Scandinavian Actuarial Journal, 2014
Vol. 2014, No.8, 690–713, http://dx.doi.org/10.1080/03461238.2012.760885

Inference in multiplicative pricing

STIG ROSENLUND*

Länsförsäkringar Alliance, Aktuarie Sak (retired), Stockholm, Sweden

In multiplicative pricing of non-life insurance, we report a simulation study of mean square errors (MSEs) of point estimates by 1) the marginal totals method, and 2) the Standard GLM method of Poisson claim numbers and gamma distributed claim severities with constant coefficient of variation. MSEs per tariff cell are summed with insurance exposures as weights to give a total MSE. This is smallest for Standard GLM under the multiplicative assumption. But with moderate deviations from parameter multiplicativity, the study indicates that the marginal totals method is typically better in the MSE sense when there are many arguments and many claims, i. e. for mass consumer insurance. A method called MVW for confidence intervals, using only the Compound Poisson model, is given for the marginal totals method. These confidence intervals are compared with the ones of Standard GLM and the Tweedie method for risk premiums in a simulation study and are found to be mostly the best. The study reports both cover probabilities, which should be close to 0.95 for 95 % confidence intervals, and interval lengths, which should be small. The Tweedie method is found to be never better than Standard GLM.

Keywords: Confidence intervals; GLM; Marginal totals method; Multiplicative tariff; Point estimates

1. Introduction

GLM models and methods for multiplicative pricing are commonly used in non-life insurance. Software is available with algorithms using the numerical Newton-Raphson method, which is very fast even with many rating factors (arguments). Appendix A.1 defines the methods MMT (Method of Marginal Totals), S-GLM (Standard GLM), and Tweedie with exponent $p \in [1, 2]$. Exponent $p = 1$, the Overdispersed Poisson model, gives the same point estimates as MMT. The S-GLM and Tweedie models yield confidence intervals for rating factor estimates. Insurance exposure weighted MSE (= Mean Square Error) is defined in Appendix A.2.1.

The purposes of this paper are to

(1) Give for MMT a variance estimate method, giving confidence intervals, that is better than using the Overdispersed Poisson model. It is described in Appendix C. We call the method

MVW = MMT Variance estimates under Weak assumptions.

(2) Compare the methods' point estimate MSEs for simulated cases under some moderate deviations from parameter multiplicativity. Results are given in Section 3.

*E-mail: stig.rosenlund@sverige.nu

Pdf by the author. Page breaks as in the printed article.

(3) Compare the methods' confidence interval accuracy for simulated cases with parameter multiplicativity and varying degrees of departure from the S-GLM assumptions. Results are given in Section 4.

All GLM theory needed for this article can be found in Ohlsson & Johansson (2010) and in Rosenlund (2010), where also references to original GLM contributions can be found.

The paper is organized this way. Section 2.1 argues for the importance to consider deviations from parameter multiplicativity and summarizes simulation results for point estimates. Section 2.2 stresses the need to consider deviations from the S-GLM claim severity variance function and summarizes simulation results for confidence intervals. Section 3 gives detailed simulation results for point estimates and Section 4 gives detailed simulation results for variance estimates and confidence intervals. After a conclusion in Section 5, the GLM methods and the new MVW method for variance estimates are summarily described in Appendix A.1. Appendix A.2 gives basic assumptions and notation. Of the appendices only A is needed to understand the simulation results. In Appendix B a detailed description is given for the MMT and S-GLM point estimate equations. This description is a starting point for the development of MVW in Appendix C. This development is hard to understand. It is necessarily a mixture of rigorous mathematics, complicated algorithms and approximations with roots in time-honoured actuarial tradition, and heuristics. MVW has been used for many years by the author and other actuaries at Länsförsäkringar Alliance. A simpler method, introduced by the author in that company 1984, is one of the building blocks of the MVW development and is also described.

