Bayesian Credibility for Excess of Loss Reinsurance Rating

By Mark Cockroft[1]

Lane Clark & Peacock LLP

GIRO conference 2004

Abstract

This paper discusses how to derive a credibility weight between exposure- and experience-driven rates for excess of loss reinsurance. The exposure rate comes from a Poisson-Pareto model, each with a prior Gamma distribution.

Keywords

Excess of loss reinsurance; rating; Bayesian credibility; Poisson; Pareto; Gamma prior

Acknowledgements and disclaimer

I would like to thank those who have assisted me in writing this paper with comments and clarifications as appropriate. Errors remain my responsibility. The views and conclusions expressed in this paper are my own.

1. Introduction

Rating excess of loss (XoL) reinsurance can be performed by considering the recent claims of the reinsured ("experience-rating") or by comparing the risk profile against a benchmark ("exposure-rating"). Rarely do the two methods agree on a rate. Sometimes there are good reasons why one method does not apply in a given circumstance (for example a significant change in the risk profile effectively invalidating the experience-rating method). But in the remaining cases there is no valid reason why a method has nil credibility and the other full credibility.

This paper builds on work published in papers by Buhlmann (1967) and Patrik & Mashitz (1990), which use least squares credibility and apply it to numbers of excess claims. The work is extended in this paper to aggregate claims to XoL layers. Aggregate loss features, such as inner aggregate deductibles, aggregate limits, reinstatement premiums and profit commissions are not considered in this paper.

The paper assumes readers are familiar with experience-rating and exposure-rating for XoL reinsurance.

[1] E-mail: mark.cockroft@lcp.uk.com. Telephone: +44 (0)20 7432 0645

The rest of this paper is in five sections. First, it formulates the problem, and introduces the notation. Then it summarises the work on least squares credibility and excess probability in claim numbers, using a Bayesian Gamma/Poisson credibility model. In Section 4, the paper considers a suitable Bayesian credibility model for claim severity. In Section 5, the paper uses claim numbers and claim severity in a combined model of Bayesian credibility. Finally, practical constraints are considered for when experience- and exposure-rating methods have already been used together.

2. The problem

Reinsurance rating tries to estimate the risk of a loss to a XoL reinsurance layer, with deductible D and upper limit U, i.e. a layer "(U-D) xs D". The risk of loss from the ground-up (FGU) is governed by a parameter θ (or parameter set Θ). In setting up the problem initially, we shall keep to the uni-parameter case, θ.

The *a priori* assumption is that the parameter has a given distribution, giving the claim distribution mean $\mu(\theta)$ and variance $\sigma^2(\theta)$, conditional on θ. There are also observed a series of claims X_i, $i = 1, \ldots, N$, which are used to give a posterior estimate, $\mu(\theta)|X_1, \ldots, X_N$. The X_i's are assumed to be independent and identically distributed. The classic estimator is the mean $E[\mu(\theta)|X_1, \ldots, X_N]$, which corresponds to the expected least squares deviation. The number of observations, N, is itself a random variable, which we assume is independent of the claim sizes and to have its own prior and posterior distributions.

The paper considers the credibility best linear estimate under quadratic loss of the Bayesian posterior mean, and then extends the problem to the excess layers. A linear estimate is one of the form $\alpha + \beta\overline{X}$.

3. Least squares credibility and Bayesian excess layer claim counts.

The following theorem and proof are taken from Buhlmann (1967):

Theorem 3.1 The best linear estimator under quadratic loss of the mean of $\mu(\theta)$ within the problem set out above is the credibility estimate: $Z\overline{X} + (1 - Z)E[\mu(\theta)]$, where

$$Z = \frac{n}{n + \rho}, \rho = \frac{E[\sigma^2(\theta)]}{Var[\mu(\theta)]}.$$

Proof: To estimate the mean of $\mu(\theta)$ using a linear estimator under quadratic loss, the amount that we are trying to minimise is $E[\alpha + \beta\overline{X} - \mu(\theta)]^2$, which can be rewritten as $E[\alpha - (1 - \beta)\mu(\theta)]^2 + E[\beta(\overline{X} - \mu(\theta))]^2$.

The first term is minimised at $\alpha = (1 - \beta)E[\mu(\theta)]$. So the total to be minimised is $(1 - \beta)^2 Var[\mu(\theta)] + \beta^2 E[(\overline{X} - \mu(\theta))]^2$.

Differentiation with respect to β and setting the derivative equal to zero gives $\beta = \dfrac{Var[\mu(\theta)]}{Var[\mu(\theta)] + E[(\overline{X} - \mu(\theta))]^2}$.

The X_i are independent and identically distributed, from which it follows that:

$$E(\overline{X} - \mu(\theta)]^2 = \frac{1}{n} E[X_1 - \mu(\theta)]^2 = \frac{1}{n} E[\sigma^2(\theta)].$$

Putting $Z = \beta$ gives the credibility estimate in the form as required.

Patrik & Mashitz (1990) apply Buhlmann's theorem to the number of claims per year adjusted for inflation, IBNR and relative exposure as follows.

It is assumed that N, the number of claims, has a Poisson distribution with parameter θ, which has a prior Gamma distribution with parameters (a, b). It is also assumed that the k years of claims have observed number of claims $n_1,\ldots, n_k$ and $m = \Sigma_j n_j$.

Immediately we can state that

- the exposure-rating for number of claims is $E[N|\theta] = E[\theta] = a/b$, and
- the experience-rating is $\overline{N} = m/k$.

Lemma 3.2 The Bayesian posterior distribution is also Gamma with parameters $(a+m, b+k)$.

Proof: $N|\theta \sim \mathrm{Poi}(\theta)$, so $f_N(n|\theta) = \dfrac{\theta^n e^{-\theta}}{n!}$, $\theta > 0$, $n = 0,1,2,\ldots$

$\theta \sim \Gamma(a, b)$, so $f_\Theta(\theta) = \dfrac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)}$, $a > 0$, $b > 0$, $\theta > 0$, where

$$\Gamma(a) = \int_0^\infty e^{-z} z^{a-1} dz$$

If $f_{\Theta|N}(\theta| n_1, \ldots,n_k)$ is the Bayesian posterior distribution of θ, then

$$\begin{aligned} f_{\Theta|N}(\theta \mid n_1,\ldots,n_k) &\propto L(n_1,\ldots,n_k \mid \theta) \times f_\Theta(\theta) \\ &= \prod_{j=1}^{k} \frac{\theta^{n_j} e^{-\theta}}{n_j!} \cdot \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} \\ &= \frac{\theta^m e^{-k\theta}}{\prod_j n_j!} \cdot \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} \propto \theta^{a+m-1} e^{-(b+k)\theta} \end{aligned}$$

This is the form of a Gamma distribution with parameters $(a+m, b+k)$ as required.

