Bayesian Extreme Value Analysis with an Application to Credibility Estimation

by R.–D. Reiss and M. Thomas

FB 6, University of Siegen

Abstract

Exact credibility (Bayesian) and linear credibility estimation of the net premium plays a central role in the actuarial business. For evaluating the risk of excess–of–loss reinsurance treaties, the linear credibility estimation was applied in a restricted Pareto model for which the Hill estimator is the MLE (e.g., Hesselager (1993) and Schnieper (1993)). One is confronted with certain deficiencies of this model (Reiss and Thomas (1997)). Bayesian parameter estimation in a suitably parametrized full Pareto model was explored in a recent paper by Reiss and Thomas (1999). The present paper deals with the exact credibility estimation of the net premium within a model of Poisson–Pareto processes.

1 Introduction

In this paper, excess–of–loss (XL) reinsurance treaties are dealt with, where the reinsurer pays the excess of a certain fixed limit for individual claim sizes. The aim is to estimate the net premium which is the expectation of the total claim amount

$$S_N = \sum_{i \le N} Z_i$$

for the next period, where N is the random number of claim excesses occurring in the period and the Z_i are the random claim excesses over the priority u. Notice that $Z_i = X_i - u$, where the X_i are the exceedances over the priority u.

Throughout, we assume that the excess times and the excesses can be jointly modeled by a Poisson process, where the excess times are described by a homogeneous Poisson process with intensity $\lambda > 0$ and the excesses have a common df F. Such a process will be called a Poisson(λ, F) process. The net premium is

$$E(S_N) = \lambda m(F),$$

where $m(F)$ denotes the mean of F. The estimator of the net premium is based on the excess times and excesses of the past T years. The Poisson process restricted to T years is denoted by Poisson(λ, F, T).

Our attention is focused on the Bayes estimation—that is, the exact credibility estimation—of the net premium. A justification for a restriction to the exact credibility estimation was formulated by S.A. Klugman (1992), page 64:

"Most of the difficulties ... are due to actuaries having spent the past half–century seeking linear solutions to the estimation problems. At a time when the cost of computation was high this was a valuable endeavor. The logical solution is to drop the linear approximation and seek the true Bayesian solution to the problem. It will have the smallest *mse* and provides a variety of other benificial properties".

One may argue that less must be known—namely, the first and second moments—about the distributions of hyperparameters in conjunction with linear credibility estimators. Yet, we merely deal with prior distributions, which are determined by such moments.

We assume that the claim excesses have a Paretian distribution which has necessarily a left endpoint equal to zero (see Section 2). Some remarks about the Bayesian estimation method within models of Poisson processes are made in Section 3. We particularly specify the likelihood function for such models. Recall that the mean of a Pareto distribution is equal to infinity if the shape parameter α is smaller or equal to 1. This yields that the support of the prior distribution must be restricted to parameters $\alpha > 1$ when Bayesian estimation of the mean is dealt with. This is the reason why truncated gamma distributions are introduced in Section 4. The subsequent sections 5–9 concern Bayesian estimation of parameters and of the net premium (credibility estimation).

2 Paretian Models

The full Paretian model of normalized excesses $Y_i = Z_i/u$ consists of Paretian dfs of the form

$$F_{\alpha,\tau}(y) = 1 - (1 + \tau y)^{-\alpha}, \qquad y \geq 0, \tag{1}$$

where $\alpha > 0$ and $\tau > 0$ are the shape and reciprocal scale parameters. The parametrization is taken in such a way that for a fixed reciprocal scale parameter $\tau = 1$ one obtains a submodel—subsequently referred to as restricted Paretian model—for which the Hill estimator is the MLE.

Within the full model (1) of Pareto dfs one gets some closer insight why the Hill estimator and related estimators are inaccurate even in such cases where a Paretian modeling is acceptable. Notice that a smaller reciprocal scale parameter $\tau < 1$—thus, a larger scale parameter—puts more weight on the upper tail of the df in the full model of Pareto dfs $F_{\alpha,\tau}$. If such a Pareto df is estimated, yet $\tau = 1$ is kept fixed in the modeling, than the heavier weight in the tail, due to a smaller parameter τ, must be compensated by a smaller estimated shape parameter α. For further details we refer to Reiss and Thomas (1997), pages 120, 149 and 213, and to Reiss and Thomas (1999).

