WORKING PAPERS ISSN 0781-4410

SUOMEN AKTUAARIYHDISTYS
The Actuarial Society of Finland

9

Matti Ruohonen

A MODEL FOR THE CLAIM NUMBER PROCESS
(1983)

A MODEL FOR THE CLAIM NUMBER PROCESS

by

Matti Ruohonen

Institute for Applied Mathematics

University of Turku, Turku, Finland

Abstract

A model for the claim number process is considered. The claim number process is assumed to be a weighed Poisson process with a three parameter gamma distribution as the structure function. Fitting of this model on 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.

1. Introduction

In this note we consider a model for the claim number process. Our model is a weighed Poisson process with three parameter gamma distribution as a structure function. 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 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 on 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 $H_o : \gamma = 0$ 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 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 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 weighed Poisson process, i.e., if the claim intensity is Λ, then the conditional process $(N_t \mid \Lambda)_{t \geq 0}$ is a Poisson process. If the intensity Λ has the distribution function U, then

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

We now assume that

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

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 (1970). From (2) it follows that the intensity has a strictly positive lower bound γ. By substituting (2) into (1) we obtain

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

From this or directly from (2) we may observe that the intensity Λ can be written as the sum $\Lambda = \gamma + \hat{\Lambda}$, where γ is a positive real number, and $\hat{\Lambda}$ 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

$\hat{\Lambda}$ = additional invidual 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 weighed Poisson process whose intensity $\hat{\Lambda}$ has the distibution $\Gamma(\alpha,\beta)$. Then

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

where $N_{1t} \sim Po(\gamma t)$ and $N_{2t} \sim NB(\alpha, \frac{\beta}{t+\beta})$. Here $\sim$ stands

we obtain

$$\hat{\beta} = 2(s^2 - \bar{x})/(\bar{x}_3 - 3s^2 + 2\bar{x}) \ ,$$

$$\hat{\alpha} = (s^2 - \bar{x})\hat{\beta}^2, \qquad (6)$$

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

Necessary and sufficient conditions for the feasibility are

$$s^2 \geqslant \bar{x}, \qquad \bar{x}_3 \geq 2s^4/\bar{x} - s^2.$$

The first condition implies that the variance has to be larger than the expected value. This is due to the presence of the negative binomial part in the model. The Poissson part gives equal variance and expected value. The second condition means that the distribution has larger third 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_o, the relative frequency of the zero class.Then we have to solve the system of equations

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

$$\alpha/\beta + \gamma + \alpha/\beta^2 = s^2 \qquad (7)$$

$$\left(\frac{\beta}{1+\beta}\right)^{\alpha} e^{-\gamma} = p_o .$$

This leads to the solution

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

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

$$\gamma = - \ln p_o + \frac{(\bar{x} - \gamma)^2}{s^2 - \bar{x}} \ln \frac{\bar{x} - \gamma}{s^2 - \gamma}. \qquad (9)$$

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

$$= \sum_{j=0}^{k} n_j \{\alpha \ln \frac{\beta}{1+\beta} - \gamma + \ln(\sum_{i=0}^{j} \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{\gamma^{j-i}}{i!(j-i)!(1+\beta)^i})\}$$

$$= n\alpha \ln \frac{\beta}{1+\beta} - n\gamma + \sum_{j=0}^{k} n_j \ln (\gamma^j \sum_{i=0}^{j} \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{1}{i!(j-i)!(\gamma(1+\beta))^i}),$$

where $n=n_o+\ldots+n_k$ 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

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

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

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

or equivalently

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

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

$$w_j(\alpha,\eta) = \sum_{i=0}^{j} \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)} \frac{1}{i!(j-i)!\eta^i},$$

for which

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

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

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

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

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

Now we have

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

and

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

We also have

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

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 $h'(y) \geqslant 0$, when $y \geqslant 0$. But this means that, if $s^2 \geqslant \bar{x}$, then $f'(\gamma) \geqslant 0$ for $0 \leqslant \hat{\gamma} \leqslant \bar{x}$. Because the condition $s^2 \geqslant \bar{x}$ is also necessary for $\hat{\alpha} \geqslant 0$, we have that the conditions

$$s^2 \geqslant \bar{x}, \quad -\bar{x} \leq p_o \leq (-\bar{x}^2/(s^2-\bar{x}))\ln(s^2/\bar{x}).$$

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.

<u>Method 3</u>. Let us assume that we have the data $n_0, n_1, \ldots, n_k$, where n_j 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

$$L(\alpha,\beta,\gamma) = \ln \prod_{j=0}^{k} (p_j(1))^{n_j} = \sum_{j=0}^{k} n_j \ln p_j(1) =$$

and

$$\frac{\partial}{\partial\eta}\tilde{L} = n\alpha((\eta-\gamma)^{-1} - \eta^{-1}) + \sum_{j=0}^{k} n_j \frac{\partial}{\partial\eta} w_j(\alpha,\eta)(w_j(\alpha,\eta))^{-1}$$

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

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 beackground 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 $(\bar{\alpha},\bar{\beta})$. If we denote by $\tilde{p}_i$ and $\bar{p}_i$ the class i probabilities given by the estimators $(\tilde{\alpha},\tilde{\beta},\tilde{\gamma})$ and $(\bar{\alpha},\bar{\beta})$, respectively, then we can form the test variable

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

For the conditions under which Y has the $\chi^2(1)$-distribution as its asymptotic distribution we refer to Rao (1973). 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 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 verified numerically that the information matrix is positive definite when $\alpha = 1$ and β is finite.

