LA DISTRIBUCIÓN DE PARETO PONDERADA POR LA FAMILIA INVERSA GAUSSIANA GENERALIZADA. APLICACIÓN A LA MODELIZACIÓN DE LOS GRANDES SINIESTROS

Cristina Lozano-Colomer
Departamento de Métodos Cuantitativos
Universidad Pontificia de Comillas (ICADE)
clozano@cee.upcomillas.es

José L. Vilar-Zanón
Departamento de Economía Financiera y Actuarial.
Universidad Complutense de Madrid

Abstract: We study new probability distributions consisting in Pareto mixed by generalized inverse Gaussian. These distributions have a right super-heavy tail, and it is well known that modelling the tail of the claim severity is very important when dealing with large claims. The lack of moments, is avoided considering a truncation. It is applied to a Spanish portfolio of civil responsibility of cars from the year 1992 to 2001 and the results are compared with the ones obtained ones from a Bayesian point estimation already studied in Vilar, Lozano 2007.

Keywords: Very heavy tail. Reinsurance, XL premiums, Bayes, Pareto distribution, Generalized Inverse Gaussian family.

Resumen: En este articulo se estudian distribuciones de probabilidad de Pareto ponderadas por distribuciones de la familia inversa Gaussiana generalizada, distribuciones que se comprueba que son de cola derecha súper-gruesa, muy útiles para la obtención de modelos de cuantías de grandes siniestros. Distribuciones que carecen de momentos, problema que se evita ponderando por una distribución truncada adecuadamente. Se aplica a una cartera española de responsabilidad civil de automóviles durante los años 1992 al 2001 y se comparan los resultados con los obtenidos a partir de una estimación puntual bayesiana.

1-Introducción

Este artículo analiza la distribución de Pareto ponderada por la familia de distribuciones inversa Gaussiana generalizada (IGG).

Esta distribución surge como distribución predictiva si se utiliza la técnica Bayesiana para estimar el parámetro de forma de la distribución de Pareto, en el caso que se tome como distribución a priori, una distribución perteneciente a la familia (IGG) y se ha demostrado (*Vilar- Lozano* (2007)), que constituye una familia conjugada para la distribución de Pareto. Las distribuciones de Pareto ponderada por IGG, contienen la distribución log-Pareto como caso particular.

En este articulo se estudian otras dos distribuciones nuevas de esta familia, Pareto ponderada por Gamma Inversa y Pareto ponderada por Inversa Gaussiana, pertenecen al grupo de las denominadas por *(Fraga Alves, de Haan, Neves (2006*) de cola súper –gruesa.

Por distribuciones de cola gruesa se entiende que son aquellas cuya función de distribución tiene una cola derecha que decrece hacia cero de forma polinómica, es decir $1-F(x) \approx 1-x^{-\alpha}$ con $\alpha > 0$ y el término súper-gruesa se reserva para colas derechas de distribuciones que decrecen hacia cero de forma mucho más lenta que las anteriores generalmente de forma logarítmica, es decir $1-F(x) \approx 1-(\log x)^{-\alpha}$ con $\alpha > 0$.

Estas distribuciones con cola súper-gruesa entre las cuales está la log-Pareto, tienen el inconveniente de carecer de momentos, planteando un problema a la hora de estimar una prima.

Reiss y Thomas (2001) se han enfrentado con este problema, condicionando la distribución a priori del parámetro , de forma que la distribución predictiva posea momentos mientras que otros (*Hesselager (1993)* y (*Vilar, Lozano* (2007)) han tomado la media de la distribución a posteriori del parámetro como estimación puntual para el parámetro de la distribución de Pareto.

En este artículo se sigue la línea de (*Reiss, Thomas* (2001)), finalmente se aplica a una cartera de responsabilidad de automóviles de ya estudiada por (*Vilar, Lozano* (2007)) y comparando los resultados de ambos trabajos.

2- Supuestos

Se consideran dos variables aleatorias N y X para el número de siniestros y su cuantía respectivamente. Fijado un umbral $a > 0$, se supone que $P^a = \Pr(x > a) > 0$ (2.1)

Entonces se definen dos nuevas variables

- N^a que representa el número de cuantías por encima de "a" que se denominará exceso y
- Z que representa este exceso sobre el umbral normalizado $Z = \frac{X-a}{a} / X > a$.

