

ON A MODEL FOR THE CLAIM NUMBER PROCESS*

BY MATTI RUOHONEN

The Sampo Group, Turku, Finland

ABSTRACT

A model for the claim number process is considered. The claim number process is assumed to be a weighted Poisson process with a three-parameter gamma distribution as the structure function. Fitting of this model to several data encountered in the literature is considered, and the model is compared with the two-parameter gamma model giving the negative binomial distribution. Some credibility theory formulae are also presented.

KEYWORDS

Claim number process; weighted Poisson; three-parameter gamma distribution.

1. INTRODUCTION

In this note we consider a model for the claim number process. Our model is a weighted Poisson process with a three-parameter gamma distribution as a structure function. This has been considered earlier by DELAPORTE (1960), see also KUPPER (1962). This is equivalent to the fact that the claim number process consists of two independent component processes, a Poisson process and a negative binomial process. The Poisson component may be thought of as the common part for all risks, and the negative binomial component as the individual contribution of a particular risk. This means that we can write the number of claims in time t , N_t as the sum of two components,

$$N_t = N_{1t} + N_{2t},$$

where N_{1t} has a Poisson distribution with the expected value γt , say, and N_{2t} has the negative binomial distribution. We consider here the fitting of our model to real data using the method of moments and the maximum likelihood estimation. Unfortunately the maximum likelihood estimators for the parameters cannot be obtained in a closed form. Hence, they are calculated via maximization of the likelihood function numerically.

We test the hypothesis $H_0: \gamma = 0$ against the one-sided alternative $H_1: \gamma > 0$. This tests the existence of the Poisson component in the model. We derive also some credibility theory formulae for our model. The corresponding formulae for

* This paper was presented to the ASTIN Colloquium at Scheveningen, the Netherlands, September 1987

the two parameter model can be found in SEAL (1969). The flavour our model gives to credibility considerations is the fact that even the best claim history, i.e. no claims at all, does not lead to zero premium in the limit. This is due to the existence of the background intensity which gives rise to the Poisson process.

2. DEFINITION OF THE MODEL

We assume that the claim number process $N_t, t \geq 0$, is a weighted Poisson process, i.e. if the claim intensity is Λ , then the conditional process $(N_t | \Lambda)_{t \geq 0}$ is a Poisson process. If the intensity Λ has the distribution function U , then

$$(1) \quad p_n(t) = P(N_t = n) = \int_0^\infty \frac{(\lambda t)^n e^{-\lambda t}}{n!} dU(\lambda).$$

We now assume that

$$(2) \quad dU(\lambda) = (\lambda - \gamma)^{\alpha-1} \beta^\alpha e^{-(\lambda-\gamma)\beta} / \Gamma(\alpha),$$

when $\lambda \geq \gamma$, and zero otherwise, with positive α, β and γ . This amounts to the fact that Λ has the three-parameter gamma distribution $\Gamma(\alpha, \beta, \gamma)$, see JOHNSON and KOTZ (1969). From (2) it follows that the intensity has a strictly positive lower bound γ . By substituting (2) into (1) we obtain

$$(3) \quad p_n(t) = \sum_{k=0}^n \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)k!} \left(\frac{\beta}{t+\beta}\right)^\alpha \left(\frac{t}{t+\beta}\right)^k \frac{(\gamma t)^{n-k} e^{-\gamma t}}{(n-k)!}.$$

Formula (3) exhibits $p_n(t)$ as the convolution of a negative binomial and a Poisson distribution.