The following six results, from Theorem 3.3 to Corollary 3.8 inclusive, are from Patrik & Mashitz (1990).

Theorem 3.3 Under the assumed circumstances described above, the best linear estimate of N is the credibility form $Z.\frac{m}{k}+(1-Z)\frac{a}{b}$, where $Z=\frac{k}{k+b}$, and this is also the mean of the Bayesian posterior distribution.

Proof: The general form of the credibility estimate follows directly from Buhlmann's theorem (Theorem 3.1 above), noting the exposure-rating and experience-rating above.

The form of Z follows from noting that the number of observed data is the number of years, k, that for a Poisson distribution with parameter θ $\mu(\theta) = \theta$ and $\sigma^2(\theta) = \theta$ and that the Gamma prior distribution with parameters (a, b) has mean a/b and variance a/b^2.

To prove that this is the mean of the Bayesian posterior, denoted $E[N|\theta, m, k, a, b]$, we note, from Lemma 3.2, that the posterior distribution is Gamma with parameters (a+m, b+k) so:

$$
\begin{aligned}
E[N|\theta,m,k,a,b] &= \frac{a+m}{b+k} = \frac{m}{b+k} + \frac{a}{b+k} \\
&= \frac{k}{b+k}\frac{m}{k} + \frac{b}{b+k}\frac{a}{b}
\end{aligned}
$$

which is the form required.

Theorem 3.4 The unconditional distribution of N is a Negative Binomial distribution with parameters (a, $b/(1+b)$).

Proof: If $f_N(n)$ represents the unconditional distribution of N, then

$$
\begin{aligned}
f_N(n) &= \int_0^\infty \frac{\theta^n e^{-\theta}}{n!}\frac{b^a e^{-b\theta}\theta^{a-1}}{\Gamma(a)}d\theta \\
&= \frac{b^a}{n!\Gamma(a)}\frac{\Gamma(a+n)}{(1+b)^{a+n}}\int_0^\infty \frac{(1+b)^{a+n}e^{-(1+b)\theta}\theta^{a+n-1}}{\Gamma(a+n)}d\theta \\
&= \frac{\Gamma(a+n)}{n!\Gamma(a)}\left(\frac{b}{1+b}\right)^a\left(\frac{1}{1+b}\right)^n\cdot 1
\end{aligned}
$$

which is the form of a Negative Binomial with parameters (a, $b/(1+b)$) as required.

To apply the credibility results to excess layers, denote N_D for the number of claims that exceed D and q_D for the probability that a given FGU claim exceeds D.

In the special case where q_D is a fixed, known ratio in terms of D, the following lemma and theorem are applicable.

Lemma 3.5 N_D is also Poisson with parameter $q_D\theta$.

Proof: $P(N_D = n) = \sum_{j=n}^{\infty} P(N_D = n | N = j) P(N = j)$.

Within the summation, the first probability inside the summation is exactly the definition of the Negative Binomial distribution with parameters (j-n, q_D). Hence:

$$P(N_D = n) = \sum_{j=n}^{\infty} \frac{j!}{n!(j-n)!}(1-q_D)^{j-n} q_D{}^n \frac{\theta^j e^{-\theta}}{j!}$$

$$= \frac{q_D{}^n e^{-\theta}}{n!} \theta^n \sum_{j=n}^{\infty} \frac{((1-q_D)\theta)^{j-n}}{(j-n)!}$$

$$= \frac{(q_D\theta)^n e^{-\theta}}{n!} e^{(1-q_D)\theta}$$

$$= \frac{(q_D\theta)^n e^{-q_D\theta}}{n!}$$

which is the form of the Poisson as required.

Define $\theta_D = q_D\theta$.

Lemma 3.6 θ_D has a Gamma distribution with parameters (a, b/q_D).

Proof: $P(\theta_D \le t) = P\left(\theta \le \frac{t}{q_D}\right) = \int_o^{t/q_D} \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} d\theta$

$\theta_D = q_D\theta$, so $\theta = 1/q_D.\theta_D$ and $d\theta = 1/q_D.d\theta_D$.

Hence:

$$P(\theta_D \le t) = \int_o^t \frac{b^a e^{-b\theta_D/q_D} \left(\theta_D/q_D\right)^{a-1}}{\Gamma(a)} \frac{1}{q_D} d\theta_D$$

$$= \int_o^t \frac{\left(b/q_D\right)^a e^{-(b/q_D)\theta_D} \theta_D{}^{a-1}}{\Gamma(a)} d\theta_D$$

as required.

Returning to the observed claim numbers, for the k years of claims denote the number of claims that exceed D in each year be $n_{D,1},\ldots, n_{D,k}$ and $m_D = \Sigma_j n_{D,j}$.

As before, we have the following estimates for N_D:

- the exposure-rating is $E[N|\theta_D] = a/(b/q_D)$, and
- the experience-rating is $\overline{N} = m_D/k$.

Theorem 3.7 Under the preceding conditions, the best linear estimate of N_D is the credibility form $Z_D.\frac{m_D}{k}+(1-Z_D)\frac{a}{b/q_D}$, where $Z_D=\frac{k}{k+b/q_D}$, and this is also the mean of the Bayesian posterior distribution.

Proof: By applying Lemmas 3.5 and 3.6 to Theorem 3.3, replacing the variable θ with θ_D and parameter b with b/q_D.

Corollary 3.8 If q_D is strictly monotonic decreasing in D, and $d_1<d_2$, then $Z_{d1}>Z_{d2}$.

Proof: Since $Z_D=\frac{k}{k+b/q_D}$ is larger for larger values of q_D.

Theorem 3.9 The unconditional distribution of numbers of claims exceeding D is Negative Binomial with parameters (a, $b/(q_D+b)$).

Proof: By applying Lemmas 3.5 and 3.6 to Theorem 3.4, replacing the variable θ with θ_D and parameter b with b/q_D.

If we relax the requirement that q_D is known and fixed in terms of just D, i.e. we let q_D become a function of some risk parameter η, then Lemmas 3.5 and 3.6 no longer apply.

Patrik & Meshitz (1990) do investigate this case though, by assuming that θ and η are independent (which is reasonable), as follows.

Lemma 3.10 If θ and η are independent, the conditional expectation of the mean and variance of θ_D are both are equal to $E[q_D(\eta)]\cdot {}^a\!/_b$.