Se considera que la variable aleatoria número de excesos N^a sigue una distribución de Poisson ponderada por inversa Gaussiana generalizada la llamada distribución de Sichel *(Klugman,, Panjer, Willmot (1998))* y por tanto en la aplicación práctica se tomaran los resultados respecto del número de excesos obtenidos en (*Vilar, Lozano* (2007)).

$$P(N=n) = \int_0^\infty \frac{\lambda^n}{n!} e^{-\lambda} u(\lambda)\, d\lambda$$

$$\hat{\lambda} = E(N) = \int_0^\infty \lambda \ u(\lambda)\, d\lambda \qquad (2.2)$$

Siendo $u(\lambda)$ una distribución de la familia inversa Gaussiana generalizada.
Se considera que el umbral es lo suficientemente alto como para que la variable aleatoria Z del exceso siga una distribución de Pareto.

El conjunto de excesos en los diferentes años $\{Z_i\}_{i\in N}$ son variables aleatorias independientes e idénticamente distribuidas con distribución condicionada y momentos

$$f(z|\alpha) = \frac{\alpha}{(1+z)^{\alpha+1}}, z > 0 \qquad (2.3)$$

$$E\{Z^r|\alpha\} = \frac{1}{\alpha - r}, \forall \alpha > r$$
$$E\{Z|\alpha\} = \frac{1}{\alpha - 1}, \forall \alpha > 1$$

Asimismo, se supone que el parámetro de forma α de la distribución es una variable aleatoria absolutamente continua A, independiente del tiempo y del parámetro de la distribución de Poisson, con función de distribución $\Pi(\alpha) = \Pr(\mathrm{A} \le \alpha)$ y función de densidad $\pi(\alpha)$.

Por tanto la distribución de probabilidad marginal es decir no condicionada, del exceso normalizado, es una distribución de Pareto ponderada, con función de densidad

$$h(z) = \int_0^{+\infty} \frac{\alpha}{(1+z)^{\alpha+1}} \pi(\alpha)\, d\alpha \qquad (2.4)$$

La distribución del exceso total sobre un umbral normalizado como una variable Poisson compuesta con Pareto

$$S^a = \sum_{i=0}^{N^a} Z_i \qquad (S^a = 0 \text{ para } N^a = 0)$$

Aplicando el principio del valor esperado y las propiedades de las distribuciones compuestas se tiene una expresión para la prima :

$$P = E[S^a] = E[N^a]\, E[Z] \qquad (2.5)$$

con

$$E[Z] = \int E[Z/\alpha]\ \pi(\alpha)\, d\alpha = \int \frac{1}{\alpha - 1}\ \pi(\alpha)\ d\alpha, \qquad \alpha > 1$$

Para que exista esta esperanza el soporte de la distribución del parámetro α, tiene que contener valores superiores a 1.

Dadas por (2.1) y (2.2) se tiene que

$$E[N^a] = \hat{\lambda}\ P^a$$

En lo que sigue, se va a considerar que la distribución de probabilidad del parámetro α , pertenece a la familia inversa Gaussiana generalizada (IGG), que es un modelo que ha sido aplicado en la ciencia actuarial para describir la heterogeneidad de una cartera (*Panjer, Willmot (1994)*.

Este modelo tiene una función de densidad que depende de tres parámetros:

$$\pi_{\beta,\psi,\chi}\left(\alpha\right)=\frac{\left(\frac{\psi}{\chi}\right)^{\beta/2}}{2K_{\beta}\left(\sqrt{\psi\chi}\right)}\alpha^{\beta-1}e^{-\frac{1}{2}\left(\frac{\chi}{\alpha}+\psi x\right)} \qquad \alpha>0,\ \beta\in R,\ \ \psi,\chi\geq 0 \quad (2.6)$$

Siendo

$$K_{\beta}(\omega)=\frac{1}{2}\int_{0}^{+\infty}y^{\beta-1}e^{-\frac{\omega}{2}\left(\frac{1}{y}+y\right)}dy,\quad \beta\in R,\ \ \omega>0 \qquad (2.7)$$

Una función de Bessel de tercera especie (véase por ejemplo *Panjer, Willmot (1994)),* que posee las siguientes propiedades:

$$K_{\beta}(y)=K_{-\beta}(y)$$

$$K_{n+\frac{1}{2}}(y)=\sqrt{\frac{\pi}{2y}}\ e^{-y}\ \sum_{i=0}^{n}\frac{(n+1)!}{(n-1)!i!}\quad (2y)^{-i},\ \ y>0, \qquad n=0,1,2...\ \ (2.8)$$

Se ha demostrado (*Vilar, Lozano* (2007)) que esta familia (IGG) es una familia conjugada de la distribución de Pareto, a partir de lo que se puede establecer un modelo de credibilidad exacta.

