بسم الله الرحمن الرحیم

دانشکده علوم ریاضی

رشته ریاضی کاربردی، گرایش ریاضی مالی

پایان‌نامه کارشناسی ارشد

توزیع پارتو وابسته چند متغیره در ریسک تجمعی

نگارنده: سیده مونا میرحجتی

استادان راهنما

دکتر احمد نزاکتی
دکتر سید مجتبی میرلوحی

تیر ۱۳۹۷

ماحصل آموخته‌هایم را تقدیم می کنم به آنان که مهر آسمانی شان آرام بخش آلام زمینی ام است

به استوارترین تکیه گاهم، دستان پر مهر پدرم

به سبزترین نگاه زندگیم، چشمان مادرم

که هرچه آموختم در مکتب عشق شما آموختم و هرچه بکوشم قطره‌ای از دریای بی کران مهربانیتان را سپاس نتوانم بگویم.

امروز هستی ام به امید شماست و فردا کلید باغ بهشتم رضای شما

ره آوردی گران سنگ تر ازاین ارزان نداشتم تا به خاک پایتان نثار کنم، باشد که حاصل تلاشم نسیم گونه غبار خستگیتان را بزداید.

بوسه بر دستان پر مهرتان

سپاس گزاری...

سپاس خداوندی را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت های او ندانند و کوشندگان، حق او را گزاردن نتوانند. و سلام و درود بر محمد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است؛ و نفرین پیوسته بر دشمنان ایشان تا روز رستاخیز...

بدون شک جایگاه و منزلت معلم، اجل از آن است که در مقام قدردانی از زحمات بی شائبه ی او، با زبان قاصر و دست ناتوان، چیزی بنگارم.

اما از آنجایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می کند و سلامت امانت هایی را که به دستش سپرده اند، تضمین؛ بر حسب وظیفه و از باب "من لم یشکر المنعم من المخلوقین لم یشکر الله عزوجل" از پدر و مادر عزیزم، این دو معلم بزرگوارم که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاوری بی چشم داشت برای من بوده اند؛

از اساتید باکمالات و شایسته؛ جناب آقایان دکتر احمد نزاکتی و دکتر سید مجتبی میرلوحی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان نامه را به عهده گرفتند؛

سیده مونا میرحججتی
تیر ۱۳۹۷

تعهد نامه

اینجانب **سیده مونا میرحجتی** دانشجوی کارشناسی ارشد رشته **ریاضی کاربردی علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **توزیع پارتو وابسته چند متغیره در ریسک تجمعی** ، تحت راهنمایی **احمد نزاکتی** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچجا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام ‘‘ دانشگاه صنعتی شاهرود ’’ یا “ Shahrood University of Technology “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

سیده مونا میرحجتی
تیر ۱۳۹۷

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

در این پایان نامه معادله‌ای برای تابع توزیع احتمالی تجمع ریسک‌ها با توزیع وابسته چند متغیری پارتو به دست می‌آوریم. با توزیع وابسته چندمتغیری پارتو نوع دوم کار می‌کنیم که کاربرد فراوانی در بیمه و تحلیل ریسک‌ها دارد. با ارائه یک مدل آغاز می‌کنیم که ساختار آن بر اساس وابستگی است و تابع چگالی احتمال آن با توزیع بتای نوع دوم انطباق دارد. ارزش در معرض ریسک و چندین شاخص سنجش ریسک را به دست می آوریم.

به بررسی جزئیات دقیق برخی از مدل‌های تجمعی می‌پردازیم که توزیعات پواسون، دو جمله‌ای منفی و هندسی را به عنوان توزیع اصلی داراست. در مدل تجمعی پارتو– پواسون، تابع چگالی احتمال تابعی از توزیع فوق هندسی کامر است و تابع چگالی احتمال پارتو– دوجمله‌ای منفی تابعی از توزیع فوق هندسی گاؤس می‌باشد. با بررسی‌هایی که روی داده‌هایی که براساس بیمه نامه های یک ساله در سال (۲۰۰۵ – ۲۰۰۴) و داده‌های شرکت بیمه ایران در سال ۹۳ انجام شد، نتیجه می گیریم که مدل های وابسته ی جمعی که با توزیع پارتو ترکیب شده‌اند نسبت به سایر مدل های جمعی بهتر و کاربردی‌تر هستند.

کلمات کلیدی:

ریسک وابسته، مدل ریسک فردی، مدل ریسک تجمعی، ساختار پارتو کلاسیک، توزیع فوق هندسی

فهرست مطالب

فهرست تصاویر

فهرست جداول

فصل ۱

مفاهیم اولیه

در این فصل برای درک و فهم بهتر پایان نامه، مختصراً به بیان برخی مفاهیم و اصطلاحات در زمینه آمار و مالی اقتصادی اولیه می‌پردازیم.

۱.۱ مروری بر برخی از مفاهیم آمار ریاضی

تعریف ۱.۱.۱. [۱] **متغیر تصادفی** : اگر S یک فضای نمونه‌ای با یک اندازه‌ی احتمال، و X یک تابع حقیقی مقدار باشد که روی عناصر S تعریف شده است، آنگاه X متغیر تصادفی نامیده می‌شود.

تعریف ۲.۱.۱. [۱] **تابع چگالی احتمال:** تابعی با مقادیر $f(x)$، که روی مجموعه‌ی تمام اعداد حقیقی تعریف شده‌است تابع چگالی احتمال یک متغیر تصادفی پیوسته‌ی X خوانده می‌شود اگر و تنها اگر، به ازای هر دو مقدار حقیقی ثابت a و b با $a \leq b$.

$$P_r(a \leq X \leq b) = \int_a^b f(x)dx$$

تعریف ۳.۱.۱. **تابع جرم احتمال** فرض کنید $X : S \to A$ که($A \subseteq R$) یک متغیر تصادفی گسسته تعریف شده در یک یک فضای نمونه S است. تابع جرم احتمال $f(x) : A \to [۰,۱]$ برای x اینگونه تعریف می‌شود:

$$f_X(x) = P_r(X = x) = P_r(s \epsilon S : X(s) = x)$$

تعریف ۴.۱.۱. [۱] **تابع توزیع تجمعی:** در مسائل زیادی، دانستن احتمال اینکه مقداری از متغیر تصادفی کوچکتر از یک مقدار حقیقی x یا برابر با آن باشد مورد توجه است، لذا احتمال این را که X مقداری کوچکتر از x یا برابر آن اختیار کند به صورت $F_X(x) = P_r[X \leq x]$ می‌نویسیم و این تابع را برای تمام اعداد حقیقی x تعریف شده است، تابع توزیع یا توزیع تجمعی متغیر تصادفیX می‌نامیم.

اگر X یک متغیر تصادفی گسسته باشد، تابعی که با

$$F_X(x) = P_r[X \leq x] = \sum_{t \leq x} P(X = t) \qquad -\infty < x < +\infty$$

داده می‌شود.

مقادیر $F_X(x)$، تابع توزیع متغیر تصادفی گسسته‌ی X، در شرایط زیر صدق می‌کنند:

•

$$F_X(-\infty) = ۰, F_X(+\infty) = ۱$$

• تابع توزیع تجمعی غیر نزولی است یعنی:

$$x_۱ \leq x_۲ \longrightarrow F_X(x_۱) \leq F_X(x_۲)$$

• تابع توزیع تجمعی از راست پیوسته است.

تعریف ۵.۱.۱. [۱] در نظریه احتمالات **امید ریاضی** [۱] ، میانگین، مقدار مورد انتظار یا ارزش مورد انتظار یک متغیر تصادفی گسسته برابر است با مجموع حاصل‌ضرب احتمال وقوع هر یک از حالات ممکن در مقدار آن حالت.

$$E(X) = \sum_{i=۱}^{\infty} x_i P_X(x_i)$$

امید ریاضی یک متغیر تصادفی پیوسته به صورت زیر تعریف می‌شود:

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) \mathrm{d}x$$

که در آن $f_X(x)$, تابع چگالی احتمال متغیر تصادفی X است.

• امید ریاضی یک عدد ثابت برابر با همان عدد ثابت است؛ یعنی اگر c عددی ثابت باشد، آن‌گاه $E(c) = c$

• امید ریاضی یک عملگر خطی است. یعنی برای هر دو متغیر تصادفی X و Y و هر عدد حقیقی a و b داریم:

$$E(aX + b) = aE(X) + b \qquad (۱.۱)$$

$$E(aX + bY) = aE(X) + bE(Y)$$

[۱] Expected Value

تعریف ۶.۱.۱. [۱] **گشتاور** [۲] معیاری کمّی برای توصیف شکل یک توزیع احتمالاتی است. گشتاور اول همان میانگین است. برای گشتاورهای مراتب بالاتر معمولاً گشتاور را حول میانگین حساب می‌کنند و آن را گشتاور مرکزی می‌نامند. kــاُمین گشتاور مرکزی متغیر تصادفی X به صورت زیر تعریف می‌شود:

$$\mu_k = E[(X-\mu)^k]$$

تعریف ۷.۱.۱. [۱] **واریانس**[۳]: گشتاور مرکزی دوم، واریانس نامیده می‌شودکه برای سنجش پراکندگی داده‌ها حول مقدار میانگین استفاده می‌شود.

اگر $\mu = E(X)$ ، امید ریاضی (میانگین) متغیر تصادفی X باشد، آن‌گاه واریانس X به‌صورت زیر تعریف می‌شود:

$$Var(X) = E[(X-\mu)^۲]$$

تعریف ۸.۱.۱. [۱] **انحراف معیار:** یکی از شاخصه های پراکندگی است که نشان می‌دهد میانگین داده‌ها چه مقدار از مقدار متوسط فاصله دارند.اگر انحراف معیار مجموعه ای از داده‌ها نزدیک به صفر باشد نشانه آن است که داده‌ها نزدیک به میانگین هستند و پراکندگی اندکی دارند. درحالیکه انحراف معیار بزرگ بیانگر پراکندگی قابل توجه داده‌ها می‌باشد. انحراف معیار از جذر واریانس بدست می‌آید.

تعریف ۹.۱.۱. **متغیر تصادفی مستقل:** متغیرهای تصادفی X و Y مستقل نامیده می‌شوند اگر و تنها اگر رابطه‌ی زیر برقرار باشد:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

که $F_X(x)$ و $F_Y(y)$ تابع توزیع تجمعی احتمال است.

تعریف ۱۰.۱.۱. **کوواریانس** [۴]: اندازه تغییرات هماهنگ دو متغیر تصادفی است.کوواریانس دو متغیر تصادفی X ، Y به صورت زیر است:

$$COV(X,Y) = E(XY) - E(X)E(Y)$$

اگرX و Y دو متغیر تصادفی مستقل باشند:

$$COV(X,Y) = ۰$$

تعریف ۱۱.۱.۱. [۱] **ضریب همبستگی:** ابزاری آماری برای تعیین نوع و درجه‌ی رابطه‌ی یک متغیر کمی با متغیر کمی دیگر است و یکی از معیارهای مورد استفاده در تعیین همبستگی دو

[۲] Moment
[۳] Variance
[۴] Covariance

متغیر می‌باشد. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین ۱− تا ۱ است و در صورت عدم وجود رابطه بین دو متغیر، برابر صفر است.

ضریب همبستگی متغیرهای X و Y که با نماد $\rho(X,Y)$ نشان داده می‌شود به صورت زیر تعریف می‌شود:

$$\rho(X,Y)=\frac{COV(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

تعریف ۱.۱.۱۲. [۱] **چگالی حاشیه ای:** اگر X و Y متغیرهای تصادفی پیوسته و $f(x,y)$ مقدار چگالی احتمال تؤام آن‌ها در (x,y) باشد، تابعی که به صورت

$$g(x)=\int_{-\infty}^{+\infty}f(x,y)dy \qquad -\infty<x<+\infty$$

داده می‌شود، چگالی حاشیه‌ای Y نام دارد. متناظراً، تابعی که به صورت

$$h(y)=\int_{-\infty}^{+\infty}f(x,y)dx \qquad -\infty<x<+\infty$$

داده می‌شود، چگالی حاشیه‌ای Y نام دارد.

تعریف ۱.۱.۱۳. **توزیع نرمال:** متغیر تصادفی X دارای توزیع نرمال است اگر و تنها اگر تابع چگالی احتمال آن به صورت زیر باشد:

$$f(x)=\frac{۱}{\sigma\sqrt{۲\pi}}e^{-\frac{۱}{۲}(\frac{x-\mu}{\sigma})^{۲}} \quad , \quad -\infty<x<+\infty$$

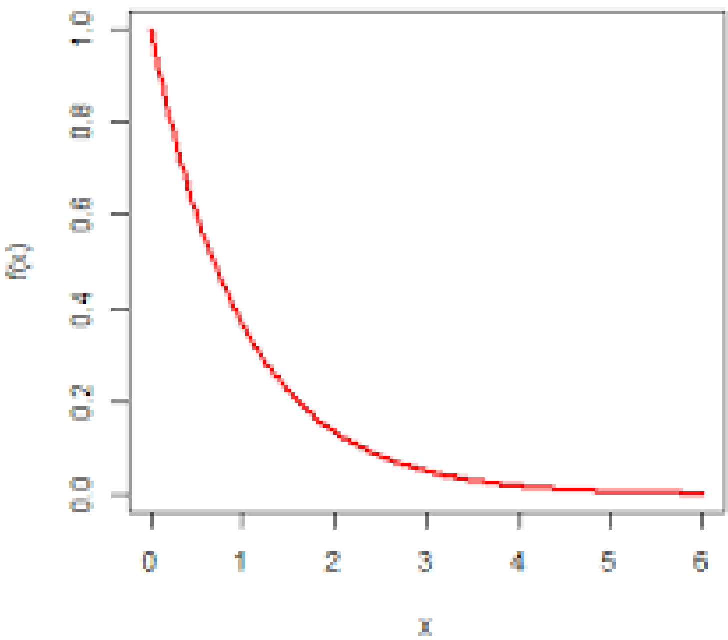
- توزیع نرمال با $\mu=۰$ ، $\sigma=۱$ را توزیع نرمال استاندارد می‌نامند.
- متغیر تصادفی X توزیع لگ_نرمال دارد هرگاه $Y=Ln(X)$ توزیع نرمال داشته باشد. بدیهی است که توزیع لگ_نرمال فقط مقادیر حقیقی مثبت را می‌گیرد.
- از مهمترین کاربردهای این تابع توزیع در دانش اقتصاد و مدیریت امروز می‌توان به مدل کردن سبد سهام‌ها در سرمایه گذاری و مدیریت منابع نام برد.

تعریف ۱.۱.۱۴. **توزیع نمایی:** متغیر تصادفی X درای توزیع نمایی با $\theta>۰$ است ، هرگاه تابع چگالی آن به صورت زیر باشد:

$$F_X(x)=\frac{۱}{\theta}e^{-\frac{x}{\theta}} \quad , \quad x>۰$$

شکل ۱.۱ نشان دهنده توزیع نمایی می‌باشد.

تعریف ۱.۱.۱۵. [۱] **توزیع گاما:** یکی از توزیع های احتمالی پیوسته است و دارای دو پارامتر مقیاس θ و پارامتر شکل k است.اگر k عددی طبیعی باشد آنگاه توزیع گاما معادل است با



شکل ۱.۱: توزیع نمایی

مجموع k متغیر تصادفی با توزیع نمایی با پارامتر $\frac{1}{\theta}$ تابع چگالی احتمال توزیع گاما به صورت زیر است.این توزیع با $X \sim \Gamma(k,\theta)$ نمایش داده می‌شود.

$$f(x;k,\theta) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k\Gamma(k)} \quad , \quad x > 0, k, \theta > 0$$

تعریف ۱۶.۱.۱. [۱] **توزیع بتا:** متغیر تصادفی X دارای توزیع بتاست، و به آن متغیر تصادفی بتا اطلاق می‌شود، اگر و تنها اگر چگالی احتمال آن به صورت

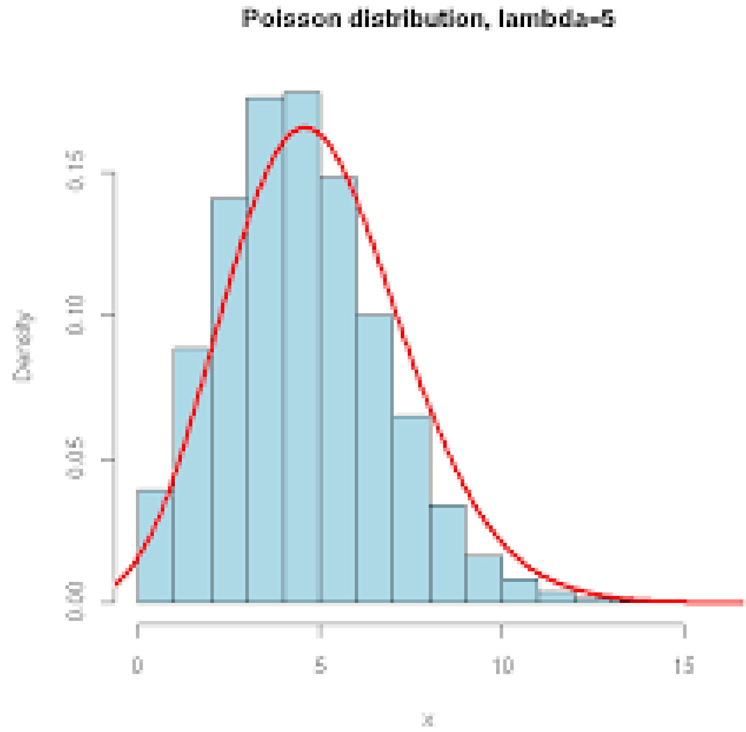
$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \quad , \quad 0 < x < 1$$

باشد، که در آن $\alpha > 0$ و $\beta > 0$ است.

تعریف ۱۷.۱.۱. [۱] **توزیع پواسون:** متغیر تصادفی X توزیع پواسون دارد و به آن متغیر تصادفی پواسون داده می‌شود، اگر و تنها اگر توزیع احتمال آن به صورت

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad k = 0, 1, 2, ..$$

که λ یک عدد حقیقی مثبت است. شکل ۲.۱ نشان دهنده توزیع پواسون است.



