ON PARETO CONJUGATE PRIORS
AND THEIR APPLICATION TO LARGE CLAIMS REINSURANCE
PREMIUM CALCULATION

BY

JOSÉ L. VILAR-ZANÓN AND CRISTINA LOZANO-COLOMER

ABSTRACT

This paper addresses the Bayesian estimation of the shape parameter of Pareto distributions, and its application to premium calculation of large claims excess of loss (XL) reinsurance contracts. It studies the use of the *generalized inverse Gaussian* (GIG) as a Pareto prior conjugate, a family that contains as a particular case the gamma distribution. An exact credibility formula is deduced allowing the calculation of individual reinsurance premiums. These are premiums suited to the excesses history of a sole portfolio. A family of predictive distributions for the excesses is derived. We apply our exact credibility model to a sample of excesses arisen in ten Spanish portfolios of liability motor insurance from year 1992 to year 2001.

KEYWORDS

Reinsurance, XL premiums, large claims, extreme value modeling, Bayes, exact credibility, Pareto distribution, generalized inverse Gaussian family.
AMS subject classification: 62P05, 60G70, 62F15.

1. INTRODUCTION

This paper addresses the Bayesian estimation of the shape parameter of Pareto distributions, and its application to the pricing of large claims XL reinsurance contracts. It is well known that modeling the tail of the claim severity distribution is of the utmost importance when dealing with large claims. *Extreme value theory* (EVT) gives clues for either determining a *threshold* defining the class of *large claims* or imposing by means of an asymptotic argument, a distribution function to the excesses over that threshold (see for instance Beirlant, Teugels, Vynckier (1996), Embrechts, Klüppelberg, Mikosch (1997), Reiss, Thomas (2001), and McNeil (1997)). This facilitates the tail modeling through a *generalized Pareto distribution* (GPD) making possible the subsequent reinsurance premium calculations. We will only focus on the most dangerous case when the GPD has a positive shape parameter $\xi > 0$, corresponding to a Pareto

distribution. In this paper we will consider XL contracts with a priority equal to the excess threshold and an unlimited cover. We will also distinguish among two kinds of reinsurance premiums. A *collective premium* is calculated starting from a sample obtained through the aggregation of the excesses occurred in many portfolios (the collective information), while an *individual premium* should be suited to the excesses history of a sole portfolio. Whenever it is possible to collect a large sample with a significant number of claim severities, it will be plausible to apply EVT techniques, making feasible the calculation of a collective premium. However, an individual premium is not so easy to obtain because an *individual sample* produced by a portfolio will always be scarce or even empty. It is at this point where the practitioner can find worth using Bayesian estimation of the Pareto index $\alpha = 1/\xi$, for it can give raise to exact credibility models able to drive the collective premium to an individual one. Therefore, the Bayesian methodology supports the idea that large claim severities arisen in different portfolios have not to be probabilistically identical because the risk selection policies carried out by each company may result in different probability tails.

The Bayesian estimation of the Pareto index has been a topic very often addressed in the literature. Fürst (1964) applies it to XL reinsurance using the gamma distribution as a Pareto prior conjugate family. Hill (1975) exploits the same conjugate family for estimation purposes. Arnold and Press (1989) include it in a more general conjugate model incorporating a scale parameter. Rytgaard (1990) uses credibility formulae to calculate reinsurance premiums corresponding to real world examples of fire and motor liability portfolios. Hesselager (1993) calculates XL premiums and predicted moments on the basis of this conjugate prior, working the same real world example as Rytgaard (1990). Reiss, Thomas (1999) study the Bayesian estimation for both the index and a scale parameter, using Rytgaard's examples among others. General presentations about Bayesian methods in extreme value modeling are found in Coles (1996, 2001) and Reiss, Thomas (2001). More recently, Pin-Hung Hsieh (2004) analyzes a statistical model for record catastrophe losses founded also on the Bayesian estimation of the Pareto index. Bühlmann, Gisler (2005) study the gamma family as a Pareto conjugate prior. In summary, the use of the gamma family as a prior when sampling a Pareto distribution has been frequently addressed through the years.

The election of this family determines the way prior information can be modeled. Reiss, Thomas (1999) point out that the lack of prior information can be implemented with the help of the *reference prior* (see for example Berger (1980)) $\pi(\alpha) = 1/\alpha, \forall \alpha > 0$, which results in a posterior belonging to the gamma family. Hesselager (1993) uses more precise prior information consisting in the knowledge of both the mean and the coefficient of variation. Pin-Hung Hsieh (2004) models different cases using a shifted gamma distribution.

It is also well known that the Pareto-gamma mixing produces an unconditioned distribution of a limited use because of its lack of finite moments. Some authors (see for instance Reiss, Thomas (2001) chapter 12 coauthored by Radtke)

try to cope with this conditioning the gamma in such a manner that the result-ing unconditioned excess distribution has a previously fixed number of finite moments (see also Pin-Hung Hsieh (2004)). A somewhat simpler approach consists in substituting a point estimate of the index into the Pareto distribution as will be done in the sequel.

In this article we will study the use of the *generalized inverse Gaussian* (GIG) as a Pareto prior conjugate, and we will refer to Jorgensen (1982) as the basic tool for deepening in this three parameter family. It contains the gamma and other distributions able to represent prior beliefs on the Pareto index. We will show how to calculate Bayesian estimates of this parameter using that conjugate family. We will also see that the GIG produces a new family of unconditioned distributions that suffers too from the non-existence of moments.

Besides, the GIG family also works as a prior conjugate when sampling a Poisson distribution (see for example Tremblay (1992), Panjer, Willmot (1992), Lemaire (1995)) with random Poisson parameter. Therefore, if we suppose the conditioned number of excesses over the threshold as Poisson distributed, it can also furnish Bayesian estimates of the mean number of those excesses. Considering the total portfolio excesses modeled by means of a compound Poisson-Pareto distribution, we will be able to calculate *collective* premiums and also to write down an exact credibility formula that will open the way to the calculation of the *individual* reinsurance premiums.

Finally, we will apply this general scheme to the Spanish liability motor insurance case. For that purpose, we will study a sample consisting of 395 claims occurred during the period 1992-2001 in ten portfolios, with a *communication level* equal to 25 millions of pesetas (approximately 150,253.02 €). We will apply standard EVT techniques to determine a threshold defining the class of large claims. We will then use our conjugate model to calculate Bayesian estimates of the Pareto index. Finally, we will obtain XL reinsurance premiums for the collective and for some of the individual portfolios aggregated in our sample.

2. ASSUMPTIONS

Let the random variables N and X stand for the claim number and the claim severities respectively. The threshold being fixed at $T > 0$, we assume that $p^T = P\{X > T\} > 0$. Then two new random variables are defined, the number of excesses over T (to be noted N^T), and the excesses over $T (X - T \,|\, X > T)$. As told in the introduction, we suppose that the threshold is large enough, so that the excesses can be considered nearly GPD distributed. Therefore, supposing a Pareto type claim severity distribution, this GPD will fall in the Pareto case. We will represent the *total normalized excesses over the threshold* as a compound random variable

$$S^T = \sum_{i=1}^{N^T} Z_i \quad (S^T = 0 \text{ when } N^T = 0) \tag{2.1}$$

where N^T is the excess number in one year, and the random variables $(Z_i)_{i=1}^\infty$ are the *normalized excesses* over the threshold:

$$Z = \frac{X - T}{T} \mid X > T \qquad (2.2)$$

Normalizing the excesses will avoid us any discussion derived from the consideration of localization and scale parameters in the excess distribution.