The restricted Paretian modeling was used by several authors also in conjunction with Bayesian and credibility estimation (e.g., Rytgaard (1990), Hesselager (1993) and Schnieper (1993)) with the consequence that just the same deficiencies occur for Bayesian estimators as for the Hill estimator. Bayesian estimators of the shape parameter α within the full Paretian model, as defined by (1), were introduced in Reiss and Thomas (1999). General remarks about the Bayesian estimation in the generalized Pareto model may be found in a paper by Pickands (1994).

3 Bayesian Estimation Within Models of Poisson–Pareto Processes

The Bayesian estimation is a likelihood–based method just as the maximum likelihood (ML) method. We give a short introduction within the rather general framework of estimating a real–valued functional parameter $T(\vartheta)$, where $\vartheta = (\vartheta_1, \ldots, \vartheta_m)$ is a parameter vector. This includes the estimation

- of the jth component ϑ_j of ϑ if $T(\vartheta) = \vartheta_j$,
- of the mean of an underlying df parametrized by ϑ.

The latter problem will be studied in conjunction with the credibility estimation.

Let X be a random variable having a distribution parametrized by ϑ. The Bayes risk of the estimator $\widehat{T}$ is the integrated MSE

$$R(p, \widehat{T}) = \int E\Big(\big(\widehat{T}(X) - T(\vartheta)\big)^2 |\vartheta\Big) p(\vartheta)\, d\nu(\vartheta),$$

where $p(\vartheta)$ is the "prior" density, with respect to some dominating measure ν. Some prior knowledge about the parameter of interest is included in the statistical modeling by means of the prior density. An estimator $\widehat{T}$ which minimizes the Bayes risk is called Bayes estimator.

The Bayesian estimate of $T(\vartheta)$ can be written as

$$T^*(x) = \int T(\vartheta) p(\vartheta|x)\, d\nu(\vartheta), \tag{2}$$

where the function

$$p(\vartheta|x) = \frac{L(\vartheta|x)p(\vartheta)}{\int L(\vartheta|x)p(\vartheta)\, d\nu(\vartheta)} \tag{3}$$

is the "posterior" density for a given outcome x and $L(\vartheta|x)$ denotes the likelihood function as a function in ϑ.

Notice that the posterior density in (3) and the Bayes estimate can be computed if the likelihood function and, thus, $L(\vartheta|x)p(\vartheta)$ as a function in ϑ is known up to a constant. Generally, we write $g(\vartheta) \propto f(\vartheta)$ if a function g is proportional to f.

Given a Poisson$(\lambda, F_\vartheta, T)$ process, where ϑ is some parameter vector, we denote the outcome by a point measure

$$\mu = \sum_{i=1}^{k} \varepsilon_{(t_i, y_i)},$$

where the t_i are the excess times and y_i the marks. Assume that the df F_ϑ has the density f_ϑ. One may prove (cf. Reiss (1993), Theorem 3.1.1) that for every μ the likelihood function $L(\lambda, \vartheta|\mu)$ satisfies the relation

$$L(\lambda, \vartheta|\mu) \,\propto\, L(\lambda|k)L(\vartheta|\boldsymbol{y}), \tag{4}$$

where "$\propto$" again denotes equality of both sides up to a constant, and

- $L(\lambda|k) = \big((\lambda T)^k/k!\big)\exp(-\lambda T)$ is the likelihood function for the Poisson random number of excesses up to time T,
- $L(\vartheta|\boldsymbol{y}) = \prod_{i\le k} f_\vartheta(y_i)$ is the likelihood function for the model of k iid random variables with common density f_ϑ.