Proof: For the conditional expectation of the mean:

$$E[\mu(\theta_D|\eta)]=E[q_D(\eta)\mu(\theta)]=E[q_D(\eta)]\cdot E[\mu(\theta)]=E[q_D(\eta)]\cdot {}^a\!/_b.$$

For the conditional expectation of the variance, $E[\sigma^2(\theta_D|\eta)]$, we note that the conditions of Lemma 3.5 apply, since we are considering θ_D conditional on η and for a given risk parameter the probability of a claim exceeding D is known. Hence, $\theta_D|\eta$ is Poisson and the variance of $\theta_D|\eta$ is the same as the mean.

The conditional variance, however, is different.

Lemma 3.11 If Y and Z are independent random variables, then:

$$\mathrm{Var}[Y.Z] = E[Y]^2.\mathrm{Var}[Z]+E[Z]^2.\mathrm{Var}[Y]+\mathrm{Var}[Y].\mathrm{Var}[Z].$$

Proof: By definition of $Var[Y.Z]$:

$$\begin{aligned} Var[Y.Z] &= E[Y^2.Z^2] - E[Y.Z]^2 \\ &= E[Y^2].E[Z^2] - E[Y]^2.E[Z]^2 \\ &= (E[Y]^2 + Var[Y]).(E[Z]^2 + Var[Z]) - E[Y]^2.E[Z]^2 \\ &= E[Y]^2.Var[Z] + E[Z]^2.Var[Y] + Var[Y].Var[Z] \\ &\quad + E[Y]^2.E[Z]^2 - E[Y]^2.E[Z]^2 \end{aligned}$$

Lemma 3.12 The conditional variance of the expected value of θ_D can be written as

$$Var[\mu(\theta_D|\eta)] = \frac{a}{b^2}\left(E[q_D(\eta)]^2 + (a+1)Var[q_D(\eta)]\right).$$

Proof: Using Lemma 3.11:

$$\begin{aligned} Var[\mu(\theta_D|\eta)] &= Var[\mu(\theta)q_D(\eta)] \\ &= E[\mu(\theta)]^2 Var[q_D(\eta)] + E[q_D(\eta)]^2 Var[\mu(\theta)] \\ &\quad + Var[\mu(\theta)]Var[q_D(\eta)] \\ &= \left(\frac{a}{b}\right)^2 Var[q_D(\eta)] + E[q_D(\eta)]^2 \frac{a}{b^2} + \frac{a}{b^2} Var[q_D(\eta)] \\ &= \frac{a}{b^2}\left(E[q_D(\eta)]^2 + (a+1)Var[q_D(\eta)]\right) \end{aligned}$$

as required.

Theorem 3.13 The best linear estimate of N_D is the credibility form

$Z_D(\eta).\frac{m_D}{k} + (1 - Z_D(\eta))\frac{a}{b}E[q_D(\eta)]$, where

$Z_D(\eta) = \frac{k}{k + b_D(\eta)}$ and $b_D(\eta) = \frac{b}{E[q_D(\eta)].(1 + (a+1)CV[q_D(\eta)]^2)}$.

Note: for a random variable Y, $CV[Y]$ is the coefficient of variation, defined as $Var[Y]^{1/2}/E[Y]$.

Proof: As before, the general form of the credibility estimate follows directly from Buhlmann's theorem (Theorem 3.1). The experience-rating is as before, and the exposure-rating follows from the factorisation of the conditional mean in Lemma 3.10.

Note that $E[\sigma^2(\theta_D|\eta)] = E[q_D(\eta)] \cdot {}^a\!/_b$ from Lemma 3.10 and, by Lemma 3.12:

$$Var[\mu(\theta_D|\eta)] = \frac{a}{b^2}\left(E[q_D(\eta)]^2 + (a+1)Var[q_D(\eta)]\right)$$
$$= \frac{a}{b^2}E[q_D(\eta)]^2\left(1+(a+1)CV[q_D(\eta)]^2\right)$$

From Theorem 3.1, the credibility factor can be written

$Z_D(\eta) = \dfrac{k}{k + b_D(\eta)}$ where

$$b_D(\eta) = \frac{E[\sigma^2(\theta_D|\eta)]}{Var[\mu(\theta_D|\eta)]}$$
$$= \frac{E[q_D(\eta)]\cdot {}^{a}\!/_{b}}{\frac{a}{b^2}E[q_D(\eta)]^2\left(1+(a+1)CV[q_D(\eta)]^2\right)}$$
$$= \frac{b}{E[q_D(\eta)].(1+(a+1)CV[q_D(\eta)]^2)}$$

as required.

Hence the credibility factor reduces as the expected value of $q_D(\eta)$ increases, i.e. with reducing excess. However, the coefficient of variation is unknown so its square may decrease faster with D than the expected value increase divided by $(a + 1)$.

This represents the limit of what can be stated reasonably using just numbers of claims – anything else requires limiting assumptions regarding the form and choice of claim severity, i.e. the $q_D(\eta)$ factor. These are explored further in Patrik & Meshitz (1990).

4. Bayesian claim severity for excess layers.

Probably the most common severity curve used in reinsurance exposure-rating is the Pareto curve. There are many versions, involving differing number of parameters and slightly differing forms of the probability density function. We shall concentrate on one, as set out below.

Further, there is the question of which parameters, or both, should be the subject of a prior distribution, and what form that distribution shall take. In the case of the Pareto, for exposure-rating reinsurance, we are concerned with the proportion of a risk that falls into a reinsurance layer. Hence, the shape parameter is our primary concern.

We are therefore in a position to state our assumptions concerning the problem outlined in Section 2, but for a fixed number of observed claims. In section 5, we shall generalise the situation for reinsurance excess of loss layers with a random number of observed claims.

- There are n claims, $X_1,\ldots, X_n$, which are independent and identically distributed with Pareto distribution parameters (ψ, λ).
- The Pareto has the form given by the distribution function:

$$F_X(x|\psi) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\psi}.$$

- ψ has a prior distribution of a Gamma variable with parameters (s, t).

Let us consider the posterior distribution of ψ and the unconditional distribution of X. The following result for FGU claims is also to be found in Hesselager (1993) although there it is specific to reinsurance layers only:

Theorem 4.1 The Bayesian posterior distribution is also Gamma with parameters ($s+n$, $t+\Sigma_i\ln(\lambda+x_i)$-$n\ln\lambda$).

Proof: $X|\psi \sim \mathrm{Par}(\psi,\lambda)$, so $f_X(x|\theta) = \frac{\psi}{\lambda}\left(\frac{\lambda}{\lambda + x}\right)^{\psi+1}$, $\lambda > 0$, $\psi > 0$, $x > 0$.