5. Credibility

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

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

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

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

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

$$E(\Lambda|n,t) = \frac{n+1}{t}\,\frac{p_{n+1}(t)}{p_n(t)} =$$

$$= \frac{n+1}{t}\,\frac{\sum_{k=0}^{n+1} \Gamma(k+\alpha)(\gamma(\beta+t))^{n+1-k}/(n+1-k)!}{\sum_{k=0}^{n} \Gamma(k+\alpha)\,(\gamma(\beta+t))^{n-k}/(n-k)!}.$$

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

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

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

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

so that

$$E(N_{t+s} - N_t \mid N_t = 0) = (\alpha/(\beta+t) + \gamma)\,s$$

$$Var(N_{t+s} - N_t \mid N_t = 0) = \alpha s^2/(\alpha+t)^2 + (\alpha/(\beta+t) + \gamma)\,s\ .$$

Further, if we let t tend to infinity, then

$$E(N_{t+s} - N_t \mid N_t = 0) \rightarrow \gamma s$$

$$Var(N_{t+s} - N_t \mid N_t = 0) \rightarrow \gamma s.$$

Equivalently we can write that

$$E(\Lambda \mid N_t = 0) = \alpha/(\beta+t) + \gamma \rightarrow \gamma$$

$$Var(\Lambda \mid N_t = 0) = \alpha/(\beta+t)^2 \rightarrow 0,$$

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 on real data

In this section we consider the fitting of our model on 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). The computations were performed by a SM-4 computer. Also (12) was computed in order to perform the likelihood ratio test.

Our first fit is on the Tröbliger (1961) data. Tröbliger fitted on 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 $\bar{\alpha} = 1.117895$, $\bar{\beta} = 7.751332$, and if our model is fitted, then $\tilde{\alpha} = 0.2766328$, $\tilde{\beta} = 3.7597937$ and $\tilde{\gamma} = 0.07064318$. The frequencies of different classes for our model and the negative binomial distribution together with the observed frequencies are given in the following table

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

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 3.84 at the 0.95-level. hence, the hypothesis H_o: $\gamma = 0$ 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 Tröbliger´s model. The estimated background intensity is 49% of the estimated mean intensity and 66% of the estimated "good" value.

We look also another example a little closer. Thyrion (1960) fitted also a three-parameter model of weighed 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 $p_o = 0.82866505$. The estimated negative binomial parameters are $\bar{\alpha} = 0.7015122$ and $\bar{\beta} = 3.2726858$. 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 the following table

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

If the three last classes and the class " $\geq$ 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 4.61 so that our model cannot be rejected. The likelihood ratio test has the $\chi^2(1)$-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 on which already the negative binomial distribution fits well. Hence the hypothesis $H_o: \gamma = 0$ 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 nebative binomial distribution has. Hence, our model leads to negative value for the background

intensity, and cannot be fitted on this data. Pesonen (1962) has data on 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 $\bar{x}$. Muff (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 Bühlmann give 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 $0.37\,\bar{x}$.

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 bad fit of negative binomial distribution cannot be corrected using this model with positive γ. However, the scarce 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, $t\beta_t$ anf 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_i, β_i and γ_i, $i=1,2$, 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

$$N_t = N_{1t}\chi + N_{2t}(1 - \chi) = (N_{11t}\chi + N_{21t}(1 - \chi)) +$$
$$+ (N_{12t}\chi + N_{22t}(1 - \chi)) = \bar{N}_{1t} + \bar{N}_{2t},$$

where N_{ijt} is the number of claims in time t in portfolio i 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 does not have any more the claim number distribution (3). In spite of this we tried this model for two composite data. We pooled Bühlmann's data with Tröbliger'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 Bühlmann's, and are not near the linear combinations of the original parameters. This can be seen in the following table

Table 3.

data	$\bar{x}$	s^2	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\gamma}/\bar{x}$
Bühlmann	.15514	.17932	.40015	4.068	.05679	0.37
Tröbliger	.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

For example the linear combinaison of the γ-parameters in the Bühlmann-Lemaire case would give 0.04887 against the obtained 0.05708.

As a last example we joined together the data of Lemaire, Thyrion, Pesonen, Tröbliger and Bühlmann and considered how our model fits with these heterogeneous data. The fitted NB-distribution had a $\chi^2(3)$-value 61.14, which means poor fit. When our model was fitted, the $\chi^2(2)$-value was 5.18, which means a moderate fit. The log likelihood was 47.55, which is a highly significant $\chi^2(1)$-value. The estimated background intensity was $\hat{\gamma}$ = 0.0654328, which is 49% of the estimated mean.

More detailed exposition of methods and results of this paper is found in a technical report Ruohonen (1983).

Acknowledgement

I would like to thank K. Loimaranta for many inspiring conversations and critical comments concerning this work.

References

Bühlmann, H.(1970). *Mathematical methods in risk theory.* Springer-Verlag, Berlin.

Delaporte, P.(1962). Sur l'efficacité des critères de tarification de l'assurance contre les accidents d'automobiles. *ASTIN Bulletin,* 2.1.,84-95

Johnson, N.I. and Kotz, S. (1969). *Discrete distributions.* Houghton Mifflin,Boston.

Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. *ASTIN Bulletin* 10.3. , 274-282.

Muff, M.(1972) The influence of the frnchise on the number of claims in motor insuransce. *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.

Seal, H.L.(1969). *Stochastic theory of a risk business.* Wiley, New York.

Snyder, D.L.(1975). *Stochastic point processes.* Wiley, New York.

Thyrion, P.(1961). Contribution à l'étude du bonus pour non sinistre en assurence automobile.*ASTIN Bulletin*1.3.,142-162.

Tröbliger, A.(1961). Mathematische Untersuchungen zur Beitragsrückgewähr in der Kraftfahrversicherung. *Bl.Deuts. Gesell.Versich.Math.*,5.,327-348.