شکل ۲.۱: توزیع پواسون

تعریف ۱۸.۱.۱. توزیع بتای نوع دوم: توزیع احتمالی که برای اعداد حقیقی بزرگتر از صفر تعریف می‌شود و دارای دو پارامتر α و β است.تابع توزیع احتمال آن به صورت زیر است:

$$f_X(x) = \frac{x^{(n-1)}}{\beta^n B(n,\alpha)(1+\frac{x}{\beta})^{(n+\alpha)}} \quad , \quad x > 0$$

که $B(\alpha,\beta)$ به صورت زیر است:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

تعریف ۱۹.۱.۱. [1] توزیع دو جمله‌ای منفی: متغیر تصادفی X توزیع دو جمله‌ای منفی دارد و به آن عنوان متغیر تصادفی دو جمله‌ای منفی داده می‌شود، اگر و تنها اگر توزیع احتمالش به ازای $k = k, k+1, K+2, ...$ به صورت

$$f_X(x;k,\theta) = \binom{x-1}{k-1}\theta^k(1-\theta)^{x-k}$$

باشد. پس شماره امتحانی که k امین پیروزی در آن رخ می‌دهد، متغیر تصادفی است که توزیع دو جمله‌ای منفی با پارامترهای k و θ دارد.

تعریف ۲۰.۱.۱. [1] توزیع هندسی: توزیعی است گسسته که بیانگر احتمال اولین پیروزی پس از $k-1$ شکست در فرآیند برنولی می‌باشد.

$$P_X(k) = P\{X = k\} = (1-P)^{k-1}P \quad , \quad k = 1,2,3,...$$

که در آن P احتمال پیروزی در یک دفعه است.

تعریف ۲۱.۱.۱. توزیع پارتو: تابع احتمال یکی از توزیع های مهم آماری توزیع پارتو است که تابع چگالی احتمال آن به صورت زیر می‌باشد:

$$f_X(x;\alpha,\beta)=\frac{\alpha}{\beta(۱+\frac{x}{\beta})^{\alpha+۱}} \quad ,x>۰$$

- اگر $x<۰$ باشد آنگاه $f(x;\alpha,\beta)=۰$ می‌باشد.
- $\alpha,\beta>۰$ می‌باشند. β پارامتر مقیاس و α پارامتر شکل می باشد.این توزیع را با $x\sim Pa(\alpha,\beta)$ نمایش می‌دهیم.
- امید ریاضی و واریانس توزیع پارتو به صورت زیر می‌باشد:

$$E(X)=\frac{\beta}{\alpha-۱}$$

$$Var(X)=\frac{\alpha\beta^{۲}}{(\alpha-۱)^{۲}(\alpha-۲)}$$

تعریف ۲۲.۱.۱. معیار اطلاعاتی آکائیک (AIC [۵]): معیار اطلاعات AIC یک میزان از کیفیت نسبی مدل آماری از یک مجموعه از داده‌ها می‌باشد. در واقع AIC ابزاری برای انتخاب مدل است. علاوه بر این AIC یک معادله بین برازش و پیچیگی مدل را توضیح می‌دهد. این آماره بر اساس پراکنش اطلاعات بنا شده است. به صورت کلیAIC را اینگونه می نویسیم:

$$AIC=۲K-۲\ln L$$

که K تعداد پارامترها در مدل آماری و L ماکزیمم تابع در درستنمایی مدل برآورده می باشد. بین چند مدل انتخابی برای داده‌ها، مدلی بهتر می باشد که کمترین مقدار AIC را داشته باشد.

۲.۱ مفاهیم مالی

تعریف ۱.۲.۱. مؤسسه مالی: به سازمانی که بر تراکنش‌های مالی همچون سرمایه‌گذاری‌ها، وام‌ها و سپرده‌ها تمرکز دارد مؤسسه مالی گفته می‌شود. مؤسسه‌های مالی به شکل مرسوم از سازمان‌هایی همچون بانک‌ها، کارگزای‌ها، شرکت‌های بیمه و ... تشکیل شده اند. تقریباً همه افراد به صورت معمول با مؤسسات مالی سروکار دارند. هرچیزی از واریز وجه گرفته تا گرفتن وام و ارز های مبادله‌ای از طریق مؤسسه مالی انجام می‌گیرد.

[۵] Akaike information criterion

تعریف ۲.۲.۱. ریسک: انحراف بازده واقعی سرمایه گذاری از بازده پیش بینی شده آن ریسک گفته می‌شود .ریسک همچنین شامل احتمال از دست رفتن همه یا بخشی از اصل سرمایه گذاری نیز می‌شود. برای سنجش ریسک سرمایه گذاری معیارهای متفاوتی وجود دارد .ریسک عدم اطمینان و آگاهی در مورد نتیجه یک عمل است. در ادبیات مالی ریسک را می‌توان به صورت رویدادهای غیرمنتظره که معمولا به صورت تغییر در ارزش دارایی‌ها یا بدهی‌ها است تعریف کرد.

تعریف ۳.۲.۱. روش های سنجش ریسک: ریسک عبارت است از احتمال نوسانات آتی نرخ بازدهی. شاخص‌های مختلفی برای تعیین نوسانات مورد استفاده قرار می‌گیرد که بعضی از آن ها بدین صورت هستند:

- دامنه‌ی تغییرات
- متوسط انحراف خطی
- واریانس
- انحراف معیار
- نیم واریانس
- نیم انحراف معیار
- شاخص بتا
- دارایی در معرض خطر

تعریف ۴.۲.۱. [۲] **ارزش در معرض ریسک (VaR** [۶]**):** ارزش در معرض ریسک یک روش آماری است که برای اندازه گیری و تعیین میزان ریسک مالی در یک شرکت و یا پرتفوی سرمایه گذاری در یک دوره زمانی مشخص استفاده می‌شود. ارزش در معرض ریسک توسط مدیران ریسک برای اندازه‌گیری و کنترل سطح ریسکی که شرکت متعهد شده است مورد استفاده قرار می‌گیرد. شرکت ها معمولا ریسک و بازده ناشی از آن را از روش های مختلف را در نظر می‌گیرند. وظیفه مدیران ریسک این است که مطمئن شوند ریسک شرکت از حدی که می‌تواند ضررهای یک نتیجه بدتر احتمالی را تحمل کند بیشتر نباشد.

بطور کلی می‌توان گفت ارزش در معرض ریسک بیشترین مقدار زیان مورد انتظار را در یک افق زمانی مشخص در سطح اطمینان معین اندازه گیری می نماید و با سه متغیر اندازه گیری می شود:

میزان ضرر و زیان بالقوه– احتمال رخداد ضرر و زیان بالقوه– بازه زمانی

[۶] Value at Risk

تعریف ۵.۲.۱. ریسک مؤسسات مالی: مؤسسات مالی با ریسک‌های مختلفی مواجه هستند که شامل موارد زیر می‌باشد:

- ریسک عدم شفافیت
- ریسک اشتباهات مدیریتی
- ریسک ورشکستگی و مالی
- ریسک قانونی
- ریسک نقدینگی
- ریسک تورم
- ریسک نوسانات نرخ بهره
- ریسک بازار
- ریسک اخلاقی
- ریسک عدم بازپرداخت تسهیلات

تعریف ۶.۲.۱. بیمه‌گر: شرکت بیمه‌ای است که مشخصات آن در بیمه‌نامه درج گردیده‌است و در ازای دریافت حق‌بیمه جبران خسارت احتمالی را طبق شرایط این بیمه‌نامه بعهده می‌گیرد.

تعریف ۷.۲.۱. بیمه‌گذار: شخص حقیقی یا حقوقی است که مالک موضوع بیمه است یا به یکی از عناوین قانونی نمایندگی مالک یا ذی‌نفع را داشته یا مسئولیت حفظ موضوع بیمه را از طرف مالک دارد و بیمه یا بیمه‌گر را منعقد می‌کند و تعهد پرداخت حق بیمه آن می‌باشد.

تعریف ۸.۲.۱. حق بیمه: مبلغی است که در بیمه‌نامه مشخص شده و بیمه‌گذار موظف است آن را هنگام صدور بیمه‌نامه یا به ترتیبی که در بیمه نامه مشخص می شود به بیمه گر پرداخت نمایید.

تعریف ۹.۲.۱. موضوع بیمه: لوازمی که مطابق کاتالوگ وسیله بیمه شده به خریدار تحویل و یا در بیمه‌نامه درج شده است نیز جزو موضوع بیمه محسوب می‌شود.

تعریف ۱۰.۲.۱. بیمه: سازوکاری است که طی آن بیمه‌گر، بنا به ملاحظاتی تعهد می‌کند که زیان احتمالی بیمه‌گذار را در صورت وقوع حادثه در یک دوره‌ی زمانی مشخص، جبران نماید یا خدمات مشخصی را به وی ارائه دهد. بنابراین بیمه یکی از روش‌های مقابله با ریسک است.

طی یک قرارداد بیمه، ریسک مشخصی از یک طرف قرارداد (که بیمه‌گذار نامیده می‌شود) به طرف دیگر (که بیمه‌گر نامیده‌می‌شود) منتقل می‌گردد.

تعریف ۱.۲.۱۱. انواع بیمه‌نامه‌ها:

- بیمه‌های اتومبیل
- بیمه‌های آتش سوزی
- بیمه‌های درمان تکمیلی
- بیمه‌های عمر و حوادث
- بیمه‌های عمر و سرمایه‌گذاری
- بیمه‌های مسئولیت
- بیمه‌های مهندسی
- بیمه‌های باربری

فصل ۲

توزیع پارتو و کاربردهای آن

مقدمه

توزیع پارتو توزیع احتمالی توانی است که بسیاری از پدیده‌های اجتماعی، علمی و ژئوفیزیکی را توصیف می‌کند. این توزیع به یاد اقتصاددان ایتالیایی ویلفردو فدریکو داماسو پارتو [۱] نام‌گذاری شده‌است. در این فصل در ابتدا به بیان تاریخچه‌ی مختصری از توزیع پارتو و کاربردهای آن پرداخته شده‌است. در ادامه به بیان رابطه بین توزیع پارتو و توزیع‌های گاما و بتای نوع دوم پرداخته شده‌است. همچنین توزیع چند متغیره وابسته پارتو و بررسی برخی خواص این توزیع ارائه شده که کاربرد فراوانی در بیمه و تحلیل ریسک دارد. [۵]،[۶]،[۸] رده‌های مختلفی از توزیع چند متغیری پارتو ارائه شده است که مهم ترین و اصلی ترین کلاس ها توسط آرنولد (۱۹۸۳ – ۲۰۱۵) که در قالب سلسله مراتب توزیع پارتو ارائه شده است. انواع دیگری نیز توسط لاندسمن [۲] (۲۰۰۹) و آسیمیت [۳] و همکارانش ارائه شده است.

[۱] Vilferedo Federico Damaso Pareto

[۲] Londsman

[۳] Asimite

۱.۲ تاریخچه توزیع پارتو:

ویلفردو فدریکو داماسو پارتو جامعه شناس، اقتصاددان، مهندس و فیلسوف ایتالیایی بود. اصل پارتو در اقتصاد به نام اوست. پارتو در شاخه های مختلف علوم انسانی نظریه پردازی کرده‌است. نظریه حکومت نخبگان در علم سیاست، قانون ۸۰ و ۲۰ علم مدیریت و اصل بهینه پارتو در علم اقتصاد از ابداعات اوست.

ویلفردو پارتو در سال ۱۸۴۸ در پاریس متولد شد. وی فرزند یک تبعیدی از ایتالیا بود. او تحصیلات خود را در فرانسه آغاز کرد ولی آن را در ایتالیا ادامه داد و در ریاضیات و ادبیات کلاسیک تخصص یافت. پارتو در سال ۱۸۶۹ از دانشکده فنی تورین [۴] فارغ التحصیل شد. سپس حدود ۲۰ سالی را بعنوان مهندس و مدیر در شرکت راه آهن ایتالیا کار کرد. وی در سال ۱۸۹۰ در سن ۴۲ سالگی به اقتصاد روی آورد و مدت ۷ سال در لوزان [۵] به تدریس اقتصاد پرداخت ولی در سال ۱۹۰۰ که ثروت زیادی به ارث برد از این کار استعفا داد و به سوئیس رفت و بقیه عمر را به مطالعه و تحقیق پرداخت.

- نقش پارتو در اقتصاد:

پارتو دومین نفر از مکتب لوزان است. مکتب لوزان توسط لئون والراس [۶] (۱۹۱۰–۱۸۳۴) پایه گذاری شد.این مکتب علاوه بر بحث درباره مقوله مطلوبیت به هزینه و عرضه نیز پرداخته است. با این بیان که اگر قیمت بالاتر از هزینه تولید قرار گیرد عرضه افزایش پیدا می کند تا قیمت را پایین آورد و بالعکس. اگر هزینه از قیمت پیشی گیرد، عرضه کاهش پیدا کرده و قیمت ها بالا می روند. پارتو یکی از بنیانگذاران اقتصاد ریاضی نیز شناخته شده است، دیدگاه او در مقوله تعادل عمومی در اقتصاد رفاه جایگاه اساسی دارد.

۱.۱.۲ توزیع پارتو و ارتباط آن با توزیع گاما

لم زیر ارائه دهنده تصویری ساده از توزیع پارتو متغیر های تصادفی می‌باشد:

لم ۱.۱.۲. [۳] :فرض کنید X و Y متغیرهای تصادفی مستقل گاما باشند، بطوریکه $X \sim \Gamma(1,1)$ و $Y \sim \Gamma(\alpha,1)$ که $\alpha > 0$ می‌باشد. اگر $\beta > 0$ داریم:

$$U = \beta \frac{X}{Y} \sim Pa(\alpha, \beta)$$

برهان. داریم:

$$f(x) = e^{-x}, \quad x > 0$$

[۴] Torin

[۵] Lausanne

[۶] Leon Warlas

$$f(y)=\frac{y^{\alpha-1}e^{-y}}{(\alpha-1)!},\quad y>0$$

$$\Rightarrow f(x,y)=\frac{y^{\alpha-1}e^{-(x+y)}}{(\alpha-1)!},\quad x,y>0$$

تغییر متغیر زیر را در نظر می‌گیریم:

$$u=\beta\frac{x}{y},\quad v=x$$

$$\Longrightarrow y=\beta\frac{v}{u},\quad u>0$$

با محاسبه ژاکوبین داریم:

$$\begin{vmatrix}\frac{dx}{du} & \frac{dx}{dv}\\ \frac{dy}{du} & \frac{dy}{dv}\end{vmatrix}=\begin{vmatrix}0 & 1\\ -\frac{\beta v}{u^2} & \frac{\beta}{u}\end{vmatrix}=\frac{\beta v}{u^2}$$

قرار می‌دهیم:

$$g(u,v)=\frac{(\frac{\beta v}{u})^{\alpha-1}e^{-(v+\frac{\beta v}{u})}}{(\alpha-1)!}\times\frac{\beta v}{u^2}$$

$$=\frac{(\frac{\beta v}{u})^{\alpha}e^{-(v+\frac{\beta v}{u})}}{u(\alpha-1)!}$$

بنابراین توزیع حاشیه‌ای u برابر است با:

$$g(u)=\int_0^\infty\frac{(\frac{\beta v}{u})^{\alpha}e^{-(v+\frac{\beta v}{u})}}{u(\alpha-1)!}dv$$

با قرار دادن $\acute{\alpha}=\alpha+1$ و $\acute{\beta}=(1+\frac{\beta}{u})^{-1}$ به صورت زیر می‌توان تابع زیر انتگرال را به توزیع گاما تبدیل کرد:

$$g(u)=\frac{\Gamma(\acute{\alpha})\acute{\beta}^{\acute{\alpha}}(\frac{\beta}{u})^{\alpha}}{\Gamma(\alpha)u}\int_0^\infty\frac{v^{\acute{\alpha}-1}e^{-(1+\frac{\beta}{u})v}}{\Gamma(\acute{\alpha})\acute{\beta}^{\acute{\alpha}}}dv$$

چون

$$\int_0^\infty\frac{v^{\acute{\alpha}-1}e^{-(1+\frac{\beta}{u})v}}{\Gamma(\acute{\alpha})\acute{\beta}^{\acute{\alpha}}}dv=1$$

در نتیجه:

$$g(u)=\frac{\Gamma(\acute{\alpha})\acute{\beta}^{\acute{\alpha}}}{\Gamma(\alpha)u}\cdot(\frac{\beta}{u})^{\alpha}$$

$$=\frac{\alpha!(\frac{\beta}{u})^{\alpha}}{(\alpha-1)!u(1+\frac{\beta}{u})^{\alpha+1}}$$

$$=(\frac{\frac{\beta}{u}}{1+\frac{\beta}{u}})^{\alpha}\times\frac{\alpha}{u(1+\frac{\beta}{u})}=(\frac{\frac{\beta}{u}}{\frac{\beta}{u}(\frac{u}{\beta}+1)})^{\alpha}\times\frac{\alpha}{(u+\beta)}$$

$$=(\frac{1}{(1+\frac{u}{\beta})^{\alpha}})\times\frac{\alpha}{\beta(\frac{u}{\beta}+1)}=\frac{\alpha}{\beta(1+\frac{u}{\beta})^{\alpha+1}}\sim Pa(\alpha,\beta)$$

□

۲.۲ توزیع پارتو و ارتباط آن با توزیع بتای نوع دوم

تعمیمی از توزیع پارتو ($f(x;\alpha,\beta)=\frac{\alpha}{\beta(1+\frac{x}{\beta})^{\alpha+1}} \quad, x>0$) در ادامه آورده می‌شود.

- یک متغیر تصادفی X را می‌توان توزیع بتا از نوع دوم دانست، اگر چگالی احتمال آن به شکل زیر باشد:

$$f(x;p,q,\beta)=\frac{x^{p-1}}{\beta^p B(p,q)(1+\frac{x}{\beta})^{p+q}} \qquad x>0$$

که $B(p,q)$ به صورت زیر می‌باشد:

$$B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

اگر $x<0$ باشد، آنگاه $f(x;p,q,\beta)=0$ و $p,q,\beta>0$ هستند [۴].

- اگر در تابع چگالی احتمال توزیع بتای نوع دوم قرار دهیم $p=1$ توزیع پارتوی $Pa(q,\beta)$ بدست می‌آید.

- توزیع بتای نوع دوم دارای نمایشی ساده و به شکل نسبت متغیر های تصادفی گاما می‌باشد.

