

A CLASS OF CONJUGATE PRIORS WITH APPLICATIONS TO EXCESS-OF-LOSS REINSURANCE

BY OLE HESSELAGER

University of Copenhagen, Denmark

ABSTRACT

We consider the problem of forecasting the total cost of claims in excess-of-loss reinsurance. The number of claims reported to the direct insurer is assumed to follow a Poisson law, and the claim severities are modelled by a Pareto distribution. The Poisson frequency as well as the Pareto parameter will be considered as random parameters in a Bayesian setting. We derive the class of conjugate joint prior distributions, which turn out to specify a (prior) dependence between the two parameters. The use of conjugate priors facilitates the mathematical analysis, and it also makes it easy to interpret the parameters of the prior distribution

KEYWORDS

Excess-of-loss quotations; limited cover, Pareto distribution; conjugate priors.

1 INTRODUCTION

Consider an excess-of-loss reinsurance cover for the layer $(b - a)$ in excess of a . That is, of a claim y exceeding the priority a , the reinsurer pays the exceeding amount $(y - a)$, limited to the maximum $(b - a)$. The reinsurer receives information about all claims exceeding some level c , say, during a fixed exposure period $[0, T]$, and the problem is to predict next years total cost of claims.

Often the reinsurer has only very sparing information about a particular contract, and may therefore want to consider also information from other similar contracts. This can be done in a formalized manner using the Bayesian paradigm. Of recent research in this direction we mention JEWELL (1990) who inspired by PATRIK and MASHITZ (1989) analyses a Bayesian model with independent prior information about the claims frequency and the severities, and suggests a linearized Bayesian forecast for the total excess-of-loss claims cost. The analysis is continued in JEWELL (1991) by showing in graphical detail the effect of using different types of data information in the prediction. RYTGAARD (1990) considers a compound Poisson model with Pareto distributed claim amounts, but with prior information only on the severity parameter.

In this paper we consider a compound Poisson model with Pareto distributed claim amounts. We assume that prior information is available about the claims

frequency as well as the severities, and present a Bayesian full-distributional analysis using conjugate priors. The use of conjugate priors makes the updating an easy matter, and also makes it easy to interpret the prior information in relation to the likelihood information. We derive the (non-linearized) Bayesian forecast for the total excess-of-loss claims cost, and also the corresponding predicted moments of 2. and 3. order

Random variables and their outcomes are denoted by upper and lower case letters, respectively. We use ordinary italics for observed or observable random quantities, and greek letters for unknown random parameters in a Bayesian setting. The shorthand $p(x)$ is used to denote the probability density function for a random variable X . Likewise, $p(x|z)$ denotes the conditional density for X given $Z = z$.

2. THE MODEL

Let $N(t)$ denote the number of claims reported to the direct insurer during $[0, t]$ and let Y_1, Y_2, \dots denote the corresponding claim amounts. Consider the following model

- Assume that $\{N(t)\}_{t \geq 0}$ is a time-homogeneous Poisson process with intensity λ , and that Y_1, Y_2, \dots are stochastically independent of $\{N(t)\}_{t \geq 0}$ and mutually independent with common distribution $F_{s, \psi}$, where

$$(2.1) \quad F_{s, \psi}(y) = 1 - (s/y)^\psi, \quad y \geq s,$$

denotes the cumulative distribution function for the Pareto distribution with parameters (s, ψ) .

For $u \geq s$, let

$$(2.2) \quad N_u(t) = \sum_{i=1}^{N(t)} I(Y_i > u)$$

denote the number of claims exceeding the level u during $[0, t]$. It appears from (2.2) that $\{N_u(t)\}_{t \geq 0}$ is a random thinning of the original claim number process, and is therefore itself a (time-homogeneous) Poisson process with intensity

$$(2.3) \quad \lambda_u = \lambda P(Y_i > u) = \lambda (s/u)^\psi, \quad u \geq s.$$

Note that λ_u depends on s as well as ψ , which has been suppressed in the notation. Note also the relation

$$(2.4) \quad \lambda_v = \lambda_u (u/v)^\psi, \quad u, v \geq s,$$

between the intensities corresponding to claim numbers exceeding different levels u and v . The claim amounts exceeding u are labeled in consecutive order, and are denoted by $Y_{u,i}$, $i = 1, \dots, N_u(t)$. It is a well-known property of the Pareto distribution that these claim amounts are again Pareto distributed, with parameters (u, ψ) .

Consider an excess-of-loss cover for the layer $b - a$ in excess of a . For a claim $Y_{a,t}$ exceeding the priority a , the reinsurer will cover the amount

$$(2.5) \quad Z_t = \min(Y_{a,t}, b) - a.$$

Since $Y_{a,t}$ is Pareto distributed with parameters (a, ψ) , it follows that Z_t has cumulative distribution function

$$(2.6) \quad H_\psi(z) = \begin{cases} F_{a,\psi}(a+z), & 0 \leq z < b-a \\ 1, & z \geq b-a \end{cases},$$

where $F_{a,\psi}$ is given by (2.1). With

$$X(t) = \sum_{i=1}^{N_a(t)} Z_i$$

denoting the reinsurer's total claims cost during $[0, t]$, it then follows that $\{X(t)\}_{t \geq 0}$ is a compound Poisson process with intensity λ_a and severity distribution H_ψ .