$\psi \sim \Gamma(s, t)$, so $f_\Psi(y) = \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)}$, $s > 0$, $t > 0$, $\psi > 0$.

$f_{\Psi|X}(\psi|\ x_1, \ldots,x_n)$, the Bayesian posterior distribution of ψ, has the form:

$$\begin{aligned}
f_{\Psi|X}(\psi \mid x_1,\ldots,x_n) &\propto L(x_1,\ldots,x_n \mid \psi) \times f_\Psi(\psi) \\
&= \prod_{i=1}^{n} \frac{\psi}{\lambda}\left(\frac{\lambda}{\lambda + x_i}\right)^{\psi+1} \cdot \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} \\
&= \frac{\psi^n \lambda^{n\psi}}{\prod_i (\lambda + x_i)^{\psi+1}} \cdot \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} \\
&= \psi^n e^{n\psi \ln(\lambda)} e^{-(\psi+1)\sum_i \ln(\lambda + x_i)} \cdot \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} \\
&\propto \psi^{s+n-1} e^{-(t+\sum_i \ln(\lambda+x_i) - n\ln\lambda)\psi}
\end{aligned}$$

This is the form of a Gamma distribution with parameters ($s+n$, $t+\Sigma_i\ln(\lambda+x_i)$-$n\ln\lambda$), as required.

The unconditional distribution has a slightly different form from the original form.

Theorem 4.2 The unconditional distribution of X is $\lambda(e^Y-1)$ where Y is a Pareto distribution with parameters (s, t).

Proof: If $f_X(x)$ represents the unconditional distribution of X, then

$$f_X(x) = \int_0^\infty \frac{\psi}{\lambda}\left(\frac{\lambda}{\lambda+x}\right)^{\psi+1}\frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)}d\psi$$

$$= \frac{t^s}{\Gamma(s)(\lambda+x)}\int_0^\infty \psi\left(\frac{\lambda}{\lambda+x}\right)^{\psi} e^{-t\psi}\psi^{s-1}d\psi$$

$$= \frac{t^s}{\Gamma(s)(\lambda+x)}\frac{\Gamma(s+1)}{\left(t+\ln\frac{\lambda+x}{\lambda}\right)^{s+1}}\int_0^\infty \frac{\left(t+\ln\frac{\lambda+x}{\lambda}\right)^{s+1}}{\Gamma(s+1)}e^{-\left(t+\ln\frac{\lambda+x}{\lambda}\right)\psi}\psi^{(s+1)-1}d\psi$$

$$= \frac{t^s}{\Gamma(s)(\lambda+x)}\frac{\Gamma(s+1)}{\left(t+\ln\frac{\lambda+x}{\lambda}\right)^{s+1}}\cdot 1 = \frac{st^s}{(\lambda+x)}\frac{1}{\left(t+\ln\frac{\lambda+x}{\lambda}\right)^{s+1}}$$

Define $y = \ln[(\lambda+x)/\lambda]$, so $x = \lambda(e^y-1)$ and $dx = \lambda e^y dy$. Hence

$$f_X(x)dx = \frac{st^s}{\lambda e^y}\frac{1}{(t+y)^{s+1}}\lambda e^y dy$$

$$= \frac{s}{t}\left(\frac{t}{t+y}\right)^{s+1} dy$$

which is the Pareto form in y as required.

The result for the unconditional distribution is hardly intuitive. It is however very practical: we now have closed forms for the distributions of the number of claims and the severity of claims given Gamma prior distributions.

Let us consider now claims to excess layers. Let X_D be the random variable defined by the excess of X over D, if any:

$$X_D = \begin{cases} X-D & \text{if } X > D \\ 0 & \text{if } X \le D \end{cases}$$

The following result is well-known and very useful: assuming a Pareto distribution, the shape of the distribution of the excess claims is the same shape as the distribution of the FGU claims. Further, once we know the position parameter of the FGU claims, the position parameter of the excess claims is the same parameter with the addition of the excess point.

Theorem 4.4 X_D has a Pareto distribution with parameters (ψ, λ+D).

Proof: From the definition of X_D:

$$P(X_D \le x) = P(X \le x + D | X > D)$$

$$= 0 + \frac{P(D < X \le x + D)}{P(X > D)}$$

$$= \frac{\left(\frac{\lambda}{\lambda + D}\right)^{\psi} - \left(\frac{\lambda}{\lambda + D + x}\right)^{\psi}}{\left(\frac{\lambda}{\lambda + D}\right)^{\psi}}$$

$$= 1 - \left(\frac{\lambda + D}{\lambda + D + x}\right)^{\psi}$$

as required.

The result means that whatever credibility estimate we can derive using FGU claim severity, we can extend to excess layers.

However, we cannot, in practice, apply the credibility estimate process in Section 3 straight to claim severity. The observed claims are not just random in size but also in the number of observations.

5. Bayesian credibility for excess layer claims

For limited excess layers, we are likely to have FGU claims above some threshold, or censor point, T. We could aim to model all FGU claims, but there is nothing to gain for rating excess of loss or credibility weighting in doing so. Therefore, where we talk about FGU claims, we are talking about only the subset of FGU claims that exceed T, with no loss of meaning.

We are now in a position to state the assumptions for the full problem for aggregate claims amounts to excess of loss reinsurance:

- The number of FGU claims a year, N, is distributed according to a Poisson distribution, parameter θ.
- θ has a prior Gamma distribution with parameter (a, b).
- The claim severities are independent, of each other and of N, and identically distributed from a Pareto with parameters (ψ, λ).
- ψ has a prior Gamma distribution with parameters (s, t).
- There are k years of claims data, with observed number of claims $n_1,\ldots,n_k$, $m = \Sigma_j n_j$ and observed claim sizes in year j are $x^{(j)}{}_1,\ldots,x^{(j)}{}_{n_j}$.

- We are trying to price a limited excess of loss layer of upper limit U and deductible D, i.e. "(U-D) xs D".