From this or directly from (2) we may observe that the intensity Λ can be written as the sum $\Lambda = \gamma + \Lambda_1$, where γ is a positive real number, and Λ_1 has the usual two-parameter gamma distribution $\Gamma(\alpha, \beta)$. The interpretation of these components is

γ = background Poisson intensity which is common for all risks

Λ_1 = additional individual intensity that varies from one risk to another

With this interpretation we can assume that the process N_t itself consists of two mutually independent component processes N_{1t} and N_{2t} , where N_{1t} is a Poisson process with intensity γ and N_{2t} is a weighted Poisson process whose intensity Λ_1 has the distribution $\Gamma(\alpha, \beta)$. Then

$$(4) \quad N_t = N_{1t} + N_{2t},$$

where $N_{1t} \sim PO(\gamma t)$ and $N_{2t} \sim NB(\alpha, \beta/(t+\beta))$. Here \sim stands for "obeys the distribution", PO means the Poisson distribution and NB means the negative binomial distribution.

The moments of N_t may be obtained from the theory of doubly stochastic Poisson processes. The stochastic intensity Λ has the moments

$$E\Lambda = \frac{\alpha}{\beta} + \gamma, \quad \text{Var}(\Lambda) = \alpha/\beta^2, \quad E((\Lambda - E\Lambda)^3) = 2\alpha/\beta^3.$$

With the help of the moments of Λ the moments of N_t can be written as

$$\begin{aligned} EN_t &= tE\Lambda \\ \text{Var}(N_t) &= t^2 \text{Var}(\Lambda) + tE\Lambda \\ E((N_t - EN_t)^3) &= t^3 E((\Lambda - E\Lambda)^3) + 3t^2 \text{Var}(\Lambda) + tE\Lambda, \end{aligned}$$

(see SNYDER 1975). By substitution we then obtain

$$(5) \quad \begin{aligned} EN_t &= (\alpha/\beta + \gamma)t \\ \text{Var}(N_t) &= (\alpha/\beta^2)t^2 + (\alpha/\beta + \gamma)t \\ E((N_t - EN_t)^3) &= (2\alpha/\beta^3)t^3 + (3\alpha/\beta^2)t^2 + (\alpha/\beta + \gamma)t. \end{aligned}$$

These could have also been obtained by using the representation (4).

3 FITTING THE DISTRIBUTION

We say that a parameter vector (α, β, γ) is feasible if all the components are positive. Analogously we say that an estimator is feasible if all three components are positive. We consider here three alternatives for fitting the distribution (3) to data. For convenience we take $t = 1$.

Method 1

We consider first the method of moments. Let the first three sample moments be \bar{x} (the sample mean), s^2 (the sample variance) and \bar{x}_3 (the third central sample moment), the two latter calculated with weights $1/(n-1)$. Equating these with the population moments (5) we obtain

$$(6) \quad \begin{aligned} \hat{\beta} &= 2(s^2 - \bar{x})/(\bar{x}_3 - 3s^2 + 2\bar{x}), \\ \hat{\alpha} &= (s^2 - \bar{x})\hat{\beta}^2, \\ \hat{\gamma} &= \bar{x} - \hat{\alpha}/\hat{\beta} \end{aligned}$$

Necessary and sufficient conditions for the feasibility are

$$s^2 > \bar{x}, \quad \bar{x}_3 > 2s^4/\bar{x} - s^2$$

The first condition implies that the sample variance has to be larger than the sample mean. This is due to the presence of the negative binomial part in the model. The Poisson part gives equal variance and mean value. The second condition means that the distribution has a larger third central sample moment than a NB-distribution with the same first two moments.

Method 2

Because the use of the third moment in estimation may give undue weight on the tail we consider here a variant of the method of moments. The idea is to fit \bar{x} , s^2 and p_0 , the relative frequency of the zero class. Then we have to solve the

system of equations

$$(7) \quad \begin{aligned} \alpha/\beta + \gamma &= \bar{x} \\ \alpha/\beta + \gamma + \alpha/\beta^2 &= s^2 \\ \left(\frac{\beta}{1+\beta}\right)^\alpha e^{-\gamma} &= p_0. \end{aligned}$$

