A SIMPLE PARAMETRIC MODEL FOR RATING AUTOMOBILE INSURANCE OR ESTIMATING IBNR CLAIMS RESERVES

By Thomas Mack

Munich Re, Munich, FRG

Abstract

It is shown that there is a connection between rating in automobile insurance and the estimation of IBNR claims amounts because automobile insurance tariffs are mostly cross-classified by at least two variables (e.g. territory and driver class) and IBNR claims run-off triangles are always cross-classified by the two variables accident year and development year. Therefore, by translating the most well-known automobile rating methods into the claims reserving situation, some known and some unknown claims reserving methods are obtained For instance, the automobile rating method of Bailey and Simon produces a new claims reserving method, whereas the model leading to the rating method called "marginal totals" produces the well-known IBNR claims estimation method called "chain ladder". A drawback of this model is the fact that it is designed for the number of claims and not for the total claims amount for which it is usually applied.

As an alternative for both, rating and claims reserving, we describe a simple but realistic parametric model for the total claims amount which is based on the Gamma distribution and has the advantage of providing the possibility of assessing the goodness-of-fit and calculating the estimation error. This method is not very well known in automobile insurance—although a satisfactory application is reported—and seems to be completely unknown in the field of claims reserving, although its execution is nearly as simple as that of the chain ladder method.

Keywords

Cross-classified data; (automobile & property) ratemaking; IBNR claims; Gamma model; maximum likelihood method.

Contents of the Paper

1. A SHORT OVERVIEW OF SOME AUTOMOBILE RATING METHODS

In the automobile insurance tariffs of many countries several tariff variables
are used, e.g the horse-power class of the car, the bonus/malus (or no claims
discount) class of the driver or the class of the territory where the car is
principally garaged. In this way the portfolio of automobile insurance policies
is cross-classified into a number of cells which are each supposed to be
homogeneous, so that all policies of the same cell pay the same premium. For
the sake of simplicity we will consider in the following only two tariff variables,
which are subdivided into m and n classes respectively. When then have mn
cells labelled (i, j), $i = 1, \ldots, m, j = 1, \ldots, n$. Now let n_{ij} be the known number
of insureds (policy years) of cell (i, j) and s_{ij} their observed total claims amount
as realization of the random variable S_{ij}. For some of the cells, n_{ij} may be so
small that it is not advisable to use s_{ij} as the only basis for the calculation of
the net premium $E(S_{ij})/n_{ij}$ of that cell. Therefore one searches for marginal
parameters $x_i, i = 1, \ldots, m$, and $y_j, j = 1, \ldots, n$, with

either $x_i y_j = E(S_{ij})/n_{ij}$ (multiplicative approach),

or $x_i + y_j = E(S_{ij})/n_{ij}$ (additive approach).

This also reduces the number of figures needed to describe the tariff premiums
from mn to $m+n$. In the following we only consider the multiplicative
approach, but the methods described can easily be translated to the additive
approach, too.

The problem of finding appropriate marginal parameters x_i and y_j is one of
the classical problems of insurance mathematics It has been known for a long
time that the simple marginal averages

$$x_i = s_{i+} / n_{i+}$$
$$y_j = (s_{+j}/n_{+j})/(s_{++}/n_{++})$$

(where a '+' indicates summation over the corresponding index) give a
satisfactory approximation of $E(S_{ij})/n_{ij}$ only if the tariff variables are indepen-
dent. But generally this is not the case. Therefore, in the last 30 years several
different methods have been proposed. We will now shortly review three of the

most well-known mainly following the description given by VAN EEGHEN/GREUP/NIJSSEN (1983) For a more comprehensive and more recent comparative analysis see JEE (1989)

The first breakthrough was achieved by BAILEY/SIMON (1960), who estimated x_i, y_j by minimizing

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} (s_{ij} - n_{ij} x_i y_j)^2 / (n_{ij} x_i y_j)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} n_{ij} (s_{ij}/n_{ij} - x_i y_j)^2 / (x_i y_j),$$

but their underlying assumption of Q having (up to a factor) the distribution of a chi-square will normally not be true (see VAN EEGHEN/GREUP/NIJSSEN 1983) Moreover, it can be shown (VAN EEGHEN/NIJSSEN/RUYGT 1982) that for the minimizing parameters x_i, y_j the inequalities

$$\sum_{j=1}^{n} n_{ij} x_i y_j \geq \sum_{j=1}^{n} s_{ij}, \qquad i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} n_{ij} x_i y_j \geq \sum_{i=1}^{m} s_{ij}, \qquad j = 1, \ldots, n,$$

hold, i e. there results an overestimation of all marginal loss amounts (in the multiplicative case only).

Therefore BAILEY (1963) and later JUNG (1968) proposed estimating x_i, y_j directly from the intuitively appealing conditions

(1a)
$$\sum_{j=1}^{n} n_{ij} x_i y_j = \sum_{j=1}^{n} s_{ij}, \qquad i = 1, \ldots, m,$$

and

(1b)
$$\sum_{i=1}^{m} n_{ij} x_i y_j = \sum_{i=1}^{m} s_{ij}, \qquad j = 1, \ldots, n,$$

which can be solved iteratively: starting with, for example, $y_j = 1$, (1a) results in $x_i = s_{i+}/n_{i+}$, which is inserted in (1b) giving new y_j etc The procedure converges quickly. This method has been called "marginal totals" If the random variables S_{ij} denote the number of claims instead of the total claims amount, then this method can be shown to be maximum likelihood under the assumption that all S_{ij} are independent and Poisson distributed with parameter $n_{ij} x_i y_j$ (see VAN EEGHEN/GREUP/NIJSSEN 1983, p. 93). But for the more important case where S_{ij} is the total claims amount one has no model from which the equations above derive and thus, for example, a statistical test of the goodness-of-fit cannot be designed either

SANT (1980) proposed estimating x_i, y_j by the method of weighted least squares, i.e. by minimizing

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (s_{ij} - n_{ij} x_i y_j)^2 / n_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} n_{ij} (s_{ij}/n_{ij} - x_i y_j)^2$$

But the powerful tools of regression analysis like the R^2 – statistic, the analysis of residuals and the estimation of the prediction error can only be applied rigorously if all S_{ij} are normally distributed with Var (S_{ij}) proportional to n_{ij}. Both assumptions are not very realistic.