We will assume the existence of an absolutely continuous random variable $\Lambda^T > 0$ (*a.s.*), independent of time with distribution function $U(\lambda) = P\{\Lambda^T \le \lambda\}$ and density function $u(\lambda)$, so that the number of excesses in different years are supposed to be conditionally independent given Λ^T,

$$\forall k \in \mathbb{N}, \forall n_1, \ldots, n_k \in \mathbb{N}_0 : P\{N_1^T = n_1, \ldots, N_k^T = n_k \mid \Lambda^T\} =$$
$$\prod_{i=1}^{k} P\{N_i^T = n_i \mid \Lambda^T\} \ (a.s.) \qquad (2.3)$$

and conditionally Poisson distributed

$$\forall i \in \mathbb{N}, \forall n \in \mathbb{N}_0 : P\{N_i^T = n \mid \Lambda^T\} = \frac{\left(\Lambda^T\right)^n}{n!} e^{-\Lambda^T} \ (a.s.) \qquad (2.4)$$

Thus the unconditioned numbers of excesses are mixed Poisson distributed:

$$\forall i \in \mathbb{N}, \forall n \in \mathbb{N}_0 : p_n^T = \Pr\{N_i^T = n\} = \int_0^{+\infty} \frac{\lambda^n}{n!} e^{-\lambda} u(\lambda) \, d\lambda \qquad (2.5)$$

We will also consider the Pareto shape parameter as an absolutely continuous random variable $\Xi > 0$ (*a.s.*), independent of time and independent with respect to the random parameter Λ^T. Then the Pareto index will be its *reciprocal* $A = \frac{1}{\Xi} > 0$ (*a.s.*). This last random variable is supposed to be distributed according to $\Pi(\alpha) = P\{A \le \alpha\}$, with density function $\pi(\alpha)$. The random variables $(Z_i)_{i \in \mathbb{N}}$ are conditionally independent given A, meaning in particular that

$$\forall n \in \mathbb{N}, \forall z_1, \ldots, z_n \in \mathbb{R} : P\{Z_1 \le z_1, \ldots, Z_n \le z_n \mid A\} =$$
$$\prod_{i=1}^{n} P\{Z_i \le z_i \mid A\} \ (a.s.) \qquad (2.6)$$

and conditionally Pareto distributed

$$\forall i \in \mathbb{N}, \forall z > 0 : P\{Z_i \le z \mid A\} = 1 - \frac{1}{(1 + z)^A} \ (a.s.) \qquad (2.7)$$

Under these assumptions, the unconditioned normalized excesses density function is mixed Pareto:

$$h(z) = \int_0^{+\infty} \frac{\alpha}{(1+z)^{\alpha+1}} \, \pi(\alpha) \, d\alpha \qquad (2.8)$$

Finally, we will assume independence of $(N_1^T, \ldots, N_k^T, \Lambda^T)$ and $(Z_1, \ldots, Z_n, A)$ for all $k, n \in \mathbb{N}$.

Given the independence between Λ^T, A, their prior joint density can be written as the product $w(\lambda, \alpha) = u(\lambda) \pi(\alpha)$. Under this condition (see Reiss, Thomas (2001) p. 204), once a number of excesses and a sample of normalized excesses have been observed during k years

$$n = \sum_{i=1}^{k} n_i, \; (z_1, \ldots, z_n) \qquad (2.9)$$

the posterior joint density for the (Λ^T, A)-random vector is:

$$w(\lambda, \alpha \,|\, n_1, \ldots, n_k; z_1, \ldots, z_n) = u(\lambda \,|\, n_1, \ldots, n_k) \, \pi(\alpha \,|\, z_1, \ldots, z_n) \qquad (2.10)$$

where

$$u(\lambda \,|\, n_1, \ldots, n_k) \propto \Pr\{N_1^T = n_1, \, \ldots, \, N_k^T = n_k \,|\, \Lambda^T = \lambda\} u(\lambda)$$
$$\pi(\alpha \,|\, z_1, \ldots, z_n) \propto g(z_1, \ldots, z_n \,|\, A = \alpha) \, \pi(\alpha) \qquad (2.11)$$

and $g(z_1, \ldots, z_n \,|\, A = \alpha)$ is the Pareto likelihood.

Taking the squared error as a loss function, it is known that the Bayes estimates for the two parameters are their posterior means

$$\lambda(n_1, \ldots, n_k) = E\{\Lambda^T \,|\, N_1^T = n_1, \, \ldots, \, N_k^T = n_k\} = \int_0^{+\infty} \lambda u(\lambda \,|\, n_1, \ldots, n_k) d\lambda \qquad (2.12)$$

$$\alpha(z_1, \ldots, z_n) = E\{A \,|\, z_1, \ldots, z_n\} = \int_0^{+\infty} \alpha \pi(\alpha \,|\, z_1, \ldots, z_n) d\alpha \qquad (2.13)$$

Now, a formula for the reinsurance premium P can be set down substituting the point estimates into the respective conditioned distribution functions, then writing down the mean of the conditioned compound random variable $(S^T \,|\, \Lambda^T = \hat{\lambda}, A = \hat{a})$:

$$P(n_1, \ldots, n_k; z_1, \ldots, z_n) = E\left\{S^T \,\middle|\, \Lambda^T = \hat{\lambda}(n_1, \ldots, n_k), A = \hat{a}(z_1, \ldots, z_n)\right\}$$
$$= E\left\{N^T \,\middle|\, \Lambda^T = \hat{\lambda}(n_1, \ldots, n_k)\right\} E\left\{Z \,\middle|\, A = \hat{a}(z_1, \ldots, z_n)\right\} \qquad (2.14)$$
$$= \frac{\hat{\lambda}(n_1, \ldots, n_k)}{\hat{a}(z_1, \ldots, z_n) - 1} p^T$$

(for $\alpha(z_1,\ldots,z_n) > 1$). Formula (2.14) is given in threshold units (T-units), while expressing it in usual monetary units gives

$$P(n_1,\ldots,n_k; z_1,\ldots,z_n) = \frac{\hat{\lambda}(n_1,\ldots,n_k)T}{\hat{a}(z_1,\ldots,z_n) - 1}\, p^T \qquad (2.15)$$