This leads to the solution

$$(8) \quad \hat{\alpha} = (\bar{x} - \hat{\gamma})^2 / (s^2 - \hat{x}), \quad \hat{\beta} = (\bar{x} - \hat{\gamma}) / (s^2 - \bar{x}),$$

with $\hat{\gamma}$ being the solution of the equation

$$(9) \quad \gamma = -\ln p_0 + \frac{(\bar{x} - \gamma)^2}{s^2 - \bar{x}} \ln \frac{\bar{x} - \gamma}{s^2 - \gamma}$$

The solution given in (8) and (9) is feasible if $\hat{\gamma}$ lies in the open interval $(0, \bar{x})$ and $s^2 > \bar{x}$. We consider next the necessary and sufficient conditions for the existence of a unique solution of (9) in this interval. For this purpose, denote

$$f(\gamma) = \gamma + \ln p_0 + \frac{(\bar{x} - \gamma)^2}{s^2 - \bar{x}} \ln \left(1 + \frac{s^2 - \bar{x}}{\bar{x} - \gamma} \right).$$

The solution of (9) is then equivalent to the solution of the equation $f(\gamma) = 0$

Now we have

$$f(0) = \ln p_0 + (\bar{x}^2 / (s^2 - \bar{x})) \ln(s^2 / \bar{x})$$

and

$$f(\bar{x}) = \bar{x} + \ln p_0$$

We also have

$$f'(\gamma) = 1 - \frac{2(\bar{x} - \gamma)}{s^2 - \bar{x}} \ln \left(1 + \frac{s^2 - \bar{x}}{\bar{x} - \gamma} \right) + \left(1 + \frac{s^2 - \bar{x}}{\bar{x} - \gamma} \right)^{-1}.$$

If we denote $y = (s^2 - \bar{x}) / (\bar{x} - \gamma)$, $h(y) = yf'(\gamma)$, then

$$h(y) = (2y + y^2) / (1 + y) - 2 \ln(1 + y)$$

From this it is easy to see that $h(0) = 0$ and $h'(y) > 0$, when $y > 0$. But this means that, if $s^2 > \bar{x}$, then $f'(\gamma) > 0$ for $0 < \hat{\gamma} < \bar{x}$. Because the condition $s^2 > \bar{x}$ is also necessary for $\hat{\alpha} > 0$, we have that the conditions

$$s^2 > \bar{x}, \quad -\bar{x} < \ln p_0 < (-\bar{x}^2 / (s^2 - \bar{x})) \ln(s^2 / \bar{x})$$

are necessary and sufficient for the existence of a unique feasible solution. These mean that the zero class probability must lie between those of a Poisson distribution and a negative binomial distribution with due first moments.

Method 3

Let us assume that we have the data n_0, n_1, \dots, n_k , where n_i is the number of risks

having had J claims in unit time. The maximum likelihood method gives us the estimator $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma})$ which maximizes the likelihood function

$$\begin{aligned} L(\alpha, \beta, \gamma) &= \ln \prod_{j=0}^k (p_j(1))^{n_j} \\ &= \sum_{j=0}^k n_j \ln p_j(1) \\ &= \sum_{j=0}^k n_j \left\{ \alpha \ln \frac{\beta}{1+\beta} - \gamma + \ln \left(\sum_{i=0}^j \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{\gamma^{j-i}}{i!(j-i)!(1+\beta)^i} \right) \right\} \\ &= n\alpha \ln \frac{\beta}{1+\beta} - n\gamma \\ &\quad + \sum_{j=0}^k n_j \ln \left(\gamma^j \sum_{i=0}^j \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{1}{i!(j-i)!(\gamma(1+\beta))^i} \right), \end{aligned}$$

where $n = n_0 + \dots + n_k$ is the total number of observed risks. To facilitate the maximization we denote $\eta = \gamma(1 + \beta)$, and substitute $(\eta - \gamma)/\gamma$ for β in L . Then the new likelihood function is

$$\tilde{L}(\alpha, \eta, \gamma) = n\alpha \ln \frac{\eta - \gamma}{\eta} - n\gamma + n\bar{x} \ln(\gamma) + \sum_{j=0}^k n_j \ln \left(\sum_{i=0}^j \frac{\Gamma(i+\alpha)}{i!(j-i)!\Gamma(\alpha)\eta^i} \right).$$