We can state the form of the Bayesian posterior joint distribution of θ and ψ, $f_{\Theta,\Psi|N,X}(\theta, \psi| n_j, x^{(j)}_1,\dots,x^{(j)}_{n_j}, j=1,\dots,k)$ as follows, using the independence of N and X:

$$f_{\Theta,\Psi|N,X}(\theta,\psi \mid n_j, x^{(j)}_1,\dots,x^{(j)}_n, j=1,\dots,k) \propto \prod_{j=1}^{k} L(n_j \mid \theta) L(x^{(j)}_1,\dots,x^{(j)}_n \mid \psi) \times f_\Theta(\theta) \times f_\Psi(\psi)$$

$$= \prod_{j=1}^{k} \frac{\theta^{n_j} e^{-\theta}}{n_j!} \prod_{i=1}^{n_j} \frac{\psi}{\lambda} \left(\frac{\lambda}{\lambda + x^{(j)}_i} \right)^{\psi+1} \cdot \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} \cdot \frac{t^s e^{-t\psi} \psi^{s-1}}{\Gamma(s)}$$

$$= \frac{\theta^m e^{-k\theta}}{\prod_j n_j!} \cdot \prod_j \frac{\psi^{n_j} \lambda^{n_j \psi}}{\prod_i (\lambda + x^{(j)}_i)^{\psi+1}} \cdot \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} \cdot \frac{t^s e^{-t\psi} \psi^{s-1}}{\Gamma(s)}$$

$$= \frac{\theta^m e^{-k\theta}}{\prod_j n_j!} \cdot \psi^m e^{m\psi \ln(\lambda)} e^{-(\psi+1)\sum_j \sum_i \ln(\lambda + x^{(j)}_i)} \cdot \frac{b^a e^{-b\theta} \theta^{a-1}}{\Gamma(a)} \cdot \frac{t^s e^{-t\psi} \psi^{s-1}}{\Gamma(s)}$$

$$\propto \theta^{a+m-1} e^{-(b+k)\theta} \cdot \psi^{s+m-1} e^{-(t + \sum_j \sum_i \ln(\lambda + x^{(j)}_i) - m \ln \lambda)\psi}$$

Hence the Bayesian posterior joint distribution is the product of the Bayesian posterior distributions of the claim numbers and the claim severities that we saw in previous sections.

For simplicity, we can write $y_{ij} = \ln((\lambda + x^{(j)}_i)/\lambda)$, so that the conjugate pairs of prior and posterior parameters are

Prior	Posterior
a	$a+m$
b	$b+k$
s	$s+m$
t	$t+\Sigma_{ij} y_{ij}$

This means we can write down the prior distribution of claim numbers to the excess layer from Lemma 3.6 as (a, b/q_D) where $q_D = P(X > D) = [\lambda/(\lambda+D)]^\psi$.

Let us define X_D as the claim severity to the layer, so

$$X_D = \begin{cases} U - D & , X > U \\ X - D & , D < X \le U \\ 0 & , X \le D \end{cases}$$

In order to create the linear best estimate of the mean of the aggregate layer claims amount Z, according to Theorem 3.1 we need to calculate expectations of the first two central moments based on the prior information.

Let Θ denote the parameter pair (θ_D, ψ), where $\theta_D = \theta[\lambda/(\lambda+D)]^{\psi}$, and write the dth central moment of the distribution of aggregate claims Z given Θ as $e_d(\Theta)$. Then we need to find $E[e_d(\Theta)]$ for $d = 1,2$.

The following result is quoted in Hesselager (1993).

Lemma 5.1 The conditional moments about the origin in this problem can be calculated recursively as

$$E[Z^d|\Theta] = \theta_D \sum_{j=1}^{d} \binom{d-1}{j-1} \mu_j(\psi) E[Z^{d-j}|\Theta], \text{ for d} = 1,2,\ldots$$

where $\mu_j(\psi) = E[X_D{}^j|\psi]$ is the jth central moment of a single claim conditional on the value of ψ.

Proof See Goovaerts *et al* (1984), p12.

Hence, we can state the first two central moments in terms of $\mu_j(\psi) = E[X_D{}^j|\psi]$.

Lemma 5.2 $e_1(\Theta) = \theta_D\mu_1(\psi)$ and $e_2(\Theta) = \theta_D\mu_2(\psi)$.

Proof $e_1(\Theta) = E[Z|\Theta] = E[N_D|\theta,\psi]E[X_D|\psi] = \theta_D\,\mu_1(\psi)$, as N_D and X_D are independent.

$e_2(\Theta) = \mathrm{E}[(Z - e_1(\Theta))^2|\Theta] = \mathrm{E}[Z^2|\Theta]- e_1(\Theta)^2$

From Lemma 5.1

$$\begin{aligned} E[Z^2|\Theta] &= \theta_D \sum_{j=1}^{2} \binom{1}{j-1} \mu_j(\psi) E[Z^{2-j}|\Theta] \\ &= \theta_D\mu_1(\psi)E[Z^1|\Theta] + \theta_D\mu_2(\psi).1 \\ &= [\theta_D\mu_1(\psi)]^2 + \theta_D\mu_2(\psi) \end{aligned}$$

Therefore $e_2(\Theta) = \theta_D\mu_2(\psi)$ as required.

Lemma 5.3 $E[e_1(\Theta)] = (a/b)E[[\lambda/(\lambda+D)]^{\psi}\,\mu_1(\psi)]$, and
$E[e_2(\Theta)] = (a/b)E[[\lambda/(\lambda+D)]^{\psi}\,\mu_2(\psi)]$.

Proof Follows from the expressions for $e_1(\Theta)$ and $e_2(\Theta)$ in Lemma 5.2, using the independence of θ and $[\lambda/(\lambda+D)]^{\psi}\,\mu_d(\psi)$ for d = 1,2.

Lemma 5.3 implies we need to concentrate on the central moments of the individual claims given the assumed prior distribution. The following results, Lemma 5.4, Lemma 5.5 and Theorem 5.6, are from Hesselager (1993). Note that the special cases of ψ taking an integer value are ignored (here and by Hesselager) because, under its prior distribution, $P(\psi = j)$ is infinitesimal for any real value j so the special cases may be safely ignored when considering moments of expectations of ψ.

Lemma 5.4 $\mu_d(\psi)$ may be expressed as

$$\mu_d(\psi) = (\lambda+D)^d \sum_{j=0}^{d} \binom{d}{j} (-1)^{d-j} \frac{j}{\psi - j}\left(1 - \left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right)$$

Proof If we condition on the ranges $X \le D$, $D < X \le U$ and $X > U$, then

$$\begin{aligned}
\mu_d(\psi) &= 0 + \int_D^U (x-D)^d \psi(\lambda+D)^{\psi}(\lambda+x)^{-(\psi+1)}dx + (U-D)^d\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi} \\
&= \int_D^U \left[(\lambda+x)-(\lambda+D)\right]^d \psi(\lambda+D)^{\psi}(\lambda+x)^{-(\psi+1)}dx + (U-D)^d\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi} \\
&= \int_D^U \sum_{j=0}^{d}\binom{d}{j}(\lambda+x)^j(-1)^{d-j}(\lambda+D)^{d-j}\psi(\lambda+D)^{\psi}(\lambda+x)^{-(\psi+1)}dx \\
&\quad + (U-D)^d\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi} \\
&= \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}(\lambda+D)^{d-j}\psi(\lambda+D)^{\psi}\int_D^U(\lambda+x)^{-(\psi-j+1)}dx \\
&\quad + (U-D)^d\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi} \\
&= \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}(\lambda+D)^{d}\psi(\lambda+D)^{\psi-j}\frac{(\lambda+D)^{-(\psi-j)}-(\lambda+U)^{-(\psi-j)}}{\psi-j} \\
&\quad + (\lambda+D)^d\left(\frac{\lambda+U}{\lambda+D}-1\right)^d\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi} \\
&= \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}(\lambda+D)^{d}\frac{\psi}{\psi-j}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right) \\
&\quad + (\lambda+D)^d\left[\sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}\left(\frac{\lambda+U}{\lambda+D}\right)^{j}\right]\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi}
\end{aligned}$$