لم ۱.۲.۲. اگر X و Y متغیرهای تصادفی مستقل گاما باشند، بطوریکه $X\sim\Gamma(p,1)$ و $Y\sim\Gamma(q,1)$ می‌باشد. اگر $\beta>0$ متغیر تصادفی جدید $U=\beta\frac{X}{Y}$ دارای تابع چگالی احتمال پارتوی تعمیم یافته است.

برهان. داریم:

$$f(x)=\frac{x^{p-1}e^{-x}}{\Gamma(p)}, \quad x>0$$

$$f(y)=\frac{y^{q-1}e^{-y}}{\Gamma(q)}, \quad y>0$$

$$f(x,y)=\frac{x^{p-1}y^{q-1}e^{-(x+y)}}{\Gamma(p)\Gamma(q)}, \quad x,y>0$$

تغییر متغیر زیر را در نظر می‌گیریم:

$$u=\beta\frac{x}{y}, \quad v=x$$

$$\Longrightarrow y=\beta\frac{v}{u}, \quad u>0$$

با محاسبه ژاکوبین داریم:

$$\begin{vmatrix} \frac{dx}{du} & \frac{dx}{dv} \\ \frac{dy}{du} & \frac{dy}{dv} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ -\frac{\beta v}{u^2} & \frac{\beta}{u} \end{vmatrix} = \frac{\beta v}{u^2}$$

قرار می‌دهیم:

$$g(u,v)=\frac{v^{p-۱}(\frac{\beta v}{u})^{q-۱}e^{-(v+\frac{\beta v}{u})}}{\Gamma(p)\Gamma(q)}\times\frac{\beta v}{u^{۲}}$$

بنابراین توزیع حاشیه‌ای u برابر است با:

$$g(u)=\int_{۰}^{\infty}\frac{v^{p-۱}(\frac{\beta v}{u})^{q-۱}e^{-(v+\frac{\beta v}{u})}}{\Gamma(p)\Gamma(q)}\times\frac{\beta v}{u^{۲}}$$

با قرار دادن $p+q=\acute{p}$ و $\acute{\beta}=(۱+\frac{\beta}{u})^{-۱}$ به صورت زیر می‌توان تابع زیر انتگرال را به توزیع گاما تبدیل کنیم:

$$g(u)=\frac{(\frac{\beta}{u})^{q}\Gamma(\acute{p})}{u\Gamma(p)\Gamma(q)}\int_{۰}^{\infty}\frac{v^{(p-۱)}v^{q}e^{-(۱+\frac{\beta}{u})v}}{\Gamma(\acute{p})}dv$$

$$=\frac{(\frac{\beta}{u})^{q}\Gamma(\acute{p})\acute{\beta}^{\acute{p}}}{u\Gamma(p)\Gamma(q)}\int_{۰}^{\infty}\frac{v^{(p-۱)}e^{-(\frac{v}{\acute{\beta}})}}{\Gamma(\acute{p})\acute{\beta}^{\acute{p}}}dv$$

چون

$$\int_{۰}^{\infty}\frac{v^{(p-۱)}e^{-(\frac{v}{\acute{\beta}})}}{\Gamma(\acute{p})\acute{\beta}^{\acute{p}}}dv=۱$$

در نتیجه:

$$g(u)=\frac{\Gamma(\acute{p})\acute{\beta}^{\acute{p}}}{u\Gamma(p)\Gamma(q)}\cdot(\frac{\beta}{u})^{q}$$

$$=\frac{u^{p-۱}(\frac{\beta}{u})^{q}\Gamma(p+q)}{u^{p-۱}u(۱+\frac{\beta}{u})^{p+q}\Gamma(p)\Gamma(q)}$$

$$=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\times\frac{u^{p-۱}}{u^{p}(۱+\frac{\beta}{u})^{p}}\times(\frac{(\frac{\beta}{u})}{(۱+\frac{\beta}{u})})^{q}$$

$$=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\times\frac{u^{p-۱}}{u^{p}(۱+\frac{\beta}{u})^{p}}\times(\frac{(\frac{\beta}{u})}{\frac{\beta}{u}(۱+\frac{u}{\beta})})^{q}$$

$$=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\times\frac{u^{p-۱}}{u^{p}(۱+\frac{\beta}{u})^{p}}\times\frac{۱}{(۱+\frac{u}{\beta})^{q}}$$

$$=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\times\frac{u^{p-۱}}{(u+\beta)^{p}}\times\frac{۱}{(۱+\frac{u}{\beta})^{q}}$$

$$=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\times\frac{u^{p-۱}}{\beta^{p}(۱+\frac{u}{\beta})^{p}}\times\frac{۱}{(۱+\frac{u}{\beta})^{q}}$$

$$=\frac{u^{p-۱}}{\beta^{p}B(p,q)(۱+\frac{u}{\beta})^{p+q}}\sim Pa(p,q)$$

□

۳.۲ توزیع چند متغیره وابسته پارتو

تعریف ۱.۳.۲. فرض کنید که $Y_۱, Y_۲, ...Y_n$ و Y_α دو به دو مستقل باشند. $Y_۱, Y_۲, ...Y_n$ متغیرهای تصادفی هم توزیع از $\Gamma(۱,۱)$ و Y_α دارای توزیع $\Gamma(\alpha, ۱)$ که $\alpha > ۰$ باشد.

توزیع چندمتغیری وابسته پارتو را به صورت

$$X = (X_۱, X_۲, ..., X_n) = (\beta\frac{Y_۱}{Y_\alpha}, \beta\frac{Y_۲}{Y_\alpha}, ..., \beta\frac{Y_n}{Y_\alpha}) \qquad (۱)$$

که $\beta > ۰$ باشد، تعریف می‌شود.

۱.۳.۲ بعضی خواص توزیع چندمتغیره پارتو

لم ۱.۳.۲. با توجه به تعریف ۱.۳.۲، داریم:

- الف. تابع چگالی احتمال بردار X به صورت زیراست:

$$g(x_۱, x_۲, ..., x_n; \alpha, n) = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)\beta^n} \times (\frac{۱}{۱ + \sum_{i=۱}^{n} \frac{x_i}{\beta}})^{\alpha+n} \qquad x_۱, ...x_n > ۰$$

- ب.

$$COV(X_i, X_j) = \frac{\beta^۲}{(\alpha - ۱)^۲(\alpha - ۲)} \qquad \forall i \neq j$$

- ج. ضریب همبستگی میان مؤلفه‌ها به صورت زیر می‌باشد:

$$\rho(X_i, X_j) = \frac{۱}{\alpha} \qquad \alpha > ۲, i \neq j$$

برهان. الف. مشابه لم ۱.۱.۲، داریم:

$$f(y_۱, ..., y_n, y_\alpha) = \prod_{i=۱}^{n} f(y_i) \cdot f(y_\alpha)$$

تغییر متغیر زیر را اعمال می‌کنیم:

$$X_۱ = \beta\frac{Y_۱}{Y_\alpha}, X_۲ = \beta\frac{Y_۲}{Y_\alpha}, ..., X_n = \beta\frac{Y_n}{Y_\alpha} \quad , \quad Z = Y_\alpha$$

با محاسبه ژاکوبین داریم:

$$g(x_۱, ..., x_n, z) = f(\frac{x_۱ z}{\beta}, \frac{x_۲ z}{\beta}, ..., \frac{x_n z}{\beta}, z) \cdot |J|$$

$$g(x_۱, ..., x_n) = \int_۰^\infty g(x_۱, ..., x_n, z)dz$$

ب.داریم:

$$COV(X_۱,X_۲)=E(X_۱X_۲)-E(X_۱)E(X_۲)$$

همچنین طبق قسمت الف لم ۱.۳.۲ داریم:

$$f(x)=\frac{\alpha}{\beta(۱+\frac{x}{\beta})^{\alpha+۱}}\Longrightarrow E(X)=\int_۰^\infty\frac{\alpha x}{\beta(۱+\frac{x}{\beta})^{\alpha+۱}}=\frac{\beta}{\alpha-۱}$$

$$f(x_۱,x_۲)=\frac{\Gamma(\alpha+۲)}{\Gamma(\alpha)\beta^۲}\times\frac{۱}{(۱+\frac{x_۱+x_۲}{\beta})^{\alpha+۲}}$$

داریم:

$$E(X_۱X_۲)=\int_۰^\infty\int_۰^\infty\frac{\Gamma(\alpha+۲)}{\Gamma(\alpha)\beta^۲}\times\frac{x_۱x_۲}{(۱+\frac{x_۱+x_۲}{\beta})^{\alpha+۲}}dx_۱dx_۲$$

$$=\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲}\int_۰^\infty\int_۰^\infty\frac{x_۱x_۲}{(\beta+x_۱+x_۲)^{\alpha+۲}}dx_۱dx_۲$$

$$=\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲}\int_۰^\infty\frac{x_۲}{\acute{\alpha}\acute{\beta}^{\acute{\alpha}}}\int_۰^\infty\frac{\acute{\alpha}x_۱}{\acute{\beta}(۱+\frac{x_۱}{\acute{\beta}})^{\acute{\alpha+۱}}}dx_۱dx_۲$$

که $\acute{\alpha}=\alpha+۱$ و $\acute{\beta}=\beta+x_۲$.

عبارت $\int_۰^\infty\frac{\acute{\alpha}x_۱}{\acute{\beta}(۱+\frac{x_۱}{\acute{\beta}})^{\acute{\alpha}+۱}}$ به امید ریاضی توزیع پارتو تبدیل شد که برابر $\frac{\acute{\beta}}{\acute{\alpha}-۱}$ است.

بنابراین:

$$\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲}\int_۰^\infty\frac{x_۲}{\acute{\alpha}\acute{\beta}^{\acute{\alpha}}}\times\frac{\acute{\beta}}{\acute{\alpha}-۱}dx_۲=$$

$$\Longrightarrow\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲}\int_۰^\infty\frac{x_۲(\beta+x_۲)}{\alpha(\alpha+۱)(\beta+x_۲)^{\alpha+۱}}dx_۲$$

فرض کنید $\acute{\alpha}=\alpha-۱$، داریم:

$$\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲\acute{\alpha}\alpha(\alpha+۱)\beta^{\acute{\alpha}}}\int_۰^\infty\frac{\acute{\alpha}x_۲}{\beta(۱+\frac{x_۲}{\beta})^{\acute{\alpha}}}dx_۲=$$

$$=\frac{\Gamma(\alpha+۲)\beta^{\alpha+۲}}{\Gamma(\alpha)\beta^۲(\alpha-۱)\alpha(\alpha+۱)\beta^{\alpha-۱}}\times\frac{\beta}{(\alpha-۲)}$$

$$=\frac{(\alpha+۱)\alpha\beta^{\alpha+۳}}{\beta^{\alpha+۱}\alpha(\alpha+۱)(\alpha-۱)(\alpha-۲)}=\frac{\beta^۲}{(\alpha-۱)(\alpha-۲)}$$

$$=\frac{(\alpha-۱)\times\beta^۲}{(\alpha-۱)\times(\alpha-۱)(\alpha-۲)}-\frac{\beta^۲\times(\alpha-۲)}{(\alpha-۱)^۲\times(\alpha-۲)}$$

$$=\frac{\alpha\beta^۲-\beta^۲-\alpha\beta^۲+۲\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}=\frac{\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}$$

ج.

$$\rho(X_i, X_j) = \frac{COV(X_i, X_j)}{\sqrt{VarX_i}\sqrt{VarX_j}} = \frac{\frac{\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}}{\sqrt{\frac{\alpha\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}}\sqrt{\frac{\alpha\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}}}$$

$$\Longrightarrow \frac{\frac{\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}}{\frac{\alpha\beta^۲}{(\alpha-۱)^۲(\alpha-۲)}} = \frac{۱}{\alpha}$$

□

فصل ۳

ارزش در معرض ریسک و مدل های ترکیبی پارتو

مقدمه

کاس [1] و همکارانش (۲۰۰۱) و کلاگمن [2] و همکارانش (۲۰۰۸) با بررسی و پژوهش به این نتیجه رسیدند که مدل های ریسک فردی و جمعی در موارد زیر از هم مستقل هستند:

- مقادیر مختلف ادعاهای خسارت
- تعداد ادعاهای خسارت و مقدار ادعاهای خسارت
- فاصله بین مقدار این ادعاها

این امر سبب تسهیل در محاسبه بسیاری از سنجه های ریسک می شود (شاخص ریسک) اما می تواند در زمینه های مختلف دارای محدودیت باشد. [۸]

اساس توزیعی که آرنولد (۱۹۸۷) و آرنولد و همکارانش (۱۹۹۳) ارائه داده اند ساختار وابستگی شرطی می‌باشد که در دو کلاس وابسته مختلف ارائه شده‌اند. بنابراین یک مدل

[1] Kass

[2] Klugman

ریسک را مورد توجه قرار می‌دهیم که اساس ساختار آن وابستگی است و چندین شاخصه آن را مورد بررسی قرار می‌دهیم.

کوزت [۳] و همکارانش (۲۰۱۳) نمایه ای از ریسک های وابسته را مورد بررسی قرار دادند که توزیع چند متغیری در آن از توزیع های ترکیبی می‌باشد.

آلبرچر [۴] و تئوگلز [۵] (۲۰۰۶) یک ساختار وابسته ی پیوسته را برای زمان میان ادعاها و اندازه ادعاهای بعدی ارائه دادند.(پیش بینی ادعاهای بعدی)

بوودریت [۶] و همکارانش (۲۰۰۶) به بررسی تعمیم مدل ریسک پواسون می پردازند که توزیع مقدار ادعای بعدی تابعی از زمان سپری شده از آخرین ادعاست.

کوزت [۷] و همکارانش تعمیم دیگری را مورد توجه قرار دادند که ارائه دهنده ساختار وابسته‌ای میان مقادیر ادعاها و زمان ادعاها می‌باشد. در ادامه توزیع مدل جمعی ترکیبی $S_N = X_1 + X_2 + ... + X_N$ را به دست می‌آوریم که توزیع فرعی از نوع پارتو است. به بررسی جزئیات دقیق برخی از مدل های تجمعی می‌پردازیم که توزیعات پواسون ، دو جمله‌ای منفی و هندسی را به عنوان توزیع اصلی داراست.

تابع چگالی احتمال مدل تجمعی پارتو_پواسون تابعی از فوق هندسی کامر [۸] است و تابع چگالی احتمال پارتو_دو جمله‌ای منفی تابعی از فوق هندسی گاؤس [۹] است.

۱.۳ ارزش در معرض ریسک

ارزش در معرض ریسک (VaR) در سطح u با $0 < u < 1$ از متغیرهای تصادفی X با توزیع تجمعی $F(X)$ به صورت زیر تعریف می‌شود:

$$VaR[X;u] = inf\{x \in R, F(X) \geq u\}$$

اگر $X \sim B2(p,q,\beta)$ باشد ، گیلن [۱۰] و همکارانش (۲۰۱۳) رابطه زیر را به دست آوردند:

$$VaR[X;u] = \beta \frac{IB(u;p,q)}{1 - IB(u;p,q)}$$

که $IB(u;p,q)$ نشان دهنده حالت معکوس تابع بتا است.

[۳] Cossette
[۴] Alberecher
[۵] Teugels
[۶] Boudreault
[۷] Cossette
[۸] Kummer
[۹] Gauss
[۱۰] Guillen

با استفاده از ۱.۳ برای توزیع تجمعی پارتو بدست می‌آوریم:

$$VaR[S_n;u] = \beta \frac{IB(u;n,\alpha)}{1 - IB(u;n,\alpha)} \qquad ,0 < u < 1$$

۲.۳ مدل ریسک فردی تحت شرایط وابستگی پارتو

در این بخش مدل ریسک فردی را مورد بررسی قرار می‌دهیم. فرض می‌کنیم که بین ریسک ها وابستگی وجود داشته باشد.

تصور کنید $(X_1, ..., X_n)$ توزیع پارتو چند متغیره که تعریف کردیم باشد. بنابراین ریسک تجمعی $S_n = X_1 + ... + X_n$ را مورد توجه قرار می‌دهیم که نتایج زیر را دربردارد:

قضیه ۱.۲.۳. [۶] تابع چگالی احتمال متغیر تصادفی مجموع S_n که مؤلفه‌های آن از نوع پارتوی تک متغیری باشد با معادله زیر تعریف می‌شود:

$$f_{S_n}(x;n,\alpha,\beta) = \frac{x^{n-1}}{\beta^\alpha B(n,\alpha)(1 + \frac{x}{\beta})^{n+\alpha}} \qquad ,x > 0$$

به این معنا که $S_n \sim B2(n,\alpha,\beta)$ است.

اگر $x < 0$ باشد $f_{S_n}(x;n,\alpha,\beta) = 0$ می‌شود.