If we put the derivative with respect to γ equal to zero we get the equation

$$(10) \quad -n\alpha/(\eta - \gamma) - n + n\bar{x}/\gamma = 0,$$

or equivalently

$$\bar{x} = \gamma + \alpha/\beta.$$

In order to handle the partial derivatives with respect to α and η we denote

$$w_j(\alpha, \eta) = \sum_{i=1}^j \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{1}{i!(j-i)!\eta^i},$$

for which

$$\frac{\partial}{\partial \alpha} w_j = \frac{1}{(j-1)!\eta} + \sum_{i=2}^j \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{\sum_{m=1}^i \prod_{m \neq i} (\alpha + m - 1)}{i!(j-i)!\eta^i},$$

and

$$\frac{\partial}{\partial \eta} w_j = \sum_{i=1}^j \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{(-i)}{i!(j-i)!\eta^{i+1}}.$$

With the help of these we have

$$\tilde{L}(\alpha, \eta, \gamma) = n\alpha \ln \frac{\eta - \gamma}{\eta} - n\gamma + n\bar{x} \ln(\gamma) + \sum_{j=0}^k n_j \ln(w_j(\alpha, \eta)),$$

and

$$(11) \quad \begin{aligned} \frac{\partial}{\partial \eta} \tilde{L} &= n\alpha((\eta - \gamma)^{-1} - \eta^{-1}) + \sum_{j=0}^k n_j \frac{\partial}{\partial \eta} w_j(\alpha, \eta)(w_j(\alpha, \eta))^{-1} \\ \frac{\partial}{\partial \alpha} \tilde{L} &= n \ln((\eta - \gamma)/\eta) + \sum_{j=0}^k n_j \frac{\partial}{\partial \alpha} w_j(\alpha, \eta)(w_j(\alpha, \eta))^{-1}. \end{aligned}$$

Because of (10) our three-dimensional maximization problem has been reduced to a two-dimensional one. This problem may be solved using an optimization method, which makes use of the gradient given in (11).

4. TESTING THE MODEL

After having fitted the model using the maximum likelihood method we can naturally test the goodness of fit of the model using a χ^2 -test.

If we have a good fit, there lies the question whether γ differs from zero significantly. The case $\gamma = 0$ corresponds to the pure negative binomial distribution, i.e., the Poisson background is absent. We need to test the null hypothesis $H_0: \gamma = 0$ against the alternative $H_1: \gamma > 0$. Under the null hypothesis the number of claims has the negative binomial distribution. This distribution is fitted to the data using the maximum likelihood method. Description of this method for negative binomial distribution can be found for example in JOHNSON and KOTZ (1969). This gives us the estimator $(\hat{\alpha}, \hat{\beta})$. If we denote by \hat{p}_i and \tilde{p}_i the class i probabilities given by the estimators $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ and $(\hat{\alpha}, \hat{\beta})$, respectively, then we can form the test variable

$$(12) \quad Y = -2 \sum_{i=0}^k n_i \ln(\hat{p}_i/\tilde{p}_i).$$

For the conditions under which a likelihood ratio has the $\chi^2(1)$ -distribution as its asymptotic distribution we refer to RAO (1973). In our case the value $\gamma = 0$ lies on the boundary of the parameter space. Hence, the asymptotic distribution is not $\chi^2(1)$ but a 50:50 mixture of $\chi^2(1)$ and a distribution degenerate at origin, as has been shown by SELF and LIANG (1987). This means that if we choose the significance level ε , the critical value will be the $(1 - 2\varepsilon)$ -fractile of the $\chi^2(1)$ distribution. The other conditions given by Rao are met by our distribution but the positive-definiteness of the information matrix. The verification of this fact seems to be a hopeless task in general. We have only shown that the determinant of the information matrix becomes zero when α and β tend to infinity with their ratio constant. This means that the results of our tests become unreliable as α or β becomes large. We have also verified numerically that the information matrix is positive definite when $\alpha = 1$ and β is finite. The applicability of our test is not rigorously verified, and the tests to be performed later are only of guiding nature.