Free program for GLM, MVW etc.: `www.stigrosenlund.se/rapp.htm`

2. Summary of simulation results

2.1. Point estimates

In real applications with many arguments one cannot avoid substantial departures from risk premium multiplicativity. Not all or even most such interactions can be handled by combining arguments, e. g. sex and age. Even so a multiplicative tariff is very often the only feasible rating rule. The degree of sensitivity to departures is important. Hence we are concerned with robustness against the multiplicativity assumption.

We report here simulation results indicating that the MMT method is often preferable to S-GLM and to the Tweedie method, even though the variance function $v(\mu) = \mu$ is too flat to be realistic for claim cost. Namely, if risk premium is not strictly multiplicative, we show, for a series of simulation cases with increasing degrees of non-multiplicativity, how the MMT method becomes more and more preferable as measured by exposure-weighted MSEs, squared deviations, of estimated risk premiums from true risk premiums. In the MSE sense the MMT method is more likely to be preferable for mass consumer insurance with many arguments and many claims than it is for insurance with few arguments and few claims. The MMT method has minimum bias 0 on the marginals, so a fair guess is that it normally has minimum bias compared with other GLM methods in the tariff cells as well, implying minimum MSE for sufficiently many claims. This guess is substantiated by our simulations. The variances of the S-GLM point estimates will however typically be smaller.

We do not here give detailed results for the method of directly minimizing the exposure-weighted sum of squared deviations of estimated multiplicative risk premiums from observed risk premiums per tariff cell. This is excruciatingly slow to use with many tariff cells. The Newton-Raphson method cannot be used for these etimates. An iterative solution like the classical non-GLM one for MMT must be used. See Mildenhall (1999). We have made simulations, which probably will not be published, for a moderate number of tariff cells, i.e. less than 5,000,000. Counter-intuitively, compared to MMT and S-GLM the method mostly yielded a *larger* expected value of the MSE as defined here, i.e. the corresponding exposure-weighted sum with true risk premiums used instead of observed ones.

2.2. Confidence intervals

The assumption that a variance function exists, i. e. that homoscedasticity (equal variance) holds for a transformation of claim cost or claim severity, is strong. It cannot be justified as more than a crude approximation of reality. What can be said is this: If a variance function $v(\mu) = \mu^p$ must be used for claim severity (= claim amount), then $p = 2$ as in S-GLM is the best choice. It means that a claim amount Z has $\mathrm{Var}[Z/\mathrm{E}[Z]]$ and thus its CV (coefficient of variation) constant. But when there are large claims, which are more frequent in some classes than in others, this assumption is likely to be considerably violated.

So there is a need for methods which give confidence intervals and are fast to use without a homoscedasticity assumption. Such methods would entail far less need to truncate large claim amounts to get not too unreasonable inference. The weaker assumptions the better. In particular this is needed for the MMT method in view of the alternative to use the unrealistic variance function $v(\mu) = \mu$. We will give such a method here. It gives confidence intervals where all classes have positive widths, whose interpretation will be explained. Intervals with width 0 for a base class per argument, like the GLM variance function based intervals, are also given in the new method. We call this the MVW method, where W stands for weak as in weak assumptions.

The Poisson assumption for claim numbers is however kept. See Rosenlund (2010) for an investigation and time-honoured references. There it is shown that it is wrong to use an Overdispersed Poisson model for claim numbers, with a dispersion parameter independent of time length and of exposure generally, due to random independent claim frequencies. All Overdispersed Poisson process claim number models are Compound Poisson. Also, for mass consumer insurance a model with random independent frequencies, giving e. g. claim numbers with negative binomial distributions, is shown to be an unnecessary complication. The simple Poisson model gives the same results for practical purposes. This is due to a result by Grigelionis (1963).

Simulation results are reported for confidence intervals with the methods MVW, the 1984 method preceding MVW, S-GLM and Tweedie with suitably chosen exponent. For S-GLM the χ^2-based Pearson dispersion parameter estimate using individual claims is used. For a motivation, see Ohlsson & Johansson (2010), Chapter 3. We report cover probabilities for 95 % intervals and mean interval widths. The closer to 0.95 the former are, and the shorter the latter are, the better. Although mean interval widths are smaller for S-GLM because of this method's smaller point estimate variances, we conclude that the MVW method is mostly the best. Also we conclude that the Tweedie method should not be used. It is never better than S-GLM and necessitates additional work to estimate the exponent.