The parameters (λ, ψ) are unknown, and will be regarded as outcomes of random parameters (A, Ψ) in a Bayesian setting. Strictly speaking, this means that the model assumptions stated above hold true conditionally given $(A, \Psi) = (\lambda, \psi)$. In accordance with (2.3) we also define the random parameter

$$(2.7) \quad A_u = A(s/u)^\psi, \quad u \geq s.$$

Whenever convenient, we may condition on (A, Ψ) or (A_u, Ψ) for arbitrary u , since there is a 1 ↔ 1 correspondence between the two sets of variables. For notational convenience, we also let $\Theta_u = (A_u, \Psi)$

At time T we wish to predict the reinsurer's total claims cost for the next year,

$$X = X(T+1) - X(T),$$

which is conditionally compound Poisson distributed, given $\Theta_a = \vartheta_a$, with Poisson parameter λ_a and severity distribution H_ψ . Let \mathcal{D} denote the observed data at time T . Ideally we would want the conditional distribution of X given \mathcal{D} . Since the Poisson process has independent increments, it follows that X is conditionally independent of \mathcal{D} , and hence that

$$(2.8) \quad p(x|\mathcal{D}) = \int p(x|\vartheta_a) p(\vartheta_a|\mathcal{D}) d\vartheta_a.$$

The compound Poisson distribution $p(x|\vartheta_a)$ may be calculated recursively using Panjer's (1981) algorithm, if we make a discrete approximation to the severity distribution H_ψ . The conditional distribution $p(x|\mathcal{D})$ may therefore (in principle) be approximated numerically from (2.8), using a discrete approximation to $p(\vartheta_a|\mathcal{D})$. This is very time-consuming, however, and more realistically

one may settle for some moments of the conditional distribution (2.8),

$$(2.9) \quad E(X^k|\mathcal{D}) = \int E(X^k|\vartheta_a) p(\vartheta_a|\mathcal{D}) d\vartheta_a.$$

The conditional moments $E(X^k|\vartheta_a)$ appearing in (2.9) may be calculated recursively (see GOOVAERTS et al. (1984), p. 12) as

$$(2.10) \quad E(X^k|\vartheta_a) = \lambda_a \sum_{j=1}^k \binom{k-1}{j-1} \mu_j(\psi) E(X^{k-j}|\vartheta_a), \quad k = 1, 2, \dots,$$

where

$$(2.11) \quad \mu_k(\psi) = E[Z_i^k|\vartheta_a] = E[Z_i^k|\psi], \quad k = 1, 2, \dots,$$

denotes the k 'th moment of the reinsurer's payment in respect of a single claim when $\Psi = \psi$ is fixed. It is demonstrated in the Appendix (formula (A.2)) that $\mu_k(\psi)$ can be written as

$$(2.12) \quad \mu_k(\psi) = ka^k \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j \frac{1 - (a/b)^{\psi-k+j}}{\psi-k+j}, \quad k = 1, 2, \dots$$

With a discrete approximation to the posterior distribution $p(\vartheta_a|\mathcal{D})$, we may then calculate $E(X^k|\mathcal{D})$ numerically from (2.9) using (2.10) and (2.12).

While the conditional moments around the origin are given by the recursive expression (2.10), we may express the central moments up to 3th order in a very convenient form. In fact, with

$$(2.13) \quad e_1(\vartheta_a) = E(X|\vartheta_a) = \lambda_a \mu_1(\psi)$$

denoting the expected total claims cost in the case of known parameters, it is well-known that

$$(2.14) \quad e_k(\vartheta_a) = E[(X - e_1(\vartheta_a))^k|\vartheta_a] = \lambda_a \mu_k(\psi),$$

for $k = 2$ and $k = 3$. Using a conjugate (joint) prior distribution for $\Theta_a = (A_a, \psi)$, which is derived in Section 3, it is an easy matter to calculate the expectation of (2.13) and (2.14), and therefore also the predicted moments $E(e_k(\Theta_a)|\mathcal{D})$.

While

$$(2.15) \quad E[e_1(\Theta_a)|\mathcal{D}] = E[X|\mathcal{D}],$$

according to (2.9) and (2.13), the reader should be aware that

$$E[e_k(\Theta_a)|\mathcal{D}] \neq E[(X - E(X|\mathcal{D}))^k|\mathcal{D}]$$

for $k = 2$ and $k = 3$. In fact, for $k = 2$ it holds that

$$(2.16) \quad \begin{aligned} \text{Var}(X|\mathcal{D}) &= \text{Var}[E(X|\Theta_a)|\mathcal{D}] + E[\text{Var}(X|\Theta_a)|\mathcal{D}] \\ &= \text{Var}[e_1(\Theta_a)|\mathcal{D}] + E[e_2(\Theta_a)|\mathcal{D}], \end{aligned}$$

where we have made use of the fact that X is conditionally independent of \mathcal{D} . Formula (2.16) shows that $\text{Var}(X|\mathcal{D})$ is composed of the average variation in case of known parameters, $E[e_2(\Theta_a)|\mathcal{D}]$, and the term $\text{Var}[e_1(\Theta_a)|\mathcal{D}]$ due to (remaining) uncertainty about Θ_a . As the amount of data increases, the posterior distribution $p(\vartheta_a|\mathcal{D})$ will concentrate its mass around the true parameter Θ_a . In this case, the term $\text{Var}[e_1(\Theta_a)|\mathcal{D}]$ will vanish, and $\text{Var}(X|\mathcal{D}) \approx E[e_2(\Theta_a)|\mathcal{D}]$.