Bringing together terms in the summations gives:

$$\begin{aligned}
\mu_d(\psi) &= \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}(\lambda+D)^{d}\left[\frac{\psi}{\psi-j}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right)+\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right] \\
&= \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}(\lambda+D)^{d}\left[1+\frac{j}{\psi-j}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right)\right] \\
&= (\lambda+D)^d\sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j} + (\lambda+D)^d\sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j}\frac{j}{\psi-j}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right)
\end{aligned}$$

The first summation is the expansion of $(1-1)^d$ and hence equals zero, as required.

Lemma 5.5 $\mu_d(\psi)$ may be further rewritten as

$$\mu_d(\psi) = (\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}(-1)^r \psi^r g_d(n-r)$$

where $g_d(m) = \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j} j^m$.

Proof We can expand $[(\lambda+D)/(\lambda+U)]^{\psi-j}$ as an exponential:

$$\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j} = e^{(\psi-j)\ln\left(\frac{\lambda+D}{\lambda+U}\right)} = \sum_{n=0}^{\infty}\left[\ln\left(\frac{\lambda+D}{\lambda+U}\right)\right]^n \frac{(\psi-j)^n}{n!}$$

$$= 1-(\psi-j)\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{(j-\psi)^{n-1}}{n!}$$

Substituting this into the form of $\mu_d(\psi)$ in Lemma 5.4 gives:

$$\mu_d(\psi) = (\lambda+D)^d \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j} j \sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{(j-\psi)^{n-1}}{n!}$$

$$= (\lambda+D)^d \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j} j \sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r} j^{n-r-1}(-1)^r\psi^r$$

$$= (\lambda+D)^d \sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r}(-1)^r\psi^r \sum_{j=0}^{d}\binom{d}{j}(-1)^{d-j} j^{n-r}$$

The last summation, for $j = 1,\ldots, d$, in the expression is the form of $g_d(n\text{-}r)$ as defined.

Note that for integer $m < d$, $g_d(m) = 0$. Therefore we can change the ranges of summation as $n = d, d+1,\ldots,\infty$ and $r = 0,1,\ldots, n\text{-}d$, as required.

Theorem 5.6 $E[e_d(\Theta)]$ for $d = 1,2$ is given by:

$$E[e_d(\Theta)] = \frac{a}{b}(\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}\frac{(-1)^r g_d(n-r)\Gamma(s+r)t^s}{\Gamma(s)\left[t+\ln\left(\frac{\lambda+D}{\lambda}\right)\right]^{s+r}}$$

Proof From Lemma 5.4:

$$\mu_d(\psi) = (\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}(-1)^r \psi^r g_d(n-r)$$

Substituting this into the expression for $E[e_d(\Theta)]$ for d = 1,2 from Lemma 5.3 gives:

$$E[e_d(\Theta)] = \frac{a}{b} E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi} (\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}(-1)^r \psi^r g_d(n-r)\right]$$

$$= \frac{a}{b}(\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}(-1)^r g_d(n-r) E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\psi^r\right]$$

ψ is Gamma distributed with parameters (s, t), so:

$$E(\psi^u e^{-v\psi}) = \frac{\Gamma(s+u)t^s}{\Gamma(s)(t+v)^{s+u}}$$

Therefore:

$$E[e_d(\Theta)] = \frac{a}{b}(\lambda+D)^d \sum_{n=d}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-d}\binom{n-1}{r}\frac{(-1)^r g_d(n-r)\Gamma(s+r)t^s}{\Gamma(s)\left[t+\ln\left(\frac{\lambda+D}{\lambda}\right)\right]^{s+r}}$$

as required.

In order to get the final expectation we need, $E[e_1(\Theta)^2]$, we need new versions of Lemmas 5.3 and 5.4.

Lemma 5.7 $e_1(\Theta)^2$ can be written as:

$$e_1(\Theta)^2 = \left(\frac{a(\lambda+D)}{b}\right)^2\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\frac{1}{(\psi-1)^2}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)^2$$

Proof From Lemma 5.3, $e_1(\Theta) = (a/b)[\lambda/(\lambda+D)]^{\psi}\mu_1(\psi)$

Substituting in the result from Lemma 5.4:

$$e_1(\Theta) = \frac{a}{b}\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\mu_1(\psi)$$

$$= \frac{a}{b}\left(\frac{\lambda}{\lambda+D}\right)^{\psi}(\lambda+D)\sum_{j=0}^{1}\binom{1}{j}(-1)^{1-j}\frac{j}{\psi-j}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-j}\right)$$

$$= 0 + \frac{a}{b}\left(\frac{\lambda}{\lambda+D}\right)^{\psi}(\lambda+D)\frac{1}{\psi-1}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)$$

which when squared gives the required result for $e_1(\Theta)^2$.

For a given function f:

$$E(f(\psi)|\psi > x; s,t) = \int_x^{\infty} f(x)\frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)}d\psi$$

If $s + u > 0$ and $t + v > 0$, the following holds true for any $j > 0$:

$$E(\psi^{-j+u}e^{-v\psi}|\psi > x; s,t) = \frac{\Gamma(s+u)t^s}{\Gamma(s)(t+v)^{s+u}}E(\psi^{-j}|\psi > x; s+u, t+v)$$