3. THE GENERALIZED INVERSE GAUSSIAN FAMILY AS A PARETO PRIOR CONJUGATE

The generalized inverse Gaussian (GIG) family is a well known probabilistic model (see for instance Jorgensen (1982)). It has been already applied in actuarial science to describe the portfolio heterogeneity (Tremblay (1992), Lemaire (1995)) giving rise to some useful mixed Poisson distributions (Panjer, Willmot (1992)). It is a three parameter family with density function:

$$\pi_{\beta,\chi,\psi}(\alpha) = \frac{(\psi/\chi)^{\beta/2}}{2K_\beta\left(\sqrt{\chi\psi}\right)}\, \alpha^{\beta-1} e^{-\frac{1}{2}\left(\frac{\chi}{\alpha} + \psi\alpha\right)}, \quad \alpha > 0,\ \beta \in \mathbb{R},\ \chi, \psi > 0 \qquad (3.1)$$

where

$$K_\beta(\omega) = \frac{1}{2}\int_0^{+\infty} y^{\beta-1} e^{-\frac{\omega}{2}\left(\frac{1}{y} + y\right)} dy, \quad \beta \in \mathbb{R},\ \omega > 0 \qquad (3.2)$$

is a Bessel function of the third kind (see for instance Panjer, Willmot (1992)). The cases where $\chi\psi = 0$ are individually considered later, see (3.10) and (3.15). The central moments and variance of the GIG are

$$E\{A^j\} = \left(\frac{\chi}{\psi}\right)^{j/2} \frac{K_{\beta+j}\left(\sqrt{\chi\psi}\right)}{K_\beta\left(\sqrt{\chi\psi}\right)}, \quad \chi\psi > 0,\ j \in \mathbb{R} \qquad (3.3)$$

$$Var\{A\} = \frac{\chi}{\psi}\left[\frac{K_{\beta+2}\left(\sqrt{\chi\psi}\right)}{K_\beta\left(\sqrt{\chi\psi}\right)} - \frac{K_{\beta+1}^2\left(\sqrt{\chi\psi}\right)}{K_\beta^2\left(\sqrt{\chi\psi}\right)}\right], \quad \chi\psi > 0 \qquad (3.4)$$

Considering the normalized Pareto (2.7) as the sampling distribution we get the following likelihood:

$$g(z_1,\ldots,z_n \mid A = \alpha) = \alpha^n e^{-(\alpha+1)\sum_{i=1}^{n}\log(z_1+1)} \qquad (3.5)$$

Supposing the random parameter distributed according to a $GIG(\beta,\chi,\psi)$ we get

$$g(z_1, \ldots, z_n \mid A = \alpha) \, \pi_{\beta, \chi, \psi}(\alpha) \propto$$

$$\propto \alpha^{n+\beta-1} \exp\left\{ -\frac{1}{2}\left[\left(2\sum_{i=1}^{n}\log(z_i+1) + \psi \right)\alpha + \frac{\chi}{\alpha} \right] \right\} \quad (3.6)$$

$$\propto \pi_{\beta', \chi', \psi'}(\alpha \mid z_1, \ldots, z_n)$$

Thus the posterior distribution of A is still a GIG with parameters:

$$\beta' = n + \beta$$
$$\chi' = \chi \quad (3.7)$$
$$\psi' = 2\sum_{i=1}^{n}\log(z_i+1) + \psi$$

We apply (3.3) to express the posterior mean of the random Pareto index A, getting

$$E\{A \mid z_1, \ldots, z_n\} = \hat{a}(z_1, \ldots, z_n) = \sqrt{\frac{\chi'}{\psi'}} \, \frac{K_{\beta'+1}\left(\sqrt{\chi'\psi'}\right)}{K_{\beta'}\left(\sqrt{\chi'\psi'}\right)} \quad (3.8)$$

where β', χ', ψ' are given by (3.7).

Whenever $\chi\psi > 0$, the unconditioned or predictive density function (2.8) is given by (see details in the appendix)

$$h_{\beta, \chi, \psi}(z) = \sqrt{\frac{\chi}{\psi}} \, \frac{K_{\beta+1}\left(\sqrt{\chi\psi}\left[1 + \frac{2}{\psi}\log(1+z)\right]^{1/2} \right)}{K_{\beta}\left(\sqrt{\chi\psi}\right)} (1+z)^{-1}$$
$$\left(1 + \frac{2}{\psi}\log(1+z)\right)^{-\left(\frac{\beta+1}{2}\right)} \quad (3.9)$$

This distribution is not easy to handle for the variable z appears inside the argument of the Bessel function. In fact, this expression is only tractable in a number of cases some of which will be listed below.

In this paper we will illustrate the choice of a prior belonging to this family through the use of the gamma, the inverse Gaussian, the reciprocal gamma and the reciprocal inverse Gaussian distributions. Details about these probability distributions are given hereafter.

Gamma (G): Choosing $\beta > 0$, $\chi = 0$, $\psi > 0$ we obtain as a limiting case of the GIG family (3.1), a gamma$(\beta, \psi/2)$ with density function

$$\pi_{\beta, 0, \psi}(\alpha) = \frac{(\psi/2)^{\beta}}{\Gamma(\beta)} \alpha^{\beta-1} e^{-\frac{\psi}{2}\alpha}, \ \alpha > 0 \qquad (3.10)$$

This is the well known conjugate model gamma-Pareto-gamma, where the predictive density is the log-Pareto distribution (see Reiss, Thomas (2001)) with density

$$h_{\beta, 0, \psi}(z) = \left(\frac{\psi}{2}\right)^{\beta} \beta(1+z)^{-1} \left(\frac{\psi}{2} + \log(1+z)\right)^{-(\beta+1)}, \ z > 0 \qquad (3.11)$$

The *reference prior*

$$\pi(\alpha) = 1/\alpha, \ \alpha > 0 \qquad (3.12)$$

models the situation where no prior information is available. Therefore, when considering it as the prior, the subsequent Bayesian analysis will only depend in the objective information, that is to say, in the sample information. A simple calculation shows that in this case, the posterior distribution is a gamma $\left(n, \sum_{i=1}^{n} \log(1+z_i)\right)$. In an attempt to explain some numerical results to be obtained in the next section, we must stress the importance of this last posterior, for in this case the Bayes and the maximum likelihood estimators (MLE) coincide (just look to the parameters of the posterior distribution). Additionally, we know that since the threshold has been previously fixed equal to one of the ordered data points, the MLE and Hill estimators share the same expression. Therefore, if we calculate the Bayes estimate by means of the reference prior (3.12), we will obtain a numerical result that could be simultaneously interpreted in three ways.

