A NOTE ON THE MULTIPLICATIVE RATEMAKING MODEL

B. AJNE
Stockholm

THE MODEL

The multiplicative ratemaking, model we have in mind is the following one. Within a certain branch of insurance we have, say for simplicity, two tarif arguments U and V. For example, in motor insurance we could think of U and V as being make of car and geographical district respectively. In fire insurance U could be class of construction for buildings and V could relate to fire defense capacities.

The arguments are of a qualitative nature and argument U has r levels, while argument V has k levels. To our disposal we have statistical experience of the business for a certain period of time, consisting of

—risk exposures n_{ij} $(i = 1 \ldots r,\ j = 1 \ldots k)$.
Risk exposure n_{ij} thus corresponds to the ith U-level and the jth V-level. It could be e.g. number of policy years or sum insured during the period of observation for objects belonging simultaneously to U-level i and V-level j.
The n_{ij}s are known non-random quantities.

—(relative) risk measures $p_{ij}(i = 1 \ldots r,\ j = 1 \ldots k)$.
Risk measure p_{ij} could be e.g. claims frequency, i.e. number of claims divided by number of policy years, or claims cost per policy year or claims cost as a percentage of sum insured. In general p_{ij} is thus the observed number or the observed amount of claims belonging simultaneously to U-level i and V-level j, divided by the corresponding risk exposure n_{ij}.
The p_{ij}s are observed values of random variables.

The multiplicative model now consists of the assumption

$$E(p_{ij}) = cu_iv_j \qquad (1)$$
$$(i = 1 \ldots r, \quad j = 1 \ldots k)$$

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press

that is, the expected values of the risk measures p_{ij}—the true risk premiums on which to found the tarif book—can be represented in the multiplicative form (1) with suitably chosen factors $c, u_1 \dots u_r, v_1 \dots v_k$.

Fitting the Model

The model has been studied by several authors, see e.g. references [1], [2], [4], [6], [7] and [8].

Several methods of graduation have been proposed and have also been implemented in EBD-systems [3], [5]. Among these is the one proposed by Jung [6] and described by the following set of equations for $c, u_1 \dots u_r, v_1 \dots v_k$

$$\begin{aligned} \sum_j n_{ij}\, cu_i v_j &= \sum_j n_{ij} p_{ij} \qquad (i = 1 \dots r) \\ \sum_i n_{ij}\, cu_i v_j &= \sum_i n_{ij} p_{ij} \qquad (j = 1 \dots k) \end{aligned} \tag{2}$$

Thus, the graduation is done so that for each U-level i the graduated "marginal" claims cost will be equal to the observed marginal claims cost and correspondingly for the V-levels. If one considers one argument at a time, the method is thus fair. As the left hand sides of (2) are the expected values of the right hand sides, one could also say that the method coincides with the method of moments. It can also be shown to coincide with the so-called modified chi-square minimum method, Jung [6]. In practice equations (2) are solved by putting c equal to the overall risk measure

$$c = \sum_{ij} n_{ij} p_{ij} / \sum_{ij} n_{ij}$$

and writing (2) in the form

$$\begin{aligned} u_i &= \sum_j n_{ij} p_{ij} / c \sum_j n_{ij} v_j \qquad (i = 1 \dots r) \\ v_j &= \sum_i n_{ij} p_{ij} / c \sum_i n_{ij} u_i \qquad (j = 1 \dots k) \end{aligned}$$

and solving for u_i, v_j by iteration.

In the following we will restrict ourselves to this method of graduation.

Properties of the Solutions

In practise you are somewhat concerned about the statistical properties of solutions $c, u_1 \dots u_r, v_1 \dots v_k$ to equations (2). If their variances are large you could obviously not put much confidence in the graduation even if the model assumption (1) is correct. Also, if you make graduations of new sets of data from the same branch of insurance, e.g. produced during consecutive years of experience, you will in that case get a strong variation in the obtained values for the factors $u_1 \ldots u_r$, $v_1 \ldots v_k$.