برهان. با استفاده از لم ۱.۱.۲ می‌دانیم توزیع پارتو از تقسیم $\Gamma(1,1)$ بر $\Gamma(\alpha,1)$ به دست می‌آید. یعنی $X_i = \beta \frac{Y_i}{Y_\alpha}$، که در آن Y_i دارای توزیع $\Gamma(1,1)$ است. با توجه به این که می‌دانیم مجموع توزیع گاماها باز هم توزیع گاما می‌باشد، بنابراین توزیع صورت کسر توزیع گاما $\Gamma(n,1)$ است. بنابرای S_n برابر است با:

$$S_n = X_1 + ... + X_n = \beta \frac{Y_1 + Y_2 + ... + Y_n}{Y_\alpha}$$

داریم:

$$f(x) = \frac{x^{n-1}e^{-x}}{\Gamma(n)}, \qquad x > 0$$

$$f(y) = \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}, \qquad y > 0$$

$$f(x,y) = \frac{x^{n-1}y^{\alpha-1}e^{-(x+y)}}{\Gamma(n)\Gamma(\alpha)}, \qquad x,y > 0$$

تغییر متغیر زیر را اعمال می‌کنیم:

$$u = \beta \frac{x}{y}, \qquad v = x$$

$$\Longrightarrow y = \beta \frac{v}{x}, \qquad u > 0$$

با محاسبه ژاکوبین داریم:

$$\begin{vmatrix} \frac{dx}{du} & \frac{dx}{dv} \\ \frac{dy}{du} & \frac{dy}{dv} \end{vmatrix} = \begin{vmatrix} ۰ & ۱ \\ -\frac{\beta v}{u^۲} & \frac{\beta}{u} \end{vmatrix} = \frac{\beta v}{u^۲}$$

سپس:

$$g(u,v) = \frac{v^{n-۱}(\frac{\beta v}{u})^{\alpha-۱}e^{-(v+\frac{\beta v}{u})}}{u\Gamma(\alpha)\Gamma(n)}dv$$

بنابراین توزیع حاشیه‌ای u برابر است با:

$$g(u) = \int_۰^\infty \frac{v^{n-۱}(\frac{\beta v}{u})^{\alpha-۱}e^{-(v+\frac{\beta v}{u})}}{u\Gamma(\alpha)\Gamma(n)}dv$$

با قرار دادن $\acute{n} = n+\alpha$ ، $\acute{\beta} = (۱+\frac{\beta}{u})^{-۱}$ به صورت زیر می‌توان تابع زیر انتگرال را به توزیع گاما تبدیل کنیم:

$$g(u) = \frac{(\frac{\beta}{u})^\alpha\Gamma(\acute{n})\acute{\beta}^{\acute{n}}}{u\Gamma(\alpha)\Gamma(n)}\int_۰^\infty \frac{v^{\acute{n}-۱}e^{-(۱+\frac{\beta}{u})v}}{\Gamma(\acute{n})\acute{\beta}^{\acute{n}}}$$

چون

$$\int_۰^\infty \frac{v^{\acute{n}-۱}e^{-(۱+\frac{\beta}{u})v}}{\Gamma(\acute{n})\acute{\beta}^{\acute{n}}} = ۱$$

در نتیجه:

$$g(u) = \frac{u^{n-۱}(\frac{\beta}{u})^\alpha\Gamma(n+\alpha)u^{n-۱}}{u^{n-۱}u\Gamma(\alpha)\Gamma(n)(۱+\frac{\beta}{u})^n+\alpha}$$

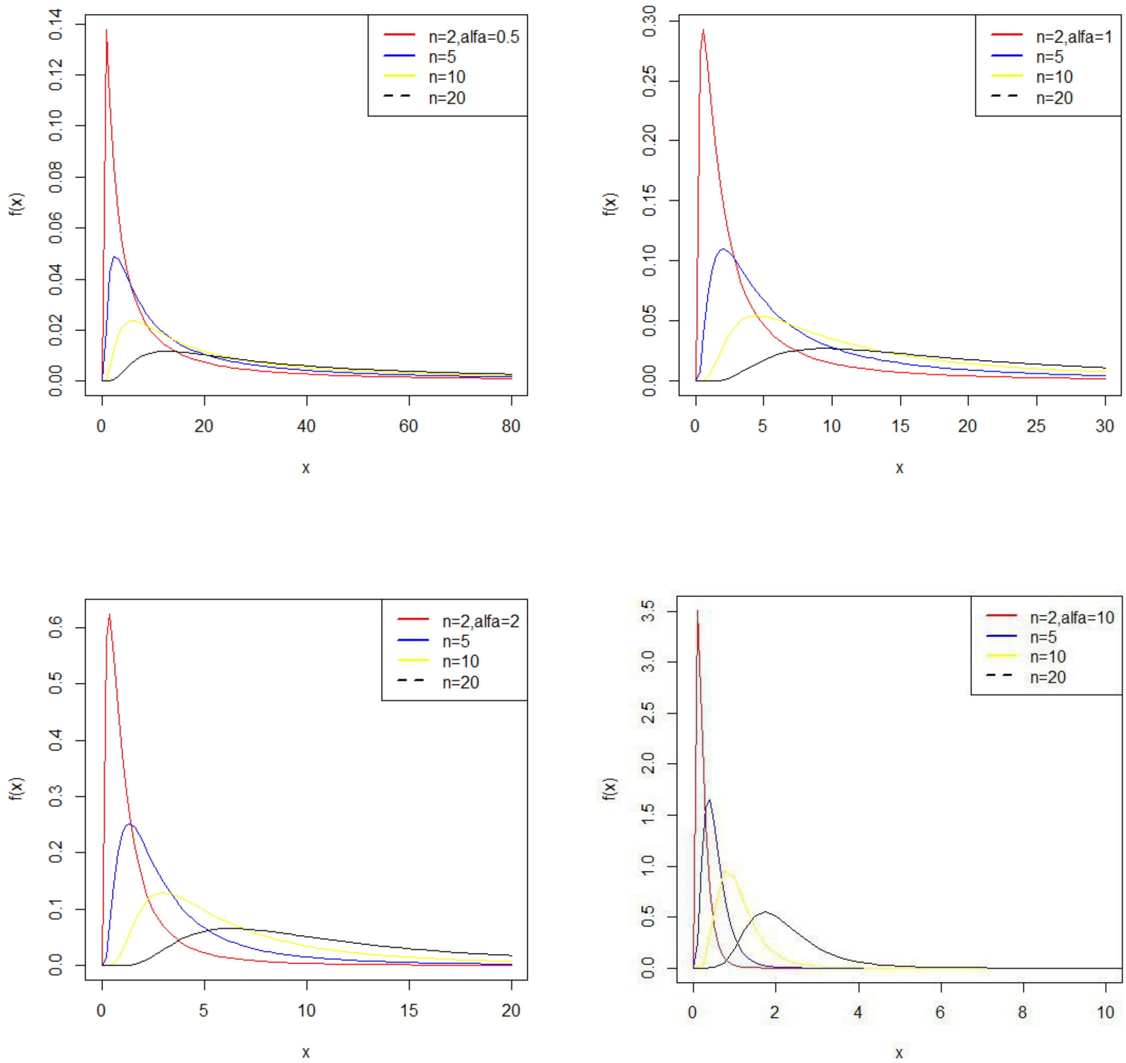
$$= \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\Gamma(n)} \times (\frac{\frac{\beta}{u}}{۱+\frac{\beta}{u}})^\alpha \times \frac{u^{n-۱}}{u^n(۱+\frac{\beta}{u})^n}$$

$$= \frac{۱}{B(n,\alpha)} \times (\frac{\frac{\beta}{u}}{\frac{\beta}{u}(۱+\frac{u}{\beta})})^\alpha \times \frac{u^{n-۱}}{(u+\beta)^n}$$

$$= \frac{u^{n-۱}}{\beta^n B(n,\alpha)(۱+\frac{u}{\beta})^{n+\alpha}} \sim Pa(n,\alpha)$$

□

- شکل ۱.۳ نشان دهنده تابع چگالی احتمال متغیر تصادفی مجموع S_n برای $\alpha = \frac{۱}{۲},۱,۲,۱۰$ و $n = ۲,۵,۱۰,۲۰$ است. با افزایش مقدار n نمودارها هموار می‌شود.



شکل ۱.۳: تابع چگالی احتمال متغیر تصادفی مجموع S_n

۳.۳ مدل ریسک تجمعی در شرایط وابستگی

در این بخش مدل ریسک تجمعی را مورد بررسی قرار می‌دهیم که بر اساس ساختار وابستگی میان مقدار ادعاهاست.

فرض کنیم N تعداد ادعاهای خسارت در سبد سهام بیمه نامه در یک دوره زمانی است.

فرض کنید X_i ، $i = ۱, ۲, ...$ مقدار iــامین ادعای خسارت و $S_N = X_۱ + ... + X_N$ مجموع ادعای خسارت‌های سبد سهام در دوره زمانی مورد نظر است.

۱.۳.۳ شاخصه های کلی

دو فرضیه را مورد توجه قرار می‌دهیم:

- الف_ فرض می‌کنیم که تمامی ادعاهای خسارت $X_۱, X_۲, ..., X_n, ...$ متغیرهای تصادفی وابسته با توزیع یکسان است.
- ب_ N متغیر تصادفی مستقل از همه مقدار ادعاهای خسارت $X_۱, X_۲, ..., X_n, ...$ مفروض است.

قضیه ۱.۳.۳. [۳] فرض کنید $X_۱, X_۲, ..., X_n, ...$ متغیرهای تصادفی وابسته و هم‌توزیع با توزیع تجمعی $F(X)$ باشند. فرض کنید N تعداد ادعاهای خسارت مشاهده شده با تابع احتمالی جرم $P_n = P_r(N = n)$ برای $n = ۰, ۱, ...$ باشد که همه X_i, $i = ۱, ۲, ...$ مستقل هستند. در نتیجه تابع توزیع تجمعی S_N به صورت زیر است:

$$F_{S_N}(x) = \sum_{n=۰}^{\infty} P_n F_X^{(n)}(x)$$

که $F_X^{(n)}(x)$ نشان دهنده تابع توزیع تجمعی پیچش و n مقادیر وابسته $(X_۱, X_۲, ..., X_n)$ می‌باشد.

برهان. داریم:

$$F_{S_N}(x) = P_r(S_N \leq x)$$

$$= \sum_{n=۰}^{\infty} P_r(S_N \leq x \mid N = n) P_r(N = n)$$

$$= \sum_{n=۰}^{\infty} P_n F_X^{(n)}(x)$$

□

۲.۳.۳ میانگین و واریانس مدل تجمعی

برای میانگین و واریانس مدل های تجمعی لم زیر را داریم:

لم ۱.۳.۳. [۱۶] میانگین و واریانس S_N در شرایط وابستگی را با رابطه زیر نمایش می‌دهیم:

- الف_

$$E(S_N) = E(N)E(X)$$

- ب_

$$Var(S_N) = E(N)Var(X) + Var(N)(E(X))^۲ + E(N(N-۱))COV(X_i, X_j)$$

برهان.

برای اثبات الف داریم:

$$E(S_N) = \sum_{n=1}^{\infty} E(S_n)P(N=n)$$

$$= \sum_{n=1}^{\infty} nE(X)P(N=n)$$

$$= E(X)\sum_{n=1}^{\infty} nP(N=n) = E(X)E(N)$$

برای اثبات ب داریم:

$$E(S_N^2) = \sum_{n=1}^{\infty} E(S_n^2)P(N=n)$$

$$= \sum_{n=1}^{\infty}(Var(S_n) + E(S_n)^2)P(N=n)$$

$$= \sum_{n=1}^{\infty}(nVar(X) + n^2E(X)^2)P(N=n) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}{}_{i\neq j} \quad X_iX_jCOV(X_i,X_j)P(N=n)$$

$$= Var(x)\sum_{n=1}^{\infty} nP(N=n) + (E(X))^2\sum_{n=1}^{\infty} n^2P(N=n) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}{}_{i\neq j} \quad X_iX_jCOV(X_i,X_j)P(N=n)$$

$$= Var(X)E(N) + (E(X))^2E(N^2) + E(N(N-1))COV(X_i,X_j)$$

$$\Longrightarrow Var(S_N) = E(S_N^2) - E(S_N)^2$$

$$= E(N)Var(X) + (E(X))^2E(N^2) + E(N(N-1))COV(X_i,X_j) - (E(N)E(X))^2$$

$$= E(N)Var(X) + (E(X))^2Var(N) + E(N(N-1))COV(X_i,X_j)$$

□

- اگر ادعاهای خسارت X_i, X_j مستقل باشند، $COV(X_i,X_j) = 0$. در نتیجه:

$$Var(S_N) = E(N)Var(X) + (E(X))^2Var(N)$$

به عبارت دیگر اگر متغیرهای تصادفی X_i وابسته باشند:

$$Var(S_N^{(1)}) < Var(S_N)$$

که $Var(S_N^{(1)})$ واریانس در نمونه مستقل و $Var(S_N)$ واریانس در نمونه وابسته است.

در حقیقت این یک نتیجه است که این ویژگی را به همراه دارد که منتهی به کوواریانسی مثبت بین X_j, X_i می‌شود.

۴.۳ توزیع ترکیبی پارتو-پواسون

مدلی را مورد توجه قرار می‌دهیم که در آن توزیع اصلی از نوع پواسون است.

قضیه ۱.۴.۳. [۸] اگر فرض کنید که توزیع پواسون با پارامتر λ به عنوان توزیع اصلی و $(X_1, ..., X_n)$ تعریف شده در ۲.۳.۰.۱ باشد. آنگاه تابع چگالی احتمال از متغیر تصادفی S_N به صورت زیر است:

$$f_{S_N}(x;\alpha,\lambda,\beta) = \frac{\alpha\lambda e^{-\lambda}}{\beta(1+\frac{x}{\beta})^{\alpha+1}} \times {}_1F_1[1+\alpha; 2, \frac{\frac{\lambda x}{\beta}}{1+\frac{x}{\beta}}] \quad x > 0$$

همچنین $f_{S_N}(0;\alpha,\lambda,\beta) = e^{-\lambda}$ است.

${}_1F_1(a;b,z)$ نشان دهنده ترکیبی از تابع فوق هندسی کامر است که به صورت زیر تعریف می‌شود:

$${}_1F_1[a;b,z] = \sum_{n=0}^{\infty} \frac{(a)_n Z^n}{(b)_n n!}$$

که $(a)_n$ نماد باز شده $(a)_n = a(a-1)...(a-n+1)$ تعریف شده است.

برهان. اگر $X = 0$ باشد :

$$f_{S_N} = P_r(N=0) = \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}$$

و اگر $X > 0$ باشد، Z را به صورت زیر در نظر می‌گیریم:

$$Z = \frac{\frac{\lambda X}{\beta}}{1+\frac{X}{\beta}}$$

بنابراین داریم:

$$f_{S_N}(x) = \sum_{n=1}^{\infty} f_x^{(n)} P_r(N=n)$$

$$= \sum_{n=1}^{\infty} \frac{x^{n-1}}{\beta^\alpha B(n,\alpha)(1+\frac{x}{\beta})^{n+\alpha}} \times \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= \frac{x^{-1}e^{-\lambda}}{\Gamma(\alpha)(1+\frac{x}{\beta})^\alpha} \sum_{n=1}^{\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)n!} (\frac{\frac{\lambda x}{\beta}}{1+\frac{x}{\beta}})^n$$

$$= \frac{x^{-1}e^{-\lambda}}{\Gamma(\alpha)(1+\frac{x}{\beta})^\alpha} \quad Z \quad \Gamma(\alpha+1){}_1F_1[1+\alpha;2;Z]$$

$$= \frac{x^{-1}e^{-\lambda}}{(\alpha-1)!(1+\frac{x}{\beta})^\alpha} \cdot \frac{\frac{\lambda x}{\beta}}{1+\frac{x}{\beta}} \cdot \alpha(\alpha-1)!{}_1F_1[1+\alpha;2;Z]$$

$$= \frac{\alpha\lambda e^{-\lambda}}{\beta(1+\frac{x}{\beta})^{\alpha+1}} {}_1F_1[1+\alpha; 2; Z]$$

□

- شکل ۲.۳ نشان دهنده تابع توزیع احتمالی پارتو_ پواسون برای $\beta = 1$، $\alpha = \frac{1}{2}, 1, 2, 10$ و $\lambda = 2, 5, 10, 20$ است.

۱.۴.۳ میانگین و واریانس مدل جمعی پارتو_ پواسون

با استفاده از فرمول میانگین و واریانس گفته شده در لم ۱.۳.۳ میانگین و واریانس مدل جمعی وابسته پارتو_ پواسون را با رابطه زیر به دست می‌آوریم:

الف_

$$E(S_N) = \frac{\lambda\beta}{\alpha - 1} \qquad , \alpha > 1$$

ب_

$$Var(S_N) = \frac{\lambda\beta^2(\lambda + 2\alpha - 2)}{(\alpha-1)^2(\alpha-2)} \qquad , \alpha > 2$$

برهان.

_ برای اثبات الف داریم:

$$E(S_N) = E(X)E(N) = \frac{\beta}{\alpha-1} \cdot \lambda = \frac{\lambda\beta}{\alpha-1}$$

_ برای اثبات ب داریم:

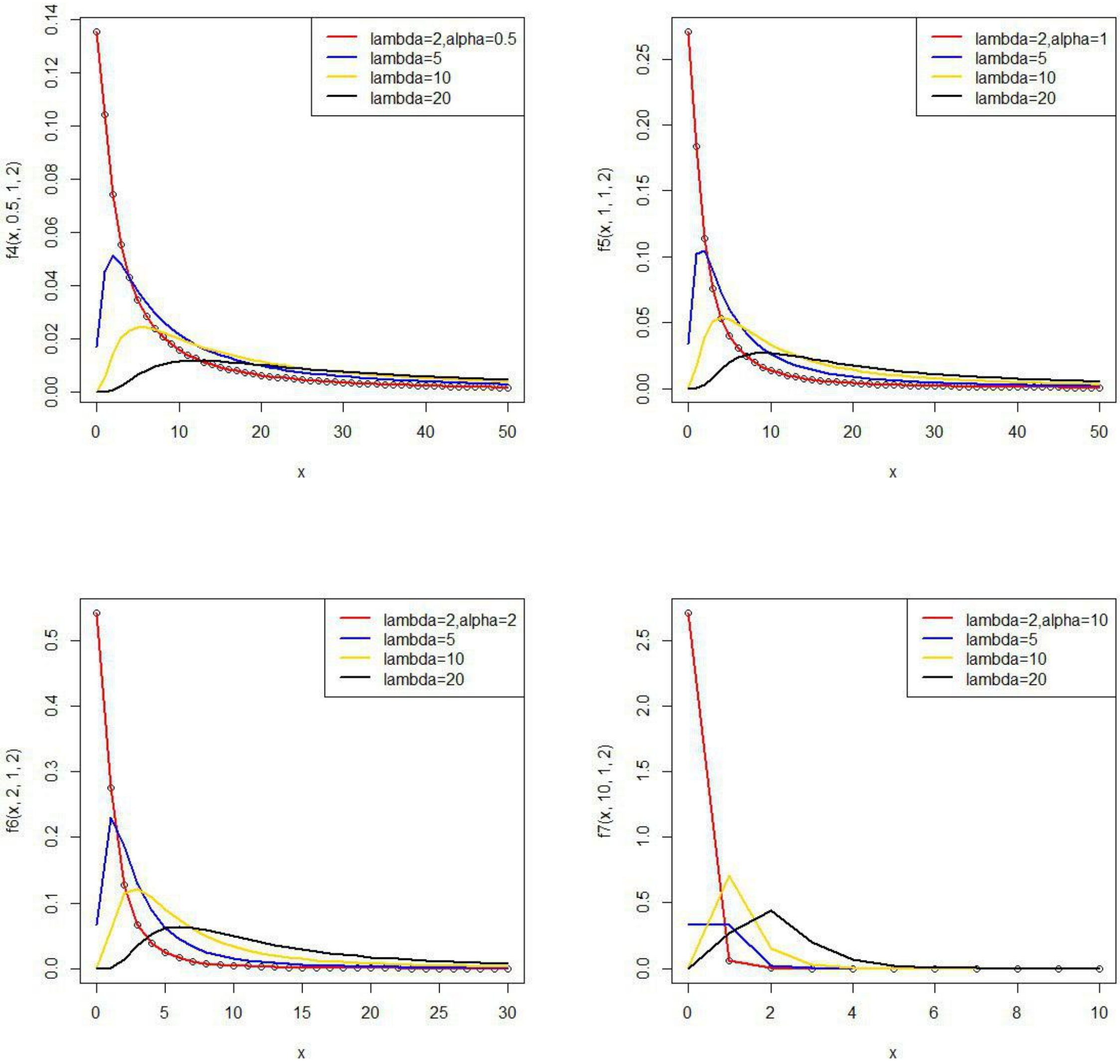
$$Var(S_N) = E(N)Var(X) + Var(N)(E(X))^2 + E(N(N-1))COV(X_i, X_j)$$

$$= \lambda \cdot \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} + \lambda \cdot \frac{\beta^2}{(\alpha-1)^2} + \lambda^2 \cdot \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

$$= \frac{\lambda\alpha\beta^2 + \lambda\alpha\beta^2 - 2\lambda\beta^2 + \lambda^2\beta^2}{(\alpha-1)^2(\alpha-2)}$$

$$= \frac{\lambda\beta^2(2\alpha - 2 + \lambda)}{(\alpha-1)^2(\alpha-2)}$$

□



شکل ۲.۳: تابع توزیع احتمالی پارتو_ پواسون

۵.۳ توزیع ترکیبی پارتو_ دو جمله‌ای منفی

تصور نمایید N توزیع دو جمله‌ای منفی با تابع احتمال جرم

$$P_r(N=n)=\frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)}\cdot P^r(1-P)^n \quad n=0,1,2... \qquad (1)$$