Inverse Gaussian (IG): It corresponds to $\beta = -\frac{1}{2}, \chi > 0, \psi > 0$. Writing $\psi = \frac{\chi}{m^2}$ ($m > 0$) we obtain an $IG(m, \chi)$ density function (see again Jorgensen (1982)) in its more familiar form

$$\pi(\alpha) = \sqrt{\frac{\chi}{2\pi}} \alpha^{-3/2} e^{\frac{\chi}{2m^2} \frac{(\alpha-m)^2}{\alpha}}, \ \alpha > 0 \qquad (3.13)$$

The unconditioned density (2.8) turns out to be (see the appendix)

$$h_{-\frac{1}{2}, \chi, \psi}(z) = \sqrt{\frac{\chi}{\psi}} (1+z)^{-1} \left(1 + \frac{2}{\psi} \log(1+z)\right)^{-1/2} \qquad (3.14)$$

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

Reciprocal gamma (RG): Considering $\beta < 0$, $\chi > 0$, $\psi = 0$, we get the *reciprocal gamma density*

$$\pi_{\beta, \chi, 0}(\alpha) = \left(\frac{2}{\chi}\right)^{\beta} \Gamma(-\beta)^{-1} \alpha^{\beta-1} e^{-\frac{\chi}{2\alpha}}, \quad \alpha > 0 \tag{3.15}$$

corresponding to the following unconditioned density function:

$$h_{\beta, \chi, 0}(z) = \left(\frac{2}{\log(1+z)}\right)^{\frac{\beta+1}{2}} \frac{\chi^{\frac{1-\beta}{2}}}{\Gamma(-\beta)(1+z)} K_{\beta+1}\left(\sqrt{2\chi \log(1+z)}\right), \quad z > 0 \tag{3.16}$$

The central moments are

$$E\{A^k\} = \left(\frac{\chi}{2}\right)^k \frac{\Gamma(-\beta-k)}{\Gamma(-\beta)}, \quad k < -\beta \tag{3.17}$$

Thus the mean and the variance are respectively

$$\begin{aligned} E\{A\} &= -\frac{\chi}{2(\beta+1)}, \quad \beta < -1 \\ Var\{A\} &= -\left(\frac{\chi}{2}\right)^2 \frac{1}{(\beta+1)^2(\beta+2)}, \quad \beta < -2 \end{aligned} \tag{3.18}$$

If we choose $\beta \in [-2, -1)$ the variance will be equal to infinity. Therefore any of these priors might be used as a representative of our lack of knowledge about the parameter when we still have some belief about its mean value. In this case we would only need to supply a *subjective* mean value μ, the β, χ, ψ parameters of the prior being obtained then through:

$$\chi = -2\mu(\beta+1), \quad \beta \in [-2, -1), \quad \psi = 0 \tag{3.19}$$

Reciprocal inverse Gaussian (RIG): It corresponds to $\beta = \frac{1}{2}$, $\chi > 0$, $\psi > 0$. The density function is

$$\pi_{\frac{1}{2}, \chi, \psi}(\alpha) = \frac{\psi^{-1/2}}{\sqrt{2\pi}} \alpha^{-1/2} e^{\frac{1}{2}\left(\chi\alpha^{-1} + \psi\alpha\right)}, \quad \alpha > 0 \tag{3.20}$$

and the unconditioned density function is (see again the appendix) $\forall z > 0$:

$$\begin{aligned} h_{\frac{1}{2}, \chi, \psi}(z) = \sqrt{\frac{\chi}{\psi}} \frac{e^{\sqrt{\chi\psi}}}{z+1} e^{-\sqrt{\chi\psi}\left[1 + \frac{2}{\psi}\log(1+z)\right]^{1/2}} \\ \left(1 + \frac{2}{\psi}\log(1+z)\right)^{-1} \left[1 + \left(\chi\psi\left(1 + \frac{2}{\psi}\log(1+z)\right)\right)^{-1/2}\right] \end{aligned} \tag{3.21}$$

Next we move to the Bayesian analysis of the number of excesses. In this case the GIG density $\pi_{\beta,\chi,\psi}(\alpha)$ will be renamed $u_{\delta,\nu,\tau}(\lambda)$. Under the hypothesis (2.3), (2.4) and (2.5) we get, given a sample $(n_1, \ldots, n_k)$ of number of excesses occurred during k years,

$$
\begin{aligned}
u(\lambda \mid n_1, \ldots, n_k) &\propto \Pr\left\{ N_1^T = n_1, \ldots, N_k^T = n_k \,\middle|\, \Lambda^T = \lambda \right\} u_{\delta,\nu,\tau}(\lambda) \\
&\propto \lambda^{\sum_{i=1}^k n_i + \delta - 1} e^{-\frac{1}{2}\left(\frac{\nu}{\lambda} + (2k+\tau)\lambda\right)}
\end{aligned}
\tag{3.22}
$$

Therefore the posterior distribution for the random variable Λ^T is again a GIG with parameters

$$
\delta' = \sum_{i=1}^k n_i + \delta, \quad \nu' = \nu, \quad \tau' = 2k + \tau
\tag{3.23}
$$

This is the well known conjugate model where the Poisson mixed with the GIG is called Sichel's distribution (see Panjer, Willmot (1992) p. 280). Recalling (3.3), the posterior mean number of excesses is

$$
\hat{\lambda}(n_1, \ldots, n_k) = E\{\Lambda^T \mid N_1^T = n_1, \ldots, N_k^T = n_k\} = \sqrt{\frac{\nu'}{\tau'}} \, \frac{K_{\delta'+1}\left(\sqrt{\nu'\tau'}\right)}{K_{\delta'}\left(\sqrt{\nu'\tau'}\right)}
\tag{3.24}
$$

This conjugate model contains the well known cases of the gamma and inverse Gaussian mixing distributions. Using any of these distributions presupposes some prior knowledge on the number of excesses phenomena, because it implies the fixation of the first and second moments of the Poisson parameter Λ^T. This could be feasible only if the complete sample of the portfolio individual losses had been available, so that the whole portfolio random parameter Λ (and not only that of the excesses, Λ^T) could be considered as $GIG(\delta, \nu, \tau)$ distributed. Recalling that , $p^T = \Pr\{X > T\}$ it can be proved that the Poisson parameter Λ^T is then distributed according to a $GIG(\delta, p^T\nu, \frac{\tau}{p^T})$ (this follows the way of reasoning that can be found in Straub (1971), and also in Patrick, Mashitz (1989)). Nevertheless, being conscious that our sample consists just in the claim costs over some communication level is far more realistic. This makes the modeling of the whole portfolio structure impossible. In brief, any prior assessment going further than the mean excess number would be hazardous and hardly defensible. This justifies the modeling of the prior beliefs on the expected number of excesses by means of a reciprocal gamma distribution with infinite variance.

Finally, substituting the posterior means (3.8) and (3.24) into (2.14) we obtain an exact credibility formula allowing the XL premium estimation depending on the historical number of excesses and the excesses of a portfolio.

4. APPLICATION

The sample was collected from ten Spanish liability motor insurance portfolios. It consists in 395 claims greater than 25 millions pesetas that aroused from year 1992 to year 2001 (we recall the exchange rate 1€ = 166.386 Pts, which approximately gives 6€ ≃ 1,000 Pts and 150,253.02€ ≃ 25 Millions Pts).

The lack of knowledge about the claim dates made impossible a deep analysis of the time series. For that reason, the claims were deflated to year 2001, trying that way to validate the stationarity property as explained in Bühlmann, Gisler (2005). Next, a non parametrical Kruskal-Wallis test was applied to these deflated claims, with the conclusion that we could not reject the hypothesis that the ten yearly samples belong to the same population, at a 1% significance level (see for instance Siegel S., Castellan N.J. (1988)). Consequently, we may assume that all the claims are identically distributed. The last step consisted in a conversion to Euros in an attempt to bring the foregoing calculations nearer to the reader's perception.