In the following we hold the model assumption (1) to be true. It is obvious that (1) does not determine the factors uniquely. We could e.g. multiply all u_is by two and divide all v_js by two without affecting the relation. We therefore impose the normalizing condition.

$$u_1 = v_1 = 1 \tag{3}$$

It is obvious that c, u_i, v_j are then uniquely determined by (1). Also, we put equations (2) in the following form

$$\begin{aligned} F_i &= cu_i \sum_j n_{ij}v_j - \sum_j n_{ij}p_{ij} = 0 \qquad (i = 2 \ldots r) \\ G_j &= cv_j \sum_i n_{ij}u_i - \sum_i n_{ij}p_{ij} = 0 \qquad (j = 2 \ldots k) \\ H &= c \sum_{ij} n_{ij}u_iv_j - \sum_{ij} n_{ij}p_{ij} = 0 \end{aligned} \tag{4}$$

That is, the difference between graduated and observed row totals should be zero for rows $2 \ldots r$, and correspondingly for columns $2 \ldots k$. Finally the difference between graduated and observed grand total should be zero. This is obviously equivalent to (2).

We now compute the jacobian matrix

$$J = \frac{\partial(H, F_2 \ldots F_r, G_2 \ldots G_k)}{\partial(c, u_2 \ldots u_r, v_2 \ldots v_k)}$$

Its first row are given by, in turn, the partial derivatives

$$\begin{aligned} \frac{\partial H}{\partial c} &= \sum_{ij} n_{ij}u_iv_j \\ \frac{\partial H}{\partial u_i} &= c \sum_j n_{ij}v_j \qquad (i = 2 \ldots r) \\ \frac{\partial H}{\partial v_j} &= c \sum_i n_{ij}u_i \qquad (j = 2 \ldots k) \end{aligned}$$

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press

Its next $(r - 1)$ rows are given by, in turn, the partial derivatives $(i = 2 \ldots r)$

$$\frac{\partial F_i}{\partial c} = u_i \sum_j n_{ij} v_j$$

$$\frac{\partial F_i}{\partial u_s} = \begin{cases} 0 \text{ if } s \neq i \\ c \sum_j n_{ij} v_j \text{ if } s = i \end{cases} \qquad (s = 2 \ldots r)$$

$$\frac{\partial F_i}{\partial v_j} = c u_i n_{ij} \qquad (j = 2 \ldots k)$$

and correspondingly for the last $(k - 1)$ rows, with the G_js instead of the F_is.

Drawing up the picture of the jacobian matrix J on a paper it is seen that its determinant, $|J|$, has the property

$$|J| = c^{r+k-2} u_2 \ldots u_r v_2 \ldots v_k |A|$$

where A is a symmetric $(r + k - 1) \times (r + k - 1)$ matrix.

Furthermore, let $a' = (z, x_2 \ldots x_r, y_2 \ldots y_k)$ be an arbitrary $(r + k - 1) \times 1$ vector. By straightforward calculation it is found that the quadratic form

$$a' A a$$

equals

$$\sum_{i=2}^{r} \sum_{j=2}^{k} n_{ij} u_i v_j (z + \xi_i + \eta_j)^2 + \sum_{i=2}^{r} n_{i1} u_i (z + \xi_i)^2 + \sum_{j=2}^{k} n_{1j} v_j (z + \eta_j)^2 + n_{11} z^2$$

where

$$\xi_i = x_i / u_i \qquad (i = 2 \ldots r)$$

$$\eta_j = y_j / v_j \qquad (j = 2 \ldots k)$$

The model naturally assumes c, u_i, v_j all to be positive, and we also assume all risk exposures n_{ij} to be positive. Thus

$$a' A a > 0$$

A is positive definite. Thus $|A|$ and $|J|$ are also positive, i.e. the jacobian matrix J is non-singular. This means that as long as the observed risk measures p_{ij} have values in a sufficiently small neighbourhood of the expected values (1), equations (4) will have a unique vector of solutions

$$\hat{f}' = (\hat{c}, \hat{u}_2 \ldots \hat{u}_r, \hat{v}_2 \ldots \hat{v}_k)$$

which converges in probability to the true values

$$f' = (c, u_2 \ldots u_r, v_2 \ldots v_k)$$

as the risk exposures n_{ij}, and hence the expected number of claims in cell (i, j), tends to infinity.