The Bayes estimate of the intensity parameter λ of the homogeneous Poisson process of excess times can be computed independently of ϑ if the prior density satisfies

$$\widetilde{p}(\lambda, \vartheta) = p_0(\lambda)p(\vartheta). \tag{5}$$

Under this condition, the posterior density is

$$\widetilde{p}(\lambda, \vartheta|\mu) = p_0(\lambda|k)p(\vartheta|\boldsymbol{y}), \tag{6}$$

where $p_0(\lambda|k) \,\propto\, L(\lambda|k)p_0(\lambda)$ and $p(\vartheta|\boldsymbol{y}) \,\propto\, L(\vartheta|\boldsymbol{y})p(\vartheta)$. Formula (6) indicates that the problem of computing Bayesian estimates splits up in computing estimates

- of the intensity parameter λ,
- of the parameter ϑ.

We have $\vartheta = \alpha$ or $\vartheta = (\alpha, \tau)$, where $\alpha, \tau > 0$ are shape and reciprocal scale parameters in Paretian models.

General references for Bayesian estimation are, e.g., Hartigan (1983) and Klugman (1992).

4 Left Truncated Gamma Distributions

Our prior distribution for the shape parameter α in the Paretian model will be specified by means of a truncated gamma distribution. Recall that the gamma density with shape parameter $s > 0$ and reciprocal scale parameter $d > 0$ is given by

$$h_{s,d}(\vartheta) = \frac{d}{\Gamma(s)}(d\vartheta)^{s-1}\exp(-d\vartheta), \qquad \vartheta > 0. \tag{7}$$

The pertaining df is denoted by $H_{s,d}$. This gamma df has the mean s/d and the variance s/d^2.

Gamma distributions, truncated left of $q > 1$, with shape and reciprocal scale parameters $s > 0$ and $d > 0$ (cf. (7)) will serve as priors. The pertaining df and density are denoted by $H_{s,d,q}$ and $h_{s,d,q}$. The gamma density $h_{s,d,q}$, truncated left of $q > 1$, is given by

$$h_{s,d,q}(\alpha) = \frac{h_{s,d}(\alpha)}{1 - H_{s,d}(q)}, \qquad \alpha > q. \tag{8}$$

It is well known that gamma distributions are conjugate priors for Paretian models. That is, the pertaining posterior densities are of the same type. This property still holds for the truncated gamma versions.

Straightforward calculations yield that such a truncated gamma distribution has the mean

$$M_{s,d,q} = \frac{s\Big(1 - H_{s+1,d}(q)\Big)}{d\Big(1 - H_{s,d}(q)\Big)} \tag{9}$$

and the variance

$$V_{s,d,q} = \frac{s\Big((s+1)(1 - H_{s+2,d}(q))(1 - H_{s,d}(q)) - s(1 - H_{s+1,d}(q))^2\Big)}{d^2(1 - H_{s,d}(q))^2}. \tag{10}$$

These formulas are useful

- to specify a truncated gamma prior with predetermined moments,
- to compute the numerical value of a Bayesian estimate.

5 Bayesian Estimation of the Shape Parameter Within the Restricted Paretian Model

First, we deal with the estimation of the shape parameter α within the restricted Paretian model with a truncated gamma density $h_{s,d,q}$ taken as a prior. The case of untruncated gamma prior is included for $q = 0$. According to (6), the computations are just the same if the Bayesian estimation is carried out within the framework of Poisson–Paretian processes.

Recall that the common df of normalized excesses within the restricted Paretian model is given by

$$F_{\alpha,1}(y) = 1 - (1 + y)^{-\alpha}, \qquad y \geq 0, \tag{11}$$

where $\alpha > 0$. These dfs are used for the modeling of normalized excesses over a priority u. If

$$p(\alpha) = h_{s,d,q}(\alpha) \tag{12}$$

is the prior, then

$$p(\alpha|\boldsymbol{y}) = h_{s',d',q} \tag{13}$$

is the posterior with $s' = s + k$ and $d' = d + \sum_{i \leq k} \log(1 + y_i)$. We see that the truncated gamma densities are conjugate priors with respect to the restricted Paretian model.

From (9) deduce that the Bayesian estimate of $\alpha > q \geq 0$ is

$$\alpha^*_{k,q}(\boldsymbol{y}) = \alpha^*_k(\boldsymbol{y})\frac{1 - H_{s'+1,d'}(q)}{1 - H_{s',d'}(q)}, \tag{14}$$

where

$$\alpha^*_k(\boldsymbol{y}) = \frac{s'}{d'} = \frac{s+k}{d + \sum_{i\leq k}\log(1+y_i)}$$

is the Bayesian estimate in the case of the gamma prior with $q = 0$. The Bayesian estimate in the case of $q = 0$ is close to the Hill estimate if s and d are small or/and k is large.