Using the additive approach, the weighted least squares method leads to the same equations for the marginal parameters x_i, y_j as the marginal totals method, which in this case is no longer the maximum likelihood estimator for Poisson distributed numbers of claims.

Altogether, in the case of S_{ij} being claims totals all three methods described above are only of a heuristic nature without an underlying realistic model.

2. SOME METHODS OF ESTIMATING IBNR CLAIMS RESERVES AND THEIR CONNECTION TO AUTOMOBILE RATING METHODS

We now turn to the problem of estimating IBNR claims reserves. For an overview see VAN EEGHEN (1981) or TAYLOR (1986). Here s_{ij} and S_{ij} respectively—we intentionally use the same symbols as before—denote the inflation-adjusted total amount of payments made in development year $j, j = 1, \ldots, n$, for accidents occurred in accident year $i, i = 1, \ldots, m$ If one works with incurred amounts, s_{ij} and S_{ij} denote the total amount of changes in valuation made in development year j on behalf of claims of accident year i. Working with incremental amounts we may assume that all S_{ij} are independent Typically, one has $n = m$ and s_{ij} is known for all $i+j \leq m+1$ (run-off triangle), and one is interested in estimating $E(S_{ij})$ for $i+j > m+1$. The known measure of exposure n_{ij} here normally only depends on the accident year i, i.e. $n_{ij} = n_i$ (number of policies or number of claims reported in the first development year) or is even ignored (i.e. $n_{ij} = 1$ for all i, j).

One of the most important ways of treating the IBNR problem is to assume a multiplicative structure of the type

$$E(S_{ij}) = x_i y_j$$

and to estimate the parameters x_i, y_j from the triangle of known data. This way was used, for example, by DE VYLDER (1978), who estimated x_i, y_j by minimizing

$$\sum_{i,j} (s_{ij} - x_i y_j)^2$$

(where the summation is for all i, j where s_{ij} is known). This is exactly the same method as was used by SANT (1980) in the context of automobile insurance if

one puts all $n_{ij} = 1$ there. Analogously each method which estimates the marginal parameters x_i, y_j for cross-classified automobile insurance data can also be translated into a method for estimating the IBNR claims reserve. One only must take the different pattern of known data (triangle instead of rectangle) into account

Let us consider as further example the method of marginal totals.

Again working with $n = m$ and $n_{ij} = 1$, we get the conditions

$$(H_i) \qquad \sum_j x_i y_j = \sum_j s_{ij}, \qquad i = 1, \ldots, m,$$

$$(V_j) \qquad \sum_i x_i y_j = \sum_i s_{ij}, \qquad j = 1, \ldots, m,$$

where the summation is for those indices where the corresponding s_{ij} are known (i.e. in the case of a full triangle j runs from 1 to $m+1-i$ and i from 1 to $m+1-j$). The same equations are also obtained if one derives the maximum likelihood equations in the Poisson case.

Because of the triangular structure, the above equations can here be solved recursively: We start with the general observation that the solution of this type of problem is only unique up to a multiplicative constant $c \# 0$ because if x_i, y_j is a solution, $x_i c, y_j/c$ is a solution as well. Therefore, without loss of generality we can put $y_1 + \ldots + y_m = 1$. Then using equation (H_1) we have $x_1 = s_{1+}$. From equation (V_m) we get $y_m = s_{1m}/x_1$. Then (H_2) yields x_2, (V_{m-1}) yields y_{m-1} etc.

But it is also possible to derive a direct formula for the unknown mean claims amount $E(S_{ij}) = x_i y_j$. For $h > m+1-i$ it can be shown (see KREMER 1985, p 133-136, or Appendix A where a shorter proof is given) that

$$x_i y_h = \left(\sum_{j=1}^{m+1-i} s_{ij} \right) \cdot f_{m+2-i} \cdot f_{m+3-i} \cdots f_{h-1} \cdot (f_h - 1)$$

where $\quad f_j = \left(\sum_{k=1}^{m+1-j} \sum_{l=1}^{j} s_{kl} \right) \bigg/ \left(\sum_{k=1}^{m+1-j} \sum_{l=1}^{j-1} s_{kl} \right), \qquad j = 2, \ldots, m.$

If one realizes that $\sum_{l=1}^{j} s_{kl}$ is the accumulated claims amount of accident year k known at the end of development year j, one sees that we have just obtained the well-known chain ladder method which is thus shown to be the same as the marginal totals method for $n_{ij} = 1$ Furthermore, from the marginal totals conditions $(H_i), (V_j)$ one easily sees that an incorporation (analogously to (1a) and (1b)) of the known exposure n_i into the estimation of the IBNR claims reserve can be dispensed with, as n_i can be amalgamated with the marginal parameter x_i (in the multiplicative approach only), whereas the

application of the chain ladder method to the claims ratios s_{ij}/n_i assumes a different model.

It is interesting to note that the analogue of the BAILEY-SIMON method seems to have never been published as a method for estimating the IBNR claims reserve.