3. Simulations of MSEs of risk premium point estimates

S-GLM risk premium estimates can be expected to have smaller variances than those of MMT in real applications. See next last paragraph of Appendix B. As exposure $\to \infty$ with a factor $c \to \infty$ common to all tariff cells, we can expect an MMT estimate to have variance k_1/c and the S-GLM estimate for the same tariff cell to have variance k_2/c, with $k_1 > k_2$.

However, if risk premium is not exactly multiplicative, the mean square deviation of a tariff cell risk premium estimate from the true value will converge to a value > 0 as $c \to \infty$, not converge to 0 along with variances. Thus we study the MSE measure defined in expression (A.1) Let the exposures be defined, like in (3.1), as

$$e_u = ce_u^0,$$

and define the MSEs

$$\mathcal{M}_M(c) = \mathcal{M}(\text{MMT}, \{ce_u^0\}),$$
$$\mathcal{M}_S(c) = \mathcal{M}(\text{S-GLM}, \{ce_u^0\}).$$

We conjecture that typically the limiting MSEs as $c \to \infty$, i.e. the squared biases, obey

$$\lim_{c\to\infty} \mathcal{M}_M(c) < \lim_{c\to\infty} \mathcal{M}_S(c).$$

If this is true, then there should be an indifference value c_0 of c where $\mathcal{M}_M(c) = \mathcal{M}_S(c)$, below which $\mathcal{M}_M(c) > \mathcal{M}_S(c)$ and above which $\mathcal{M}_M(c) < \mathcal{M}_S(c)$. If variances k_1/c for MMT and k_2/c for S-GLM, where $k_1 > k_2$, hold asymptotically also for $c \to 0$, then this would follow from the identity "(mean square error) = variance + (square of bias)" generalized to the collection of all tariff cells. (We disregard the possibility that the equation $\mathcal{M}_M(c) = \mathcal{M}_S(c)$ might have several solutions $c > 0$.) We give simulation results of four cases corroborating this conjecture. It is shown that the more deviations from risk premium multiplicativity in the case, the smaller is the indifference value c_0. Judging by the expected total numbers of claims, the MMT method should be preferable in a typical consumer insurance application in a middle-size to large company.

The inequality above is not, however, universally true. We have made calculations for the simple 2×2 case, and in one of about 10 cases we found the reverse inequality. This case had multiplicative mean claim, which speaks for S-GLM, and non-multiplicative claim frequency.

We describe the cases in detail. Thus our simulations can be repeated and checked. There are six arguments denoted T_1, T_2, T_3, T_4, T_5 and T_6, each $\in \{1, \ldots, 13\}$. Define arrays

k	1	2	3	4	5	6	7	8	9	10	11	12	13
f_k	0.50	0.58	0.67	0.75	0.83	0.92	1.00	1.08	1.17	1.25	1.33	1.42	1.50
g_k	0.54	0.62	0.69	0.77	0.85	0.92	1.00	1.08	1.15	1.23	1.31	1.38	1.46

$$h_{i,k} = \begin{cases} f_k, & i \leq 6, \\ g_k, & i \geq 7, \end{cases} \quad \text{and} \quad d^{(1)}_{i,k} = 1/(1 + |i-k|) \text{ correctly rounded to four decimals.}$$

$$e_u = ce^0_u = c\ 3.74\ d^{(1)}_{T_1,T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6} = \text{exposure in tariff cell } u = (T_1, \ldots, T_6). \quad (3.1)$$

That is, exposure is non-multiplicative in (T_1, T_2) and otherwise multiplicative.

Claim frequencies and mean claims are defined in four cases.