Furthermore, let

$$R_i = \sum_j n_{ij} p_{ij} \qquad (i = 2 \ldots r)$$

$$C_j = \sum_i n_{ij} p_{ij} \qquad (j = 2 \ldots k)$$

$$T = \sum_{ij} n_{ij} p_{ij}$$

and put

$$t' = (T_1 R_2 \ldots R_r, C_2 \ldots C_k)$$

Then, asymptotically as all n_{ij} tend to infinity

$$\hat{f} = f + J^{-1}(t - E(t))$$

As, under the usual Poisson assumption, the p_{ij}s are asymptotically normal with variances of the order of magnitude n_{ij}^{-1}, we see that $\hat{f}$ is asymptotically normal with mean vector f and variances and covariances of an order of magnitude corresponding to the reciprocals of the n_{ij}s.

NUMERICAL ILLUSTRATION

We have not yet made any theoretical investigations as to the statistical properties of the estimates

$$\hat{f}' = (\hat{c}, \hat{u}_2 \ldots \hat{u}_r, \hat{v}_2 \ldots \hat{v}_k)$$

for finite sizes of the riskexposures n_{ij}. We have however, made a simulation experiment. We would like to report on some findings from this experiment, as it illustrates the asymptotic theory and might give some clues for the finite theory.

The experiment was actually carried out for three tarif arguments U, V and W, with two, three and ten levels respectively. The risk

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press

exposures n_{ijk} were chosen to be proportionate to the following numbers

$k = 1, 2, 3, 4$

i/j	1	2	3
1	600	500	400
2	1,600	1,500	1,400

$k = 5, 6, 7, 8$

i/j	1	2	3
1	400	500	600
2	400	500	600

$k = 9, 10$

i/j	1	2	3
1	100	150	200
2	50	50	50

Risk measures p_{ijk} were simulated assuming p_{ijk} to be normally distributed with mean

$$E(p_{ijk}) = cu_i v_j w_k$$

and variance

$$\text{Var}\,(p_{ijk}) = cu_i v_j w_k / n_{ijk}$$

The multiplicative model was thus assumed to be true, and with the following values for the factors

$c = 0.05$

$(u_1, u_2) = (1, 1.4)$

$(v_1, v_2, v_3) = (1, 1.2, 1.4)$

$(w_1, w_2, w_3, \ldots, w_9, w_{10}) = (1, 1.15, 1.30, \ldots.\ 2.20, 2.35)$

The experiment was carried out 100 times. After each simulation estimates $\hat{c}$, $\hat{u}_1$, $\hat{v}_j$, $\hat{w}_k$ were computed from equations (2), or—to be exact—from their analogues for three tarif arguments.

The whole procedure was repeated four times corresponding to four choices of proportionality factor for the risk exposures n_{ijk}, namely

$$1/81 \quad 1/9 \quad 1 \quad 1/0.09$$

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press

Thus e.g. n_{111} was given in the four repetitions the respective values

$$600/81 \quad 600/9 \quad 600 \quad 600/0.09$$

Note: The same basic set of random numbers were used in all four repetitions. The different sizes of n_{ijk}s were taken into account in the transformation to the normal distribution for the p_{ijk}s.

Factors w_k should be most critical as they are supported by the smallest marginal risk exposures. Following are results for three w-factors, one from each exposure-size group.

Factor	True value		Observed mean value and standard deviation of estimate in 100 simulations Proportionality factor for exposures 1/81	1/9	1	1/0, 09
w_3	1.30	mean	1.474	1.333	1.307	1.301
		st.d	0.75	0.23	0.075	0.023
w_7	1.90	mean	2.230	1.946	1.908	1.901
		st.d	1.20	0.38	0.12	0.037
w_{10}	2.35	mean	3.739	2.530	2.369	2.354
		st.d	2.42	0.81	0.28	0.083

Asymptotic unbiasedness is well illustrated. So is the inverse relationship between variances and risk exposures, at least when, as in this case, the latter tend to infinity at the same rate.

For finite exposures we seem to have a positive bias (this goes for the other u-, v- and w-factors not shown here, too). The dependance of this bias, as well as the variances, on total and marginal exposures might be worth studying.