A threshold was fixed by means of standard graphical techniques (see for instance Reiss, Thomas (2001), Beirlant, Teugels, Vynckier (1996), or Embrechts, Klüppelberg, Mikosch (1997)). For this aim we have generated the exponential quantile-quantile plot (QQ-plot) and plotted the sample mean excesses (figures 1 and 2).
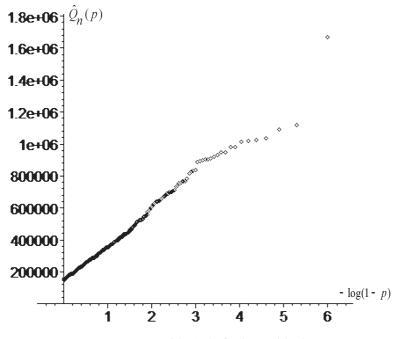

FIGURE 1: Exponential QQ-plot for the 395 claim data.

In figure 1 we can see a nearly linear plot denouncing the trace of an exponential distribution over an interval of claim severities with left end point at 150,253.02€. Figure 2 allows us choosing as the threshold $X^*_{395-18} = 888,310.66€$ this value being the one where the mean excesses begin to increase. Once the threshold has been fixed, we adjust by least squares a straight line to the first 377 data, obtaining

$$y = 218969.1049\, x + 150253.0261 \tag{4.1}$$

The linear correlation coefficient is .99753753 (figure 3).

Therefore we may model the claim sizes distribution between 150,253.0261€ and 888,310.6632€ by means of a shifted and truncated exponential distribution with estimated parameter $\hat{\theta} = .4566854308e - 5$. For claim severities $x > 150253.0261$ the shifted exponential would be

$$f(x) = \hat{\theta} \exp\{-\hat{\theta}\,(x - 150253.0261)\} \tag{4.2}$$

According to (4.2), the cumulated probability in the interval [150253.0261, 888310.6632] is equal to .9656310486, while the tail probability is

$$p^T = P\{X > 888310.6632\} = .0343689514$$

denouncing that about 3.5% of the claim severities greater than 150,253.02 € will be above the threshold. Given a threshold fixed at $X^*_{377} = 888,310.6632$ the claim severities above it consist in the eighteen data points whose distribution through the portfolios and over the years is summarized in table 1.

Next, we calculate the Hill estimates $\hat{\alpha} = 1/H_{k,395}$ (see figure 4) where

$$H_{k,n} = \frac{1}{k} \sum_{j=1}^{k} \log(X^*_{n-j+1}) - \log(X^*_{n-k}),\ k < n \tag{4.3}$$

Corresponding to a number of excesses $k = 18$ we obtain

$$\hat{\alpha} = 1/H_{18,395} = 8.495929855 \tag{4.4}$$

This estimation coincides with the MLE one since we are working in the case where a threshold $T = X^*_{395-18} = 888,310.6632$ has been previously fixed, so the two estimators result in the same expression (4.3).

We begin the Bayesian analysis of the Pareto index calculating its collective estimation. For this sake we can model the lack of prior information in two ways. We may consider the reference prior (3.12) if we want to represent a total default of information. Or we may use a reciprocal gamma distribution (3.15) with infinite variance and finite mean to model vague prior belief about
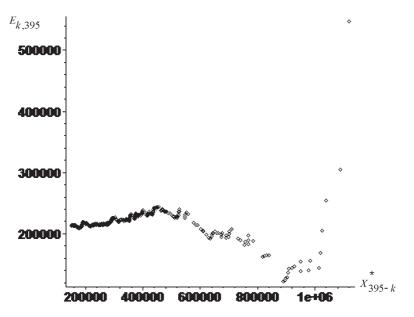
FIGURE 2: Abscissas: ordered sample. Ordinates: sample mean excesses $E_{k,395}$, $k = 1, \ldots, 394$. The lowest point is attained at $k = 18$ with a size of 888,310.66€.
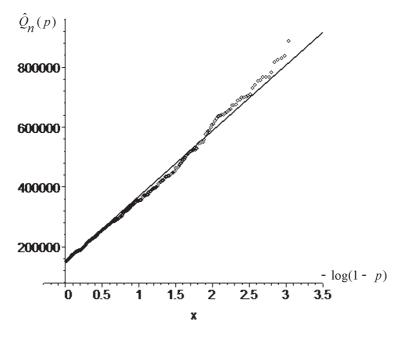

FIGURE 3: Line adjusted by least squares to the QQ-plot of the first 377 data.

TABLE 1

THE EIGHTEEN CLAIMS ROUNDED TO THE NEAREST UNIT, OVER THE THRESHOLD $X_{377}^* = 888.310.66€$ (LAST COLUMN IN BOLD CHARACTERS), GROUPED BY PORTFOLIOS (COLUMNS A, C, E, F, H) AND YEARS (ROWS 1992,…, 2001). NO CLAIM ABOVE THE THRESHOLD WAS COMMUNICATED EITHER IN THE OTHER FIVE INSURANCE COMPANIES OR IN YEAR 2001.

	A	C	E	F	H
1992	949,459		1,667,634		
1993	905,909		1,119,830		
1994		1,014,673			
1995			978,507		949,494
1996	1,018,229				**888,310** 929,092
1997		894,461	903,656 1,088,642		1,038,262
1998			908,713 1,023,955	898,126	
1999		918,619			
2000		982,315			
2001					

the parameter expectation. Having in mind that a mean 2.5 has been generally assumed in motor liability insurance (see the example worked out in Hesselager (1993), Reiss, Thomas (1999)), we have considered $\mu = 2.5$ jointly with $\beta = -2$ in (3.19). The resulting calculations are summarized in table 2 where we have

TABLE 2

THE COLLECTIVE PARETO INDEX. ROWS: A 95% PROBABILITY α-INTERVAL, POSTERIOR MEAN, STANDARD DEVIATION, COEFFICIENT OF VARIATION AND SKEWNESS. COLUMNS: SUCCESSIVE PRIORS ARE CHOSEN AS RECIPROCAL GAMMA (MEAN = 2.5, STANDARD DEVIATION = ∞), GAMMA REFERENCE, INVERSE GAUSSIAN (MEAN = 2.5, COEFFICIENT VARIATION = 0.3), RECIPROCAL INVERSE GAUSSIAN (MEAN = 2.5, COEFFICIENT VARIATION = 0.3), AND GAMMA (MEAN = 2.5, COEFFICIENT VARIATION = 0.3).