3. CONJUGATE PRIORS

The reinsurer is assumed to receive information about all claims exceeding a fixed limit $c > s$; the data capture level. Thus, the observed data at time T are

$$\mathcal{D} = \{N_c(T) = n, \quad Y_{c,i} = y_i, \quad i = 1, \dots, n\}.$$

Strictly speaking, the reinsurer would also know the occurrence times of the claims. However, under the model assumptions $N_c(T)$ is a sufficient statistic. For fixed $(A, \Psi) = (\lambda, \psi)$, the number of claims $N_c(T)$ is Poisson distributed with mean $T\lambda_c$, where λ_c is given by (2.3), and the claim amounts $Y_{c,i}$ are identically Pareto distributed with parameters (c, ψ) . The likelihood function is therefore given by,

$$\begin{aligned} (3.1) \quad l(\lambda_c, \psi|\mathcal{D}) &= \frac{(T\lambda_c)^n}{n!} e^{-T\lambda_c} \prod_{i=1}^n (\psi/c) (y_i/c)^{-(\psi-1)} \\ &\propto (\lambda_c^n e^{-T\lambda_c}) (\psi^n e^{-\psi z}), \end{aligned}$$

where \propto means “is proportional to as a function of (λ_c, ψ) ”, and

$$(3.2) \quad z = \sum_{i=1}^n \ln(y_i/c).$$

Recall that the gamma distribution with shape parameter γ and inverse scale parameter δ has probability density function

$$(3.3) \quad g_{\gamma,\delta}(x) = \frac{\delta^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-\delta x}, \quad \gamma, \delta > 0.$$

The term appearing within the first bracket in (3.1) is recognized as the essential part of a gamma density for λ_c with shape parameter $(n+1)$ and inverse scale parameter T . The second bracket contains the essential part of a gamma density for Ψ with shape parameter $(n+1)$ and inverse scale parameter z . This observation suggests that we should consider prior distributions for (A_c, Ψ) , where A_c and Ψ are independently gamma distributed. That is,

$$(3.4) \quad \{p(\lambda_c, \psi) = g_{v,\tau}(\lambda_c) g_{\gamma,\zeta}(\psi) | v, \tau, \gamma, \zeta \geq 0\}.$$

By virtue of (2.3), the class (3.4) induces a class of joint prior distributions for the basic parameters (A, Ψ) . When A_c is gamma distributed with parameters

(ν, τ) , it follows that $\lambda = \lambda_c(c/s)^\psi$, for fixed $\Psi = \psi$, is gamma distributed with parameters $(\nu, \tau(s/c)^\psi)$. Thus, the class of joint prior distributions for (λ, Ψ) , induced by (3.4), is given by

$$(3.5) \quad \{p(\lambda, \psi) = g_{\nu, \tau(s/c)^\psi}(\lambda|\psi) g_{\gamma, \zeta}(\psi) | \nu, \tau, \gamma, \zeta \geq 0\}.$$

The posterior distribution corresponding to a prior from (3.4) (or (3.5)) is obtained upon multiplying $p(\lambda_c, \psi)$ by the likelihood function (3.1), and normalizing to obtain a probability density function. Thus,

$$p(\lambda_c, \psi | \mathcal{D}) \propto \ell(\lambda_c, \psi | \mathcal{D}) p(\lambda_c, \psi) \\ \propto (\lambda_c^{\nu+n-1} e^{-\lambda_c(T+\tau)}) (\psi^{\nu+\gamma-1} e^{-\psi(z+\zeta)}),$$

which gives a new member of the class (3.4) with parameters being updated according to the following rule:

$$(3.6) \quad \begin{array}{ccccc} \text{prior} & \nu & \tau & \gamma & \zeta \\ \text{posterior} & \nu+n & \tau+T & \gamma+n & \zeta+z \end{array}.$$

This shows that (3.4) is a conjugate class of priors for (λ_c, Ψ) , and (3.5) is consequently also a conjugate class of priors for (λ, Ψ) . It appears from (3.6) that the three parameter family of prior distributions obtained from (3.4) by letting $\gamma = \nu$ is also closed under sampling, and would therefore also qualify as a conjugate class of priors. The updating rule (3.6) also shows that the prior parameters can be given a natural interpretation. The prior knowledge corresponding to $(\nu, \tau, \gamma, \zeta)$ can be viewed as the information obtained by observing a similar contract during τ years, with ν and/or γ being the observed number of claims exceeding the data capture level c , and ζ being the corresponding value of the statistic (3.2). If one insists on this interpretation, the parameters ν and γ should in fact be taken equal.

Note that λ_c and Ψ are stochastically independent prior as well as posterior to data, whereas λ and Ψ are stochastically dependent, according to (3.5). The (marginal) distribution of the basic Poisson rate λ depends in a complicated manner on all the parameters $(\nu, \tau, \gamma, \zeta)$. This is in contrast to Jewell's (1990) model, in which the prior information is expressed in terms of λ and the severity parameter. In that model the two parameters are independent prior to data, but, posterior dependent due to the shape of the likelihood function (3.1). In this author's opinion it seems more reasonable that the reinsurer should express his prior opinion in terms of the rate of observed claim numbers $N_c(t)$, rather than $N(t)$ which will never be observed.

4. PREDICTING THE FUTURE CLAIMS COST

We wish to calculate the posterior expectation at time T of quantities like (2.13) and (2.14). If the prior distribution is taken from the conjugate class (3.4), it suffices to calculate $E e_k(\theta_a)$, where expectation refers to a generic member of the class (3.4). The posterior expectations are then obtained by inserting the updated parameters (3.6) for $(\nu, \tau, \gamma, \zeta)$. Thus, assume that λ_c and

Ψ are independently gamma distributed with parameters (ν, τ) and (γ, ζ) , respectively

From (2.13), (2.14) and (2.4) we first note that

$$(4.1) \quad \begin{aligned} E e_k(\Theta_a) &= E[A_a \mu_k(\Psi)] = E[A_c (c/a)^\Psi \mu_k(\Psi)] \\ &= \frac{\nu}{\tau} E[(c/a)^\Psi \mu_k(\Psi)], \end{aligned}$$

since A_c is independent of Ψ with expectation ν/τ . For calculating the expectation (4.1), it is convenient to use the expression (see (A.7) in the Appendix)

$$(4.2) \quad \mu_k(\psi) = b a^{k-1} \sum_{n=k}^{\infty} p(n) \sum_{i=0}^{n-k} \binom{n-1}{i} (-1)^i \psi^i g_k(n-i),$$

where

$$(4.3) \quad p(n) = \frac{(\ln(b/a))^n}{n!} e^{-\ln(b/a)}, \quad n = 0, 1, \dots,$$

denotes the Poisson probabilities corresponding to the parameter $\ln(b/a)$, and

$$(4.4) \quad g_k(m) = \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} j^m.$$