5. CREDIBILITY

We now look at what some credibility theory formulae look like for our model. We denote

$$p_{l|n}(s | t) = P(N_{t+s} - N_t = l | N_t = n),$$

the conditional probability of l claims in time s after having had n claims in time t . Now we have

$$p_{l|n}(s | t) = \binom{l+n}{n} \left(\frac{t}{t+s}\right)^n \left(\frac{s}{t+s}\right)^l p_{l+n}(t+s) / p_n(t),$$

(see SEAL, 1969 p. 27). For example the probability of no claims after having had no claims in time t is

$$p_{0|0}(s | t) = \left(\frac{\beta+t}{\beta+t+s}\right)^\alpha e^{-\gamma s}.$$

The conditional expectation of the intensity Λ after n claims in time t is

$$\begin{aligned} E(\Lambda | n, t) &= \frac{n+1}{t} \frac{p_{n+1}(t)}{p_n(t)} \\ &= \frac{n+1}{\beta+t} \frac{\sum_{k=0}^{n+1} \Gamma(k+\alpha)(\gamma(\beta+t))^{n+1-k} / ((n+1-k)!k!)}{\sum_{k=0}^n \Gamma(k+\alpha)(\gamma(\beta+t))^{n-k} / ((n-k)!k!)}. \end{aligned}$$

Further the conditional density of Λ after n claims in time t can after some manipulation be written as

$$dU(\Lambda | n, t) = \frac{(\beta+t)^\alpha (\lambda-\gamma)^{\alpha-1} e^{-(\lambda-\gamma)(\beta+t)}}{\Gamma(\alpha)} \frac{(\lambda t)^\alpha p_0(t)}{n! p_n(t)} d\lambda,$$

for $\lambda > \gamma$. The first factor here is the density function of the distribution $\Gamma(\alpha, \beta+t, \gamma)$. Especially after claim-free time t we have

$$(\Lambda | N_t = 0) \sim \Gamma(\alpha, \beta+t, \gamma)$$

so that

$$\begin{aligned} E(N_{t+s} - N_t | N_t = 0) &= (\alpha / (\beta+t) + \gamma) s \\ \text{Var}(N_{t+s} - N_t | N_t = 0) &= \alpha s^2 / (\beta+t)^2 + (\alpha / (\beta+t) + \gamma) s. \end{aligned}$$

Further, if we let t tend to infinity, then

$$\begin{aligned} E(N_{t+s} - N_t | N_t = 0) &\rightarrow \gamma s \\ \text{Var}(N_{t+s} - N_t | N_t = 0) &\rightarrow \gamma s \end{aligned}$$

Equivalently we can write that

$$\begin{aligned} E(\Lambda | N_t = 0) &= \alpha / (\beta+t) + \gamma \rightarrow \gamma \\ \text{Var}(\Lambda | N_t = 0) &= \alpha / (\beta+t)^2 \rightarrow 0, \end{aligned}$$

as $t \uparrow \infty$. This means that $(\Lambda | N_t = 0) \rightarrow \gamma$ in probability, so that a claimless risk will approach a risk with pure Poisson claim process. This means also that the credibility premium would converge to γ and not to zero. Similar to the preceding results, various results concerning the bonus class systems can be presented in a computable form in our case.