En este caso la familia de distribuciones predictivas tendrán unas funciones de densidad Pareto ponderadas por inversa gaussiana generalizada dadas para $\chi,\ \psi>0\ \ \beta\in R$ y $z>0$ por:

$$h_{\beta,\chi,\psi}(z)=\sqrt{\frac{\chi}{\psi}}\,\frac{K_{\beta+1}\left(\sqrt{\chi\psi}\left[1+\frac{2}{\psi}\log(1+z)\right]^{1/2}\right)}{K_{\beta}\left(\sqrt{\chi\psi}\right)(1+z)}\left(1+\frac{2}{\psi}\log(1+z)\right)^{-\left(\frac{\beta+1}{2}\right)} \quad (2.9)$$

En lo que sigue se estudian algunas de estas distribuciones, para diferentes valores de los parámetros, analizando y comprando el comportamiento de sus colas derechas.

3. Distribuciones de Pareto ponderadas por Inversa Gaussiana – Generalizada

La distribución (2.9) constituye una amplia familia de distribuciones de Pareto ponderadas, por la familia IGG, una de las cuales es la distribución denominada log-Pareto *(Reiss, Thomas (2001)*, que resulta de ponderar la distribución de Pareto por la distribución Gamma.

Haciendo uso de la propiedad (2.8) es posible obtener la expresión analítica de todas las Pareto ponderadas por distribuciones de la familia IGG con parámetro $\beta = n + \frac{1}{2}, \quad n = 0,1...$

Aquí estudiaremos los casos Pareto- Gamma, Pareto- inversa Gaussiana y Pareto- Gamma inversa. Compararemos las colas de estas distribuciones entre si y con la distribución de Pareto

3.1 Pareto-Gamma: Tomando $\beta > 0, \chi = 0, \psi > 0$, se obtiene como caso límite en la familia IGG, una distribución gamma ($\beta, \psi / 2$)

$$\pi_{\beta,0,\psi}(\alpha) = \frac{(\psi/2)^{\beta}}{\Gamma(\beta)} \alpha^{\beta-1} e^{-\frac{\psi}{2}\alpha} \quad \alpha > 0 \qquad (3.1)$$

En este caso se tiene la función de densidad predictiva conocida como log-Pareto *(Reiss, Thomas (2001))* con función de densidad:

$$h_{\beta,0,\psi}(z) = \left(\frac{\psi}{2}\right)^{\beta} \beta (1+z)^{-1} \left(\frac{\psi}{2} + \log(1+z)\right)^{-(\beta+1)} \quad z > 0 \qquad (3.2)$$

Se comprueba fácilmente que la función de distribución es:

$$H_{\beta,0,\psi}(z) = 1 - \left(1 + \frac{2}{\psi}\log(1+z)\right)^{-\beta} \quad z > 0 \qquad (3.3) \qquad\qquad (3.3)$$

Esta variable aleatoria posee una distribución que se denomina de cola súper-gruesa ya que se caracteriza por la propiedad de que su logaritmo posee una distribución de cola gruesa.

3.2 Pareto-Gamma inversa: Si se toma $\beta < 0, \chi > 0, \psi = 0$ en la familia IGG, se obtiene la función de densidad gamma inversa

$$\pi_{\beta,\chi,0}(\alpha) = \left(\frac{2}{\chi}\right)^{\beta} \Gamma(-\beta)^{-1} \alpha^{\beta-1} e^{-\frac{\chi}{2\alpha}}, \quad \alpha > 0 \qquad (3.4)$$

En este caso la distribución predictiva tiene la forma:

$$h_{\beta,\chi,0}(z) = \left(\frac{2}{\log(1+z)}\right)^{\frac{\beta+1}{2}} \frac{\chi^{\frac{1-\beta}{2}}}{\Gamma(-\beta)(1+z)} K_{\beta+1}(\sqrt{2\chi\log(1+z)}), \quad z > 0 \quad (3.5)$$

Para poder expresar analíticamente la función de Bessel utilizando la propiedad dada por (2.8) con $\beta < 0$, tomamos $\beta = -\frac{1}{2}$, $\psi > 0$, obteniéndose la siguiente densidad predictiva

$$h_{-1/2,0,\psi}(z) = \sqrt{\frac{\chi}{2}}(1+z)^{-1} \log(1+z)^{-\frac{1}{2}} \; e^{-(2\chi\log(1+z))^{1/2}}, \quad z > 0 \qquad (3.6)$$

En este caso la función de distribución es:

$$H_{-1/2,\chi,0}(z) = 1 - e^{-[2\chi\log(1+z)]^{1/2}}, \quad z > 0 \qquad (3.7)$$

En este caso la distribución ponderada es tipo log-Weibull, también con una cola súper- gruesa. *(Fraga Alves, de Haan, Neves (2006)*).