With (4.2) inserted into in (4.1), it appears that we need to calculate expectations of the form $E(\Psi^m e^{-s\Psi})$. When Ψ is gamma distributed with parameters (γ, ζ) , it holds that

$$(4.5) \quad E(\Psi^m e^{-s\Psi}) = \frac{\Gamma(\gamma+m)\zeta^\gamma}{\Gamma(\gamma)(\zeta+s)^{\gamma+m}},$$

and we arrive at the expression,

$$(4.6) \quad E e_k(\Theta_a) = \frac{\nu b a^{k-1}}{\tau} \sum_{n=k}^{\infty} p(n) \sum_{i=0}^{n-k} \binom{n-1}{i} \frac{(-1)^i \Gamma(\gamma+i)\zeta^\gamma g_k(n-i)}{\Gamma(\gamma)(\zeta+\ln(a/c))^{\gamma+i}}.$$

The predicted moments $E(e_k(\Theta_a)|\mathcal{A})$, and in particular the predicted future claims cost $E(X|\mathcal{A}) = E(e_1(\Theta_a)|\mathcal{A})$, are now obtained by inserting the updated parameters (3.6) into (4.6)

One may also want to calculate separate estimates for the expected number of claims in excess of a , and the average reinsurance compensation for the layer $b-a$ in excess of a . The expected number of claims exceeding the priority a is

$$(4.7) \quad E[A_a] = \frac{\nu}{\tau} E[(c/a)^\Psi] = \frac{\nu}{\tau} \frac{\zeta^\gamma}{(\zeta + \ln(a/c))^\gamma},$$

because of (4.5). From (4.2) and (4.5) we also obtain the average reinsurance compensation,

$$(4.8) \quad E \mu_k(\Psi) = ba^{k-1} \sum_{n=k}^{\infty} p(n) \sum_{i=0}^{n-k} \binom{n-1}{i} (-1)^i \frac{\Gamma(\gamma+i) g_k(n-i)}{\Gamma(\gamma) \zeta^i}$$

The expressions (4.6) and (4.8) involve an infinite sum. In practice, the first (approximately) 10 terms will give a sufficient degree of accuracy. It would of course be tempting to use the updated parameters (3.6) to obtain a posterior estimate of Ψ ,

$$\tilde{\Psi} = E(\Psi|\mathcal{O}) = \frac{\gamma+n}{\zeta+z},$$

and then base the quotations on $\mu_1(\tilde{\Psi})$. However, it is shown in the Appendix (formula (A.8)) that the function $\mu_k(\psi)$ are strictly convex, which according to Jensen's inequality implies that

$$E[\mu_1(\Psi)|\mathcal{O}] > \mu_1(\tilde{\Psi}).$$

This less sophisticated approach will thus lead to an underestimation of the expected future claims cost

It is interesting to rewrite (4.1) as

$$(4.9) \quad \begin{aligned} E e_k(\Theta_a) &= \frac{v}{\tau} E[(c/a)^\Psi \mu_k(\Psi)] \\ &= \frac{v}{\tau} \{ \text{Cov} [(c/a)^\Psi, \mu_k(\Psi)] + E[(c/a)^\Psi] E \mu_k(\Psi) \} \\ &= \frac{v}{\tau} \text{Cov} [(c/a)^\Psi, \mu_k(\Psi)] + E A_a E \mu_k(\Psi). \end{aligned}$$

It is shown in the Appendix (formula (A.8)) that $\mu_k(\psi)$ is strictly decreasing. The covariance appearing in (4.9) is therefore positive if the function $(c/a)^\Psi$ is decreasing, and negative if it is increasing. This shows that

$$(4.10) \quad E e_k(\Theta_a) \begin{cases} < \\ = \\ > \end{cases} E A_a E \mu_k(\Psi), \quad a \begin{cases} < \\ = \\ > \end{cases} c.$$

In particular, if the total claims cost X is predicted by predicting the claims frequency A_a and the average compensation $\mu_1(\Psi)$ separately, one will underestimate the true expected claims cost if the priority a exceeds the data capture level c , and overestimate otherwise. In practice one would presumably in most cases have $a > c$.

5. AN EXAMPLE

The data presented in Table 1 are taken from RYTGAARD (1990). The reinsurer has access to information about all claims exceeding $c = 1.5$ million during a period of $T = 5$ years.

TABLE 1
CLAIMS EXCEEDING $c = 1.5$ MILLION DURING $T = 5$ YEARS

Year no				
1	2	3	4	5
2 495	1 985	3 215	—	19 180
2 120	1 810	2 105	—	1 915
2 095	1 625	1 765	—	1 790
1 700	—	1 715	—	1 755
1 650	—	—	—	—

The observed average number of claims pr. year is 3.2 and the statistic z in (3.2) equals 6.48 in this case. We assume that Ψ is gamma distributed with mean $\nu/\tau = 2$ and a coefficient of variation equal to $1/\sqrt{\nu} = 0.3$, and that A_c (with $c = 1.5$ million) is gamma distributed with mean $\gamma/\zeta = 3$ and also a coefficient of variation equal to $1/\sqrt{\gamma} = 0.3$. This specifies a (joint) conjugate prior distribution from the class (3.4), with parameters $(\nu, \tau, \gamma, \zeta) = (11.1, 3.7, 11.1, 5.6)$. This prior corresponds to the information obtained by observing a similar contract which gives rise to $\nu = \gamma = 11.1$ claims in excess of 1.5 million during a period of $\tau = 3.7$ years. The corresponding value of the statistic in (3.2) is $\zeta = 5.6$. From (3.6) we obtain the updated parameters $(\nu, \tau, \gamma, \zeta) = (27.1, 8.7, 27.1, 12.1)$. Thus, posterior to data, we expect $27.1/8.7 = 3.1$ claims pr. year in excess of 1.5 million, and the posterior expectation of Ψ equals $27.1/12.1 = 2.24$. The updated coefficient of variation is $1/\sqrt{27.1} = 0.19$.