6 FITTING THE MODEL TO REAL DATA

In this section we consider the fitting of our model to some data that can be found in the actuarial literature. We calculate the maximum likelihood estimates for α and β in the case when $\gamma = 0$, and for α , β and γ in the general case. To get started we solve γ from (9) using $\gamma = \bar{x}/2$ as the first guess. Then we use this γ together with α and β obtained from (8) as the initial guess for the calculation of the maximum likelihood estimation. These estimates were computed using the Davidon–Fletcher–Powell method, see RAO (1978). Also (12) we compute in order to perform the likelihood ratio test.

Our first fit is to the TROBLIGER (1961) data. Trobliger fitted to his data a model in which the risks were classified into two classes “the good” and “the bad”. The fit was good with $\chi^2(1) = 0.44$. These data give $\bar{x} = 0.14421976$, $s^2 = 0.1638699$ and $p_0 = 0.872949$. If the negative binomial distribution is fitted, then $\hat{\alpha} = 1.117895$, $\hat{\beta} = 7.751332$, and if our model is fitted, then $\hat{\alpha} = 0.2766328$, $\hat{\beta} = 3.7597937$ and $\hat{\gamma} = 0.07064318$. The frequencies of different classes for our model and the negative binomial distribution together with the observed frequencies are given in Table 1.

If the three last classes and the class “ ≥ 7 ” are joined together, the $\chi^2(1)$ -value for goodness of fit test of our model is 0.0042. This extremely low value is due to the fact that three parameters were fitted. The likelihood ratio test has now the $\chi^2(1)$ -value 3.93 which exceeds the critical value 2.706 at the 0,95-level. Hence, the hypothesis $H_0: \gamma = 0$ is rejected. We now have the estimate 0.071 for the background intensity. This may be compared with the mean intensity $\bar{x} = 0.144$ and the “good” intensity 0.109 in Trobliger’s model. The estimated background intensity is 49% of the estimated mean intensity and 66% of the estimated “good” value.

WILLMOT (1988) has fitted an extended negative binomial distribution to this data. The χ^2 value was 0.0282 which indicates a very good fit.

TABLE 1

No. of claims	Observed	Our model	NB
0	20592	20591.87	20596.76
1	2651	2651.45	2631.03
2	297	296.42	318.37
3	41	41.12	37.81
4	7	6.70	4.45
5	0	1.18	0.52
6	1	0.21	0.06

We look also at another example a little closer. THYRION (1960) fitted also a three-parameter model of weighted Poisson type. This model has a reasonable fit. The estimation was not maximum likelihood, and so no χ^2 -test is available. The estimated parameters are $\bar{x} = 0.2143537$, $s^2 = 0.2889314$ and $p_0 = 0.82866505$. The estimated negative binomial parameters are $\tilde{\alpha} = 0.7015122$ and $\tilde{\beta} = 3.2726858$. The estimated parameters of our model are $\tilde{\alpha} = 0.2006137$, $\tilde{\beta} = 1.6665135$ and $\tilde{\gamma} = 0.09397439$. The calculated and observed frequencies are collected in Table 2.

If the three last classes and the class " ≥ 8 " are joined together, the goodness of fit test for our model has the $\chi^2(2)$ value 4.12. This is below the 90%-value 4.605 so that our model cannot be rejected. The likelihood ratio test has the test-value 9.53, which exceeds even the 0.995-level. The hypothesis $H_0: \gamma = 0$ is then rejected. The estimator for the background intensity $\tilde{\gamma} = 0.094$ is about 44% of the estimated mean intensity \bar{x} .