3.3 Pareto-Inversa Gaussiana Tomado $\beta=-\frac{1}{2},\ \chi>0,\ \psi>0$ en la familia IGG, se obtiene la distribución inversa Gaussiana con función de densidad:

$$\pi(\alpha)=\sqrt{\frac{\chi}{2\pi}}\alpha^{-3/2}e^{-\frac{\psi}{2}\frac{\left(\alpha-\sqrt{\frac{\chi}{\psi}}\right)^2}{\alpha}}\quad \alpha>0 \qquad (3.8)$$

La función de densidad predictiva se obtiene sustituyendo en (2.9) el valor de los parámetros

$$h_{-1/2,\chi,\psi}(z)=\sqrt{\frac{\chi}{\psi}}(1+\frac{2}{\psi}\log(1+z))^{-\frac{1}{2}}\ \frac{e^{\sqrt{\chi\psi}\,[1-(1+\frac{2}{\psi}\log(1+z))^{1/2}]}}{(1+z)},\ \ \chi,\psi>0\ \ z>0\ \ (3.9)$$

Con función de distribución:

$$H_{-1/2,\chi,\psi}=1-e^{\sqrt{\chi\psi}\left[1-(1+\frac{2}{\psi}(1+\log(1+z))^{1/2}\right]},\ \ \chi,\psi>0\ \ z>0 \quad (3.10)$$

Es una transformada de una log-Weibull.

3.4 Comparación de las colas de las distribuciones de Pareto con log-Pareto, log-Weibull y Pareto ponderada por inversa Gaussiana.

Vamos a comparar el comportamiento de las colas de las anteriores distribuciones entre si y con la distribución de Pareto y corroborar que poseen colas mas gruesas que esta.

Como se indica en *(Klugman., Panjer, Willmot (1998)),* dadas dos variables aleatorias X e Y con funciones de distribución $F_X(x)$ y $F_Y(x)$ respectivamente se dice que tienen colas equivalentes si $\lim_{x\to\infty}\frac{1-F_X(x)}{1-F_Y(x)}=1$, si este límite es cero entonces la variable Y tiene una cola mas gruesa que la variable X, y si el límite es infinito la variable Y tiene una cola más fina que la variable X.

3.4.1 Comparación de las colas de las distribuciones de Pareto y Log-Pareto.

Tomando la expresión de la función de distribución correspondiente a la función de densidad dada por (2.3) y la función de distribución Log-Pareto dada por (3.2) con $\alpha, \chi, \psi > 0$ se tiene:

$$\lim_{x\to\infty} \frac{(1+z)^{-\alpha}}{[1+\frac{2}{\psi}\log(1+z)]^{-\beta}} = \lim_{x\to\infty} e^{-\alpha\log(1+z)+\beta\log\left[1+\frac{2}{\psi}\log(1+z)\right]} = 0$$

Esto supone que la distribución Log-Pareto posee una cola más gruesa que la distribución de Pareto.

3.4.2 Comparación de las colas de las distribuciones de Pareto y de Pareto ponderada por inversa Gaussiana

Ahora comparamos las funciones de distribución dadas por (2.3) y (3.10) con $\alpha, \chi, \psi > 0$:

$$\lim_{x\to\infty} \frac{(1+z)^{-\alpha}}{e^{-[\chi\psi+\chi\log(1+z)]^{1/2}}} = \lim_{x\to\infty} e^{-\alpha\log(1+z)+[\chi\psi+\chi\log(1+z)]^{1/2}} = 0$$

Por tanto la distribución Pareto ponderada por inversa gaussiana tiene la cola más gruesa que la distribución de Pareto.