So, we can build an expression for $E(\psi^{j}|\psi > x;\, s,\, t)$, with integer $j > 0$, iteratively as follows:

$$E(\psi^{0}\big|\psi > x; s,t) = \int_x^\infty 1 \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} d\psi = 1 - \Gamma(x;s,t)$$

where $\Gamma(x;\, s,\, t)$ denotes the gamma cumulative distribution function:

$$\Gamma(x;s,t) = \int_0^x \frac{t^s e^{-tz} z^{s-1}}{\Gamma(s)} dz$$

and, for integer $j > 0$:

$$\begin{aligned}
E(\psi^{-j}\big|\psi > x; s,t) &= \int_x^\infty \psi^{-j} \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} d\psi \\
&= \left[\frac{1}{-j+s}\frac{t^s e^{-t\psi}\psi^{-j+s}}{\Gamma(s)}\right]_x^\infty + \frac{t}{-j+s}\int_x^\infty \psi^{-(j-1)} \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} d\psi \\
&= \frac{1}{j-s}\frac{t^s e^{-tx} x^{-j+s}}{\Gamma(s)} - \frac{t}{j-s} E(\psi^{-(j-1)}\big|\psi > x; s,t)
\end{aligned}$$

The above collapses if s is an integer less than or equal to j, when the iteration will involve the integral:

$$\begin{aligned}
E(\psi^{-s}\big|\psi > x; s,t) &= \int_x^\infty \psi^{-s} \frac{t^s e^{-t\psi}\psi^{s-1}}{\Gamma(s)} d\psi \\
&= \frac{t^s}{\Gamma(s)} \int_x^\infty \frac{e^{-t\psi}}{\psi} d\psi
\end{aligned}$$

In this case, substituting $z = t\ln\psi$ results in the pdf of a Gumbel distribution, which can be evaluated, and therefore the iteration should start at $j = s$ rather than $j = 0$.

Let us denote $\Omega_j(x;\, s,\, t)$ for $E(\psi^{j}|\psi > x;\, s,\, t)$ defined iteratively as above for integer $j > 0$. So we can state:

$$E(\psi^{-j+u} e^{-v\psi}\big|\psi > x; s,t) = \frac{\Gamma(s+u)t^s}{\Gamma(s)(t+v)^{s+u}} \Omega_j(x; s+u, t+v).$$

Theorem 5.8 $E[e_1(\Theta)^2]$ can be written as:

$$\begin{aligned}
&E[e_1(\Theta)^2] \\
&= \left(\frac{a(\lambda+D)}{b}\right)^2 \frac{t^s}{\Gamma(s)} \sum_{n=1}^{\infty} \left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!} \sum_{r=0}^{n-1} \binom{n-1}{r} (-1)^r \\
&\times \left\{ \sum_{j=0}^{\infty} \Gamma(1;s,t)\Gamma(s+j+r)\left(-\frac{\Gamma[1; s+j+r, t+2\ln\lambda_D]}{[t+2\ln\lambda_D]^{s+j+r}}\right)\right. \\
&+ \sum_{j=0}^{\infty} \Gamma(1;s,t)\Gamma(s+j+r)\left(\frac{\lambda+U}{\lambda+D}\right)\frac{\Gamma[1; s+j+r, t+\ln\lambda_D+\ln\lambda_U]}{[t+\ln\lambda_D+\ln\lambda_U]^{s+j+r}} \\
&+ \sum_{j=1}^{\infty} [1-\Gamma(1;s,t)]\Gamma(s+r)\frac{\Omega_j[1; s+r, t+2\ln\lambda_D]}{[t+2\ln\lambda_D]^{s+r}} \\
&\left. - \sum_{j=1}^{\infty} [1-\Gamma(1;s,t)]\Gamma(s+r)\left(\frac{\lambda+U}{\lambda+D}\right)\frac{\Omega_j[1; s+r, t+\ln\lambda_D+\ln\lambda_U]}{[t+\ln\lambda_D+\ln\lambda_U]^{s+r}} \right\}
\end{aligned}$$

where $\lambda_D = (\lambda + D)/\lambda$ and $\lambda_U = (\lambda + U)/\lambda$.

Proof From Lemma 5.7:

$$E[e_1(\Theta)^2] = \left(\frac{a(\lambda+D)}{b}\right)^2 E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi} \frac{1}{(\psi-1)^2}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)^2\right]$$

$$= \left(\frac{a(\lambda+D)}{b}\right)^2 \left\{E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi} \frac{1}{(\psi-1)^2}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\right]\right.$$

$$\left. - E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1} \frac{1}{(\psi-1)^2}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\right]\right\}$$

For $\frac{1}{(\psi-1)^2}$, we need to factorise the expression as $\frac{1}{\psi-1}\cdot\frac{1}{\psi-1}$ and expand one of the factors. To do this, we need to consider the cases $\psi < 1$ and $\psi > 1$ separately.

If $\psi < 1$: $\frac{1}{\psi-1} = -\sum_{j=0}^{\infty}\psi^j$

and if $\psi > 1$: $\frac{1}{\psi-1} = \frac{1}{\psi}\cdot\frac{1}{1-\frac{1}{\psi}} = \frac{1}{\psi}\sum_{j=0}^{\infty}\psi^{-j} = \sum_{j=1}^{\infty}\psi^{-j}$.

Conditioning on the value of ψ:

$$E[e_1(\Theta)^2]$$

$$= \left(\frac{a(\lambda+D)}{b}\right)^2 \Gamma(1;s,t)\sum_{j=0}^{\infty}\left\{-E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^j\frac{1}{\psi-1}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\middle|\psi<1\right]\right.$$

$$\left.+E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\psi^j\frac{1}{\psi-1}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\middle|\psi<1\right]\right\}$$

$$+\left(\frac{a(\lambda+D)}{b}\right)^2[1-\Gamma(1;s,t)]\sum_{j=1}^{\infty}\left\{E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^{-j}\frac{1}{\psi-1}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\middle|\psi>1\right]\right.$$

$$\left.-E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\psi^{-j}\frac{1}{\psi-1}\left(1-\left(\frac{\lambda+D}{\lambda+U}\right)^{\psi-1}\right)\middle|\psi>1\right]\right\}$$

We can expand $[(\lambda+D)/(\lambda+U)]^{(\psi-1)}$ as an exponential:

$$\left(\frac{\lambda+D}{\lambda+U}\right)^{(\psi-1)} = e^{(\psi-1)\ln\left(\frac{\lambda+D}{\lambda+U}\right)} = \sum_{n=0}^{\infty}\left[\ln\left(\frac{\lambda+D}{\lambda+U}\right)\right]^n\frac{(\psi-1)^n}{n!}$$

$$= 1-(\psi-1)\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{(1-\psi)^{n-1}}{n!}$$