Case	Claim frequency	Mean claim
1	$0.2 f_{T_1} f_{T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$	$5000 f_{T_1} f_{T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
2	$0.2 f_{T_1} h_{T_1,T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$	$5000 f_{T_1} h_{T_1,T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
3	$0.2 f_{T_1} h_{T_1,T_2} h_{T_2,T_3} f_{T_4} f_{T_5} f_{T_6}$	$5000 f_{T_1} h_{T_1,T_2} h_{T_2,T_3} f_{T_4} f_{T_5} f_{T_6}$
4	$0.2 f_{T_1} h_{T_1,T_2} h_{T_2,T_3} h_{T_3,T_4} f_{T_5} f_{T_6}$	$5000 f_{T_1} h_{T_1,T_2} h_{T_2,T_3} h_{T_3,T_4} f_{T_5} f_{T_6}$

In Case 1 claim frequency and mean claim are strictly multiplicative. In Case 2 there is a deviation from multiplicativity such that, if $T_1 \leq 6$ then the factor series f_k applies for T_2, but if $T_1 \geq 7$ then the slightly less steep factor series g_k applies for T_2. In Case 3, in addition to Case 2, the less steep factor series applies for T_3 if $T_2 \geq 7$. In Case 4, in addition to Case 3, the less steep factor series applies for T_4 if $T_3 \geq 7$. The cases are thus ordered with respect to increasing degree of deviation from multiplicativity.

The claim severity distribution is that claims are exponentially distributed except when $T_1 = 5$ or $T_2 = 9$. This exponential distribution function is $1 - e^{-\beta x}$ with β = 1/(mean claim). When $T_1 = 5$ or $T_2 = 9$, then with probability $p_0 = 0.95$ and with a = 0.1 the claim has distribution function $1 - e^{-(\beta/a)x}$ and with probability $1 - p_0$ the claim takes the value (mean claim)$(1 - p_0 a)/(1 - p_0)$ = (mean claim)18.1. This leaves mean claim unchanged, but introduces large claims for some tariff cells. For realism the claims should not be too well behaved.

In Table 1 estimates $\widehat{\mathcal{M}}_M(c)$ and $\widehat{\mathcal{M}}_S(c)$ are given with standard errors and the best method that can be inferred. "Best" means here the method with smallest MSE. A question mark indicates that the standard errors are too large relative to the difference in MSE for large certainty of the best method. E.g. for Case 4, line 2, we can compute confidence intervals $139,381 \pm 1.96 \times 704$ for MMT and $139,719 \pm 1.96 \times 736$ for S-GLM. They overlap substantially, so no conclusion as to the best method can be drawn. If a stated best method precedes the question mark, as for Case 3, lines 2 and 3, then we have a moderate certainty of the best method. This is somewhat vague, but we do not want to burden the account in Table 1 with the formal apparatus of tests.

The lines with $c = \infty$ in Table 1 were obtained by setting exactly (number of claims) = (claim frequency)(exposure) and (claim amount) = (mean claim), not by simulating Poisson claim numbers and random claim amounts.

Table 1. MSEs

Case	Number of repetitions	c	MMT $\widehat{\mathcal{M}}_M(c)$	Standard error	S-GLM $\widehat{\mathcal{M}}_S(c)$	Standard error	Best method	Average number of claims
1	200	1.00	9,081	147	7,906	140	S-GLM	1,624,001
1	1	∞	0		0		S-GLM	
2	800	0.40	25,781	176	24,767	159	S-GLM	644,932
2	800	0.50	21,634	148	21,099	140	S-GLM	806,203
2	2000	0.60	18,745	75	18,654	76	S-GLM?	967,413
2	800	0.70	16,697	103	16,921	97	MMT?	1,128,747
2	800	0.80	15,063	87	15,497	88	MMT	1,289,873
2	1	∞	4,448		6,383		MMT	
3	800	0.05	170,200	1455	159,064	1363	S-GLM	80,190
3	800	0.10	89,694	679	87,604	671	S-GLM?	160,410
3	200	0.15	61,440	807	63,970	972	MMT?	240,575
3	800	0.20	48,641	353	51,898	379	MMT	320,780
3	200	0.25	39,973	559	43,490	643	MMT	400,998
3	1	∞	7,880		16,040		MMT	
4	800	0.03	269,793	2182	257,429	2064	S-GLM	47,847
4	2000	0.06	139,381	704	139,719	736	?	95,692
4	800	0.08	106,412	871	111,121	910	MMT	127,594
4	800	0.10	87,159	641	94,227	766	MMT	159,478
4	200	0.20	50,829	652	60,971	841	MMT	319,013
4	1	∞	11,491		25,844		MMT	