Another interesting point is the fact that in the context of IBNR claims estimation only the multiplicative approach seems to have been used, although several applications to automobile rating indicate that there the additive approach might give a better fit (see e.g. CHANG/FAIRLEY 1979). A special feature of the additive approach is that it may lead to negative values $E(S_{ij}) = x_i + y_j$. This would make no sense in the ratemaking situation but in the case of claims reserving it can be very realistic (settlement gains)

Clearly, also in the context of claims reserving the least squares method and the marginal totals method (and, of course, the BAILEY-SIMON method) could be carried through with the additive approach, too, both producing an identical set of equations for x_i, y_j as has already been mentioned in the section on automobile rating

There is a natural connection between the multiplicative and the additive approaches because, through the log-transformation,

$$s_{ij}/n_{ij} \approx x_i y_j$$

becomes

$$\log(s_{ij}/n_{ij}) \approx \log(x_i) + \log(y_j).$$

This means that an estimate for $E(S_{ij}/n_{ij})$ can be established by applying an additive approach to the log-transformed data $\log(s_{ij}/n_{ij})$ and by transforming back the obtained solution $\log(x_i)$, $\log(y_j)$ using the exponential function. This was done by CHANG/FAIRLEY (1979) for automobile rating and by KREMER (1982) (see also ZEHNWIRTH 1989) for claims reserving (with $n_{ij} = 1$). For the solution of the transformed (additive) problem, both used the method of (weighted) least squares (here giving the same result as the marginal totals method) in order to estimate the marginal parameters $\log(x_i)$, $\log(y_j)$.

As ZEHNWIRTH (1989) points out, this procedure contains an implicit distributional assumption: In order to fulfill the conditions of normality and homoscedasticity for the least squares estimation of the parameters $\log(x_i)$ and $\log(y_j)$, it has to be assumed that $\log(S_{ij}/n_{ij})$ has a normal distribution with mean value $\log(x_i) + \log(y_j)$ and a variance which is proportional to $1/n_{ij}$ This implies that S_{ij}/n_{ij} is assumed to have a lognormal distribution. CHANG/FAIRLEY and KREMER did not take this implicit distributional assumption into account Therefore, they systematically underestimated $E(S_{ij}/n_{ij})$ as they used $x_i y_j = \exp(\log(x_i) + \log(y_j))$, which is the median of the lognormal distribution whereas the expected value is $x_i y_j \exp(\sigma_{ij}^2/2)$ with $\sigma_{ij}^2 = \text{Var}(\log(S_{ij}/n_{ij}))$. As stated above, we have homoscedasticity if we assume that $\sigma_{ij}^2 = \sigma^2/n_{ij}$, where σ^2 can be estimated by

$$\sum_{i,j} n_{ij}(\log(s_{ij}/n_{ij}) - \log(x_i y_j))^2/(c - m - n + 1),$$

which is just the expression to be minimized by the least squares method. Here c denotes the number of cells where s_{ij} is known.

Unfortunately, we have lost the multiplicative structure, as generally $E(S_{ij}/n_{ij}) = x_i y_j \exp(\sigma_{ij}^2/2)$ cannot be cast into the form $E(S_{ij}/n_{ij}) = \tilde{x}_i \tilde{y}_j$ anymore.

Whereas all the models discussed before have been shown to be only of a heuristic nature both in automobile rating and in claims reserving, the lognormal model relies on a parametric assumption for S_{ij}, and the instruments of regression analysis can be used to check this assumption against the data. In the next section another method is given which relies on a reasonable distributional model and therefore also allows the application of various important and useful statistical tools. This model has two advantages over the lognormal model First, it is not just any model for S_{ij} but can be traced back to a micro-model for the total claims amount of each single insured unit and can therefore be expected to be realistic. Second, we can choose either the multiplicative or the additive structure for $E(S_{ij}/n_{ij})$, whereas the lognormal model yields neither of these structures.

3. A PARAMETRIC MODEL FOR RATING AUTOMOBILE INSURANCE OR ESTIMATING IBNR CLAIMS RESERVES

We use the same notations as before, i e. we have mn cells labelled (i,j), each with known measure of exposure n_{ij} (possibly independent of j in the case of claims reserving) and with total claims amount variable S_{ij} (realization s_{ij}). In the case of claims reserving we know the realizations s_{ij} in the run-off triangle only. We now assume, following TER BERG (1980), that the total claims amount R_{ijk} of each unit $k = 1, \ldots, n_{ij}$ of cell (i,j) has a Gamma distribution with mean value m_{ij} (independent of k) and shape parameter α (independent of i, j, k), i.e. with probability density function

$$f_{ij}(z) = \exp(-\alpha z/m_{ij}) z^{\alpha-1} (\alpha/m_{ij})^{\alpha}/\Gamma(\alpha)$$

(here the usual representation of the Gamma density has been reparametrized in order to implement the mean value m_{ij} directly as a parameter). Because in practice many units k will have a realization $r_{ijk} = 0$ of R_{ijk}, the shape parameter α has to be conceived of as smaller than 1 in order to attribute a high probability to the neighbourhood of $z = 0$ (for instance, we have prob $(R_{ijk} \leq m_{ij}/10) = 0.79$ for $\alpha = 0.05$) Assuming that all n_{ij} units of cell (i,j) are independent, our distributional assumption implies that $S_{ij} = R_{ij1} + R_{ij2} + \ldots$ also has a Gamma distribution but with mean value $n_{ij}m_{ij}$ and shape parameter $n_{ij}\alpha$ And this is the distribution we shall work with in the following, because we usually know only the realizations s_{ij} of S_{ij} and not those of R_{ijk}. The assumption that the shape parameter α is the same for the units of all cells may seem questionable in some cases But this should be detected by testing the goodness-of-fit (see next Section).