For $q = 0$, Bayesian estimates of α were computed by Rytgaard (1990), Hesselager (1993) and Schnieper (1993).

6 Bayesian Estimation of Parameters Within the Full Paretian Model

In contrast to the preceding lines, we assume that the normalized excesses are distributed according to a Pareto df

$$F_{\alpha,\tau}(y) = 1 - (1+\tau y)^{-\alpha}, \qquad y \geq u, \tag{15}$$

where $\alpha, \tau > 0$ (cf. (1)). If

$$p(\alpha, \tau) = h_{s,d,q}(\alpha) f(\tau) \tag{16}$$

is the prior, then

$$p(\alpha, \tau|\boldsymbol{y}) = h_{s',d'(\tau),q}(\alpha)\tilde{f}(\tau) \tag{17}$$

is the posterior, where $s' = s + k$, $d'(\tau) = d + \sum_{i\leq k}\log(1+\tau y_i)$, and

$$\tilde{f}(\tau) \propto \tau^k (d'(\tau))^{-s'} \exp\Big(-\sum_{i\leq k}\log(1+\tau y_i)\Big) f(\tau). \tag{18}$$

As the Bayesian estimate of $\alpha > q \geq 0$ and $\tau > 0$ one gets

$$\hat{\alpha}_{k,q}(\boldsymbol{y}) = \int \frac{s'}{d'(\tau)} \frac{(1 - H_{s'+1,d'(\tau)}(q))}{(1 - H_{s',d'(\tau)}(q))} \tilde{f}(\tau)\, d\tau, \qquad q \geq 0, \tag{19}$$

and, independently of q,

$$\hat{\tau}_{k,q}(\boldsymbol{y}) = \hat{\tau}_k(\boldsymbol{y}) = \int \tau \tilde{f}(\tau)\, d\tau. \tag{20}$$

The integrals in (19) and (20) must be numerically evaluated.

Specifically, one may take a gamma density $f = h_{a,b}$ as a prior for the parameter τ. In view of the fact that the restricted Paretian model is a submodel for $\tau = 1$, it can be plausible to choose the parameters a and b such that $h_{a,b}$ is centered around $\tau = 1$. One possibility to achieve this goal is to take $a = b$. Such a gamma distribution has the mean 1 and variance $1/b$. It is approximately equal to a normal density with mean 1 and variance $1/b$ as b approaches infinity.

The special case of $q = 0$ was dealt with in Reiss and Thomas (1999).

7 Exact Credibility Estimation Within the Restricted Poisson–Paretian Model

As outlined in the introduction, we want to estimate the net premium for the next period based on the recorded claims of the past T periods. We distinguish two different cases, namely, that the common excess df F belongs (a) to the restricted Paretian model and (b) to the full Paretian model. We start with the restricted Paretian model.

The net premium for the excesses over the priority is given by

$$m(\lambda, \alpha) = \frac{\lambda u}{\alpha - 1}, \qquad \alpha > 1. \tag{21}$$

With the likelihood function as in (4) and a prior density $\widetilde{p}(\lambda, \alpha) = p_0(\lambda)p(\alpha)$ one gets the posterior density $\widetilde{p}(\lambda, \alpha|\mu) = p_0(\lambda|k)p(\alpha|\boldsymbol{y})$ according to (6). Therefore, a Bayesian estimate of $m(\lambda, \alpha)$ is given by

$$\begin{aligned} m_k^*(\mu) &= \int \frac{\lambda u}{\alpha - 1} p_0(\lambda|k)p(\alpha|\boldsymbol{y})\, d\lambda d\alpha \\ &= \left(\int \lambda p_0(\lambda|k)\, d\lambda\right)\left(\int \frac{u}{\alpha-1} p(\alpha|\boldsymbol{y})\, d\alpha\right). \end{aligned} \tag{22}$$