As for the asymptotic normality, we have tried to illustrate it by four histograms for w_{10} shown at the end of the paper.

Acknowledgement

Sincere thanks are due to Mrs. M. Lindberg for carrying through the simulations and to Messrs. E. Elvers and P. Lindstroem for valuable discussions concerning the implications of them.

References

[1] Almer, B. (1957), *Risk Analysis in Theory and practical statistics*, 15th International Congress of Actuaries, Vol. 2.

[2] ALMER, B. (1962), "Individual Risk Theory and Risk statistics as applied to Fire Insurance", *The ASTIN Bulletin,* Vol. II, Part III.

[3] ANDREASSON, G. and WENANDER, M.-L. (1970), *Modifications in the Multiplicative Ratemaking Model Program.* Unpublished stencile.

[4] BAILEY, R. A. and SIMON, L. J. (1960), "Two Studies in Automobile Insurance Ratemaking", *The ASTIN Bulletin,* Vol. I, Part IV.

[5] BOEHM, C., KLINGEN, N. and MEHRING, J. (1968), *Ein maschinelles Verfahren für Statistik- und Tarifaufgaben,* 18. Internationaler Kongress der Versicherungsmathematiker, Band II.

[6] JUNG, J. (1965), "On Automobile Insurance Ratemaking," *The ASTIN Bulletin,* Vol. V, Part I.

[7] MEHRING, J. (1964), *Ergebnisse einer Stichprobenuntersuchung in der deutschen Kraftfahrt — Haftpflichtversicherung,* 17th International Congress of Actuaries, Vol. III, Part II.

[8] SEAL, H. L. (1968), *The Use of Multiple Regression in Risk Classification Based on Proportionate Losses,* 18. Internationaler Kongress der Versicherungsmathematiker, Band II.