باشد. قضیه زیر را دایم:

قضیه ۱.۵.۳. [۹] فرض کنیم که یک توزیع دو جمله‌ای منفی با پارامتر r و پارامتر P و تابع احتمال جرم (۱) به عنوان توزیع اصلی ارائه شود و $(X_1,...,X_n)$ تعریف شده در ۱.۳.۲ باشد،

تابع توزیع احتمالی متغیر تصادفی S_N به صورت زیر به دست می‌آید:

$$f_{S_N}(x;\alpha,\beta,r,P)=\frac{r\alpha(1-P)P^r}{\beta(1+\frac{x}{\beta})^{\alpha+1}}{}_2F_1[1+r,1+\alpha;2,\frac{(1-P)\frac{x}{\beta}}{1+\frac{x}{\beta}}]\quad ,x>0$$

و $f_{S_N}=(0;\alpha,\beta,r,P)=P^r$.

${}_2F_1(a,b;c;Z)$ نشان دهنده تابع فوق هندسی گاؤس است، که به صورت زیر تعریف می‌شود:

$${}_2F_1[a,b;c;Z]=\sum_{n=0}^{\infty}\frac{(a)_n(b)_nZ^n}{(c)_n n!}$$

که $(a)_n$ نشان دهنده نماد باز شده $(a)_n=a(a-1)...(a-n+1)$ می‌باشد.

برهان. اگر $x=0$ باشد:

$$f_{S_N}(0)=\frac{\Gamma(r)}{\Gamma(1)\Gamma(r)}P^r(1-P)^0=P^r$$

و اگر $x>0$ باشد،

Z را به صورت زیر در نظر می‌گیریم:

$$Z=\frac{(1-P)\frac{X}{\beta}}{1+\frac{X}{\beta}}$$

بنابراین:

$$f_{S_N}(x)=\sum_{n=1}^{\infty}f_x^{(n)}P_r(N=n)$$

$$\Longrightarrow\sum_{n=1}^{\infty}\frac{x^{n-1}}{\beta^nB(n,\alpha)(1+\frac{x}{\beta})^{n+\alpha}}\cdot\frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)}\cdot P^r(1-P)^n$$

$$=\sum_{n=1}^{\infty}\frac{x^{n-1}\Gamma(n+\alpha)}{\beta^n\Gamma(n)\Gamma(\alpha)(1+\frac{x}{\beta})^{n+\alpha}}\cdot\frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)}\cdot P^r(1-P)^n$$

$$=\frac{x^{-1}P^r}{\Gamma(\alpha)\Gamma(r)(1+\frac{x}{\beta})^{\alpha}}\sum_{n=1}^{\infty}\frac{\Gamma(n+\alpha)\Gamma(n+r)x^n(1-P)^n}{\Gamma(n+1)\Gamma(r)\beta^n(1+\frac{x}{\beta})^n}$$

$$=\frac{x^{-1}P^r}{\Gamma(\alpha)\Gamma(r)(1+\frac{x}{\beta})^{\alpha}}\sum_{n=1}^{\infty}\frac{\Gamma(n+\alpha)\Gamma(n+r)Z^n}{n!\Gamma(r)}$$

$$=\frac{x^{-1}P^r}{\Gamma(\alpha)\Gamma(r)(1+\frac{x}{\beta})^{\alpha}}\cdot Z\cdot\Gamma(\alpha+1)\Gamma(r+1)$$

$$=\frac{x^{-1}P^r}{\Gamma(\alpha)\Gamma(r)(1+\frac{x}{\beta})^{\alpha}}\cdot\frac{(1-P)x}{\beta(1+\frac{X}{\beta})}\cdot\Gamma(\alpha+1)\Gamma(r+1)$$

$$=\frac{P^r(1-P)\alpha r}{\beta(1+\frac{x}{\beta})^{\alpha+1}}{}_2F_1[1+r,1+\alpha;2;\frac{\frac{x(1-P)}{\beta}}{1+\frac{x}{\beta}}]$$

□

۱.۵.۳ میانگین و واریانس مدل جمعی پارتو_دو جمله‌ای منفی

لم ۱.۵.۳. دوباره با استفاده از میانگین و واریانس گفته شده در لم ۱.۳.۳ میانگین و واریانس مدل جمعی پارتو_دو جمله‌ای منفی را به دست می‌آوریم:

- الف_

$$E(S_N) = \frac{r(1-P)\beta}{P(\alpha-1)} \qquad \alpha > 1$$

- ب_

$$Var(S_N) = \frac{r(1-P)\beta^2(1+P)(\alpha-1) + r(1-P)}{P^2(\alpha-1)^2(\alpha-2)}$$

برهان. برای اثبات الف داریم:

$$E(S_N) = E(X)E(N)$$

$$= \frac{\beta}{\alpha-1} \cdot \frac{r(1-P)}{P} = \frac{r(1-P)\beta}{P(\alpha-1)}$$

برای اثبات ب داریم:

$$Var(S_N) = E(N)Var(X) + Var(N)(E(X))^2 + E(N(N-1))COV(X_i, X_j)$$

$$= \frac{r(1-P)}{P} \cdot \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} + \frac{r(1-P)}{P^2} \cdot \frac{\beta^2}{(\alpha-1)^2} + \frac{r(1-P)}{P^2} \cdot \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} +$$

$$\frac{r^2(1-P)^2}{P^2} \cdot \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} - \frac{r(1-P)}{P} \cdot \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

$$= \frac{rP(1-P)\alpha\beta^2 + r(1-P)(\alpha-2)\beta^2 + r(1-P)\beta^2 + r^2\beta^2(1-P)^2 - r\beta^2P(1-P)}{P^2(\alpha-1)^2(\alpha-2)}$$

$$= \frac{r(1-P)\beta^2[P\alpha + \alpha - 2 + 1 + r(1-P) - P]}{P^2(\alpha-1)^2(\alpha-2)}$$

$$= \frac{r(1-P)\beta^2[P\alpha - P + \alpha - 1] + r(1-P)}{P^2(\alpha-1)^2(\alpha-2)}$$

$$= \frac{r(1-P)\beta^2[(P+1)(\alpha-1)] + r(1-P)}{P^2(\alpha-1)^2(\alpha-2)}$$

□

۶.۳ توزیع ترکیبی پارتو_ هندسی

در این قضیه تابع توزیع احتمال را در توزیع پارتو_ هندسی به دست می‌آوریم.

قضیه ۱.۶.۳. [۳] فرض کنیم یک توزیع هندسی با پارامتر و تابع احتمال جرم

$$P_r(N=n)=P(1-P)^n \quad , \quad n=0,1,..$$

به عنوان توزیع اصلی و $(X_1,...,X_n)$ تعریف شده در ۱.۳.۲ باشد، تابع توزیع احتمالی متغیر تصادفی S_N به صورت زیر به دست می‌آید:

$$f_{S_N}(x;\alpha,\beta,P)=\frac{\alpha P(1-P)}{\beta(1+\frac{Px}{\beta})^{\alpha+1}} \quad ,x>0$$

اگر $x=0$ باشد، $f_{S_N}(0;\alpha,\beta,P)=P$.

برهان. اگر $x=0$ باشد،

$$f_{S_N}(0)=P_r(N=0)=P(1-P)^0=P$$

اگر $x>0$ باشد ، Z را به صورت زیر در نظر می‌گیریم:

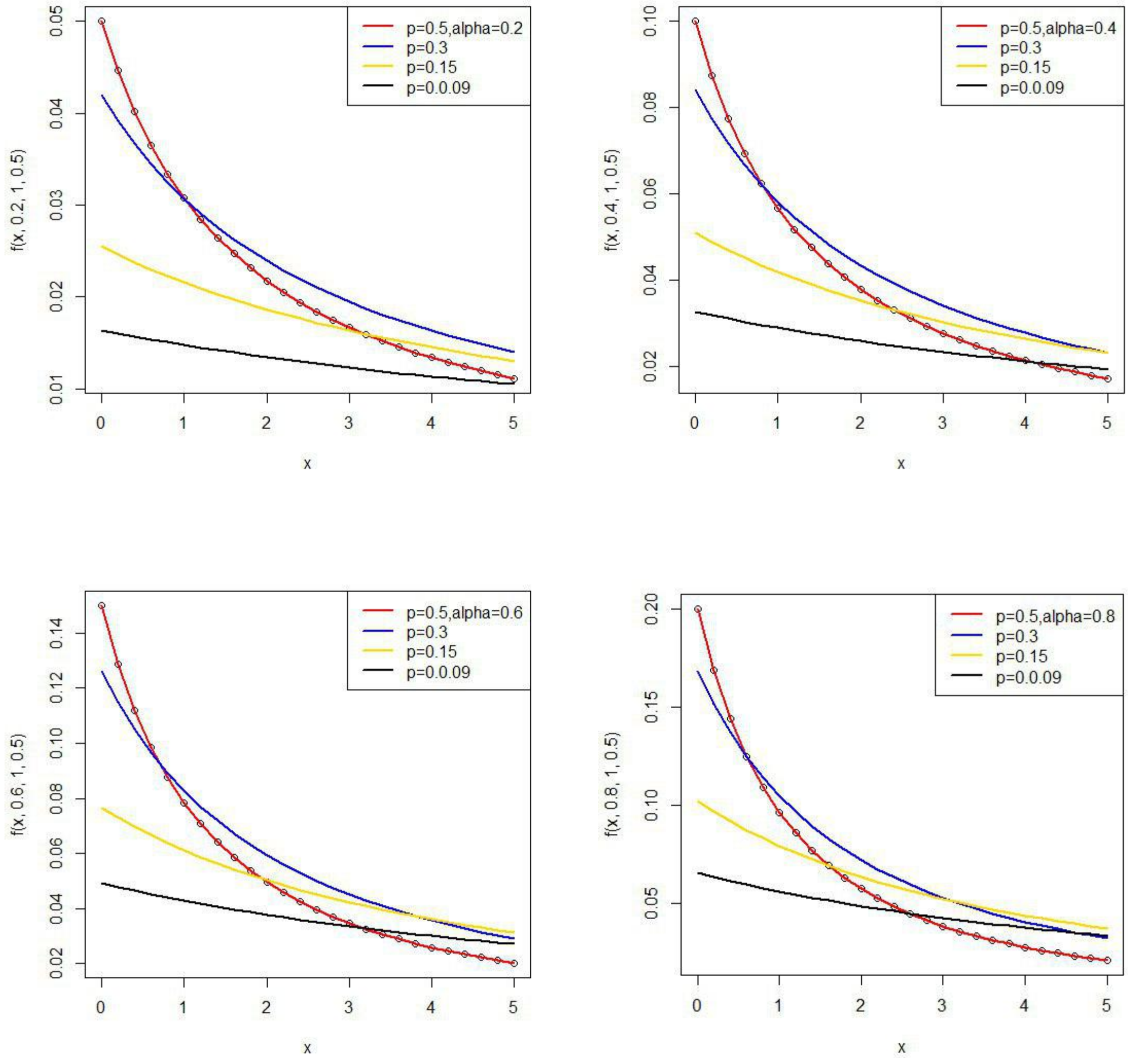
$$Z=\frac{\frac{x(1-P)}{\beta}}{1+\frac{x}{\beta}}$$

داریم:

$$\begin{aligned}
f_{S_N}(X)&=\sum_{n=1}^{\infty}f_X^{(n)}P_r(N=n)\\
&=\sum_{n=1}^{\infty}\frac{x^{n-1}}{\beta^n B(n,\alpha)(1+\frac{x}{\beta})^{n+\alpha}}\cdot P(1-P)^n\\
&=\frac{x^{-1}P}{\Gamma(\alpha)(1+\frac{x}{\beta})^{\alpha}}\sum_{n=1}^{\infty}\frac{x^n\Gamma(n+\alpha)(1-P)^n}{\Gamma(n)(1+\frac{x}{\beta})^n\beta^n}\\
&=\frac{x^{-1}P}{\Gamma(\alpha)(1+\frac{x}{\beta})^{\alpha}}\cdot Z\cdot\Gamma(\alpha+1)\\
&=\frac{x^{-1}P}{\Gamma(\alpha)(1+\frac{x}{\beta})^{\alpha}}\cdot\frac{\frac{x(1-P)}{\beta}}{1+\frac{x}{\beta}}\cdot\Gamma(\alpha+1)\\
&=\frac{\alpha P(1-P)}{\beta(1+\frac{x}{\beta})^{\alpha+1}}, \quad x>0.
\end{aligned}$$

□

شکل ۳.۳ نشان دهنده تابع توزیع احتمالی پارتو_ هندسی برای $\beta=1$، $P=0/09,0/15,0/3,0/35$ و $\alpha=0/2,0/4,0/6,0/8$ است.



شکل ۳.۳: تابع توزیع احتمالی پارتو ـ هندسی

فصل ۴

کاربرد عددی از تعمیم مدل‌های جمعی

جهت مقایسه عملکرد مدل‌های ارائه شده در این پایان نامه به بررسی داده های یکساله در شرکت بیمه خودرو در سال (۲۰۰۵ – ۲۰۰۴) پرداختیم.
این مجموعه داده‌ها در دانشکده اقتصاد و بازرگانی، دانشگاه مک کواری [۱] (سیدنی_ استرالیا) جمع آوری شده‌اند که توسط بسته $insuranceData$ در نرم افزار آماری R فراخوانی شده‌اند.

https://cran.r-project.org/web/packages/insuranceData/index.html

۱۰۰ مشاهده اول از این مجموعه داده ها را با مؤلفه‌های از چپ به راست : تعداد بیمه‌نامه ، تعداد ادعاهای خسارت و اندازه ادعاها ی خسارت در جدول ۱.۴ آورده‌ایم.
کل پرتفولیو دارای ۶۷۸۵۶ بیمه‌نامه است که ۴۶۲۴ مورد از آن ها حداقل نشان دهنده یک ادعای خسارت است.

برخی از آماره های توصیفی در جدول ۲.۴ آورده‌ایم:

- انحراف معیار بسیار زیاد است که بدین معناست که محاسبه حق بیمه تنها بر اساس میانگین اندازه ادعاهای خسارت کافی نیست.
- کوواریانس میان تعداد ادعاهای خسارت و اندازه ادعاهای خسارت دارای شکلی مثبت است و مقدار آن برابر ۱۴۱.۵۷۴است.

شکل ۱.۴ نشان دهنده تعداد کامل ادعاهای خسارت و مقدار کل آن ها می‌باشد.