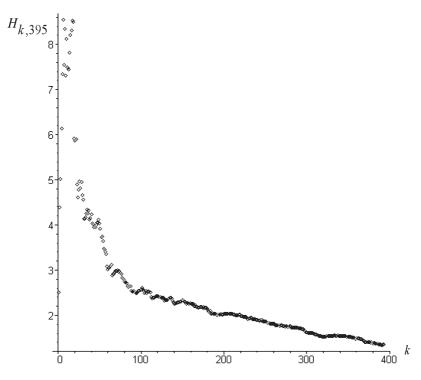|	Reciprocal gamma	Gamma reference	Inverse Gaussian	Rec. inverse Gaussian	Gamma
95% Prob. interval	(4.5, 11.88)	(3.94, 12.09)	(2.52, 6.49)	(2.54, 6.43)	(2.92, 5.64)
μ	7.714613	8.495929	4.736500	4.707062	4.435569
σ	1.888458	2.002509	0.975273	0.959308	0.822091
σ/μ	0.244789	0.235702	0.205905	0.203802	0.185340
γ	0.499636	0.471404	0.464210	0.464210	0.370681

FIGURE 4: Inverses of the Hill estimates corresponding to the number of excesses $k = 1, ..., 395$. For $k = 18$ we get $\hat{\alpha} = 1/H_{18,395} = 8.495929855$. These could be considered as well as the MLE in case of a previously fixed threshold $T = X^*_{395-18} = 888,310.6632$.

written in the first and second columns a 95%-probability A-interval followed by its mean, standard deviation, coefficient of variation and skewness, all given by the posterior distributions.

Another possibility is to assume some prior knowledge about the first and second moments of α. While we choose the same value for the mean, we will set the coefficient of variation equal to 0.3 as it has been established for motor liability insurance in the references cited above. Then we can choose as a prior some of the distributions reviewed in section 3. The same characteristics calculated by means of their respective posterior distributions are quoted in the third, fourth and fifth columns of table 2.

The next step is the estimation of the individual Pareto index for each of the portfolios contributing to the sample of claim severities. For the sake of concision, this will be done only for portfolios A, E and H (see table 1). The estimation using no prior knowledge (gamma reference) or vague information (reciprocal gamma) is summarized in table 3.

Assuming for the prior distribution (in each of the three cases gamma, inverse Gaussian and reciprocal inverse Gaussian) a mean value of 2.5 and a coefficient of variation of 0.3, we get the results listed in table 4.

TABLE 3

THE POSTERIOR INDIVIDUAL PARETO INDEXES. ROWS: FOR EACH PORTFOLIO (A, E, H) WE CONSIDER THE CASES OF VAGUE (RECIPROCAL GAMMA WITH MEAN = 2.5 AND STANDARD DEVIATION = ∞) AND NO PRIOR KNOWLEDGE (GAMMA REFERENCE). COLUMNS: A 95% PROBABILITY α-INTERVAL, POSTERIOR MEAN, STANDARD DEVIATION, COEFFICIENT OF VARIATION AND SKEWNESS.

Portf.	α Prior	95% Prob. Interval	μ	σ	σ/μ	γ
A	Rec. gamma	(1, 17.70)	7.070347	4.973582	0.703442	1.690130
	Gamma ref.	(1, 28.46)	13.471733	7.777909	0.577350	1.154700
E	Rec. gamma	(1,7.38)	4.238813	1.680509	0.396457	0.873800
	Gamma ref.	(1, 8.82)	5.210403	1.969347	0.377964	0.755928
H	Rec. gamma	(1, 15.33)	6.137050	4.191834	0.680374	1.659250
	Gamma ref.	(1, 23.81)	11.215884	6.475493	0.577350	1.154700

TABLE 4

THE INDIVIDUAL PARETO INDEXES WITH SOME PRIOR KNOWLEDGE. COLUMNS: A 95% PROBABILITY α-INTERVAL, POSTERIOR MEAN, STANDARD DEVIATION, COEFFICIENT OF VARIATION AND SKEWNESS. ROWS: PORTFOLIOS A, E, H, WITH PRIORS GAMMA, INVERSE GAUSSIAN AND RECIPROCAL INVERSE GAUSSIAN. ALL THESE CONSIDERED WITH MEAN 2.5 AND COEFFICIENT OF VARIATION 0.3.

Portf.	α Prior	95% Prob. Interval	μ	σ	σ/μ	γ
A	Gamma	(1.46 , 4.14)	3.023507	0.804878	0.266206	0.532413
	Inverse Gaussian	(0.76, 4.65)	3.045151	0.875818	0.287610	0.793319
	Rec. inverse Gaussian	(0.80, 4.63)	3.043348	0.869265	0.285622	0.766482
E	Gamma	(1.73 , 4.17)	3.129127	0.735276	0.234978	0.469956
	Inverse Gaussian	(1.20, 4.57)	3.143024	0.792164	0.252038	0.647198
	Rec. inverse Gaussian	(1.23, 4.56)	3.142330	0.787341	0.250559	0.629559
H	Gamma	(1.45 , 4.10)	2.994767	0.797227	0.266206	0.532413
	Inverse Gaussian	(0.77, 4.59)	3.011321	0.862422	0.286393	0.790522
	Rec. inverse Gaussian	(0.80, 4.57)	3.010001	0.856504	0.284552	0.764241

In table 2 we see that the probability A-intervals are wider in the cases of no prior information than in the prior information ones (as could be expected). The same observation applies for the individual Pareto indexes (compare the intervals found in table 3 — no prior information — with those found in table 4 — prior information). As an illustration, in figure 5 we can find the plots of the posterior A-densities in the case of the reciprocal gamma prior.

In figure 6 we plot, for the collective normalized excesses, its sampling distribution together with the Pareto distributions resulting from substitution of α by its Bayes estimations when choosing as a prior each one of the cases reviewed in this paper. Even if the main foundation of our modeling lays on
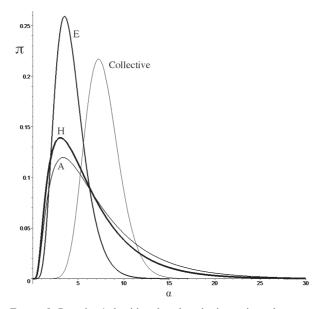
FIGURE 5: Posterior A-densities when the prior is a reciprocal gamma:
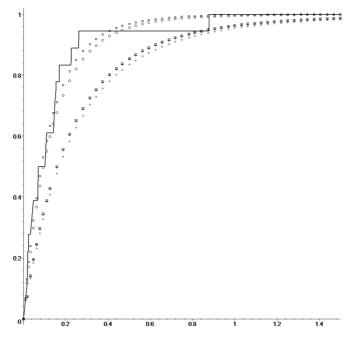portfolios A, H, E and the collective.

FIGURE 6: Excess distributions in the following cases (from left to right): sampling distribution
(continuous), and Pareto distributions obtained through substitution of the Bayes α-estimate
corresponding to the following priors: the reference prior (box), reciprocal gamma (circle), inverse
Gaussian (box), reciprocal inverse Gaussian (cross), gamma (diamond).

an asymptotic argument (extreme value theory) that drive us to the assumption that the severity distribution should be of Pareto type, we see it useful to strengthen this assumption by conducting a goodness-of-fit test on the Pareto family as a suitable model for our normalized excesses. For that aim we will use the Pareto distribution were α has been substituted by the Bayes estimate resulting from the reference prior. Remember that this case takes only into account the sample information (there is no subjective one), and the three statistics (Bayes with quadratic loss function, MLE and Hill) result in the same expression (see the comments after (3.12)).