Taking the special priors $p_0(\lambda) = h_{r,c}(\lambda)$ and $p(\alpha) = h_{s,d,q}(\alpha)$ with $q > 1$, one obtains the Bayesian estimate

$$m_{k,q}^*(\mu) = \lambda_k^* m_{k,q}^{**}(\boldsymbol{y}) \tag{23}$$

of the net premium $m(\lambda, \alpha)$ with $\lambda_k^* = (r + k)/(c + T)$ and

$$m_{k,q}^{**}(\boldsymbol{y}) = \int_q^\infty \frac{u}{\alpha - 1} h_{s',d',q}(\alpha)\, d\alpha. \tag{24}$$

This integral must be evaluated numerically.

8 Exact Credibility Estimation Within the Full Poisson–Paretian Model

In this section, we deal with the estimation of the parameters α and τ if a truncated gamma density $h_{s,d,q}$, with $q > 1$, is taken as a prior.

Now, the net premium is given by

$$m(\lambda, \alpha, \tau) = \frac{\lambda u}{\tau(\alpha - 1)}, \qquad \alpha > 1. \tag{25}$$

With the likelihood function as in (4) and a prior $\widetilde{p}(\lambda, \alpha, \tau) = p_0(\lambda)p(\alpha, \tau)$ one gets the posterior density $\widetilde{p}(\lambda, \alpha, \tau|\mu) = p_0(\lambda|k)p(\alpha, \tau|\boldsymbol{y})$ according to (6).

Therefore, a Bayesian estimate of $m(\lambda, \alpha, \tau)$ can be written as

$$\begin{aligned} \hat{m}_k(\mu) &= \int \frac{\lambda u}{\tau(\alpha-1)} p_0(\lambda|k) p(\alpha, \tau|\boldsymbol{y})\, d\lambda d\tau \\ &= \left(\int \lambda p_1(\lambda|k)\, d\lambda \right) \left(\int \frac{u}{\tau(\alpha-1)} p_2(\alpha, \tau|\boldsymbol{y})\, d\tau \right) . \end{aligned} \tag{26}$$

Taking the special priors $p_0(\lambda) = h_{r,c}(\lambda)$ and $p(\alpha, \tau) = h_{s,d,q}(\alpha) f(\tau)$ as in (16) and (17), one obtains the Bayesian estimate

$$\hat{m}_{k,q}(\mu) = \lambda_k^* \widetilde{m}_{k,q}(\boldsymbol{y}) \tag{27}$$

of the net premium $m(\lambda, \alpha, \tau)$ with $\lambda_k^* = (r+k)/(c+T)$ and

$$\widetilde{m}_{k,q}(\boldsymbol{y}) = \int_0^\infty \frac{1}{\tau} \Big(\int_q^\infty \frac{u}{\alpha - 1} h_{s', d'(\tau), q}(\alpha)\, d\alpha \Big) \tilde{f}(\tau)\, d\tau. \tag{28}$$

Again, one may take a gamma density $f = h_{a,b}$ as a prior for the parameter τ (cf. Section 6).

9 An Application to Fire Reinsurance Data

We analyze the fire reinsurance data given in the paper by Schnieper (1993) and further discussed in Reiss and Thomas (1997), (1999). The original claim sizes are detrended, indexed and made anonymous.

Table 1. Claims sizes over 22.0 million NKr from 1983 to 1992.

year	claim size	year	claim size
1983	42.719		23.208
1984	105.860	1990	37.772
1986	29.172		34.126
	22.654		27.990
1987	61.992	1992	53.472
	35.000		36.269
1988	26.891		31.088
1989	25.590		25.907
	24.130		

Using the Hill estimator the estimate $\alpha = 2.219$ is obtained by Schnieper. The visual insight, gained from mean excess functions, speaks in favor of a lighter Paretian tail. A related remark holds with respect to the collective claim degree distribution with a gamma prior as also presented in Schnieper (1993).

Besides the Hill estimate, Table 2 also contains the MLE and a moment estimate of α which are estimation procedures designed for the generalized Pareto model. Estimates of the reciprocal scale parameter τ are added (taken from Reiss and Thomas (1997)). The pertaining estimated net premiums and a nonparametric estimate of the net premium are included. A fixed value $\lambda = 1.7$ is used for the intensity.