In the multiplicative approach we assume furthermore that m_{ij} can be displayed in the form $m_{ij} = x_i y_j$ with unknown parameters x_i, y_j, which we shall estimate with the maximum likelihood method.

Assuming that all S_{ij} are independent, the likelihood function on the basis of the realizations $s_{ij} > 0$ is given by

$$L = \prod_{i,j} \exp\left(-\frac{\alpha s_{ij}}{x_i y_j}\right)\left(\frac{\alpha s_{ij}}{x_i y_j}\right)^{n_{ij}\alpha}\bigg/\left(s_{ij}\,\Gamma(n_{ij}\alpha)\right).$$

Therefore the loglikelihood function is

$$\log(L) = \sum \{-\alpha s_{ij}/(x_i y_j) + n_{ij}\alpha \log(\alpha s_{ij}) - n_{ij}\alpha \log(x_i y_j) - \log(s_{ij}\Gamma(n_{ij}\alpha))\}$$

(where the summation is for all i,j where s_{ij} is known). The maximum likelihood estimator are those values x_i, y_j, α which maximize L or equivalently $\log(L)$ They are given by the equations

$$0 = \partial \log(L)/\partial x_i = \alpha \sum_j (s_{ij}/(x_i^2 y_j) - n_{ij}/x_i), \qquad i = 1, \ldots, m,$$

$$0 = \partial \log(L)/\partial y_j = \alpha \sum_i (s_{ij}/(x_i y_j^2) - n_{ij}/y_j), \qquad j = 1, \ldots, n,$$

which show that the last condition $\partial \log(L)/\partial\alpha = 0$ is not needed for the calculation of the likelihood estimator for x_i, y_j, which can immediately be seen to be given by

$$(2) \quad \begin{cases} x_i = \dfrac{1}{n_{i+}} \sum_j \dfrac{s_{ij}}{y_j}, & i = 1, \ldots, m, \\[3mm] y_j = \dfrac{1}{n_{+j}} \sum_i \dfrac{s_{ij}}{x_i}, & j = 1, \ldots, n \end{cases}$$

These equations have a high intuitive appeal. For, considering the goal of approximating s_{ij} by $n_{ij} x_i y_j$, we see that this amounts to approximating $s_{ij}/(n_{ij} y_j)$ by x_i and therefore the n_{ij}-weighted mean of $s_{ij}/(n_{ij} y_j)$, $j = 1, \ldots, n$, should be a reasonable estimator for x_i

Also, equations (2) are not new. They have already been used by VAN EEGHEN/NIJSSEN/RUYGT (1982). They call them the "direct method" and write (on page 111)·

"This set of equations are a direct translation of the intuitive calculations presented ... by F. K. GREGORIUS. In fact, a solution is found when iteratively calculating the values x_i and y_j by means of the formulae given in (2) by letting $y_j = 1$ $(j = 1, \ldots, n)$ be the starting value The procedure converges rapidly

We may rewrite (2) as

$$\sum_j n_{ij} x_i = \sum_j s_{ij}/y_j, \qquad i = 1, \ldots, m,$$

$$\sum_i n_{ij} y_j = \sum_i s_{ij}/x_i, \qquad j = 1, \ldots, n,$$

which is similar but not equivalent to (1a) and (1b).

As yet, we have not been able to find an argument why a 'satisfactory' solution should (approximately) satisfy (2) ...

The method was more or less developed as a first try and we were surprised to see, that, once formalized, it produced practically the same results as the method of marginal totals."

So much for the quotation from VAN EEGHEN/NIJSSEN/RUYGT (1982).

One year later the Dutch actuaries found an argument for their method because the booklet of VAN EEGHEN, GREUP and NIJSSEN (1983) contains on page 109 a small hint saying that the assumption of a Gamma distribution for R_{ijk} would lead to the "direct method". But there, as in TER BERG (1980), a much more general regression model is considered, of which our simple cross-classified situation is just a special case. Moreover, these authors have concentrated on ratemaking, whereas we want to emphasize the applicability to claims reserving, too.

Finally, it is interesting to note that the likelihoood equations for the additive approach

$$\sum_j \left(s_{ij}/(x_i+y_j)^2 - n_{ij}/(x_i+y_j) \right) = 0, \qquad i = 1, \ldots, m,$$

$$\sum_i \left(s_{ij}/(x_i+y_j)^2 - n_{ij}/(x_i+y_j) \right) = 0, \qquad j = 1, \ldots, n,$$

must be solved with the help of, for example, the NEWTON-RAPHSON numerical method. Moreover, these equations are different from those suggested by the "direct method":

$$x_i = \sum_j (s_{ij} - n_{ij} y_j)/n_{i+}, \qquad i = 1, \ldots, m,$$

$$y_j = \sum_i (s_{ij} - n_{ij} x_i)/n_{+j}, \qquad j = 1, \ldots, n.$$

4. STATISTICAL ANALYSIS OF THE MODEL

This parametric approach with a realistic distributional assumption enables us to use many tools for the statistical analysis, as has been clearly set out by

ALBRECHT (1983), who describes the case $\alpha = 1$ in considerable detail but again as a general regression model. Besides the consistent and (asymptotically) efficient estimation of the model parameters, we have the possibility of testing the significance of the tariff variables with the likelihood ratio test (see ALBRECHT (1983) for details), we can calculate the error variances of the parameter estimators and we can check the goodness-of-fit We first consider the goodness-of-fit · According to our model, S_{ij} has a Gamma distribution with $E(S_{ij}) = n_{ij} m_{ij}$ and $\mathrm{Var}\,(S_{ij}) = n_{ij} m_{ij}^2/\alpha$. The higher the shape parameter $n_{ij}\alpha$ of this distribution, the closer it is to the normal distribution If all S_{ij} are approximately normally distributed the statistic

$$\sum_{i,j} \frac{(S_{ij} - E(S_{ij}))^2}{\mathrm{Var}\,(S_{ij})} = \alpha \sum_{i,j} \frac{(S_{ij}/(x_i y_j) - n_{ij})^2}{n_{ij}}$$

$$= \alpha \sum_{i,j} n_{ij} \left(\frac{S_{ij}}{n_{ij} x_i y_j} - 1 \right)^2$$

is, under the hypothesis of our model, approximately at chi-square with $c - m - n$ degrees of freedom, where c is the number of cells where s_{ij} is known