The MSEs for the Tweedie method for risk premium with exponent $p \in \{1.0, 1.1, \ldots, 2.0\}$ have also been studied, but only for $c = \infty$. The reason for excluding studies for finite c is obvious from Table 2. For Case 1 the MSEs are 0. For $p = 1$ the values for MMT are reproduced, since the point estimates are here the same as those of MMT. For $p = 1.5$ the values for S-GLM are nearly exactly reproduced. The series from 1.0 to 2.0 is increasing for each Case 2-4. So the closer the Tweedie method is to MMT, the better it is with many claims and the worse with few claims. Also it is obvious that exponent 2, the "direct" method, should not be used.

We summarize the results of Table 1 in Table 3 with a rough estimate of the indifference value c_0 of c where $\mathcal{M}_M(c) = \mathcal{M}_S(c)$ and an approximate expected number

Table 2. Asymptotic MSEs for the Tweedie method

p	Case 2	Case 3	Case 4
1.00	4,448	7,880	11,491
1.10	4,711	8,751	12,958
1.20	5,023	9,887	14,917
1.30	5,398	11,380	17,534
1.40	5,841	13,369	21,076
1.50	6,383	16,046	25,860
1.60	7,050	19,663	32,360
1.70	7,882	24,548	41,231
1.80	8,925	31,190	53,338
1.90	10,254	40,208	69,801
2.00	11,952	52,369	92,151
S-GLM	6,383	16,040	25,844

Table 3. Limits for best method

Case	Indifference value $\hat{c}_0$	Average number of claims
1	∞	
2	0.65	1,048,000
3	0.12	192,500
4	0.06	95,692

of claims at c_0. The more deviations from multiplicativity, the smaller are c_0 and the expected number of claims at c_0. This implies that, for mass consumer insurance with many arguments and many claims, the MMT method is best in the sense of this section. The ability to obtain variance estimates and confidence intervals for the MMT estimates under weak assumptions with the MVW method, which will be described in Appendix C, is another and unrelated reason for preferring the MMT method in many situations.

4. Simulations of confidence intervals

4.1. Overview

We are here studying 95 % level confidence intervals in the GLM form, i.e. zero width for a base class with factor 1 is applied to make possible a comparison of the methods. Base class 1 is used. For MVW the form is expression (C.30) using Equation (C.27).

Claim frequency and mean claim are multiplicative, since we do not want to burden the comparisons by having to ascertain what parameters – different for different methods – we are estimating. See the end of Appendix C.1.

We compare MVW, the 1984 method, S-GLM and Tweedie.

In Section 4.3 we study six simulation cases where the claim frequency and mean claim factor series are equal. The Tweedie model with exponent $p = 1.5$ can here be shown to be true when the S-GLM model is true. For its confidence intervals we used the Pearson χ^2 dispersion parameter estimate, since the simulation cases resemble Case 4 in Section 4.2 of Rosenlund (2010). The results were the same as for S-GLM. Exponents 1 and 2 were also studied. They gave worse results than S-GLM regardless of whether the Pearson estimate or the estimate of Equation (8) of Rosenlund (2010) were used. The Tweedie results are not tabulated.

We also studied a case with constant claim severity CV where all mean claim factors are 1, which can be shown to imply that the Tweedie model with exponent $p = 1$ (ODP = Overdispersed Poisson) is true. Also we studied a case with constant claim severity CV where all claim frequency factors are equal to 1, implying that the Tweedie model with exponent $p = 2$ is true. The results were as S-GLM with the correct p and the proper dispersion parameter estimate used. Using $p = 1.5$ gave worse results. S-GLM was equal to MVW in the first case and better than MVW in the second case. These cases are not tabulated.