3.4.3 Comparación de las colas de las distribuciones Log-Pareto y Pareto ponderada por inversa Gaussiana

Ahora comparamos las funciones dadas por (3.2) y por (3.10) con $\alpha, \chi, \psi > 0$:

$$\lim_{x\to\infty}\frac{\left[1+\frac{2}{\psi}\log\left(1+z\right)\right]^{-\beta}}{e^{-[\chi\psi+\chi\log(1+z)]^{1/2}}}=\lim_{x\to\infty}e^{-\beta\log\left[1+\frac{2}{\psi}\log(1+z)\right]+[\chi\psi+\chi\log(1+z)]^{1/2}}=\infty$$

Esto significa que la distribución de Pareto ponderada por inversa Gaussiana tiene la cola mas fina que la distribución Log-Pareto.

3.4.4 Comparación de las colas de las distribuciones Log-Pareto y Pareto ponderada por gamma inversa con $\beta=-\frac{1}{2}$

Tomando las funciones de distribución dadas por (3.2) y (3.7) se tiene:

$$\lim_{x\to\infty}\frac{e^{-[2\chi\log(1+z)]^{1/2}}}{e^{-[\chi\psi+\chi\log(1+z)]^{1/2}}}=1$$

Por tanto tienen la cola equivalente.

Se pueden ordenar las distribuciones según el grueso de la cola en orden creciente de la forma Pareto < Pareto con Gamma inversa ~Pareto con inversa Gaussiana < log-Pareto.

Se han obtenido dos ejemplos de distribuciones con colas muy gruesas pero algo menos gruesas que la Log-Pareto, constituyendo una nueva herramienta para el estudio de los sucesos con grandes pérdidas.

Se demuestra que estas distribuciones denominadas de colas súper gruesa, no poseen momentos de ningún orden y si se quieren utilizar para tarificar en un contrato de reaseguro no será posible utilizar el principio del valor esperado.

Dado que se han obtenido estas distribuciones como resultado de ponderar la distribución de Pareto por la distribución de parámetro de forma, y como se ve en una manera posible de obtener la esperanza de la distribución de Pareto ponderada está dada por la expresión (2.5), que tiene la restricción de que el soporte sea $\alpha>1$, consideremos que las distribuciones del parámetro son distribuciones Inversa generalizada truncada. Comprobando en el apartado siguiente que también constituye una familia conjugada para la distribución de Pareto, y obteniéndose su distribución predictiva.

4-La Distribución inversa gaussiana generalizada truncada como familia Pareto conjugada.

La familia inversa gaussiana generalizada (IGG) truncada en un valor $b \geq 0$ tiene función de densidad

$$\pi^{b}_{\beta,\chi,\psi}(\alpha) = \frac{\pi_{\beta,\chi,\psi}(\alpha)}{1-\Pi_{\beta,\omega,\eta}(b)}, \quad \alpha > b > 0,\ \beta \in \Re, \quad \psi,\chi \geq 0$$

$$\Pi_{\beta,\omega,\eta}(b) = \int_0^b \pi_{\beta,\chi,\psi}(\alpha)\ d\alpha \tag{4.1}$$

Siendo $\pi_{\beta,\chi,\psi}(\alpha)$ la función de densidad de (IGG) dada en (3.1) y $\Pi_{\beta,\omega,\eta}(b)$ su función de distribución en b.

Es fácil comprobar que también constituye una familia conjugada para la distribución de Pareto ya que la única diferencia con la familia IGG es una constante de proporcionalidad.

La media de la distribución truncada es:

$$E_b\{\alpha\} = \frac{1}{1-\Pi_{\beta,\chi,\psi}(b)} \int_0^b \alpha \frac{\left(\frac{\psi}{\chi}\right)^{\beta/2}}{2K_\beta\left(\sqrt{\psi\chi}\right)}\ \alpha^{\beta-1} e^{-\frac{1}{2}\left(\frac{\chi}{\alpha}+\psi x\right)} d\alpha$$

$$= \frac{\left(\frac{\psi}{\chi}\right)^{1/2}}{1-\Pi_{\beta,\omega,\eta}(b)} \frac{K_{\beta+1}\left(\sqrt{\psi\chi}\right)}{K_\beta\left(\sqrt{\psi\chi}\right)} \int_b^\infty \frac{\left(\frac{\psi}{\chi}\right)^{-\frac{\beta+1}{2}}}{2K_{\beta+1}\left(\sqrt{\psi\chi}\right)} \alpha^\beta e^{-\frac{1}{2}\left(\frac{\chi}{\alpha}+\psi x\right)} d\alpha$$