The special form of this statistic allows its application without having estimated α. For this purpose we fix α in such a way that the value of the statistic is just below the (say) 0.95-fractile of the chi-square distribution. If using this value of α a normality condition like "$n_{ij}\alpha > 10$" is fulfilled for nearly all cells, we may be satisfied with the goodness-of-fit of the model. But we have to realize that this goodness-of-fit test only checks the fit of aggregated figures and cannot test the distributional assumptions within the cells

Applying this procedure to SANT'S (1980) collision data (126 cells) we get $\alpha\,(<) = 0.021$ and the three lowest values of $n_{ij}\alpha$ turn out to be 6.8, 9.4 and 11 5, so we may accept the multiplicative Gamma model. Using CHANG/FAIR-LEY'S (1979) combined compulsory data (105 cells), we get $\alpha\,(<) = 0.0094$ and have 9 cells were the resulting value of $n_{ij}\alpha$ is lower than 10, the lowest being 4.5, so the fit is less satisfactory.

A simple formula for an estimator of α is given by the method of moments, i e. by equating the variances

$$\sum_{i,j} (s_{ij} - n_{ij} x_i y_j)^2 = \sum_{i,j} n_{ij} (x_i y_j)^2/\alpha.$$

This yields $\alpha = 0.014$ for Sant's data and $\alpha = 0.0093$ for Chang/Fairley's data.

Strictly speaking we should use the likelihood estimator for α We then must solve the likelihood equation

$$0 = \partial \log\,(L)/\partial \alpha = \sum_{i,j} n_{ij}\{\log\,(\alpha s_{ij}) - \log\,(x_i y_j) - \psi\,(n_{ij}\alpha)\}$$

Here equations (2) have been used to obtain $\Sigma n_{ij} = \Sigma s_{ij}/(x_i y_j)$. $\psi(z) = \Gamma'(z)/\Gamma(z)$ denotes the digamma function, for which the asymptotic approximation

$$\psi(z) \approx \log(z) - (2z)^{-1} - z^{-2}/12$$

exists which even for arguments as low as $z \geq 4$ is exact to 4 decimal places. This approximation yields as the solution of the likelihood equation

$$\alpha \approx (c + \sqrt{c^2 + ab})/a$$

with

$$a = 4 \sum_{i,j} n_{ij} \log(n_{ij} x_i y_j /s_{ij}) > 0,$$

$$b = \sum_{i,j} (3 n_{ij})^{-1}$$

$$c = \sum_{i,j} 1 = \text{number of cells where } s_{ij} \text{ is known}$$

Applied to Sant's data this yields $\alpha \approx 0.0202$. For Chang/Fairley's data we get $\alpha \approx 0.0097$. If we have some small exposures n_{ij} such that $n_{ij}\alpha < 4$, we should refine the approximation of the digamma function by using the recursion $\psi(z) = \psi(z+1) - 1/z$ and by including more terms of the approximation series. Then a direct formula for α cannot be given anymore We must therefore solve the likelihood equation iteratively with the NEWTON-RAPHSON method.

Having estimated α, we are also in the position to calculate the estimation error of the estimators for x_i and y_j. This is done in Appendix B.

According to the experience of the Dutch actuaries, the results of applying the "direct method" to automobile insurance data are rather close to the results obtained by the marginal totals method. Translated to the IBNR claims reserving problem this means that the "direct method" results will be similar to the chain ladder results. But with the "direct method" we can additionally make use of the aforementioned advantages. Moreover, the formulae provide the possibility of taking the exposure n_i of accident year i into account (which is different from the situation with the chain ladder). And perhaps the goodness-of-fit statistic or the size of the likelihood function gives an indication to answering the question "additive or multiplicative?" Because of these advantages of the parametric method we believe that before using a rather heuristic method like BAILEY/SIMON or chain ladder one should examine whether the parametric method fits the data.

5. IMPROVEMENT OF THE MODEL IN THE CASE OF KNOWN CLAIMS NUMBERS

Especially in the claims reserving situation we will often have difficulties in finding an adequate measure n_{ij} of exposure

Therefore mostly $n_{ij} = n_i$ or even $n_{ij} = 1$ is taken. However, this is not satisfactory because the exposure to further payments or changes in valuation varies in fact rather strongly over the development years. Therefore, a more meaningful measure of exposure will be the number t_{ij} of those claims of accident year i where there is a change in amount during development year j. These data $t_{ij}, i+j \le m+1$, are often available in practice.

Rating in property insurance presents a similar problem. There, even the risks of the same cell vary greatly with respect to their size, which is usually measured by the sum insured. Therefore, the number of risks is not a good measure for the exposure of a cell (i, j), and the sum insured is taken instead. But then an assumption of our micro-model is not fulfilled anymore because the "units" of sum insured are not independent, as a single risk consists of several such units. We therefore must abandon our micro-model and try directly whether the Gamma model for S_{ij} with mean value $E(S_{ij}) = n_{ij} x_i y_j$ and shape parameter $n_{ij} \alpha$ fits the data if n_{ij} is the sum insured. The parameter α then does not have a specific interpretation anymore But if we know additionally the total number t_{ij} of claims of cell (i, j) we can apply the following stepwise approach which assumes a Gamma distribution (with shape parameter α) not for the total claims amount per risk unit but for the amount of each single claim. Of course, this procedure can also be applied in automobile ratemaking if the number t_{ij} of claims is available.