The prior and posterior distribution for (A_a, Ψ) are shown in Figure 1 for $a = c = 1.5$ million, and in Figures 2 and 3 for $a = 0.8$ and $a = 2.2$ million, respectively. Since $(c/a)^\Psi$ is increasing for $a < c$, it follows that $A_a = A_c(c/a)^\Psi$ is positively correlated with Ψ for $a < c$, in accordance with Figure 2. Similarly, A_a and Ψ are negatively correlated for $a > c$, which is confirmed by Figure 3.

Table 2 shows the predicted total claims cost prior and posterior to data for the layer 5 million in excess of $a = 0.8, 1.5, 2.2$ million, respectively, $E e_1(\Theta_a)$ and $E[e_1(\Theta_a)|\mathcal{D}]$, and also the corresponding moments of 2nd and 3th order. In Table 2 we have also calculated the first 3 moments of X (prior and posterior to data), by integrating the conditional moments (2.10) with respect to a discrete approximation to the prior and posterior distributions. In each case the approximation involves 60×60 vales of (λ_a, ψ) , and there-

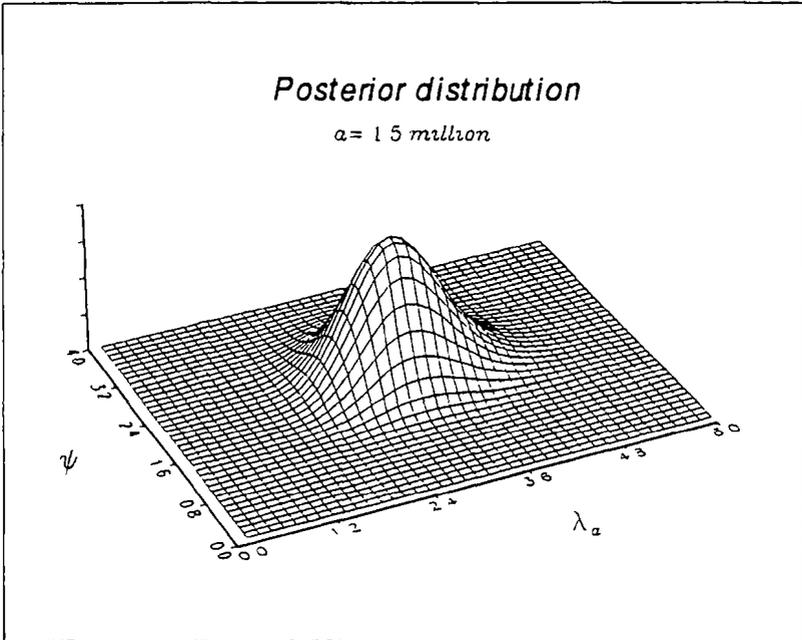
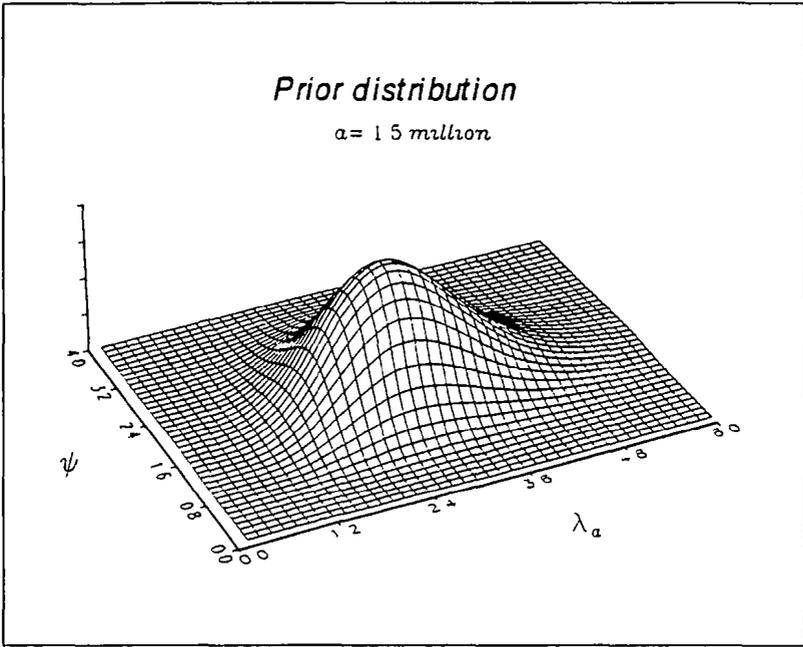
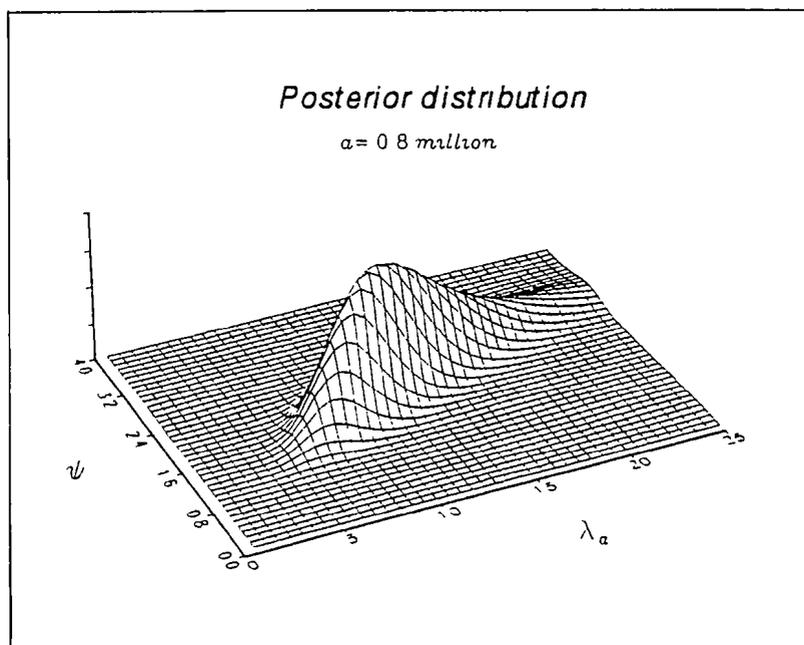
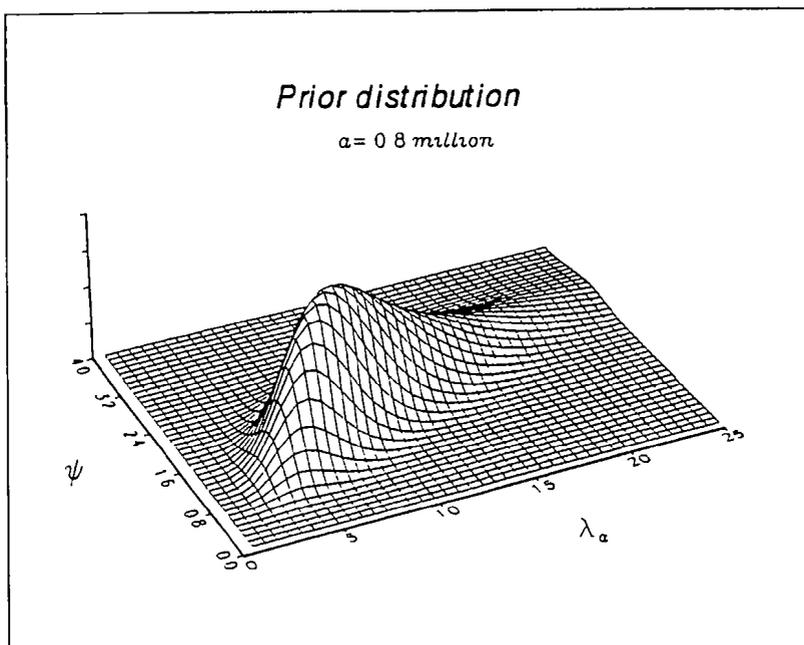