$$= \frac{1-\Pi_{\beta+1,\chi,\psi}(b)}{1-\Pi_{\beta,\chi,\psi}(b)}\ E\{\alpha\}$$

$$E\{\alpha\} = \frac{K_{\beta+1}\left(\sqrt{\psi\chi}\right)}{K_\beta\left(\sqrt{\psi\chi}\right)} \left(\frac{\psi}{\chi}\right)^{1/2} \tag{4.2}$$

y con

$$1-\Pi_{\beta+1,\chi,\psi}(b)=\int_b^{\infty}\frac{\left(\frac{\psi}{\chi}\right)^{-\frac{\beta+1}{2}}}{2K_{\beta+1}\left(\sqrt{\psi\chi}\right)}\alpha^{\beta}e^{-\frac{1}{2}\left(\frac{\chi}{\alpha}+\psi x\right)}d\alpha$$

De manera similar se obtiene la expresión de la varianza de la distribución truncada

$$\mathrm{Var}_b\{\alpha\}=\frac{1-\Pi_{\beta+2,\chi,\psi}(b)}{1-\Pi_{\beta,\chi,\psi}(b)}E\{\alpha^2\}-\left(\frac{1-\Pi_{\beta+1,\chi,\psi}(b)}{1-\Pi_{\beta,\chi,\psi}(b)}\right)^2E\{\alpha\}^2 \qquad (4.3)$$

Estas expresiones son útiles para especificar una distribución a priori con momentos predeterminados *(Reiss, Thomas (2001))*.

Para obtener la distribución predictiva Pareto ponderada por IGG truncada, suponemos $\chi,\psi>0$ y tomamos los nuevos parámetros

$$\omega=\sqrt{\chi\psi},\quad \eta=\sqrt{\frac{\chi}{\psi}}:$$

$$h^b_{\beta,\chi,\psi}(z)=\int_b^{+\infty}f(z/\alpha)\ \pi^b_{\beta,\chi,\psi}(\alpha)\,d\alpha$$

$$=\frac{\eta^{-\beta}}{2K_\beta(\omega)(1+z)(1-\Pi_{\beta,\chi,\psi}(b))}\int_b^{+\infty}\alpha^{\beta}e^{-\frac{\omega}{2}\left(\frac{\eta}{\alpha}+\alpha\left[\frac{1}{\eta}+\frac{2}{\omega}\log(1+z)\right]\right)}d\alpha=\ldots$$

Tomando entonces

$$A=\omega\left(1+\frac{\eta}{\omega}2\log(z+1)\right)^{1/2}$$

$$B=\eta\left(1+\frac{\eta}{\omega}2\log(z+1)\right)^{-1/2}$$

Y multiplicando y dividiendo por $\dfrac{2K_{\gamma+1}\left(A\right)}{B^{-(\gamma+1)}}$, se obtiene:

$$
\begin{aligned}
h^{b}_{\beta,\omega,\eta}(z) &= \frac{\eta^{-\beta}\, 2K_{\beta+1}\left(A\right)}{2(1+z)K_{\beta}\left(\omega\right)\left(1-\Pi_{\gamma,\omega,\eta}\left(b\right)\right)B^{-(\beta+1)}} \int_{b}^{\infty} \frac{B^{-(\beta+1)}}{2K_{\beta+1}\left(A\right)}\alpha^{\beta} e^{-\frac{A}{2}\left(\frac{B}{\alpha}+\frac{\alpha}{B}\right)} d\alpha \\
&= \frac{\eta}{(1+z)} \frac{K_{\beta+1}\left(A\right)\left(1-\Pi_{\beta+1,A,B}\left(b\right)\right)}{K_{\beta}\left(\omega\right)\left(1-\Pi_{\beta,\omega,\eta}\left(b\right)\right)} \left(1+\frac{\eta}{\omega} 2\log(z+1)\right) \\
&= \frac{\eta K_{\beta+1}\left(\omega\left(1+2\dfrac{\eta}{\omega}\log(1+z)\right)^{1/2}\right)}{\left(1+z\right)K_{\beta}\left(\omega\right)\left(1+2\dfrac{\eta}{\omega}\log(1+z)\right)^{(\beta+1)/2}} \frac{\left(1-\Pi_{\beta+1,A,B}\left(b\right)\right)}{\left(1-\Pi_{\beta,\omega,\eta}\left(b\right)\right)} \\
&= \frac{\left(1-\Pi_{\beta+1,A,B}\left(b\right)\right)}{\left(1-\Pi_{\beta,\omega,\eta}\left(b\right)\right)}\; h_{\beta,\omega,\eta}(z), \;\; z>0 \qquad (4.4)
\end{aligned}
$$

Como se puede ver la función de densidad de la distribución predictiva Pareto ponderada por IGG truncada es proporcional a la función de densidad de la distribución predictiva Pareto ponderada por IGG sin truncar.