Placing ourselves in the case where the Pareto index is unknown (Case II, see D'Agostino, Stephens (1986) p. 134, 254) and considering the MLE estimation $\hat{\alpha} = 8.495929$ (see table 2, second column) we obtain a Cramér-von-Mises statistic equal to 0.03551, which is lower than 0.222 the critical value at a 5% significance level. It follows that the Pareto family should not be rejected as a model for our sample. If we try to conduct Goodness-of-fit tests for the other mentioned Pareto distributions, the situation would be easier as we may place ourselves in a Case 0 (see D'Agostino, Stephens (1986) p. 104-105) where the distribution is fully specified. Then calculating the Cramér-von-Mises statistic for each of the considered Pareto distributions, we would check that the heavier is the subjective information used in the Pareto index estimation,
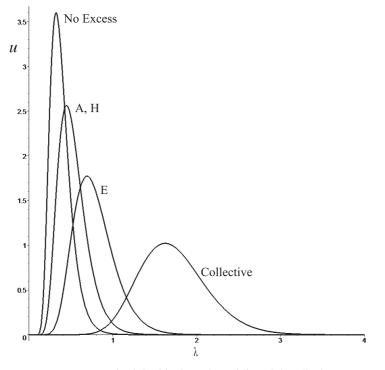

FIGURE 7: Posterior λ-densities for each portfolio and the collective.

TABLE 5

Portfolio	Posterior λ	95% Prob. Interval	XL Premium	E.SF. 50% level	α Prior distribution
Collective	1.722697	(0.80, 2.44)	7,833	14,223	Reciprocal gamma
			7,016	12,713	Gamma reference
A	0.524104	(0.04, 0.84)	2,636	5,192	Reciprocal gamma
			1,283	2,530	Gamma reference
E	0.781445	(0.15, 1.22)	7,366	14,387	Reciprocal gamma
			5,666	11,372	Gamma reference
H	0.524104	(0.04, 0.84)	3,115	6,244	Reciprocal gamma
			1,566	3,145	Gamma reference
No excess	0.381603	(0.01, 0.61)	1,735	3,458	Reciprocal gamma
			1,554	3,121	Gamma reference

the higher is the evidence for rejecting the model. This is not surprising as the goodness-of-fit test relies only in the sample information, and it does not go in the detriment of the usefulness of these distributions for we may look to them as representatives of conservative models reflecting worse cases than the one supported by the sample information. Nevertheless we have to point out that the Pareto distribution coming from the Bayes estimation using the reciprocal gamma prior (Case 0 with Pareto index $\alpha = 7.714613$, see table 2, third column) gives a Cramér-von-Mises statistic equal to 0.04873 still lower than 0.461, the critical value at a 5% significance level.

We still need estimates of the posterior mean number of excesses to calculate reinsurance premiums. Having in mind the comments made at the end of section 3, this can be achieved modeling the prior knowledge of the mean number of excesses as vague information, i.e. with the help of a reciprocal gamma prior. This can be taken with a mean equal to 2 (see Hesselager (1993)) and infinite variance. Figure 7 draws the posterior Λ-densities corresponding to the portfolios and the collective. The resulting Bayes estimations are listed in the second column of table 5, where the case of a portfolio having not suffered from any excess has been also included. In its third column we report 95% probability Λ-intervals.

TABLE 6

Portfolio	α Prior distribution	E.SF. 50% level	XL Premium
Collective	Gamma	28,204	15,309
	Inverse Gaussian	25,765	14,076
	Reciprocal inverse Gaussian	25,771	14,188
A	Gamma	15,829	7,908
	Inverse Gaussian	15,666	7,824
	Reciprocal inverse Gaussian	15,659	7,830
E	Gamma	22,486	11,205
	Inverse Gaussian	22,182	11,133
	Reciprocal inverse Gaussian	22,748	11,136
H	Gamma	16,329	8,022
	Inverse Gaussian	15,836	7,955
	Reciprocal inverse Gaussian	15,920	7,961
No excess	Gamma	6,729	3,391
	Inverse Gaussian	6,140	3,118
	Reciprocal inverse Gaussian	6,387	3,143

Now reinsurance premiums calculations can be achieved. These consist in the substitution of the λ and α estimates into formula (2.15). Tables 5 and 6 summarize the cases of inexistent, vague and precise prior knowledge respectively. Portfolios lacking of excesses constitute a special case, because there is no information giving raise to the α-posterior. We have performed the premium calculation using each portfolio λ-estimate together with the collective α-estimate.

In table 5 (fifth column) and table 6 (third column) we have calculated estimations of the expected shortfall (E.SF.) to give an impression of the risk in each case (see for instance Acerbi, Tasche (2002)). This could be useful when adding a risk-dependent security margin to the net XL premium.

5. SUMMARY AND CONCLUSIONS

We have seen in the preceding sections how the generalized inverse Gaussian family works as a Pareto prior conjugate and we have deduced Bayesian estimates of the Pareto index for the collective and the individual portfolios. Knowing that the GIG also works as a Poisson prior conjugate, we have written down

an exact credibility formula (2.14) (or (2.15)) filled with (3.8) and (3.24), to calculate XL reinsurance premiums suited to the excesses history of one portfolio. For this scheme to be successfully applied, we must have at hand a large sample of excesses allowing the previous fixation of the threshold defining the class of large claims. Ours consisted in the claim severities (greater than a communication level 150,253€) arisen in ten Spanish portfolios of liability motor insurance from 1992 to 2001.

In our general Bayesian analysis the traditional use of a gamma prior appears as a particular case among others, some of which have been reviewed. These were the reciprocal gamma, the inverse Gaussian and the reciprocal inverse Gaussian distributions.

The family (3.9) of predictive distributions for the excesses over the threshold has been derived. This family contains as a particular case the log-Pareto distribution (2.25) as well as many other distributions like for instance (3.14), (3.16) and (3.21).

The reference prior (3.12) is useful when the only information available is the one contained in the sample. This is an important case indeed, for the Bayes, MLE and Hill statistics resume in the same expression, due in part to the fact that the threshold was previously fixed in a data point of the original ordered sample. Substituting the resulting Bayes estimation into the Pareto index, we have obtained a model for the collective excesses with a fairly good fit, confirming the previously assumed hypothesis of Pareto distributed excesses over the threshold.

The reciprocal gamma distribution allows the modeling of vague prior information. This happens when we are not able to assign a value to the second moment of the Pareto index though we still want to introduce into the Bayesian analysis some guess about its expectation. The Pareto distribution obtained through substitution of the Pareto index by this Bayes estimation also results in a good fit. This way, varying the prior A-mean would furnish successive Pareto distributions (for the collective) the lower the prior mean the more conservative the model, though still providing good fits.