FIGURE 1 Prior and posterior distributions for (λ_a, ψ) , $a = 15 \text{ million}$

FIGURE 2 Prior and posterior distributions for (λ_a, ψ) , $a = 0.8 \text{ million}$

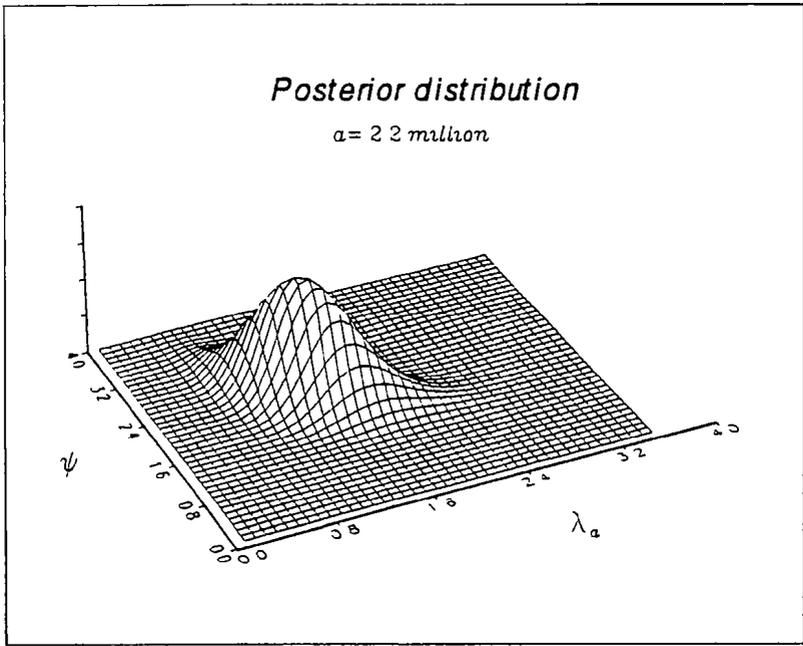
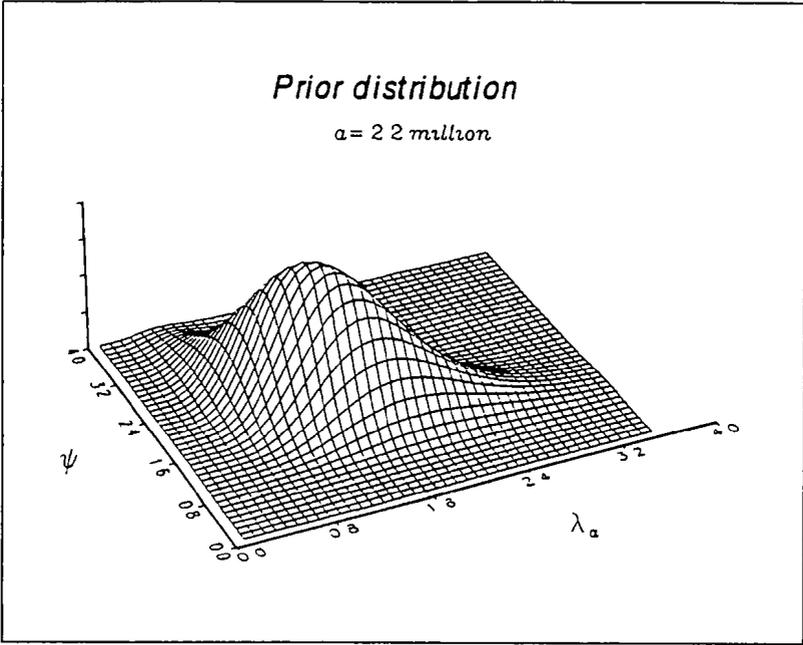


