌سازی داده‌های خسارتی کمک کند.

نتیجه‌گیری و پیشنهادها

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

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

با توجه به اینکه نتایج تحقیق نشان می‌دهد مدل‌سازی انجام‌شده 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 به‌تبع آن تعین حق بیمهٔ مناسب برای آنها از این مدل استفاده کنند.

البته در این رابطه لازم است که اقدام‌های اساسی از سوی بیمهٔ مرکزی جمهوری اسلامی ایران، به عنوان متولی امر بیمه در کشور انجام شود. در واقع نظام تعرفه‌ای موجود، که مشکلات عدیدهٔ آن یکی از عمده دلایل عدم رشد صنعت بیمه در کشور است، باعث می‌شود انگیزهٔ چندانی برای استفاده از مدل‌های اکچوئری در شرکت‌های بیمهٔ ایرانی وجود نداشته باشد. از این رو لازم است که اولاً بیمهٔ مرکزی و شورای‌عالی 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.

Archive of SID

www.SID.ir

از بعد پیشنهاد برای تحقیقاتی آتی لازم به توضیح است که در تحقیق حاضر ساختار وابستگی در سراسر توزیع زیان‌ها یکسان فرض شد، حال آنکه بر پایهٔ برخی مطالعات، لزوماً ساختار وابستگی در سراسر توزیع یکسان نیست، از این رو محققان علاقه‌مند می‌توانند ساختار وابستگی را در مقادیر مرکزی توزیع و مقادیر حدی آن به‌صورت مجزا مدل‌سازی کنند و به تخمین‌های بهتری از تابع توزیع توأمان دست یابند.

فهرست منابع

باجلان، س.، راعی، ر.، محمدی، ش. (۱۳۹۵). مدل‌سازی تابع زیان بیمه‌ای با استفاده از ترکیب توزیع تی استودنت چوله هایپربولیک تعمیم‌یافته و نظریهٔ 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.

Bouyè, 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., Lucianco, E., Vecchiato, W. (2004). *Copula Methods in Finance.* John Wiley & Sons, Sussex, England.

Archive of SID

www.SID.ir

Debbie, J, D., Jones, B, L. (2006). Multivariate Extreme Value Theory and Its Usefulness in Understanding Risk. *North American Acutarial Jouranl,* 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 Distribution*s. 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 perspectiv*e. 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 Statististics,* 3,119-131.

Archive of SID

www.SID.ir

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

www.SID.ir