Table 2. Estimated net premium for fire reinsurance data.

	α	τ	$m(\lambda, \alpha, \tau)$
nonparametric			26.98
Hill	2.219	1	30.70
MLE	3.939	0.467	27.23
moment	3.407	0.535	29.00

The MLE and the moment estimate correspond to the insight gained from nonparametric plots (e.g., sample mean and median excessfunctions). The conjectured heavy–tailed distribution is confirmed, yet with a significantly larger shape parameter than indicated by the Hill estimate. This example also illustrates the remark made in Section 2 about the interdependence of the shape and reciprocal scale parameter. Because the actual reciprocal scale parameter τ is considerably smaller than 1, one obtains the inaccurate estimate of the shape parameter by using the Hill procedure.

On the other hand, the pertaining estimates of the net premium are closely related to each other which is due to the fact that the net premium (cf. (25)) only depends on α and τ by means of the product $\tau(\alpha - 1)$.

In the following lines, we estimate the net premium using the Bayesian estimators within the restricted and the full Paretian model.

- In the restricted model, Schnieper applies the Bayesian principle with the hyperparameters $s = 30$ and $d = 16$. Putting the truncation point q equal to 1.1 and using a truncated gamma prior with the same mean and variance as $h_{30,16}$, one obtains the estimates $\alpha^*_{k,q} = 1.99$ and $m^*_{k,q} = 42.13$. The same shape parameter $\alpha^*_k = 1.99$ is estimated for an untruncated gamma prior ($q = 0$).

- In the full Paretian model, we use a truncated gamma prior with mean and variance equal to 4. The prior $h_{a,b}$ for τ is chosen with $a = 1.5$ and $b = 1.2$. This is a gamma density with a mode centered at $\tau = 1$. Table 3 lists the parameters of the truncated gamma prior $h_{s,d,q}$ and the estimated values $\hat{\alpha}_{k,q}$, $\hat{\tau}_{k,q}$ and $\widehat{m}_{k,q}$ (taken with $\lambda^*_k = 1.7$) for different values of q.

Although we started with a prior for the reciprocal scale parameter τ centered at 1, we obtain Bayesian estimates with significantly smaller values. Thus, the Bayesian values react to the data in an expected manner.

Table 3. Estimated net premium for fire reinsurance data using Bayesian estimator within the full Paretian model.

q	s	d	$\hat{\alpha}_{k,q}$	$\hat{\tau}_{k,q}$	$\widehat{m}_{k,q}$
1.05	3.634	0.932	3.384	0.771	32.539
1.2	3.431	0.895	3.346	0.786	31.705
1.5	2.817	0.789	3.237	0.837	30.451

References

Hartigan, J.A. (1983). *Bayes Theory.* Springer, New York.

Hesselager, O. (1993). A class of conjugate priors with applications to excess–of–loss reinsurance. ASTIN Bulletin 23, 77–90.

Klugman, S.A. (1992). *Bayesian Statistics in Actuarial Science.* Kluwer, Boston.

Pickands, J. III (1994). Bayes quantile estimation and threshold selection for the generalized Pareto family. In: Extreme Value Theory and Applications, 123–138, J. Galambos et al. (eds.), Kluwer, Dordrecht.

Reiss, R.–D. (1993). *A Course on Point Processes.* Springer, New York.

Reiss, R.–D. and Thomas, M. (1997). *Statistical Analysis of Extreme Values.* Birkhäuser, Basel (with Xtremes on CD–ROM).

Reiss, R.-D. and Thomas, M. (1999). A new class of Bayesian estimators in Paretian excess–of–loss reinsurance. ASTIN Bulletin 29, 339–349.

Rytgaard, M. (1990). Estimation in the Pareto distribution. ASTIN Bulletin 20, 201–216.

Schnieper, R. (1993). Praktische Erfahrungen mit Grossschadenverteilungen. Mitteil. Schweiz. Verein Versicherungsmath., 149–165.

FB 6, Mathematik
Universität Siegen
D–57068 Siegen
Germany