FIGURE 3 Prior and posterior distributions for (λ_a, ψ) , $a = 2.2$ million

TABLE 2
 PREDICTED MOMENTS FOR THE LAYER 5 MILLION IN EXCESS OF a MILLION,
 ACCORDING TO (4.6) AND (2.9), (2.10)

Priority a	$E e_1(\Theta_a)$	$E(e_1(\Theta_a) \mathcal{D})$	EX	$E(X \mathcal{D})$
0.8 million	7.63	7.69	7.50	7.66
1.5 million	3.75	3.26	3.72	3.26
2.2 million	2.45	1.92	2.44	1.92
	$E e_2(\Theta_a)$	$E(e_2(\Theta_a) \mathcal{D})$	$\text{Var } X$	$\text{Var}(X \mathcal{D})$
0.8 million	16.70	14.27	21.39	15.47
1.5 million	11.05	8.45	13.55	8.92
2.2 million	8.26	5.83	10.16	6.26
	$E e_3(\Theta_a)$	$E(e_3(\Theta_a) \mathcal{D})$	$E(X - EX)^3$	$E[(X - E(X \mathcal{D}))^3 \mathcal{D}]$
0.8 million	59.37	46.02	105.67	61.04
1.5 million	43.86	31.28	83.50	43.55
2.2 million	34.61	23.13	64.49	31.61

for the computation of 3600 conditional moments. Except for inaccuracies due to the discretization, it should hold true that $E e_1(\Theta_a) = EX$ and $E[e_1(\Theta_a)|\mathcal{D}] = E(X|\mathcal{D})$, according to (2.15). For the second order moments we note that $\text{Var } X > E e_2(\Theta_a)$ and $\text{Var}(X|\mathcal{D}) > E[e_2(\Theta_a)|\mathcal{D}]$, according to (2.16), and, as noted at the end of Section 2, $\text{Var}(X|\mathcal{D}) \approx E[e_2(\Theta_a)|\mathcal{D}]$ as the amount of data increases.

In Table 3 we have calculated the expected number of excess-of-loss claims (4.7), and the expected reinsurance compensation (4.8) in respect of a single claim, prior and posterior to data. These are shown together with the estimate $E A_a E \mu_1(\Psi)$, and $E e_1(\Theta_a)$ from Table 2, with expectation taken prior and posterior to data. A comparison of the estimates $E e_1(\Theta_a)$ and $E A_a E \mu_1(\Psi)$ (prior and posterior to data) confirms the relation (4.10), and indicates that $E A_a E \mu_1(\Psi)$ significantly overestimates the future claims cost when $a < c$. In

TABLE 3

EXPECTED NUMBER OF CLAIMS IN EXCESS OF a MILLION, ACCORDING TO (4.7) EXPECTED EXCESS-OF-LOSS COMPENSATION FOR THE LAYER 5 MILLION IN EXCESS OF a MILLION, ACCORDING TO (4.8)

Priority a	$E A_a$	$E(A_a \mathcal{D})$	$E \mu_1(\Psi)$	$E(\mu_1(\Psi) \mathcal{D})$
0.8 million	11.39	13.33	0.78	0.62
1.5 million	3.00	3.12	1.25	1.05
2.2 million	1.43	1.33	1.62	1.40
	$E A_a E \mu_1(\Psi)$	$E(A_a \mathcal{D}) E(\mu_1(\Psi) \mathcal{D})$	$E e_1(\Theta_a)$	$E(e_1(\Theta_a) \mathcal{D})$
0.8 million	8.91	8.27	7.63	7.69
1.5 million	3.75	3.26	3.75	3.26
2.2 million	2.31	1.86	2.45	1.92

this case, by 16% based on the prior distribution, and by 7.5% based on the posterior distribution. For $a > c$, the effect is less marked; in this case an underestimation by 6.1% and 3.8% in the case of prior and posterior distributions, respectively.

Since the posterior mean of Ψ , 2.24, is greater than the prior mean, 2, we expect that the distribution of single claim amounts is less heavy-tailed after having seen the data. This explains why the expected excess-of-loss compensation $E\mu_1(\Psi)$ shows a decrease when moving from prior to posterior distributions. For $a = c = 1.5$ million the expected rate of claims in excess of a is only affected slightly by the data (from 3.00 to 3.12). Since $E\lambda_a = E\lambda_c E(c/a)^\Psi$, and the posterior estimate of Ψ exceeds the prior estimate, we should expect the posterior estimate of λ_a to exceed the corresponding prior estimate when $c > a$. Conversely, of course, when $c < a$. This is confirmed by the result in Table 3.

ACKNOWLEDGEMENT

The author would like to thank an anonymous referee and Søren Scheuer, Northwestern National Life, International Reinsurance, for useful comments on the paper.