In these situations we should use t_{ij}—the corresponding random variable is denoted by T_{ij}—as an additional measure of exposure and adopt the following three-steps-approach, which follows the ideas of ALBRECHT (1983): In the first step we take the observed number t_{ij} of claims of cell (i, j) as the measure of exposure and assume that the size of each corresponding amount has a Gamma distribution with mean value $m_{ij} = x_i y_j$ and shape parameter α. Then we are in our original model (with n_{ij} replaced with t_{ij}) leading to the direct method This yields smoothed average claims amounts $x_i y_j$. In the second step we smooth the t_{ij} by assuming that all T_{ij} are independent of each other and that each T_{ij} has a Poisson distribution with parameter $n_{ij} v_i w_j$ (here using the 'old' measure of exposure). Then the maximum likelihood estimator of v_i, w_j on basis of the realizations t_{ij} is given by the equations (1a) and (1b) with x_i, y_j, s_{ij} replaced with v_i, w_j, t_{ij} respectively. This yields smoothed numbers $n_{ij} v_i w_j$ of claims. In the last step, $E(S_{ij})$ is estimated by $n_{ij} v_i w_j x_i y_j$ implying that in each cell the number of claims is independent of the average claims amount.

6 FINAL REMARK

In the context of this paper we should point out the following further connection between rating methods and claims reserving methods. Another important rating method which smoothes the claims experience of several tariff classes is the Bühlmann-Straub credibility model. It also uses a cross-classifying approach by the two dimensions 'tariff classes' and 'observation years'.

Therefore, one will presume that it could also be translated into a method for estimating IBNR claims reserves. But there is a difficulty because the Bühlmann-Straub model assumes that the average claims amount S_{ij}/n_{ij} of tariff class i has the same expected value over all years j, whereas in the run-off triangle the expected value of the average claims amount S_{ij}/n_i of accident year i and development year j varies in a certain but unknown pattern over the development years. However, this difficulty can be overcome in such a way that the Bühlmann-Straub model can directly be used for claims reserving, too (see MACK 1990).

<div align="center">

APPENDIX A

PROOF THAT THE CHAIN LADDER METHOD CAN BE DERIVED FROM THE MARGINAL TOTALS CONDITIONS (AND THEREFORE IS MAXIMUM LIKELIHOOD IN THE POISSON CASE)

</div>

We show that the chain ladder method

$$x_i y_h = \left(\sum_{j=1}^{m+1-i} s_{ij} \right) \cdot f_{m+2-i} \cdot f_{m+3-i} \cdots f_{h-1} \cdot (f_h - 1), \qquad h > m+i-1,$$

with

$$f_j = \left(\sum_{k=1}^{m+1-j} \sum_{l=1}^{j} s_{kl} \right) \bigg/ \left(\sum_{k=1}^{m+1-j} \sum_{l=1}^{j-1} s_{kl} \right), \qquad j = 2, \ldots, m,$$

can be deduced from the marginal conditions

$$(H_i) \qquad \sum_{j=1}^{m+1-i} x_i y_j = \sum_{j=1}^{m+1-i} s_{ij}, \qquad i = 1, \ldots, m,$$

$$(V_j) \qquad \sum_{i=1}^{m+1-j} x_i y_j = \sum_{i=1}^{m+1-j} s_{ij}, \qquad j = 1, \ldots, m.$$

Let $c_{ij} = \sum_{l=1}^{j} x_i y_l$ and $b_{ij} = \sum_{l=1}^{j} s_{il}$ $(i+j \le m+1)$ denote the expected and the observed accumulated claims amount of accident year i at the end of development year j respectively. Then conditions (H_i) can be written shortly as $c_{i, m+1-i} = b_{i, m+1-i}$. For $h > m+1-i$ we have

$$c_{ih} = c_{i, m+1-i} \cdot \frac{c_{i, m+2-i}}{c_{i, m+1-i}} \cdots \frac{c_{ih}}{c_{i, h-1}}.$$

Therefore

$$x_i y_h = c_{ih} - c_{i, h-1}$$

$$= c_{1,\,m+1-\iota}\ \frac{c_{1,\,m+2-\iota}}{c_{1,\,m+1-\iota}}\cdots\frac{c_{1,\,h-1}}{c_{1,\,h-2}}\left(\frac{c_{\iota h}}{c_{1,\,h-1}}-1\right)$$

$$=\left(\sum_{J=1}^{m+1-\iota}s_{IJ}\right)\cdot\frac{c_{1,\,m+2-\iota}}{c_{1,\,m+1-\iota}}\cdots\frac{c_{\iota,\,h-1}}{c_{\iota,\,h-2}}\cdot\left(\frac{c_{\iota h}}{c_{\iota,\,h-1}}-1\right)$$

and we have only to show that $c_{IJ}/c_{1,\,J-1}=f_J$. Because of

$$\frac{c_{IJ}}{c_{1,\,J-1}}=\frac{\sum\limits_{l=1}^{J}y_{l}}{\sum\limits_{l=1}^{J-1}y_{l}}=\frac{\left(\sum\limits_{k=1}^{m+1-J}x_{k}\right)\left(\sum\limits_{l=1}^{J}y_{l}\right)}{\left(\sum\limits_{k=1}^{m+1-J}x_{k}\right)\left(\sum\limits_{l=1}^{J-1}y_{l}\right)}=\frac{\sum\limits_{k=1}^{m+1-J}c_{kJ}}{\sum\limits_{k=1}^{m+1-J}c_{k,\,J-1}}$$

and of

$$f_{J}=\left(\sum_{k=1}^{m+1-J}b_{kJ}\right)\bigg/\left(\sum_{k=1}^{m+1-J}b_{k,\,J-1}\right)$$

it is enough to show that

$$(A_{J})\qquad\qquad\sum_{k=1}^{m+1-J}c_{kJ}=\sum_{k=1}^{m+1-J}b_{kJ}$$

and

$$(B_{J})\qquad\qquad\sum_{k=1}^{m+1-J}c_{k,\,J-1}=\sum_{k=1}^{m+1-J}b_{k,\,J-1}$$

hold for $J=2,\ ..,m$ We show this by recursion from $J=m$ to $J=2\cdot$ (A_{m}), ı e $c_{1m}=b_{1m}$, holds because of (H_{1}).