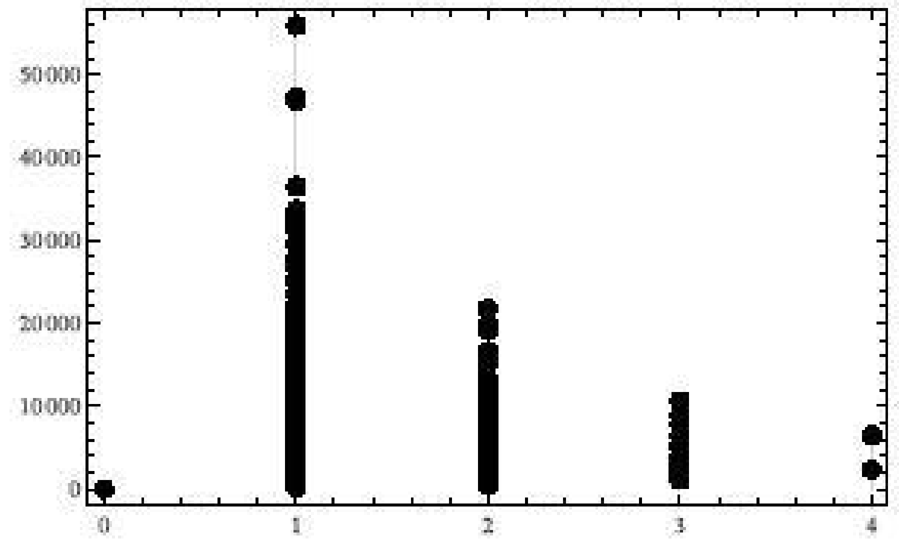
[۱] macquarie

از چپ به راست : تعداد بیمه‌نامه ، تعداد ادعاهای خسارت و اندازه ادعاها ی خسارت											
۱	۰	۰	۲۶	۰	۰	۵۱	۰	۰	۷۶	۰	۰
۲	۰	۰	۲۷	۰	۰	۵۲	۰	۰	۷۷	۰	۰
۳	۰	۰	۲۸	۰	۰	۵۳	۰	۰	۷۸	۰	۰
۴	۰	۰	۲۹	۰	۰	۵۴	۰	۰	۷۹	۰	۰
۵	۰	۰	۳۰	۰	۰	۵۵	۰	۰	۸۰	۰	۰
۶	۰	۰	۳۱	۰	۰	۵۶	۰	۰	۸۱	۰	۰
۷	۰	۰	۳۲	۰	۰	۵۷	۰	۰	۸۲	۰	۰
۸	۰	۰	۳۳	۰	۰	۵۸	۰	۰	۸۳	۰	۰
۹	۰	۰	۳۴	۰	۰	۵۹	۰	۰	۸۴	۰	۰
۱۰	۰	۰	۳۵	۰	۰	۶۰	۰	۰	۸۵	۰	۰
۱۱	۰	۰	۳۶	۰	۰	۶۱	۰	۰	۸۶	۰	۰
۱۲	۰	۰	۳۷	۰	۰	۶۲	۰	۰	۸۷	۰	۰
۱۳	۰	۰	۳۸	۰	۰	۶۳	۰	۰	۸۸	۰	۰
۱۴	۰	۰	۳۹	۰	۰	۶۴	۰	۰	۸۹	۰	۰
۱۵	۱	۶۶۹.۵۱	۴۰	۰	۰	۶۵	۱	۵۴۳۴.۴۴	۹۰	۰	۰
۱۶	۰	۰	۴۱	۲	۱۸۱۱.۷۱	۶۶	۱	۸۶۵.۷۹	۹۱	۰	۰
۱۷	۱	۸۰۶.۶۱	۴۲	۰	۰	۶۷	۰	۰	۹۲	۰	۰
۱۸	۱	۴۰۱.۸۰	۴۳	۰	۰	۶۸	۰	۰	۹۳	۰	۰
۱۹	۰	۰	۴۴	۰	۰	۶۹	۰	۰	۹۴	۰	۰
۲۰	۰	۰	۴۵	۰	۰	۷۰	۱	۸.۴۶۰.۰۰۰	۹۵	۰	۰
۲۱	۰	۰	۴۶	۰	۰	۷۱	۰	۰	۹۶	۱	۱۱۰۵.۷۷
۲۲	۰	۰	۴۷	۰	۰	۷۲	۰	۰	۹۷	۰	۰
۲۳	۰	۰	۴۸	۰	۰	۷۳	۰	۰	۹۸	۰	۰
۲۴	۰	۰	۴۹	۰	۰	۷۴	۰	۰	۹۹	۱	۲۰۰
۲۵	۰	۰	۵۰	۰	۰	۷۵	۰	۰	۱۰۰	۰	۰

جدول ۱.۴: ۱۰۰ مشاهده اولیه

برخی از اطلاعات توصیفی ادعا و اندازه ادعا برای مجموعه داده		
مقدار کل ادعاها	تعداد ادعاها	
۱۳۷.۲۷	۰.۰۷۲	میانگین
۱۰۵۶.۳۰	۰.۲۷۸	انحراف معیار
۰	۰	مینیمم
۵۵۹۲۲.۱۰	۴	ماکسیمم

جدول ۲.۴: آماره‌های توصیفی



شکل ۱.۴: تعداد و مقدار کل ادعاهای خسارت

- به روشنی واضح است که بیشترین مقدار ادعاها در نمونه ای مجزا وجود دارد در حالی که با افزایش تعداد ادعاهای خسارت شاهد کاهش مقدار آن‌ها خواهیم بود.
مدل تلفیق شده توزیع پواسون با توزیع نمایی را مورد بررسی قرار می‌دهیم:
شاید شناخته شده ترین مدل ادعاهای خسارت جمعی را زمانی بتوان بدست آورد که توزیع اصلی و فرعی توزیعات پواسون و نمایی باشند.
عمدتاً این امر به دلیل پیچیدگی مدل ریسک تجمعی تحت توزیعات احتمال های دیگر است، مانند توزیع پارتو و لگ_ نرمال

در این حالت رولسکی [۲] و همکارانش (۱۹۹۹) نشان دادند که تابع چگالی متغیر تصادفی توزیع پواسون_ نمایی در مجموع داده ها را می‌توان با معادله زیر به دست آورد:

$$f_S(x;\alpha,\lambda)=\sqrt{\frac{\lambda\alpha}{x}}exp(-\lambda-\alpha x)I_1(2\sqrt{\lambda\alpha x}),\quad x>0.$$

که $f_S(0;\alpha,\lambda)=exp(-\lambda)$ است.
در اینجا $\lambda>0$ و $\alpha>0$ به ترتیب پارامترهای توزیع های پواسون و نمایی هستند، و

$$I_\nu(z)=\sum_{k=0}^{\infty}\frac{(\frac{z}{2})^{2k+\nu}}{\Gamma(k+1)\Gamma(\nu+k+1)}\quad z,\nu\in R$$

این معادله نشان دهنده تابع اصلاح شده بسل [۳] نوع اول است.
علاوه بر این توزیع دو جمله‌ای منفی با پارامتر $r>0$ و $0<p<1$ را می‌توان به عنوان توزیع اصلی و توزیع نمایی را به عنوان توزیع فرعی دانست، که در این حالت تابع چگالی احتمال در متغیر تصادفی مجموع ادعاها را با معادله زیر نشان می‌دهیم:

$$f_S(x;\alpha,r,p)=\alpha rp^r(1-p)exp(-\alpha x)\,{}_1F_1[1+r;2;\alpha(1-p)x]\qquad ,x>0$$

که ${}_1F_1[.;.;.]$ تابع فوق هندسی و $f_S(0;\alpha,r,p)=p^r$ است.

- اگر در معادله بالا $r=1$ قرار دهیم، تابع چگالی احتمال مقدار کل ادعاهای خسارت بدست می‌آید که توزیع اصلی هندسی و توزیع فرعی از نوع نمایی می‌باشد.

$$f_S(x;\alpha,p)=\alpha p(1-p)exp(-\alpha px)\qquad ,x>0$$

که $f_S(0;\alpha,p)=p$ می‌باشد.

- در نهایت مدلی را مورد بررسی قرار می‌دهیم که توزیع اصلی در آن ترکیبی از توزیع دو جمله‌ای منفی با پارامتر های β و $\beta(\beta+\mu)$ ، $\mu>0$ ، $\beta>0$ و اینکه تابع چگالی احتمالی با پارامتر

[۲] Rolski
[۳] Bessel

میانگین μ ، در میان بیمه کنندگان سبد سهام بیمه،تابع توزیع پارتو با تابع چگالی احتمال به دست آمده ۱.۳.۲ در نظر می‌گیریم. مجدداً توزیع فرعی را توزیع نمایی با پارامتر $\lambda > ۰$ در نظر می‌گیریم. تابع توزیع احتمال برای متغیرهای تصادفی $S_N = X_۱ + ... + X_N$ مقدار کل ادعاهای خسارت را با معادله زیر نشان می‌دهیم:

$$f_S(x;\alpha,\beta,\lambda) = \frac{\alpha\beta\lambda exp(-\lambda x) {}_۱F_۱(\beta+۱;\alpha+\beta+۲,\lambda x)}{(\alpha+\beta+۱)(\alpha+\beta)} \qquad ,\alpha > ۰$$

که $f_S(۰;\alpha,\beta) = \frac{\alpha}{(\alpha+\beta)}$ می‌باشد.
برآوردها را می توان با حل مجموعی از معادله های ساده بدست آورد. علاوه بر این پارامترهای مدل‌های مختلف در رابطه را با مقدار کل ادعاهای خسارت را می‌توان از طریق بیشترین احتمال برآورد کرد.
نمونه پارتو_ هندسی را بررسی می‌کنیم:
برای انجام این کار نمونه ی تصادفی $x_۱, ..., x_n$ را بکار می بریم.تابع لگ_احتمال را می‌توان به شکل زیر نوشت:

$$f(x,\alpha,\beta,P) = \frac{\alpha P(۱-P)}{\beta(۱+\frac{Px}{\beta})^{\alpha+۱}} \quad ,x > ۰ \qquad ,f(۰;\alpha,\beta,P) = P$$

اگر $n_۰$ مشاهدات اولیه ، $n - n_۰$ تعداد نمونه های مشاهده شده و n اندازه نمونه باشد، آنگاه داریم:

$$f(x_۱,...,x_n) = \prod_{i=۱}^{n-n_۰} \frac{\alpha P(۱-P)}{\beta(۱+\frac{Px_i}{\beta})^{\alpha+۱}} \cdot \prod_{i=۱}^{n_۰} P$$

$$= \frac{\alpha^{n-n_۰} P^{n-n_۰} (۱-P)^{n-n_۰}}{\beta^{n-n_۰}} \prod_{i=۱}^{n-n_۰} \frac{۱}{(۱+\frac{Px_i}{\beta})^{\alpha+۱}} \cdot P^{n_۰}$$

از طرفین لگاریتم می‌گیریم:

$$\ell(\alpha,\beta,p) = n_۰ \log p + (n-n_۰)[\log\alpha + \log p$$
$$+ \log(۱-p) + \alpha\log\beta] - (\alpha+۱)\sum_{x_i>۰}\log(\beta + px_i)$$

به این ترتیب به منظور ساده سازی محاسبات مقدار کل ادعاهای خسارت تقسیم بر ۱۰۰۰ است.
در نتیجه‌ی این معادله نمی‌توانیم برآورد حداکثر احتمال را به راحتی حل کنیم، آن ها باید به صورت عددی حل شوند یا به صورتی که اینجا به حل آن پرداختیم به طور مستقیم با حداکثر رساندن تابع لگ_ احتمال حل شوند.
از آنجا که برای حداکثر سطح لگ_احتمال ضمانتی وجود ندارد، اختلاف مقدار اولیه پارامتر فضای مورد بررسی به صورت نقطه‌ای محاسبه می‌شود.

الگوریتم های حداکثرسازی مختلف مانند نیوتن [۴] و پرنسی پلکسس [۵] و کوآسی نیوتن [۶] برای اطمینان از اینکه نتایج ما درست هستند مورد استفاده قرار گرفته است.
خلاصه ای از نتایج را می‌توانید در جدول ۳.۴ ببینید.

توزیع اصلی	توزیع فرعی	$\hat{r}$	$\hat{p}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	AIC
پواسون	نمایی			۰.۵۰۰۶۵	۰.۰۰۸۱۷		۵۳۳۷۱.۸
هندسی	نمایی		۰.۹۹۹۰۱		۰.۶۸۷۹۴		۵۳۶۰۸.۲۲
دو جمله‌ای منفی	نمایی	۰.۰۱۸۳	۰.۳۵۰۲		۰.۹۶۰۰		۵۳۵۵۱.۷۳
دوجمله‌ای منفی-پارتو	نمایی			۰.۸۰۳۴	۲.۰۱۲۳	۰.۰۱۲۰	۵۱۱۵۷.۵۳
پواسون	پارتو			۰.۵۷۵۲	۳.۰۲۳۶	۰.۷۲۷۶	۴۸۷۵۸.۵۱
هندسی	پارتو		۰.۹۹۱۵		۲.۲۵۶	۱.۵۸۹	۳۴۵۷۰.۱۴
دو جمله‌ای منفی	پارتو	۰.۱۰۲۶	۰.۱۱۰۰۰		۰.۳۹۰۰	۱.۰۰۰۱	۴۸۷۲۸.۱۹

جدول ۳.۴: برآورد ارزش پارامترها با ۶۷۸۵۶ داده

در جدول ۳.۴ برآورد ارزش پارامترها همراه و در جدول ۴.۴ خطای استاندارد آن‌ها نشان داده شده، همچنین AIC را نیز نشان می‌دهند.
AIC معیاری برای سنجش نیکویی برازش است. با توجه به زمانی که چند مدل رقیب با توجه به مقدار AIC رتبه‌بندی شوند، مدل دارای کمترین AIC بهترین است. بوذوگان [۷] (۱۹۸۷) مشاهده نمود که AIC وابستگی مستقیمی به اندازه نمونه ندارد.
نتایج ما نشان می‌دهند که مدل ترکیبی پارتو دارای عملکرد بهتری از مدل‌های دیگری که در اینجا استفاده شده است دارد.
نمودارهای ۲.۴ چهار مدلی می‌باشد که با استفاده از برآوردهای جدول ۳.۴ ترسیم شده‌است.
همانگونه که مشاهده می‌کنید مدل‌های ترکیب شده با پارتو در مقایسه با مدل‌های دیگری از

[۴] Newton
[۵] principalaxis
[۶] Quasinewton
[۷] Bozdogan

توزیع اصلی	توزیع فرعی	$\hat{r}$	$\hat{p}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$
پواسون	نمایی			۰.۰۰۱۰۴	۰.۰۰۷۵۹	
هندسی	نمایی		۰.۰۰۰۹۷		۰.۰۰۷۸۵	
دو جمله‌ای منفی	نمایی	۰.۰۰۰۰	۰.۰۰۰۰۰		۰.۰۰۴۷۱	
جمله‌ای منفی پارتو–دو	نمایی			۰.۰۱۵۹۸	۰.۱۲۲۹۱	۰.۰۰۹۵۸
پواسون	پارتو			۰.۰۰۱۰۲	۰.۰۰۹۷۴	۰.۰۴۸۷۹
هندسی	پارتو		۰.۰۰۰۹۷		۰.۰۸۸۲۸	۰.۱۲۴۰۷
دو جمله‌ای منفی	پارتو	۰.۲۱۵۰۵	۰.۱۲۰۵۲		۰.۰۹۰۰۰	۰.۱۷۳۲۴

جدول ۴.۴: خطای استاندارد

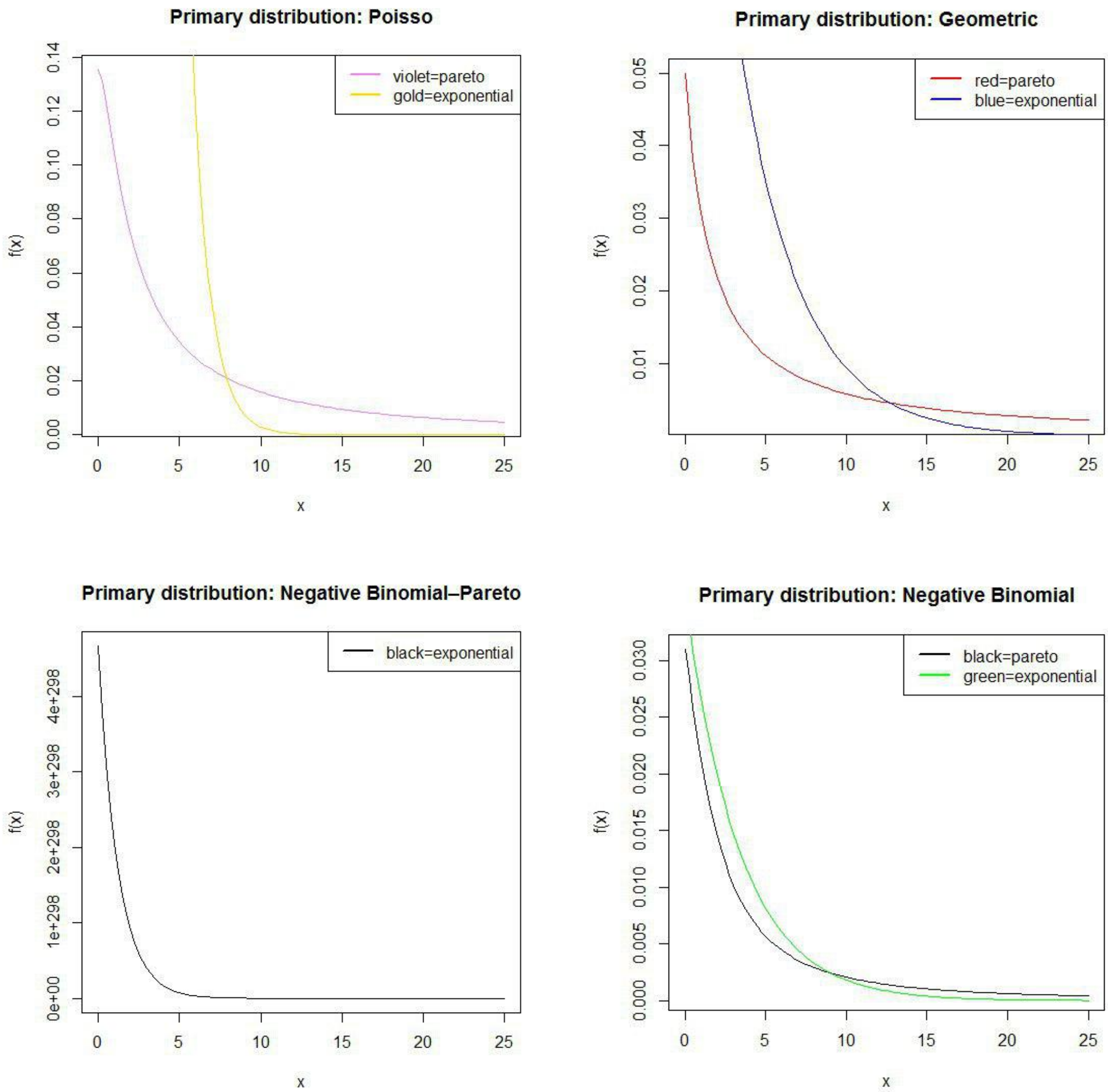
کشیدگی بیشتری برخوردار است. علاوه بر این مدل‌های ترکیبی با پارتو در مقایسه با مدل‌هایی که بر اساس توزیع فرعی نمایی می‌باشند با روند آرامتری به سمت صفر می‌روند.