We have considered several other data from traffic insurance. We shall review them here only briefly to save space. LEMAIRE (1979) gives data to which already the negative binomial distribution fits well. Hence the hypothesis $H_0: \gamma = 0$ is not rejected. In spite of this the maximum likelihood estimator for the background intensity is 40% of the estimated mean intensity \bar{x} . DELAPORTE (1962) gives data, which has the tail shorter than the fitted negative binomial distribution has. Hence, our model leads to a negative value for the background intensity, and cannot be fitted to this data. PESONEN (1962) has data to which already the negative binomial model fits well, and the hypothesis of zero background intensity is not rejected. Again, however, the estimated background intensity is a large percentage, 60%, of the estimated mean intensity \bar{x} . MUIF (1972) gives two sets of data, *A* and *B*. The data *A* lead to a similar situation as that of Delaporte, and the data *B* similar to those of Pesonen and Lemaire. Finally BUHLMANN (1970) gives data for which the null hypothesis of zero background intensity is rejected with a high χ^2 -value. On the other hand $\tilde{\gamma}$ is as low as $0.37 \bar{x}$. GOSSIEUX and LEMAIRE (1981) have also considered the same data and they have found that the best fit among four distributions was given by a mixture of two Poisson distributions.

As a conclusion we must admit that the model presented here is not a general solution to the problem of determining the claim number distribution. If the data have a long tail then this model is worth considering. If the tail is short then the

TABLE 2

No. of claims	Observed	Our model	NB
0	7840	7837.40	7847.01
1	1317	1326.16	1288.36
2	239	222.76	256.53
3	42	52.68	54.07
4	14	15.08	11.71
5	4	4.66	2.58
6	4	1.50	0.57
7	1	0.50	0.13

bad fit of the negative binomial distribution cannot be corrected using this model with positive γ . However, the knowledge we have of fitting this model indicates that in most of the cases the background intensity is somewhere around the half of the mean, approximately between $0.4\bar{x}$ and $0.6\bar{x}$. Additionally this model can be used to build up a bonus-malus system with some definite lower boundary for the premium.

7 ADDITIONAL TOPICS

Several Years' Data

Let the same portfolio be observed during a period of several years. Let us assume that our model is the true one. Let the α_t , β_t and γ_t be the parameters α , β and γ , if t is selected to be the time unit. Equating the first three moments for the number of claims in time t calculated using time units 1 and t , respectively, we obtain

$$\alpha_t = \alpha_1, \quad \beta_t = \beta_1/t, \quad \gamma_t = t\gamma_1$$

This means that if our model is the true one, then the observed values of α_t , $t\beta_t$ and γ_t/t should be fairly constant during the observation period.

Two Portfolios

Let us join two portfolios which have the distribution (3) for the number of claims with parameters α_t , β_t and γ_t , $t = 1, 2$, respectively. Let the sizes of the portfolios be in ratio $p/(1-p)$. Let, further,

$$\chi = \begin{cases} 1, & \text{if the risk is from the portfolio 1} \\ 0, & \text{if the risk is from the portfolio 2.} \end{cases}$$

Then for a randomly chosen risk we have

$$\begin{aligned} N_t &= N_{1t}\chi + N_{2t}(1-\chi) = (N_{11t}\chi + N_{21t}(1-\chi)) \\ &\quad + (N_{12t}\chi + N_{22t}(1-\chi)) = \bar{N}_{1t} + \bar{N}_{2t}. \end{aligned}$$

where N_{jt} is the number of claims in time t in portfolio j due to the component j as in (4). Then N_t is divided into two components the first of which is a mixture of two Poisson distributions and the second a mixture of two negative binomial distributions. Hence, the combined portfolio no longer has the claim number distribution (3). In spite of this we tried this model for two composite data. We pooled Buhlmann's data with Trobliger's data, I, and then with Lemaire's data, II. The fit was excellent in both cases, and the null hypothesis of zero background intensity was rejected with great significance. The interesting feature is that the parameters obtained are close to those of Buhlmann's, and are not near the linear combinations of the original parameters. This can be seen in Table 3. For example the linear combination of the γ -parameters in the Buhlmann-Lemaire case would give 0.04887 against the obtained 0.05708.