Se obtienen las diferentes expresiones para las distintas distribuciones de la familia IGG sustituyendo los valores de los parámetros.

Esta distribución tiene los mismos inconvenientes que la distribución predictiva sin truncar la distribución del parámetro, pero tiene la ventaja de poseer momentos de orden $r > b$. Momentos que son más fáciles de calcular si se utilizan las propiedades de las distribuciones ponderadas.

En este caso, se tiene que para la esperanza matemática,

(4

$$
\begin{aligned}
E[Z] &= \int_{b}^{+\infty} E[Z|\alpha]\; \pi^{b}_{\beta,\chi,\psi}(\alpha)\; d\alpha \\
&= \int_{b}^{+\infty} \frac{1}{\alpha-1}\; \pi^{b}_{\beta,\chi,\psi}(\alpha)\; d\alpha, \quad \alpha > b > 1 \quad (4.5)
\end{aligned}
$$

Esta expresión (4.5) es la que se utilizará para estimar la prima de reaseguro.

El problema ahora está en encontrar el valor b en el cual truncar la distribución del parámetro. Si lo que se busca es la media de la distribución predictiva es necesario que el truncamiento b sea superior a 1, si es la varianza tendría que ser superior a 2 y así sucesivamente para momentos de orden superior.

Para fijar un valor se va utilizar el criterio de tomar el extremo inferior de un intervalo de credibilidad para el parámetro al γ %, con un valor de γ lo mayor posible, obtenido a partir de la distribución a posteriori de dicho parámetro, siempre y cuando este intervalo contenga exclusivamente valores superiores a 1, si no es así habrá que tomar un intervalo con una credibilidad menor hasta conseguirlo.

5-Aplicación

La muestra corresponde a unas carteras de seguros de responsabilidad civil de automóviles en España de 1992 a 2001, estudiada en (*Vilar, Lozano* (2007)).

Se van a utilizar los resultados de dicho trabajo correspondientes a todo el colectivo. El umbral fijado en este trabajo ha sido

$$a = 888.310{,}66 \text{ €}$$

La probabilidad estimada de que la cuantía sea superior a este umbral

$$P^{a} = \Pr(x > 888.310{,}66\ ) = 0{,}03436$$

El número de excesos sobre el umbral correspondiente es k=18.

El estimador $\hat{\lambda}$ del parámetro de Poisson se ha obtenido a partir del modelo conjugado Poisson- inversa Gaussiana generalizada $IGG(\delta', \nu', \tau')$, siendo la media de la distribución a posteriori del número de excesos, dada por (*Panjer, Willmot (1994))*:

$$\hat{\lambda} = \sqrt{\frac{\nu'}{\tau'}} \frac{K_{\delta'+1}(\sqrt{\nu'\tau'})}{K_{\delta'}(\sqrt{\nu'\tau'})} = 1.722$$

Se obtienen las cuantías de las diferentes primas netas de reaseguro para todas las distribuciones a posteriori obtenidas en (*Vilar, Lozano* (2007)).

Así como los valores de los extremos inferiores de los intervalos de credibilidad al 95% obtenido en dicho trabajo, para las diferentes distribuciones a priori consideradas.

Utilizando para la esperanza de la cuantía del exceso la expresión (4.5) y para la obtención de la prima la expresión (2.5).

Los resultados obtenidos con la distribución predictiva comparándolos con los obtenidos anteriormente se resumen en la tabla 1

		Prima(XL)	
Distribución a priori de α	Truncamiento(b)	(Predictiva)	Estimación puntual
Gamma de referencia	3,948	7.533	7.016
Gamma recíproca	4,5	8.166	7.833
Gamma	2,92	15.106	15.309
Inversa Gaussiana	2,52	14.638	14.076

Tabla 1

6. Conclusiones

Se han obtenido las expresiones de las funciones de distribución de unas distribuciones nuevas con cola muy gruesa, cuyo decrecimiento en la cola es más lento que las distribuciones de tipo polinómico, las llamadas de variación regular las distribuciones tipo Pareto, pero menos gruesas que la conocida log-Pareto. Distribuciones que se pueden utilizar para modelizar grandes cuantías y que se pueden englobar dentro de las denominadas súper gruesas. Así mismo se ha realizado un estudio comparativo entre el grueso de sus colas.