استفاده از داده های حقیقی

به بررسی داده‌های یک ساله در شرکت بیمه ایران در سال ۹۳ می‌پردازیم.
به دلیل محرمانه بودن اطلاعات بیمه با رایزنی‌هایی که صورت گرفت فقط به تعداد ۱۰۰ داده دسترسی پیدا کردیم. روندهای انجام شده برای ۶۷۸۵۶ داده را روی این ۱۰۰ داده انجام میدهیم و نتایج را مقایسه می‌کنیم.
جدول ۵.۴ ۱۰۰ داده از شرکت بیمه ایران می‌باشد.
به ترتیب از چپ به راست : تعداد بیمه‌نامه ، تعداد ادعاهای خسارت و اندازه ادعاهای خسارت می‌باشد.
در جدول ۶.۴ برآورد ارزش پارامترهای شرکت بیمه ایران آورده شده است.
همانطور که در جدول ۶.۴ ملاحظه می‌کنید مدل های ترکیب شده با توزیع پارتو دارای کمترین AIC و در نتیجه دارای عملکرد بهتری نسبت به مدل‌های دیگری که در اینجا استفاده شده می‌باشد.
نمودار ۳.۴ چهار مدلی می‌باشد که با استفاده از برآوردهای شرکت بیمه ایران ترسیم شده‌است.

از چپ به راست : تعداد بیمه‌نامه ، تعداد ادعاهای خسارت و اندازه ادعاها ی خسارت											
۱	۲	۳۶.۸۱۲.۰۰۰	۲۶	۰	۰	۵۱	۱	۸.۴۸۹.۰۰۰	۷۶	۰	۰
۲	۰	۰	۲۷	۰	۰	۵۲	۰	۰	۷۷	۱	۵.۸۸۲.۰۰۰
۳	۰	۰	۲۸	۰	۰	۵۳	۱	۷.۳۵۵.۰۰۰	۷۸	۰	۰
۴	۰	۰	۲۹	۰	۰	۵۴	۰	۰	۷۹	۰	۰
۵	۰	۰	۳۰	۰	۰	۵۵	۰	۰	۸۰	۰	۰
۶	۰	۰	۳۱	۰	۰	۵۶	۰	۰	۸۱	۱	۱۱.۴۴۸.۰۰۰
۷	۱	۸.۵۵۰.۰۰۰	۳۲	۰	۰	۵۷	۰	۰	۸۲	۰	۰
۸	۰	۰	۳۳	۰	۰	۵۸	۰	۰	۸۳	۰	۰
۹	۰	۰	۳۴	۰	۰	۵۹	۰	۰	۸۴	۰	۰
۱۰	۰	۰	۳۵	۰	۰	۶۰	۰	۰	۸۵	۰	۰
۱۱	۰	۰	۳۶	۰	۰	۶۱	۰	۰	۸۶	۰	۰
۱۲	۰	۰	۳۷	۰	۰	۶۲	۰	۰	۸۷	۰	۰
۱۳	۰	۰	۳۸	۰	۰	۶۳	۰	۰	۸۸	۱	۱۹.۰۵۳.۰۰۰
۱۴	۱	۱۰.۷۵۵.۰۰۰	۳۹	۰	۰	۶۴	۰	۰	۸۹	۰	۰
۱۵	۰	۰	۴۰	۰	۰	۶۵	۱	۵.۵۶۲.۰۰۰	۹۰	۰	۰
۱۶	۰	۰	۴۱	۰	۰	۶۶	۱	۴.۵۶۸.۰۰۰	۹۱	۰	۰
۱۷	۰	۰	۴۲	۰	۰	۶۷	۰	۰	۹۲	۰	۰
۱۸	۰	۰	۴۳	۰	۰	۶۸	۰	۰	۹۳	۰	۰
۱۹	۰	۰	۴۴	۱	۱۱۵.۹۸۱.۰۰۰	۶۹	۰	۰	۹۴	۰	۰
۲۰	۰	۰	۴۵	۰	۰	۷۰	۱	۸.۴۶۰.۰۰۰	۹۵	۰	۰
۲۱	۰	۰	۴۶	۱	۲۳.۷۸۰.۰۰۰	۷۱	۰	۰	۹۶	۰	۰
۲۲	۰	۰	۴۷	۰	۰	۷۲	۰	۰	۹۷	۰	۰
۲۳	۰	۰	۴۸	۰	۰	۷۳	۰	۰	۹۸	۰	۰
۲۴	۰	۰	۴۹	۰	۰	۷۴	۰	۰	۹۹	۰	۰
۲۵	۰	۰	۵۰	۰	۰	۷۵	۰	۰	۱۰۰	۲	۸.۵۳۰.۰۰۰

جدول ۵.۴: داده‌های شرکت بیمه ایران



شکل ۲.۴: تابع چگالی احتمال توزیع‌های ترکیبی با برآورد پارامترها

تقریباً همانند نمودارهای ۲.۴ مدل‌های ترکیب شده با پارتو در مقایسه با مدل‌های دیگر کشیده‌تر می‌باشد و با روند آرامتری به سمت صفر می‌روند. این مقدار ناچیز خطا هم به دلیل کم بودن تعداد داده‌های ما و تعداد زیاد صفر در داده‌ها می‌باشد.

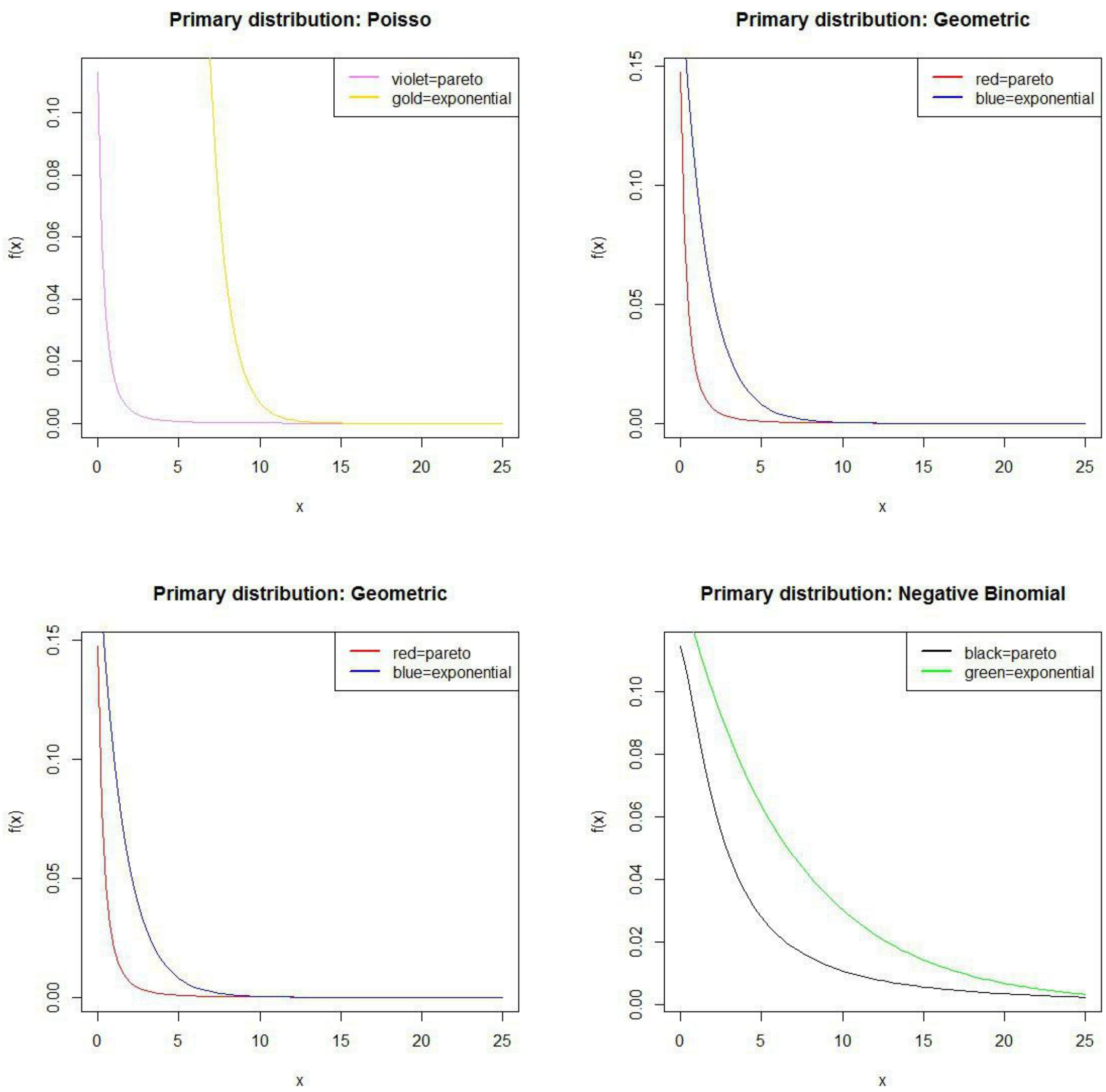
نتیجه‌گیری

سرانجام با بررسی هایی که روی داده‌هایی که براساس بیمه‌نامه‌های یک ساله در سال (۲۰۰۵ – ۲۰۰۴) و داده‌های شرکت بیمه ایران در سال ۹۳ انجام شد، که دقیقا نتایج یکسانی از این بررسی‌ها گرفتیم، به این نتیجه رسیدیم که مدل‌های جمعی وابسته عملکرد مطلوبتری

توزیع اصلی	توزیع فرعی	$\hat{r}$	$\hat{p}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	AIC
پواسون	نمایی			۰.۸۷۲۲	۰.۰۱۲۳۲		۷۲۷۶.۶۰
هندسی	نمایی		۰.۸۶۰۰۰		۱.۰۱۷۴۴		۱۱۲۷۷.۱۶
دو جمله‌ای منفی	نمایی	۲.۲۱۹۴	۰.۰۲۱۰۳		۲.۰۳۰۰۰		۱۵۱۷۲۱.۴
دوجمله‌ای منفی-پارتو	نمایی			۰.۰۳۵۲۴	۲.۸۴۳۷	۰.۲۰۷۲	۹۸۴۹.۸۵
پواسون	پارتو			۰.۰۶۰۰	۲.۰۲۸۳۱	۰.۰۳۸۸۵۶	-۱۹۹۰.۶۹
هندسی	پارتو		۰.۹۲۴۹۴		۲.۹۲۰۸۳	۰.۰۶۰۷۴	-۱۵۸۷.۰۴
دو جمله‌ای منفی	پارتو	۲.۰۳۶۵	۰.۷۲۳۷		۲.۰۴۰۲	۱.۹۱۰۳	-۳۴۲۶.۴۲

جدول ۶.۴: برآورد ارزش پارامترها با داده‌های شرکت بیمه ایران

نسبت به مدل‌های جمعی دیگر دارند.