The supposition of a deeper prior information resumes to the knowledge of the mean and the coefficient of variation (equivalently the second moment) of the prior distribution. Within this approach, a larger standard deviation would correspond to gradually less confidence on the value chosen for the mean. Complete prior information has been modeled by means of three distributions, namely the inverse Gaussian, the reciprocal inverse Gaussian and the gamma, all belonging to the GIG family. As a general rule we find that the A-estimations found in this case are lower than the ones corresponding to the lackness or vague prior knowledge. This is caused by the prior information we introduced, that was voluntarily chosen to prevent a more pessimistic real situation than the one squeezed out strictly from the sole sample (i.e. when using the reference prior). Consequently, we conclude that trying lower A-means with more narrow standard deviations (for the prior distribution) furnishes lower posterior Bayes A-estimates, thus resulting in a more conservative Pareto

distribution (once the substitution of the index by its estimation is done). Inside this gradation, the extreme (more conservative) case is represented by the gamma-Pareto-gamma scheme that gave the lowest A-estimations. From the point of view of the fits (collective case), we recall that the evidence for rejecting the models grows as the prior knowledge gets more precise, shadowing the objective information resumed in the sample.

Thus widening the Pareto prior family by using the GIG family has allowed us to produce a variety of Pareto distributions for the excesses representative of successively more conservative models. Correspondingly, we have been able to obtain reinsurance premiums which are ranged from the lowest ones (corresponding to the lack of prior information) to the highest ones (where more precise prior information is denouncing a lower Pareto index). This could be useful for it could make possible the calculation of successively higher premiums (either collective or individual) while knowing at the same time to what kind of overview (mean and coefficient of variation) they may correspond.

Future research could focus on the study of the predictive distribution for the excesses (3.9), a generalization of the well known log-Pareto distribution (3.11). Some of its simplest particular cases have been already reviewed in (3.14), (3.16) and (3.21). For instance, it could be interesting to introduce a shift parameter denoting the practitioner's prior belief on the number of finite moments of the unconditioned excesses distribution.

ACKNOWLEDGEMENTS

REFERENCES

ACERBI, C. and TASCHE, D. (2002) On the coherence of expected shortfall. *Journal of Banking and Finance*, **26(7)**, 1487-1503.

D'AGOSTINO, R.B. and STEPHENS, M.A. (1986) *Goodness-of-fit techniques*. Statistics: textbooks and monographs Vol. 68. Marcel Dekker Inc. New York and Bassel.

ARNOLD, B.C. and PRESS, S.J. (1989) Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association. Theory and Methods*, **84(408)**, 1079-1084.

BEIRLANT, J., TEUGELS, J.L. and VYNCKIER, P. (1996) *Practical analysis of extreme values*. Leuven University Press.

BERGER, J.O. (1980) *Statistical decision theory and Bayesian analysis*. Springer-Verlag. London.

BÜHLMANN, H. and GISLER, A. (2005) *A course in credibility theory and its applications*. Springer.

COLES, S. and POWELL, E.A. (1996) Bayesian methods in extreme value modeling: a review and new developments. *International Statistical Review*, **64(1)**, 119-136.

COLES, S. (2001) *An introduction to statistical modeling of extreme values*. Springer-Verlag. London.

EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997) *Modeling extremal events for insurance and finance*. Springer.

FÜRST, D. (1964) Formulation Bayesienne du problème des valeurs extrêmes en relation à la réassurance en excédent des sinistres. *Astin Bulletin*, **3(2)**, 153-162.

HESSELAGER, O. (1993) A class of conjugate priors with applications to excess-of-loss reinsurance. *Astin Bulletin*, **23(1)**, 77-93.

HILL, B.M. (1975) A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, **3(5)**, 1163-1174.

JORGENSEN (1982) *Statistical properties of the generalized inverse Gaussian distribution*. Springer-Verlag. New York.

LEMAIRE, J. (1995) *Bonus-malus systems in automobile insurance*. Kluwer Academic Publishers.

MCNEIL, A.J. (1997) Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin*, **27(1)**, 117-137.

PANJER, H.H. and WILLMOT, G.E. (1992) *Insurance risk models*. Society of Actuaries.

PATRICK, G. and MASHITZ, I. (1989) *Credibility for reinsurance excess pricing*. 21st Astin Colloquium. New York.

PIN-HUNG HSIEH (2004) A data record analytic method for forecasting next record catastrophe loss. *Journal of Risk and Insurance*, **71(2)**, 309-322.

REISS, R.D. and THOMAS, M. (1999) A new class of Bayesian estimators in paretian excess-of-loss reinsurance. *Astin Bulletin*, **29(2)**, 339-349.

REISS, R.D. and THOMAS, M. (2001) *Statistical analysis of extreme values*. 2nd edition. Birkhäuser.

RYTGAARD, M. (1990) Estimation in the Pareto distribution. *Astin Bulletin*, **20(2)**, 201-216.

SIEGEL, S. and CASTELLAN, N.J. (1988) *Nonparametric statistics for the behavioural sciences*. 2nd edition. New York. McGraw-Hill.

STRAUB, E. (1971) Estimation of the number of excess claims by means of the credibility theory. *Astin Bulletin*, **5(3)**, 388-392.

TREMBLAY, L. (1992) Using the Poisson inverse Gaussian in bonus-malus systems. *Astin Bulletin*, **22(1)**, 97-106.

JOSÉ LUIS VILAR-ZANÓN
Dept. Economía Financiera I.
Facultad Ciencias Económicas y Empresariales.
Universidad Complutense de Madrid. Campus de Somosaguas.
Pozuelo de Alarcón, 28223, España.
E-mail: jlvilarz@ccee.ucm.es

CRISTINA LOZANO-COLOMER
Dept. Métodos Cuantitativos
Universidad Pontificia de Comillas (ICADE). Madrid.
E-mail: clozano@cee.upcomillas.es

APPENDIX

We obtain the unconditioned distribution (3.9) substituting (3.1) into (2.8) then integrating. First we reparametrize the GIG as follows:

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

Assuming $\chi\psi > 0$, (2.8), becomes:

$$h(z) = \frac{\eta^{-\beta}}{2K_\beta(\omega)(1+z)} \int_0^{+\infty} \alpha^\beta e^{-\frac{\omega}{2}\left(\frac{\eta}{\alpha} + \alpha\left[\frac{1}{\eta} + \frac{2}{\omega}\log(1+z)\right]\right)} d\alpha$$

Considering:

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

we obtain

$$h(z) = \frac{\eta^{-\beta}}{2K_\beta(\omega)(1+z)} \int_0^{+\infty} \alpha^\beta e^{-\frac{A}{2}\left(\frac{B}{\alpha} + \frac{\alpha}{B}\right)} d\alpha$$

Recalling that (3.2) is the Bessel function of third kind we finally get

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

Expressions (3.14) and (3.21) are straightforward considering (3.9) and the following two properties of the Bessel function of the third kind (see for instance Panjer, Willmot (1992) pages 413, 414):

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

$$K_\beta(\omega) = K_{-\beta}(\omega), \quad \beta \in \mathbb{R}, \quad \omega > 0$$

while (3.11) (respectively (3.16)) can be easily obtained filling (2.8) with (3.10) (respectively (3.15)) then integrating it.