Estas distribuciones se han obtenido como distribuciones predictivas en un proceso de estimación bayesiana, con una distribución de Pareto y distribución del parámetro inversa Gaussiana generaliza, por tanto estas distribuciones son el resultado de ponderar una distribución de Pareto por tanto es posible aplicar las propiedades de las distribuciones ponderadas sobre todo en lo que se refiere a la obtención de los momentos a partir de las dos distribuciones la variable y la del parámetro de esta forma se ha podido estimar la prima en el caso de un contrato de reaseguro truncando la distribución del parámetro.

Se han obtenido las expresiones analíticas de las funciones de distribución de una colección de estas distribuciones de Pareto ponderadas, la ya conocida log-Pareto obtenida ponderando la distribución de Pareto por la distribución Gamma, pero también Pareto ponderada por Gamma inversa y Pareto ponderada por inversa Gaussiana.

El conocimiento de la expresión analítica de esta función de distribución permite obtener de manera sencilla cualquier cuantil, por supuesto la mediana de la distribución, es decir cualquier medida de posición de la distribución. También esta expresión analítica de la funciones de distribución pueden utilizarse para realizar de manera sencilla ajustes de Kolomogorov–Smirnov de grandes cuantías a estas distribuciones de colas súper gruesas. Finalmente en la aplicación practica que se ha realizado con los mismos datos que en el trabajo de(*Vilar, Lozano* (2007)) como cabía esperar las primas obtenidas con la distribución predictiva truncando la distribución del parámetro son superiores a las primas obtenidas en el trabajo de (*Vilar, Lozano* (2007)), dado que se toma una media ponderando por todos los posibles valores de $\alpha > 1$ mientras que en el caso de dicho trabajo se utilizaba un único valor para α , un estimador puntual, la media de la distribución a posteriori. Esto no se verifica para la distribución Gamma.

Referencias

- Antal, P. (2003): Quantitative methods in reinsurance. Swiss Re.
- Arnold B.C., Press S.J. (1989): Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association. Theory and Methods*. Vol. **84**, nº408, pp.1079-1084.
- Beirlant J., Teugels J.L., Vynckier P. (1996): Practical analysis of extreme values. Leuven University Press.
- Embrechts P., Klüppelberg C., Mikosch T. (1997): Modeling extremal events for insurance and finance. Springer.
- Fraga Alves, M.I,, de Haan, L. y Neves , C. (2006). Statistical inference for heavy tailed distribution. Disponible en http:// people.few.eur.nl/ ldehaan

- Hesselager O. (1993): A class of conjugate priors with applications to excess-of-loss reinsurance. *Astin Bulletin*, Vol. **23**, nº1, pp 77-93.
- Hill B.M. (1975): A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, Vol. **3**, nº5, pp.1163-1174.
- Hogg R.V., Klugman S.A. (1984): Loss distributions. Wiley. New York.
- Jorgensen (1982): Statistical properties of the generalized inverse Gaussian distribution. Springer-Verlag. New York.
- Klugman S.A., Panjer H.H., Willmot G.E. (1998): Loss models. From data to decisions. Wiley. New York.
- Lozano-Colomer C. (2005): Análisis Bayesiano de los excesos. Aplicación el reaseguro XL. Tesis doctoral (sin publicar).
- McNeil, A.J. (1997): Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin*. Vol. **27**, nº1, pp. 117-137.
- Pin-Hung Hsieh (2004): A data record analytic method for forecasting next record catastrophe loss. *Journal of Risk and Insurance*, June 2004, Vol. **71**, nº2, pp. 309-322.
- Reiss R.D., Thomas M. (1999): A new class of Bayesian estimators in Paretian excess-of-loss reinsurance. *Astin Bulletin* Vol. **29** nº2, pp. 339-349.
- Reiss R.D., Thomas M. (2001): Statistical analysis of extreme values. 2nd edition. Birkhäuser.
- Rytgaard, M. (1990): Estimation in the Pareto distribution. *Astin Bulletin*, Vol. **20**, nº2, pp. 201-216.
- Vilar-Zanón J.L., Lozano-Colomer C. (2007): On Pareto conjugate priors and their application to large claim reinsurance premium calculation. *Astin Bulletin*, Vol. 37(2), pp. 405-428.