شکل ۳.۴: تابع چگالی احتمال توزیع‌های ترکیبی با برآورد پارامترهای شرکت بیمه ایران

پیوست آ

کدنویسی با نرم افزار R

```
#### baravord Pareto-geometric######
rm(list=ls())
library(maxLik)
library(insuranceData)
data(dataCar)
x=as.matrix(dataCar)
x1=x[,4]
z=c()
logLikFun2<- function(param) {
p<- param[1]
alfa<- param[2]
beta<- param[3]
x1=x[,4]
c=as.numeric(x1)
N=length(c)
```

```
n0=length(which(c==0))
z<-n0*log(p)+(N-n0)*(log(alfa)+log(p)+log(1-p)+alfa*log(beta))-(alfa+1)*sum(log(beta+p*
z
}
mle1<- maxLik(logLik = logLikFun2,start = c(p=0.01, alfa =0.1, beta=0.9))
summary(mle1)
AIC(mle1)
#######baravorde namaee-hendesi###############
rm(list=ls())
library(maxLik)
require(stats4)
library(insuranceData)
data(dataCar)
x=as.matrix(dataCar)
x1=as.matrix(x)
z=c()
logLikFun <- function(param) {
p<- param[1]
alfa<- param[2]
x1=x[,4]
c=as.numeric(x1)
N=length(c)
n0=length(which(c==0))
z<-n0*log(p)+(N-n0)*log(alfa)+(N-n0)*log(1-p)+(N-n0)*log(p)-(alfa*p)*sum(c)
z
}
mle <- maxLik(logLik = logLikFun,start = c(p=0.5, alfa = 0.82))
summary(mle)
AIC(mle)
###plot compound poisson-pareto###
f5(x,1,1,2)
k <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
```

```
labels <- c("lambda=2,alpha=1", "lambda=5", "lambda=10", "lambda=20")
# plot Z first
plot(x, f5(x,1,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f5(x,1,1,k[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
#######

x <- seq(0, 30)
f6=function(x,a,b,k){
g=((k*x)/(b))/(1+x*b)
h= hypergeom1F1(g, a+1, 2)
s=((a*k*exp(-k))/(b*(1+(x)/b))^(a+1))*h
s
}

f6(x,2,1,2)
k <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("lambda=2,alpha=2", "lambda=5", "lambda=10", "lambda=20")
# plot Z first
plot(x, f6(x,2,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f6(x,2,1,k[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
```

```
##########
x <- seq(0, 10)
f7=function(x,a,b,k){
g=((k*x)/(b))/(1+x*b)
h= hypergeom1F1(g, a+1, 2)
s=((a*k*exp(-k))/(b*(1+(x)/b))^(a+1))*h
s
}

f7(x,10,1,2)
k <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("lambda=2,alpha=10", "lambda=5", "lambda=10", "lambda=20")
# plot Z first
plot(x, f7(x,10,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f7(x,10,1,k[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
################################3
###########fig3##################
###############################################
library(BMS)
x <- seq(0, 5,0.2)
f1=function(x,a,b,p){
k=(a*p*(1-p))/(b*(1+((p*x)/b))^(1+a))
}
p<- c(0.2, 0.4, 0.6,0.8)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
```

```
labels <- c("p=5,alpha=3", "p=3", "p=1.5", "p=0.5")
# plot Z first
plot(x, f1(x,5,1,0.2))
lines(x, f1(x,5,1,0.2), lwd=2, col=colors[i])
plot(x, f1(x,5,1,0.4))
lines(x, f1(x,5,1,0.4), lwd=2, col=colors[i])
lines(x, f1(x,5,1,0.6), lwd=2, col=colors[i])
lines(x, f1(x,5,1,0.8), lwd=2, col=colors[i])

# add lines representing t-distribution
for (i in 1:4){
lines(x, f1(x,5,1,p[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
```

```
library(ggfortify)
library(extraDistr)
x <- seq(0,80)
# try different degrees of freedom
n <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("n=2,alpha=0.5", "n=5", "n=10", "n=20")
# plot Z first
plot(x, f(x,0.5,1,n))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.5,1,n[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright", title="t-Distribution",
```

```
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
######################################################
plot(x, dpios(x), type="l", lty=2, xlab="x value",
ylab="Density", main="Comparison of t Distributions")
#########################################################
x <- seq(0, 80)
f=function(x,a,b,n){
g=beta(n,a)
d=(x^(n-1))/((b^n)*g*((1+(x/b))^(n+a)))
d
}
n <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("n=2,alpha=0.5", "n=5", "n=10", "n=20")
# plot Z first
plot(x, f(x,0.5,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.5,1,n[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
#################################
x <- seq(0, 30)
f=function(x,a,b,n){
g=beta(n,a)
d=(x^(n-1))/((b^n)*g*((1+(x/b))^(n+a)))
d
}
n <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
```

```
labels <- c("n=2,alpha=1", "n=5", "n=10", "n=20")
# plot Z first
plot(x, f(x,1,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,1,1,n[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
############################################
x <- seq(0, 20)
f=function(x,a,b,n){
g=beta(n,a)
d=(x^(n-1))/((b^n)*g*((1+(x/b))^(n+a)))
d
}
f(x,2,1,2)
f(x,2,1,5)
f(x,2,1,10)
f(x,2,1,20)
n <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("n=2,alpha=2", "n=5", "n=10", "n=20")
# plot Z first
plot(x, f(x,2,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,2,1,n[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
```

```
###############################
################################
x <- seq(0, 10)
f=function(x,a,b,n){
g=beta(n,a)
d=(x^(n-1))/((b^n)*g*((1+(x/b))^(n+a)))
d
}
n <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
# plot Z first
plot(x, f(x,10,1,2))
lines(x, f(x,10,1,2), lwd=2, col="red")
plot(x, f(x,10,1,5))
lines(x, f(x,10,1,5), lwd=2, col="blue")
plot(x, f(x,10,1,10))
lines(x, f(x,10,1,10), lwd=2, col="gold")
plot(x, f(x,10,1,20))
lines(x, f(x,10,1,20), lwd=2, col="black")
# add lines representing t-distribution
# add legend
colors <- c("red", "blue", "gold", "black")
labels <- c("n=2,alpha=10", "n=5", "n=10", "n=20")
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)

###############################################
##############FIGUR222###############
##############################################
require("hypergeo")
require(graphics)
library(hypergeo)
library(CharFun)
x <- seq(0, 50)
```

```
f4=function(x,a,b,k){
g=((k*x)/(b))/(1+x*b)
h= hypergeom1F1(g, a+1, 2)
s=((a*k*exp(-k))/(b*(1+(x)/b))^(a+1))*h
s
}

f4(x,0.5,1,2)
k <- c(2, 5, 10, 20)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("lambda=2,alpha=0.5", "lambda=5", "lambda=10", "lambda=20")
# plot Z first
plot(x, f4(x,0.5,1,2))
# add lines representing t-distribution
for (i in 1:4){
lines(x, f4(x,0.5,1,k[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)
########
x <- seq(0, 50)
f5=function(x,a,b,k){
g=((k*x)/(b))/(1+x*b)
h= hypergeom1F1(g, a+1, 2)
s=((a*k*exp(-k))/(b*(1+(x)/b))^(a+1))*h
s
}


}
else{
z=NA
```

```
}
sum(z)
}
mle1<- maxLik(logLik = logLikFun2,start = c(p=0.93, alfa =2.04, beta=2.03))
summary(mle1)
AIC(mle1)


rm(list=ls())
library(maxLik)
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)
z=c()
logLikFun <- function(param) {
p<- param[1]
alfa<- param[2]
x=x1[,1]
c=as.numeric(x)
N=length(c)
n0=length(which(c==0))
for(i in 1:N){
if(p>=0&&alfa>=0){
z[i]<-N*log(p)+(N-n0)*(log(alfa)+log(1-p))-(alfa*p)*sum(c)
}
else{
z[i]=NA
}}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(p= 0.92, alfa = 0.53))
summary(mle)
AIC(mle)
```

```
#########negativebinomial-expossional#########
rm(list=ls())
library(maxLik)
require("CharFun")
require("stats4")
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)
z=c()
logLikFun <- function(param) {
p<- param[1]
alfa<- param[2]
r<- param[3]
x1=x[,1]
c=as.numeric(x1)
N=length(c)
N0=length(which(c==0))
for(i in 1:N){
if(p>=0&&alfa>=0&&r>=0){
#h=hypergeom1F1(r+1,2,r*(1-p)*c)
z[i]<-(N-N0)*(log(alfa)+log(r)-log(1-p))-N0*r*log(p)
}
else {z[i]=NA}
}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(p= 0.5, alfa = 2.03, r=2.12))
summary(mle)
AIC(mle)


######pareto-negative binomial#################
rm(list=ls())
library(maxLik)
require("BMS")
```

```
require("hypergeo")
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)
z=c()
logLikFun <- function(param) {
r<- param[1]
alfa<- param[2]
beta<- param[3]
p<-param[4]
x1=x[,1]
c=as.numeric(x1)
N=length(c)
n0=length(which(c==0))
#g=((1-p)*(c/beta))/(1+(c/beta))
#h=hypergeo(r+1,alfa+1,2,g)
#r=abs(log(h))
#q=sum(log(1/(1+(N/beta))^(alfa+1)))
#d=abs(q)
for(i in 1:N){
if(p>=0&&alfa>=0&&r>=0&&beta>=0){
z[i]<-N*r*log(p)+(N-n0)*(log(alfa)+log(1-p)-log(beta))
}
else{z[i]=NA}
}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(r=0.31, alfa = 2.04, beta=1.91,p=0.81 ))
summary(mle)
AIC(mle)


#######################
#### piosson- namaee###########
rm(list=ls())
```

```
library(maxLik)
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)

z=c()
logLikFun <- function(param) {
lambda<- param[1]
alfa<- param[2]
x=x1[,1]
c=as.numeric(x)
N=length(c)
N0=length(which(c==0))
k=10
a=gamma(k+1)
b=gamma(k+2)
for(i in 1:N){
if(alfa>=0&&lambda>=0){
#f=sum(log(((alfa*lambda)^k)/(a*b)))
z[i]<-(N-N0)*(-lambda-alfa*c)-N0*lambda
}
else{z[i]=NA
}
}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(lambda= 0.12, alfa = 0.87))
summary(mle)
AIC(mle)


##########pareto-poisson#######
rm(list=ls())
library(maxLik)
require("CharFun")
```

```
require("stats4")
###############
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)
######################TREUE
z=c()
logLikFun <- function(param) {
lambda<- param[1]
alfa<- param[2]
beta<- param[3]
x1=x[,1]
c=as.numeric(x1)
N=length(c)
n0=length(which(c==0))
#q=sum(log(1/(1+(c/beta))^(alfa+1)))
#g=(((lambda*c)/beta)/(1+c*beta))
#f=qhyper(p, m, n, k, lower.tail = TRUE, log.p = TRUE)
#h=hypergeom1F1(g,alfa+1,2)
for(i in 1:N){
if(lambda>=0&&alfa>=0&&beta>=0){
z[i]<-(N-n0)*(log(alfa)+log(lambda)-log(beta))-(lambda*N)
}
else{z[i]=NA
}
}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(lambda= 0.06, alfa = 2.03, beta=2.12))
summary(mle)
AIC(mle)


####negative pareto expotional#########
rm(list=ls())
```

```
library(maxLik)
require("BMS")
require("hypergeo")
require("CharFun")
require("stats4")
x=read.csv("C:/bime.csv",head=FALSE)
x1=as.matrix(x)
x2=x1[,1]
z=c()
logLikFun <- function(param) {
lambda<- param[1]
alfa<- param[2]
beta<- param[3]
x=x1[,1]
c=as.numeric(x)
N=length(c)
N0=length(which(c==0))
for(i in 1:N){
if(lambda>=0&&alfa>=0){
#b=sum(hypergeom1F1(beta+1,alfa+beta+2,lambda*c))
z[i]<-N*(log(alfa)-log(alfa+beta))+(N-N0)*(log(beta)-log(alfa+beta+1))-sum(lambda*c)
}
else{ z[i]=NA
}
}
sum(z)
}
mle <- maxLik(logLik = logLikFun,start = c(lambda= 0.82, alfa = 2.84, beta=0.207))
summary(mle)
AIC(mle)


#####poisson- pareto#####
library(hypergeo)
```

```
library(CharFun)
x <- seq(0,25,1)
f=function(x,a,b,k){
g=((k*x)/(b))/(1+x*b)
h= hypergeom1F1(g, a+1, 2)
s=((a*k*exp(-k))/(b*(1+(x)/b))^(a+1))*h
s
}
f(x,0.5,1,2)
f(x,2.04,2.13,0.07)
curve(f(x,2.04,2.13,0.07)
,from=0,to=25, col = "violet",main="Primary distribution: Poisso")
curve(f(x,0.5,1,2)
,from=0,to=25, col = "violet",main="Primary distribution: Poisso",ylab="f(x)")

colors <- c("violet", "gold")
labels <- c("violet=pareto","gold=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)
#######poisson-namaee############
x <- seq(0,25,1)
f1=function(x,alfa,lambda,k){
g=gamma(k+1)
d=gamma(k+2)
v=sum(((lambda*alfa*x)^k+1/2)/g*d)
s=sqrt((alfa*lambda)/x)*exp(lambda-alfa*x)*v
s
}
curve(f1(x,0.87,0.12,1)
,from=0,to=25, col = "gold",add=TRUE)


#######################################
########geometric#############
```

```
###############################
##############geometric-pareto#######
x <- seq(0, 25,1)
f3=function(x,a,b,p){
s=(a*p*1-p)/(b*((1+(p*x)/b)^(a+1)))
s
}
f3(x,2.04,2.13,0.07)

curve(f(x,2.04655,2.05481,0.93)
,from=0,to=25, col = "red",main="Primary distribution: Geometric",ylab="f(x)")
curve(f(x,0.5,1,2)
,from=0,to=25, col = "violet",main="Primary distribution: Poisso",ylab="f(x)")

colors <- c("red", "blue")
labels <- c("red=pareto","blue=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)
######geometric-exponentioal############
x <- seq(0, 25,1)
f4=function(x,alfa,p){
s=alfa*p*(1-p)*exp(-alfa*p*x)
s
}
f4(x,1.005,0.93)
curve(f4(x,1.005,0.93)
,from=0,to=25, col = "blue",add=TRUE)


##########################################
##########negative -binomial########
###################################################
#######negative -binomial-pareto######
x <- seq(0, 25,1)
```

```
f5=function(x,r,alfa,beta,p){
v=hypergeo(1+r,1+alfa,2,(((1-p)*x)/beta)/(1+x/beta))
e=((r*alfa*(1-p)*(p^r))/(beta*((1+x/beta)^alfa+1)))*v
e
}

curve(f5(x,0.31749,2.05542,1.91539,0.80067)
,from=0,to=25, col = "black",main="Primary distribution: Negative Binomial",ylab="f(x)")

colors <- c("black", "green")
labels <- c("black=pareto","green=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)
#######negative -binomial-expontial######
library(CharFun)
x <- seq(0, 25,1)
f6=function(x,r,alfa,p){
v=hypergeom1F1(alfa*(1-p)*x,2, 1+r)
e=alfa*r*(p^r)*(1-p)*exp(-alfa*x)*v
e
}
curve(f6(x,0.51168,0.55250,0.87090),from=0,to=25, col = "green",add=TRUE)


#########################################
########negative-binomial-pareto-exp######
#############################################
x <- seq(0, 25,1)
f7=function(x,beta,alfa,lambda){
v=hypergeom1F1(beta+1,alfa+beta+2,lambda*x)
e=(alfa*beta*exp(-lambda*x)*v)/((alfa+beta+1)*(alfa+beta))
e
}
```

```
curve(f7(x,0.20723,2.84501,0.82117)
,from=0,to=25, col = "black",main="Primary distribution: Negative Binomial-
Pareto",ylab="f(x)")
curve(f7(x,0.2,2.8,0.82)
,from=0,to=25, col = "black",main="Primary distribution: Negative Binomial-
Pareto",ylab="f(x)")


colors <- c("black")
labels <- c("black=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)


###########fig3##################
###############################################
#####alfa=0.2###############
rm(list=ls())
x <- seq(0, 5,0.2)
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}

p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.2", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.2,1,0.5))
lines(x, f(x,0.2,1,0.5), lwd=2, col="red")
lines(x, f(x,0.2,1,1.5), lwd=2, col="red")

# add lines representing t-distribution
```

```
for (i in 1:4){
lines(x, f(x,0.2,1,p[i]), lwd=2, col=colors[i])
}
# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=color


################################
#####alfa=0.4###############
x <- seq(0, 5,0.2)
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}

p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.4", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.4,1,0.5))
lines(x, f(x,0.4,1,0.5), lwd=2, col="red")
lines(x, f(x,0.4,1,1.5), lwd=2, col="red")

lines(x, f(x,0.4,1,0.5), lwd=2, col="red")

# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.4,1,p[i]), lwd=2, col=colors[i])
}
####################################
#####alfa=0.6###############
x <- seq(0, 5,0.2)
```

```
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}

p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.4", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.6,1,0.5))
lines(x, f(x,0.6,1,0.5), lwd=2, col="red")
lines(x, f(x,0.6,1,1.5), lwd=2, col="red")

lines(x, f(x,0.6,1,0.5), lwd=2, col="red")

# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.6,1,p[i]), lwd=2, col=colors[i])
}

# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)

####################
#####alfa=0.8##############

x <- seq(0, 5,0.2)
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}
```

```
p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.4", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.8,1,0.5))
lines(x, f(x,0.8,1,0.5), lwd=2, col="red")
lines(x, f(x,0.8,1,1.5), lwd=2, col="red")

lines(x, f(x,0.8,1,0.5), lwd=2, col="red")

# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.8,1,p[i]), lwd=2, col=colors[i])
}

# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)


x <- seq(0, 5,0.2)
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}

p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.6", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.6,1,0.5))
lines(x, f(x,0.6,1,0.5), lwd=2, col="red")
```

```
lines(x, f(x,0.6,1,1.5), lwd=2, col="red")

lines(x, f(x,0.6,1,0.5), lwd=2, col="red")

# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.6,1,p[i]), lwd=2, col=colors[i])
}

# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)

####################
x <- seq(0, 5,0.2)
f=function(x,alfa,beta,p){
v=(alfa*p*(1-p))/(beta*(1+(p*x)/beta)^(alfa+1))
v
}

p<- c(0.5, 0.3, 0.15, 0.09)
# set up colours and labels for the functions and plot legend
colors <- c("red", "blue", "gold", "black")
labels <- c("p=0.5,alpha=0.8", "p=0.3", "p=0.15", "p=0.0.09")
# plot Z first
plot(x, f(x,0.8,1,0.5))
lines(x, f(x,0.8,1,0.5), lwd=2, col="red")
lines(x, f(x,0.8,1,1.5), lwd=2, col="red")

lines(x, f(x,0.8,1,0.5), lwd=2, col="red")

# add lines representing t-distribution
for (i in 1:4){
lines(x, f(x,0.8,1,p[i]), lwd=2, col=colors[i])
```

```
}

# add legend
legend("topright",
labels, lwd=2, lty=c(1, 1, 1, 1, 2), col=colors)


##########################################
##########negative -binomial########
###################################################
#######negative -binomial-pareto######
x <- seq(0, 25,1)
f5=function(x,r,alfa,beta,p){
v=hypergeo(1+r,1+alfa,2,(((1-p)*x)/beta)/(1+x/beta))
e=((r*alfa*(1-p)*(p^r))/(beta*((1+x/beta)^alfa+1)))*v
e
}

curve(f5(x,0.31,2.05,1.91,0.8)
,from=0,to=25, col = "black",main="Primary distribution: Negative Binomial",ylab="f(x)")

colors <- c("black", "green")
labels <- c("black=pareto","green=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)
#######negative -binomial-expontial######
library(CharFun)
x <- seq(0, 25,1)
f6=function(x,r,alfa,p){
v=hypergeom1F1(alfa*(1-p)*x,2, 1+r)
e=alfa*r*(p^r)*(1-p)*exp(-alfa*x)*v
e
}
curve(f6(x,0.5,0.4,0.8),from=0,to=25, col = "green",add=TRUE)
```

```
#######geometric############
###############################
#############geometric-pareto#######
x <- seq(0, 25,1)
f3=function(x,a,b,p){
s=(a*p*(1-p))/(b*((1+(p*x)/b)^(a+1)))
s
}
f3(x,2.04,2.13,0.07)

##curve(f3(x,2.04655,2.05481,0.93)
###,from=0,to=25, col = "red",main="Primary distribution: Geometric",ylab="f(x)")
curve(f3(x,0.2,1,0.5)
,from=0,to=25, col = "red",main="Primary distribution: Geometric",ylab="f(x)")

colors <- c("red", "blue")
labels <- c("red=pareto","blue=exponential")
legend("topright",
labels, lwd=2, lty=c(1, 1), col=colors)
######geometric-exponentioal###########
x <- seq(0, 25,1)
f4=function(x,alfa,p){
s=alfa*p*(1-p)*exp(-alfa*p*x)
s
}
curve(f4(x,0.53,0.5)
,from=0,to=25, col = "blue",add=TRUE)
```

مراجع

[۱] فروند ج، ۲۰۰۵، عنوان کتاب آمار ریاضی و کاربردهای آن، ترجمه محمدقاسم وحیدی اصل و علی عمیدی، انتشارات مرکز نشر دانشگاهی

[۲] هال ج، ۱۹۴۵، عنوان کتاب مبانی مهندسی مالی و مدیریت ریسک ، ترجمه سجاد سیاح و علی صالح آبادی، انتشارات گروه رایانه تدبیر پرداز

[3] Arnold, B.C., Castillo, E., Sarabia, J.M., 1993. Multivariate distributions with generalized Pareto conditionals. Statist. Probab. Lett. 17, 361–368

[4] Arnold, B.C., 1987. Bivariate distributions with Pareto conditionals. Statist. Probab. Lett. 5, 263–266.

[5] Arnold, B.C., 2015. Pareto Distributions, second ed. Chapman, Hall/CRC Monographs on Statistics , Applied Probability, Boca Ratón, FL.

[6] Arnold, B.C., Castillo, E., Sarabia, J.M., 1993. Multivariate distributions with generalized Pareto conditionals. Statist. Probab. Lett. 17, 361–368.

[7] Arnold, B.C., Castillo, E., Sarabia, J.M., 2001. Conditionally specified distributions: An introduction (with discussion). Statist. Sci. 16, 151–169.

[8] Asimit, A.V., Furman, E., Vernic, R., 2010. On a multivariate Pareto distribution. Insurance Math. Econom. 46, 308–316.

[9] Bozdogan, H., 1987. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 52, 345–370.

[10] Chiragiev, A., Landsman, Z., 2009. Multivariate flexible Pareto model: Dependency structure, properties and characterizations. Statist. Probab. Lett. 79, 1733–1743.

[11] Cossette, H., Coté, M.-P., Marceau, E., Moutanabbir, K., 2013. Multivariate distribution defined with Farlie–Gumbel–Morgenstern copula and mixed Erlang marginals: Aggregation and capital allocation. Insurance Math. Econom. 52, 560–572.

[12] Cossette, H., Landriault, D., Marceau, E., 2004. Compound binomial risk model in a Markovian environment. Insurance Math. Econom. 35, 425–443.

[13] Cossette, H., Marceau, E., Marri, F., 2008. On the compound Poisson risk model with dependence based on a generalized Farlie Gumbel Morgenstern copula. Insurance Math. Econom. 43, 444–455.

[14] Esary, J.D., Proschan, F., Walkup, D.W., 1967. Association of random variables, with applications. Ann. Math. Statist. 38, 1466–1474.

[15] Gerber, H.U., 1988. Mathematical fun with the compound binomial process. ASTIN Bull. 18, 161–168. Elsevier, pp154-163

[16] Genest, G., Marceau, E., Mesfioui, M., 2003. Compound Poisson approximations for individual models with dependent risks. Insurance Math. Econom. 32, 73

[17] José María Sarabia ,Emilio Gómez-Déniz, Faustino Prieto, Vanesa Jordá (2016), "Risk aggregation in multivariate dependent Pareto distributions"

[18] Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M., 2001. Modern Actuarial Risk Theory. Kluwer Academic Publishers, Boston.

[19] Klugman, S.A., Panjer, H.H., Willmot, G.E., Loss Models. From Data to Decisions, third ed. John Wiley, New York.

[20] Ramsay, C.M., 2006. The distribution of sums of certain . Pareto variates.

[21] Sarabia, J.M., Guillén, M., 2008. Joint modelling of the total amount and the number of claims by conditionals. Insurance Math. Econom. 43, 466–473. Statist. Theory Methods 35, 395–405.

[22] Shengwang, M., Wei, Y., Whitmore, G.A., 1999. Accounting for individual overdispersion in a bonus-malus automobile insurance system. ASTIN Bull. 29 (2), 327–337.

Aabstract

In this dissertation we obtain closed expressions for the probability distribution function of aggregated risks with multivariate dependent Pareto distributions. We work with the dependent multivariate Pareto type II proposed by Arnold (1983, 2015), which is widely used in insurance and risk analysis. We begin with an individual risk model, where the probability density function corresponds to a second kind beta distribution, obtaining the Value at Risk and several other risk measures. Then, we consider a collective risk model based on dependence, where several general properties are studied. We study in detail some relevant collective models with Poisson, negative binomial and geometric distributions as primary distributions. In the collective Pareto–Poisson model, the probability density function is a function of the Kummer confluent hypergeometric function, and the density of the Pareto–negative binomial is a function of the Gauss hypergeometric function.With reviews on data based on one-year insurance policies annually $(2004 - 2005)$ And the data of Iran's insurance company was carried out in 93, we conclude that our collective dependent models outperform other collective models considered in the actuarial literature in terms of AIC statistics.
Keywords: Dependent risks Individual risk model Collective risk model Classical Pareto distribution Hypergeometric functions

Shahrood University of Technology

Faculty of Mathematical Sciences

MSc Thesis in: Financal Mathematical

Risk aggregation in multivariate dependent Pareto distributions

By: Seyede Mona MirHojati

Supervisors

Dr. Ahmad Nezakati
Dr. Seyed Mojtaba Mirlohi

June 2018