(B_{J}) follows from (A_{J}) and (V_{J}) as

$$\sum_{k=1}^{m+1-J}c_{k,\,J-1}=\sum_{k=1}^{m+1-J}(c_{kJ}-x_{k}y_{J})=\sum_{k=1}^{m+1-J}c_{kJ}-\sum_{k=1}^{m+1-J}x_{k}y_{J}$$

$$=\sum_{k=1}^{m+1-J}b_{kJ}-\sum_{k=1}^{m+1-J}s_{kJ}=\sum_{k=1}^{m+1-J}(b_{kJ}-s_{kJ})=\sum_{k=1}^{m+1-J}b_{k,\,J-1}$$

Finally, (A_{J-1}) follows from (B_{J}) and (H_{m+2-J}) as

$$\sum_{k=1}^{m+2-J}c_{k,\,J-1}=\sum_{k=1}^{m+1-J}c_{k,\,J-1}+c_{m+2-J,\,J-1}$$

$$=\sum_{k=1}^{m+1-J}b_{k,\,J-1}+b_{m+2-J,\,J-1}=\sum_{k=1}^{m+2-J}b_{k,\,J-1}$$

This completes the proof.

APPENDIX B
ESTIMATION OF THE (ASYMPTOTIC) ERROR VARIANCES

We have estimated the marginal parameters x_i, y_j with

either $x_i y_j = E(S_{ij}/n_{ij})$ (multiplicative approach)

or $x_i + y_j = E(S_{ij}/n_{ij})$ (additive approach).

by the maximum likelihood method and now want to know how precise these estimates are, i.e. we want to calculate $\mathrm{Var}(X_i)$, $\mathrm{Var}(Y_j)$, $\mathrm{Var}(X_i Y_j)$ or $\mathrm{Var}(X_i + Y_j)$ where X_i and Y_j denote the random variables corresponding to the estimators for x_i and y_j respectively. A standard result of maximum likelihood theory states that under certain regularity conditions which are fulfilled here, the following holds true: If a parameter vector $\Theta = (\Theta_1, \ldots, \Theta_r)$ is estimated by the maximum likelihood method, the obtained estimator $\hat{\Theta}$ has asymptotically a normal distribution with mean value Θ and with a covariance matrix which is equal to the inverse of the information matrix

$$I(\Theta) = E\left(- \frac{\partial^2 \log(L)}{\partial \Theta_i \, \partial \Theta_j} \right)_{i,j}$$

where $L = L(\Theta)$ is the likelihood function.

In our case we have $\Theta = (x_2, \ldots, x_m, y_1, \ldots, y_n)$ where we have omitted x_1 without loss of generality in order to obtain a unique solution of the likelihood equations and have considered α as being known (For the case of α being included in Θ, TER BERG (1980) has shown that this does not change the calculation of $\mathrm{Var}(X_i)$, $\mathrm{Var}(Y_j)$ and $\mathrm{Cov}(X_i, Y_j)$). We now have

$$\mathrm{Cov}(X_2, \ldots, X_m, Y_1, \ldots, Y_n) \approx I(x_2, \ldots, x_m, y_1, \ldots, y_n)^{-1} =: I^{-1}$$

$$\approx I(\hat{x}_2, \ldots, \hat{x}_m, \hat{y}_1, \ldots, \hat{y}_n)^{-1} = \cdot \hat{I}^{-1}$$

where $\hat{x}_2, \ldots, \hat{y}_n$ denote the estimated values of the true parameters $x_2, \ldots, y_n$ From $\hat{I}^{-1}$ we directly obtain asymptotic approximative values for $\mathrm{Var}(X_i)$, $\mathrm{Var}(Y_j)$ and $\mathrm{Cov}(X_i, Y_j)$. This also gives immediately an approximation for

$$\mathrm{Var}(X_i + Y_j) = \mathrm{Var}(X_i) + 2\,\mathrm{Cov}(X_i, Y_j) + \mathrm{Var}(Y_j)$$

which we want to know in the additive approach. In order to obtain $\mathrm{Var}(X_i Y_j)$ for the multiplicative approach, we make use of a general theorem on the higher moments of normally distributed variables (see e.g RICHTER 1966, p 369) to get

$$\mathrm{Var}(X_i Y_j) \approx \mathrm{Var}(X_i)\,\mathrm{Var}(Y_j) + (\mathrm{Cov}(X_i, Y_j))^2 + \mathrm{Var}(X_i)\,(E(Y_j))^2 +$$
$$+ 2E(X_i)\,\mathrm{Cov}(X_i, Y_j)\,E(Y_j) + (E(X_i))^2\,\mathrm{Var}(Y_j)$$

(which holds exactly if X_i and Y_j are normally distributed) This can be calculated from $\hat{I}^{-1}$ and from $E(X_i) \approx \hat{x}_i$, $E(Y_j) \approx \hat{y}_j$.