Substituting in this expansion gives:

$$E[e_1(\Theta)^2]$$
$$=\left(\frac{a(\lambda+D)}{b}\right)^2\Gamma(1;s,t)\sum_{j=0}^{\infty}\left\{-E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^j\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{(1-\psi)^{n-1}}{n!}\middle|\psi<1\right]\right.$$
$$\left.+\left(\frac{\lambda+U}{\lambda+D}\right)E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\left(\frac{\lambda}{\lambda+U}\right)^{\psi}\psi^j\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{(1-\psi)^{n-1}}{n!}\middle|\psi<1\right]\right\}$$
$$+\left(\frac{a(\lambda+D)}{b}\right)^2[1-\Gamma(1;s,t)]\sum_{j=1}^{\infty}\left\{E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^{-j}\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{(1-\psi)^{n-1}}{n!}\middle|\psi>1\right]\right.$$
$$\left.-\left(\frac{\lambda+U}{\lambda+D}\right)E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\left(\frac{\lambda}{\lambda+U}\right)^{\psi}\psi^{-j}\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{(1-\psi)^{n-1}}{n!}\middle|\psi>1\right]\right\}$$

Expanding out $(1-\psi)^{n-1}$:

$$E[e_1(\Theta)^2]$$
$$=\left(\frac{a(\lambda+D)}{b}\right)^2\Gamma(1;s,t)\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r}(-1)^r$$
$$\times\sum_{j=0}^{\infty}\left\{-E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^{j+r}\middle|\psi<1\right]+\left(\frac{\lambda+U}{\lambda+D}\right)E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\left(\frac{\lambda}{\lambda+U}\right)^{\psi}\psi^{j+r}\middle|\psi<1\right]\right\}$$
$$+\left(\frac{a(\lambda+D)}{b}\right)^2[1-\Gamma(1;s,t)]\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r}(-1)^r$$
$$\times\sum_{j=1}^{\infty}\left\{E\left[\left(\frac{\lambda}{\lambda+D}\right)^{2\psi}\psi^{-j+r}\middle|\psi>1\right]-\left(\frac{\lambda+U}{\lambda+D}\right)E\left[\left(\frac{\lambda}{\lambda+D}\right)^{\psi}\left(\frac{\lambda}{\lambda+U}\right)^{\psi}\psi^{-j+r}\middle|\psi>1\right]\right\}$$

ψ is Gamma distributed with parameters (s, t), so:

$$E(\psi^u e^{-v\psi}\,|\,\psi<x)=\frac{\Gamma(s+u)t^s}{\Gamma(s)(t+v)^{s+u}}\Gamma(x;s+u,t+v)$$

provided that $s+u>0$ and $t+v>0$.

So, using $\Omega_j(x; s, t)$ defined as before and with $\lambda_D = (\lambda + D)/\lambda$ and $\lambda_U = (\lambda + U)/\lambda$:

$$E[e_1(\Theta)^2]$$
$$= \left(\frac{a(\lambda+D)}{b}\right)^2 \Gamma(1;s,t)\frac{t^s}{\Gamma(s)}\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n \frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r}(-1)^r\sum_{j=0}^{\infty}\left\{\Gamma(s+j+r)\right.$$
$$\left.\times\left(-\frac{\Gamma[1;s+j+r,t+2\ln\lambda_D]}{[t+2\ln\lambda_D]^{s+j+r}}+\left(\frac{\lambda+U}{\lambda+D}\right)\frac{\Gamma[1;s+j+r,t+\ln\lambda_D+\ln\lambda_U]}{[t+\ln\lambda_D+\ln\lambda_U]^{s+j+r}}\right)\right\}$$
$$+\left(\frac{a(\lambda+D)}{b}\right)^2[1-\Gamma(1;s,t)]\frac{\Gamma(s+r)t^s}{\Gamma(s)}\sum_{n=1}^{\infty}\left[\ln\left(\frac{\lambda+U}{\lambda+D}\right)\right]^n\frac{1}{n!}\sum_{r=0}^{n-1}\binom{n-1}{r}(-1)^r$$
$$\times\sum_{j=1}^{\infty}\left(\frac{\Omega_j[1;s+r,t+2\ln\lambda_D]}{[t+2\ln\lambda_D]^{s+r}}-\left(\frac{\lambda+U}{\lambda+D}\right)\frac{\Omega_j[1;s+r,t+\ln\lambda_D+\ln\lambda_U]}{[t+\ln\lambda_D+\ln\lambda_U]^{s+r}}\right)$$

which can be rearranged as required.

Now we have all the elements to a credibility weighted rate. Recall from Theorem 3.1 that the best linear estimate under quadratic loss is the Bayesian credibility estimate $Z\overline{X}+(1-Z)E[\mu(\theta)]$, where

$$Z=\frac{k}{k+\rho},\rho=\frac{E[\sigma^2(\theta)]}{Var[\mu(\theta)]}.$$

In terms of $e_1(\Theta)$ and $e_2(\Theta)$, the first two central moments of Θ, this is equivalent to $Z\overline{X}+(1-Z)E[e_1(\Theta)]$, where

$$Z=\frac{k}{k+\rho},\rho=\frac{E[e_2(\Theta)]}{E[e_1(\Theta)^2]-E[e_1(\Theta)]^2}.$$

$E[e_1(\Theta)]$ and $E[e_2(\Theta)]$ are set out in Theorem 5.6 and $E[e_1(\Theta)^2]$ is set out in Theorem 5.8.

6. Practical considerations

We have an exact formulation to calculate the Bayesian credibility estimate from exposure and experience rates for an excess of loss layer. It is not simple or elegant and, with five separate infinite series summations for $E[e_1(\Theta)^2]$ alone, it can never be computed exactly. However, it does allow significant precision using a macro, calculating to as many terms as needed. In practice, the macro can be stopped after about 10 terms for each summation with no significant loss of precision.

However, in the real world there are many instances when exposure and experience methods do interact already, blurring the credibility weighting, including:

- The exposure curve has been calibrated from revalued historic claims, which may include claims data from the contract being rated.

- One way of rating higher layers is by applying ILF factors from a lower layer that has been rated using experience methods, so involving a combination of experience and exposure methods.

- The experience methods may use an initial estimate of the loss ratio from an exposure method.

- The exposure method may adjust the rate to allow for an estimate of the insurer's profitability based on a reserving exercise.

Anyone using the credibility weighting factors set out here will need to keep these practical considerations in mind.

References

Buhlmann, Hans; "*Experience Rating and Credibility*"; ASTIN Bulletin, Volume 4.3 (1967)

Hesselager, Ole; "*A Class of Conjugate Priors with Applications to Excess-of-Loss Reinsurance*"; ASTIN Bulletin, Volume 23.1 (1993)

Goovaerts, M.J., de Vylder, F. and Haezendonck, J.; "*Insurance Premiums*"; North-Holland, Amsterdam-New York (1984)

Patrik, Gary and Mashitz, Isaac; "*Credibility for Treaty Reinsurance Excess Pricing*"; CAS Discussion Papers (1990)