REFERENCES

- GOOVAERTS, M. J., DE VYLDER, F. and HAEZENDONCK, J. (1984) *Insurance premiums*. North-Holland, Amsterdam-New York.
- HARTIGAN, J. A. (1983) *Bayes theory*. Springer-Verlag, New York.
- JEWELL, W. S. (1990) Up the misty staircase with credibility theory. *Mittel Schweiz Verein Versich Math*, **90**, 281-312.
- JEWELL, W. S. (1991) The value of information in forecasting excess losses. *Paper presented to the 23rd Astin Colloquium, Stockholm*.
- PATRIK, G. and MASHITZ, I. (1989) Credibility for reinsurance excess pricing. *Paper presented at the 21st Astin Colloquium, New York*.
- PANJER, H. H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* **11**, 22-26.
- RYTGAARD, M. (1990) Estimation in the Pareto distribution. *ASTIN Bulletin* **20**, 201-216.

A. APPENDIX

Consider Z , defined in (2.5), with cumulative distribution function H_ψ from (2.6). The non-central moments $\mu_k(\psi) = E[Z_i^k|\psi]$ can be calculated as follows,

$$\begin{aligned} \mu_k(\psi) &= \int_a^b (y-a)^k \psi a^\psi y^{-(\psi+1)} dy + (b-a)^k (b/a)^{-\psi} \\ &= \psi a^\psi \int_a^b \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} y^j a^{k-j} y^{-(\psi+1)} dy + (b-a)^k (a/b)^\psi \\ &= \psi a^\psi \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} a^{k-j} \frac{a^{j-\psi} - b^{j-\psi}}{\psi-j} + a^k (b/a-1)^k (a/b)^\psi \\ &= a^k \left\{ \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} \frac{\psi}{\psi-j} (1-(a/b)^{\psi-j}) \right. \\ &\quad \left. + \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} (a/b)^{\psi-j} \right\} \end{aligned}$$

$$(A.1) \quad = a^k \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} j \frac{1-(a/b)^{\psi-j}}{\psi-j}$$

$$(A.2) \quad = ka^k \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j \frac{1-(a/b)^{\psi-k+j}}{\psi-k+j}$$

An alternative expression for $\mu_k(\psi)$, which is more convenient for calculating posterior expectations, is obtained by expanding the exponential term $(a/b)^{\psi-j}$ appearing in (A.1),

$$\begin{aligned} (A.3) \quad \frac{1-(a/b)^{\psi-j}}{\psi-j} &= - \sum_{n=0}^{\infty} \frac{(\ln(a/b))^{n+1}}{(n+1)!} (\psi-j)^n \\ &= \sum_{n=0}^{\infty} \frac{(\ln(b/a))^{n+1}}{(n+1)!} (j-\psi)^n \\ &= (b/a) \sum_{n=0}^{\infty} p(n+1) \sum_{l=0}^n \binom{n}{l} (-1)^l \psi^l j^{n-l}, \end{aligned}$$

where

$$(A.4) \quad p(n) = \frac{(\ln(b/a))^n}{n!} e^{-\ln(b/a)}, \quad n = 0, 1, \dots,$$

denotes the Poisson probabilities corresponding to the parameter $\ln(b/a)$. By

inserting (A.3) into (A.1) we obtain the expression

$$(A.5) \quad \mu_k(\psi) = ba^{k-1} \sum_{n=0}^{\infty} p(n+1) \sum_{l=0}^n \binom{n}{l} (-1)^l \psi^l \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} j^{n-l+1}.$$

Finally, let

$$(A.6) \quad g_k(a) \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} j^a.$$

Since $g_k(a) = 0$ for $a = 0, \dots, k-1$, we may change the range of summation in (A.5) to $n = k-1, \dots, \infty$ and $l = 0, \dots, n-k+1$, and so we may write

$$(A.7) \quad \mu_k(\psi) = ba^{k-1} \sum_{n=k}^{\infty} p(n) \sum_{l=0}^{n-k} \binom{n-1}{l} (-1)^l \psi^l g_k(n-l).$$

The moments $\mu_k(\psi)$, regarded as functions of ψ , are strictly decreasing and convex. We shall prove that

$$(A.8) \quad (-1)^m \frac{d^m}{d\psi^m} \mu_k(\psi) > 0, \quad k = 1, 2, \dots$$

This is seen by writing

$$(A.9) \quad \mu_k(\psi) = \int_a^{\infty} h_k(y) (\psi/y) (a/y)^\psi dy,$$

where

$$h_k(y) = (\min(y, b) - a)^k > 0, \quad y > a,$$

for $a < b$. By differentiating (A.9) we obtain that

$$\frac{d}{d\psi} \mu_k(\psi) = \int_a^{\infty} h_k(y) (1/y) (a/y)^\psi [1 + \psi \ln(a/y)] dy,$$

and it follows by induction that

$$(A.10) \quad \frac{d^m}{d\psi^m} \mu_k(\psi) = \int_a^{\infty} h_k(y) (1/y) (\ln(a/y))^{m-1} (a/y)^\psi [m + \psi \ln(y/a)] dy$$

By a change of variable we may rewrite (A.10) as

$$(A.11) \quad \begin{aligned} \frac{d^m}{d\psi^m} \mu_k(\psi) &= (-1)^m \int_0^{\infty} h_k(ae^z) z^{m-1} e^{-\psi z} [\psi z - m] dz \\ &= (-1)^m \frac{\Gamma(m)}{\psi^m} \int_0^{\infty} h_k(ae^z) (\psi z - m) g_{m,\psi}(z) dz, \end{aligned}$$

where $g_{m,\psi}(z)$ is the gamma density (3.3). Since $h_k(ae^z)$ is an increasing function, it follows that $E[h_k(ae^Z)(\psi Z - m)] > E(h_k(ae^Z)) E[\psi Z - m]$, and $E[\psi Z - m] = (\psi(m/\psi) - m) = 0$, when Z is gamma distributed with parameters (m, ψ) . This, together with (A 11), verifies (A 8).

OLE HESSELAGER

*Laboratory of Actuarial Mathematics, Universitetsparken 5,
University of Copenhagen, DK-2100 Copenhagen Ø*