4.2. Variations in distributional form

The cases 1, 3, 5 below have exponential claim severity distribution with constant CV. For the cases 2, 4, 6 this distribution is modified with a large claim probability for some argument classes. The distributional form might conceivably influence the results.

Therefore we have also studied cases equal to case 1, except that the claim severity Γ-distribution has mean θ_u and variance θ_u^2/α for $\alpha \in \{0.5, 1.5, 2\}$. When $\alpha = 1$ it is the exponential distribution. The results were not different from the results for the corresponding cases with exponential distribution. Mean interval widths and cover probabilities were not affected.

We studied also cases equal to cases 3 and 5 respectively, except that claims Z are distributed as $z_0 e^T$, where T has distribution function $1 - e^{-\alpha t}$, with $\alpha = 2.1$. That is, (European) Pareto distributed claims with distribution function $1 - (z/z_0)^{-\alpha},\ z > z_0$. Here the CV is constant, but the distribution is far more heavy-tailed than the Γ-distribution. Distributions in applications have mostly heavier tails than the Γ-distribution. If the Pareto distribution has parameter $\alpha \leq 2$, it has no variance. So the value 2.1 is the smallest multiple of 0.1 that gives the claim amounts finite variance. The Γ- and Pareto distributions are examples of about the most light-tailed (excluding the trivial case of constant claim amounts) and heavy-tailed distributions with constant CV that are realistic in applications.

The Pareto results, for mean interval widths and cover probabilities, of comparisons between the MVW, 1984 and S-GLM methods were not very different from the results for the corresponding cases with exponential distribution. Cover probabilities were the same. Mean width for MVW in percent of the one for S-GLM was somewhat smaller, namely 108.9 instead of 116.2 for Case 3 and 117.8 instead of 120.5 for Case 5. The form of claim severity distribution is not important for the S-GLM confidence intervals to be valid, only the assumption of constant CV.

The case as 3, except that Pareto was used, was studied also for Tweedie with exponent 1.5. This gave too small cover probabilities and interval widths. Mean cover probability was 0.943, which is significantly below 0.95, and width 97.7 percent of S-GLM.

4.3. Six cases where claim frequency factor equals mean claim factor

The cases studied resemble Case 1 of Section 3, but exposure is fixed for every case. We use the notation of Section 3 and also

$d_{i,k}^{(2)} = 1/(1 + 4(i-k)^2)$,

correctly rounded to four decimals, for instance $d_{2,7}^{(2)} = 1/(1 + 4(2-7)^2) = 1/101 = 0.0099$. We define exposure by means of this expression in some cases, in order to introduce strongly non-multiplicative exposure. This will test the MVW method.

Confidence intervals on the 95 % level for risk premium factors $\gamma_{jk}^B = \gamma_{jk}/\gamma_{j1}$ ($k = 2, \ldots, 13$) were studied. See Section A.2 for the definition of risk premium factors. All cases have six arguments with 13 classes each, as in Section 3. Claim frequency and mean claim are as in Case 1 of Section 3. The claim severity distribution functions are as follows.

F_1 $1 - e^{-\beta x}$ with $\beta = 1/(\text{mean claim})$. This satisfies the S-GLM assumptions.

F_2 as in Section 3, i. e. with a large claim probability if $T_1 = 5$ or $T_2 = 9$. The CV is not constant, so S-GLM is not satisfied.

The cases are defined by the following table. Half of them satisfy the S-GLM assumptions. One third have strongly non-multiplicative exposure in two thirds of the arguments. Our experience is that real applications typically satisfy the S-GLM assumptions to a lesser degree and have more multiplicative exposure.

The total expected number of claims is between 161,000 and 169,000.

For Case 1, the three confidence interval methods MVW, 1984, and S-GLM can be expected to give cover probabilities 0.95 for all risk premium factors on the 95 % level.