Therefore, the only thing left to do is the calculation of I and I^{-1} Concentrating again on the multiplicative approach, the loglikelihood function is

$$\log(L) = -\sum_{i,j} (\alpha S_{ij}/(x_i y_j) + \alpha n_{ij} \log(x_i y_j) + g(\alpha, n_{ij}, S_{ij}))$$

and yields (using $E(S_{ij}) = n_{ij} x_i y_j$ and the Kronecker symbol δ_{ij} with $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise)

$$A_{ik} := E\left(-\frac{\partial^2 \log(L)}{\partial x_i \partial x_k}\right) = \frac{\alpha n_{i+}}{x_i^2} \delta_{ik}, \qquad 2 \le i, k \le m,$$

$$B_{ij} := E\left(-\frac{\partial^2 \log(L)}{\partial x_i \partial y_j}\right) = \frac{\alpha n_{ij}}{x_i y_j}, \qquad 2 \le i \le m, 1 \le j \le n,$$

$$C_{lj} := E\left(-\frac{\partial^2 \log(L)}{\partial y_l \partial y_j}\right) = \frac{\alpha n_{+j}}{y_j^2} \delta_{lj}, \qquad 1 \le l, j \le n,$$

(where n_{+j} includes n_{1j}). With the matrices $A = (A_{ik})$, $B = (B_{ij})$, $C = (C_{lj})$ the information matrix I can be represented as partitioned matrix

$$I = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix}$$

where A and C are diagonal matrices.

Unfortunately, an explicit formula for the inverse matrix I^{-1} is not available. One therefore must apply a numerical inversion method. But the dimension of the inversion problem can be reduced with the help of the following result for the inverse of a partitioned matrix (which can be verified by calculating $I^{-1}I$ and II^{-1}):

$$I^{-1} = \begin{pmatrix} D^{-1} & -D^{-1}BC^{-1} \\ -C^{-1}B^t D^{-1} & C^{-1} + C^{-1}B^t D^{-1}BC^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} A^{-1} + A^{-1}BF^{-1}B^t A^{-1} & -A^{-1}BF^{-1} \\ -F^{-1}B^t A^{-1} & F^{-1} \end{pmatrix}$$

with

$$D = A - BC^{-1}B^t,$$

$$F = C - B^t A^{-1}B.$$

A straightforward calculation yields for the elements of D and F

$$D_{ik} = \alpha(\delta_{ik} n_{i+} - p_{ik})/(x_i x_k), \qquad 2 \le i, k \le m,$$

$$F_{lj} = \alpha(\delta_{lj} n_{+j} - q_{lj})/(y_l y_j), \qquad 1 \le l, j \le n,$$

with

$$p_{ik} = \sum_{j=1}^{n} \frac{n_{ij} n_{kj}}{n_{+j}}, \qquad q_{lj} = \sum_{i=2}^{m} \frac{n_{il} n_{ij}}{n_{i+}}.$$

Therefore, only the smaller matrices D and F must be inverted in order to obtain I^{-1} and also $\hat{I}^{-1}$.

REFERENCES

ALBRECHT, P (1983) Parametric Multiple Regression Risk Models *Insurance Mathematics & Economics*, 49–66, 69–73 and 113–117
BAILEY, R A (1963) Insurance Rates with Minimum Bias *Proceedings of the Casualty Actuarial Society*, 4–11
BAILEY, R A and SIMON, L J (1960) Two Studies in Automobile Insurance Rate Making *ASTIN Bulletin* 1, 192–217
CHANG, L and FAIRLEY, W B (1979) Pricing Automobile Insurance under Multivariate Classification of Risks Additive versus Multiplicative *Journal of Risk and Insurance*, 75–98
DE VYLDER, F (1978) Estimation of IBNR Claims by Least Squares *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 249–254
JEE, B (1989) A Comparative Analysis of Alternative Pure Premium Models in the Automobile Risk Classification System *Journal of Risk and Insurance*, 434–459
JUNG, J (1968) On Automobile Insurance Ratemaking *ASTIN Bulletin* 5, 41–48.
KREMER, E (1982) IBNR-Claims and the Two-way Model of ANOVA *Scandinavian Actuarial Journal*, 47–55
KREMER, E (1985) *Einführung in die Versicherungsmathematik* Vandenhoek & Ruprecht, Gottingen
MACK, Th (1990) Improved Estimation of IBNR Claims by Credibility Theory *Insurance Mathematics & Economics*, 51–57
RICHTER, H (1966) *Wahrscheinlichkeitstheorie* Springer, Heidelberg
SANT, D T (1980) Estimating Expected Losses in Auto Insurance *Journal of Risk and Insurance*, 133–151
TAYLOR, G C (1986) *Claims Reserving in Non-life Insurance* North Holland, Amsterdam
TER BERG, P (1980) On the loglinear Poisson and Gamma model *ASTIN Bulletin* 11, 35–40
VAN EEGHEN, J (1981) *Loss Reserving Methods* Nationale-Nederlanden N V, Rotterdam
VAN EEGHEN, J, GREUP, E K and NIJSSEN, J A (1983) *Rate Making* Nationale-Nederlanden N V, Rotterdam
VAN EEGHEN, J, NIJSSEN, J A and RUYGT, F A M (1982) Interdependence of Risk Factors Application of Some Models New Motor Rating Structure in the Netherlands, 105–119 *ASTIN-groep*, Nederland
ZEHNWIRTH, B (1989) *The Chain Ladder Technique — A Stochastic Model* Claims Reserving Manual Vol 2, 2–9 Institute of Actuaries, London

Dr THOMAS MACK
Munchener Ruckversicherungs-Gesellschaft, Koniginstr. 107,
D-8000 Munchen 40, FRG.