TABLE 3

Data	\bar{x}	s^2	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\gamma}/\bar{x}$
Buhlmann	15514	17932	40015	4 068	05679	0 37
Trobliger	14422	16387	27663	3 760	07064	0 49
mixture I	15334	17679	37838	4 018	05918	0 39
Lemaire	10108	10745	58881	9 641	04001	0 40
mixture II	12965	14615	31966	4 405	05708	0 44

As a last example we joined together the data of Lemaire, Thysson, Pesonen, Trobliger and Buhlmann and considered how our model fits with these heterogeneous data. The fitted NB -distribution had a χ^2 (3)-value 61.14, which means poor fit. When our model was fitted, the χ^2 (2)-value was 5.18, which means a moderate fit. The likelihood ratio test value was 47.55 which is a highly significant value. The estimated background intensity was $\hat{\gamma} = 0.0654328$, which is 49% of the estimated mean.

A more detailed exposition of methods and results of this paper is found in a technical report RUOHONEN (1983).

ACKNOWLEDGEMENT

I should like to thank K. Loimaranta for many inspiring conversations and critical comments concerning this work. I should also like to thank Dr. B. Sundt and Dr. G. Willmot for pointing out errors in the original manuscript.

REFERENCES

- BUEHLMANN, H. (1970) *Mathematical Methods in Risk Theory*. Springer-Verlag, Berlin.
- DE LAPORTE, P. (1960) Un probleme de tarification de l'assurance accidents d'automobiles examiné par la statistique mathématique. *Transactions of the International Actuarial Congress, Subject B1*.
- DE LAPORTE, P. (1962) Sur l'efficacité des critères de tarification de l'assurance contre les accidents d'automobiles. *ASTIN Bulletin* 2 (1), 84-95.
- GOSSILUX, A. and LEMAIRE, J. (1981) Méthodes d'ajustement de distribution de sinistres. *Bulletin of the Association of Swiss Actuaries* 81, 87-95.
- JOHNSON, N. L. and KOTZ, S. (1969) *Discrete Distributions*. Houghton Mifflin, Boston.
- KLUPPER, J. (1962) Wahrscheinlichkeitstheoretische Modelle in der Schadenversicherung. *Bl. Deuts. Gesell. Versich. Math.* Band V, Heft 4, 451-503.
- LEMAIRE, J. (1979) How to define a bonus-malus system with an exponential utility function. *ASTIN Bulletin* 10 (3), 274-282.
- MUTH, M. (1972) The influence of the franchise on the number of claims in motor insurance. *ASTIN Bulletin* 6 (3), 191-194.
- PESONEN, E. (1962) A numerical method of finding a suitable bonus scale. *ASTIN Bulletin* 2(1), 102-108.
- RAO, C. R. (1973) *Linear Statistical Inference and its Applications*. Wiley, New York.
- RAO, S. S. (1978) *Optimization Theory and Applications*. Wiley Eastern, New Delhi.
- RUOHONEN, M. (1983) Fitting a convolution of Poisson and negative binomial distributions on data. *Rep. Inst. Appl. Math. Univ. Turku* No. 118.
- SRINIVASAN, H. L. (1969) *Stochastic Theory of a Risk Business*. Wiley, New York.

- SEITZ, S. G. and LIANG, K.-Y. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82** (398), 605–610
- SNYDER, D. L. (1975) *Random Point Processes*. Wiley, New York
- THYRION, P. (1960) Contribution a l'etude du bonus pour non sinistre en assurance automobile. *ASTIN Bulletin* **1** (3), 142–162
- TROBLIGER, A. (1961) Mathematische Untersuchungen zur Beitragsrückgewahr in der Kraftfahrversicherung. *Bf. Deuts. Gesell. Versich. Math.* **5**, 327–348
- WILFMOT, G. (1988) Sundt and Jewell's family of discrete distributions. *ASTIN Bulletin* **18** (1), 17–29

MATTI RUOHONEN

The Sampo Group, P.O. Box 216, SF-20101 Turku, Finland.