Case	df	CV	Exposure multiplicativity	Exposure expression
1	F_1	Constant	Multiplicative	$0.1 f_{T_1} f_{T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
2	F_2	Non-constant	Multiplicative	$0.1 f_{T_1} f_{T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
3	F_1	Constant	Non-multiplicative	$0.374 d^{(1)}_{T_1,T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
4	F_2	Non-constant	Non-multiplicative	$0.374 d^{(1)}_{T_1,T_2} f_{T_3} f_{T_4} f_{T_5} f_{T_6}$
5	F_1	Constant	Strongly non-multiplicative	$0.8156\, d^{(2)}_{T_1,T_2} d^{(2)}_{T_3,T_4} f_{T_5} f_{T_6}$
6	F_2	Non-constant	Strongly non-multiplicative	$0.8156\, d^{(2)}_{T_1,T_2} d^{(2)}_{T_3,T_4} f_{T_5} f_{T_6}$

4.3.1. Cover probabilities

In Case 1, 1000 repetitions showed that all cover probabilities were 0.95 as expected. We made 8000 repetitions for each of the remaining cases. If a cover probability for an argument and a class is 0.95, then the cover number = (number of experiments N where the true parameter is in the confidence interval) is binomially distributed (8000,0.95). By the normal approximation, the hypothesis that the cover probability is 0.95 is rejected with significance level 0.05 against a two-sided alternative, if $|N/8000 - 0.95| \geq 1.96\sqrt{(0.95\ 0.05)/8000} = 0.00478$. If the hypothesis is true, the probability is 0.0000406 that $|N/8000 - 0.95| \geq 0.01$, i. e. that the cover frequency is $<$ 0.94 or $>$ 0.96. We report here the arguments with those deviating cover frequencies, which are highly significant as well as of some importance. We give the cover percent $= 100N/8000$.

Case 2. The higher claim severity CV for $T_1 = 5$ and $T_2 = 9$ results in too low cover percents implying too narrow S-GLM confidence intervals for these classes, and too high cover percents implying too wide S-GLM intervals for other classes in arguments n:o 1 and 2. Other arguments for S-GLM, and all arguments for the MVW and 1984 methods, gave results $\in [94, 96]$.

Case 3. One deviating cover percent 93.95 was found, for the 1984 method.

Case 4. Same phenomenon as in Case 2 for S-GLM. The MVW method gave results all $\in [94, 96]$, while the 1984 method gave six results $\in [93.8, 94)$.

Case 5. Since the S-GLM assumptions are satisfied, its results were all $\in [94, 96]$. Strongly non-multiplicative exposure gave results $\notin [94, 96]$ for the MVW and 1984 methods. The MVW method sometimes understated and sometimes overstated confidence intervals somewhat, while the 1984 method understated them considerably. This

Table 4. Case 2 S-GLM cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	96.9	96.8	96.7	86.5	96.8	96.8	96.6	96.9	96.8	96.9	97.0	97.1
T_2	98.0	98.4	98.4	98.2	98.4	98.4	98.4	92.4	98.6	98.2	98.4	98.3

Table 5. Case 4 S-GLM cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	98.5	98.4	98.3	88.0	97.9	97.7	97.8	97.2	97.9	98.2	98.0	97.9
T_2	98.0	97.8	97.5	96.7	97.8	98.0	98.3	91.2	98.2	98.0	98.2	98.5

Table 6. Case 5 MVW cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	93.5	93.1	92.9	93.4	94.0	94.5	94.7	95.8	96.1	96.6	96.8	97.4
T_2	93.1	92.5	92.8	93.5	93.5	94.3	94.9	95.8	96.1	96.4	96.9	97.1
T_3	93.3	93.1	92.9	93.2	94.0	94.5	95.5	95.8	96.0	96.6	97.2	97.3
T_4	93.7	93.3	92.9	93.7	94.3	94.6	95.2	95.8	96.1	96.6	96.9	97.2

Table 7. Case 5 1984 cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	89.6	85.9	84.3	82.7	81.5	80.8	80.3	79.8	79.4	78.6	78.0	76.8
T_2	89.3	86.1	83.7	83.1	81.0	80.4	79.9	80.0	78.9	78.4	77.9	77.0
T_3	89.4	86.1	84.2	83.2	82.3	81.2	80.5	80.1	79.0	78.7	77.7	77.9
T_4	89.7	86.4	84.3	83.3	82.2	82.0	80.4	80.4	79.8	79.3	78.9	77.8