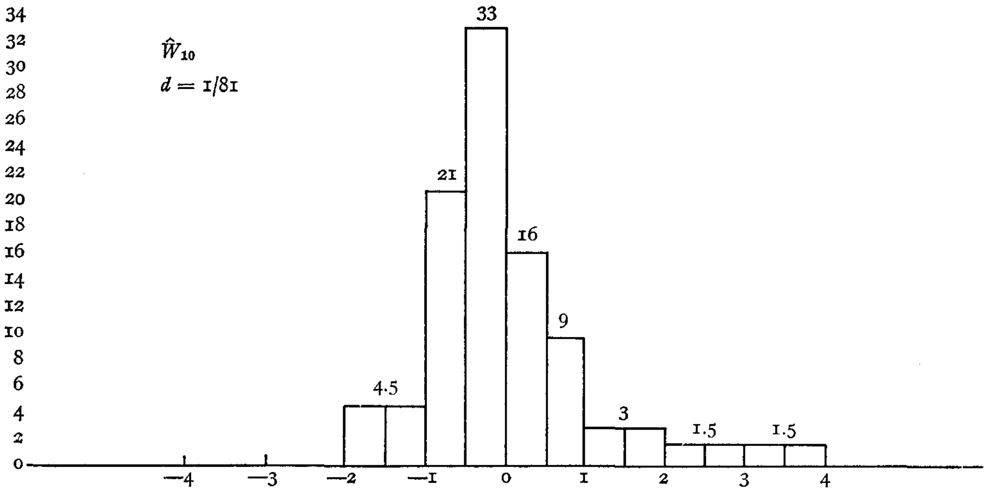

Fig. 1

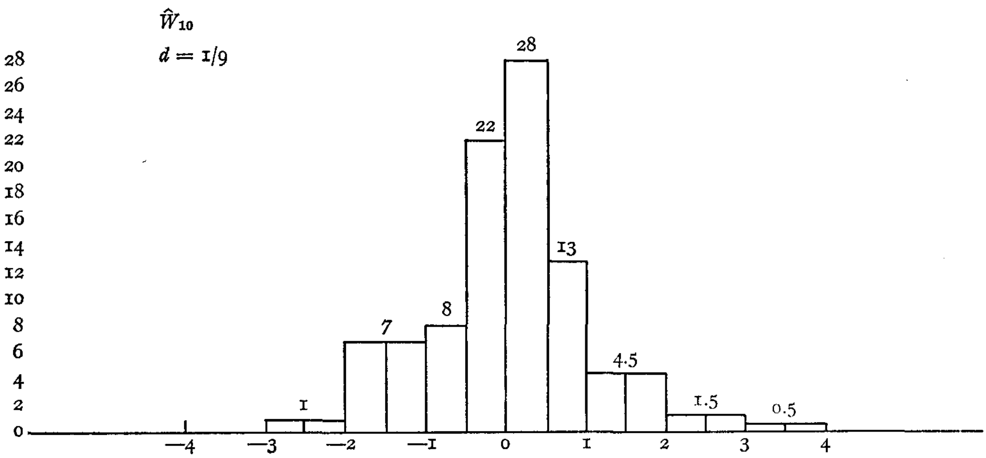

Fig. 2

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press

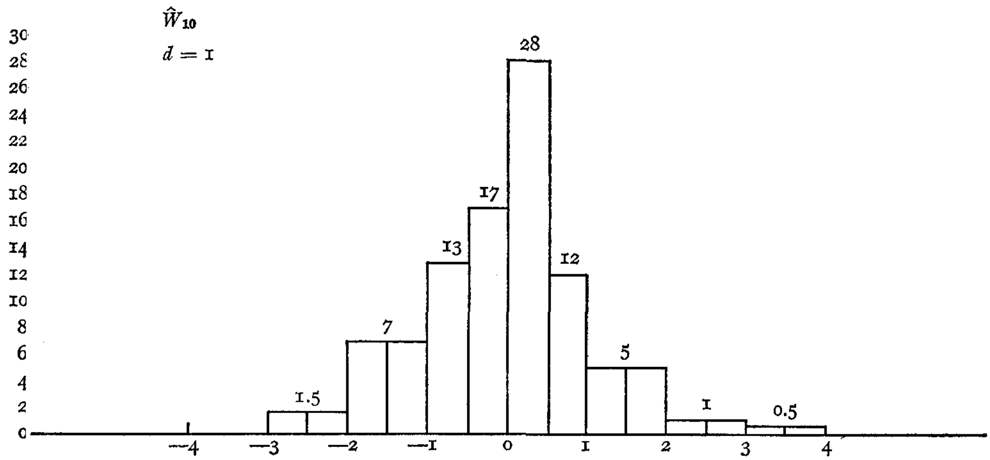

Fig. 3

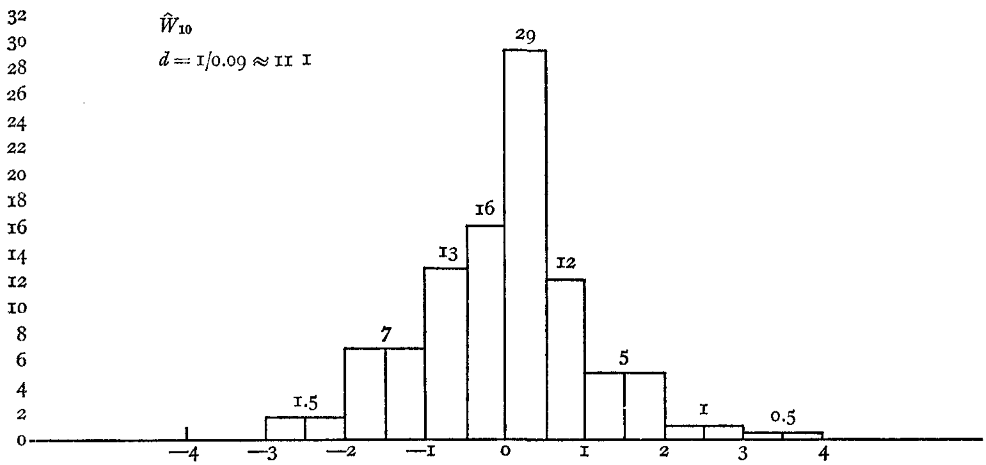

Fig. 4

https://doi.org/10.1017/S0515036100009296 Published online by Cambridge University Press