Table 8. Case 6 MVW cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	92.3	90.6	90.4	99.1	91.5	92.9	96.5	99.8	96.4	94.1	93.9	94.0
T_2	92.1	91.7	93.5	98.4	93.2	91.4	91.1	99.975	91.6	91.9	93.1	93.4
T_3	93.0	92.4	92.6	93.2	93.6	94.7	95.6	96.1	96.4	97.2	97.3	97.7
T_4	93.1	92.6	92.7	93.4	93.7	94.6	95.2	96.0	96.2	96.8	97.5	97.7

Table 9. Case 6 1984 cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	88.6	83.3	79.6	87.2	77.5	76.6	76.9	80.7	73.9	72.8	72.1	70.9
T_2	88.4	84.2	82.6	86.5	78.4	77.3	76.0	81.7	74.0	73.2	73.1	72.4
T_3	89.6	86.0	83.8	82.6	81.7	81.1	80.4	79.4	78.7	78.5	77.6	77.2
T_4	89.6	86.3	84.4	82.9	81.5	81.0	80.9	79.8	80.0	79.1	78.6	77.8

Table 10. Case 6 S-GLM cover percentages

	2	3	4	5	6	7	8	9	10	11	12	13
T_1	99.2	99.0	98.7	93.6	98.3	97.9	97.9	97.1	97.6	97.9	97.9	97.7
T_2	99.0	98.4	97.7	96.2	97.4	97.6	97.5	95.1	97.7	97.7	97.8	97.8

Table 11. Mean confidence interval widths and cover percentages

	Mean width % of S-GLM		Mean cover percentages		
Case	MVW	1984	MVW	1984	S-GLM
1	113.6	113.6	95.50	95.50	95.22
2	103.9	103.9	94.85	94.85	95.82
3	116.2	115.7	94.94	94.82	95.06
4	104.8	104.2	94.78	94.65	95.70
5	120.5	86.6	94.99	86.16	94.96
6	106.3	74.2	94.57	85.04	95.96

applies to the arguments T_1, T_2, T_3 and T_4 where exposure is non-multiplicative. The other arguments gave results $\in [94, 96]$. T_1 - T_4 have the same cover probability series due to symmetry.

Case 6. S-GLM assumptions are not satisfied and exposure is strongly non-multiplicative. Here all methods fail to some degree in arguments T_1 and T_2. The MVW method fails to some degree also in T_3 and T_4. The 1984 method fails in the same way as in Case 5.

Overall the MVW method gave the best results in Case 6, if measured by the mean percentages in Table 11. However, the method can be improved in view of the high values in Table 8 for classes 5 and 9 in arguments T_1 and T_2, respectively. The highest percentage 99.975 corresponds to an interval width about twice as large as it should be.

4.3.2. Mean widths and cover percentages

We give mean widths as percentages of the ones of S-GLM, averaged over all arguments and classes ≥ 2 per case, and mean cover percentages. Cases 1, 3, 5 satisfy the S-GLM assumptions.

5. Conclusions

The simulations of Section 3 show that with strict parameter multiplicativity the S-GLM point estimates are best in the MSE sense. Allowing moderate deviations from parameter multiplicativity the MMT point estimates are best with sufficiently many arguments and claims.

The simulations of Section 4 show that the MVW variance estimate method is mostly the best, due to the weaker assumptions imposed by it.

The Tweedie method for risk premium is never better than S-GLM and sometimes worse. So it should not be used. This holds when the exponent p is known and hence à fortiori when it must be estimated. As a corollary, Overdispersed Poisson (Tweedie with exponent 1) and the "direct" method (Tweedie with exponent 2) should not be used.

There is room for improvement of the estimate $\hat{v}_{\text{MVW}}(\hat{\rho}_{jk})$ in Equation (C.26) and as a corollary $\hat{v}_{\text{MVW}}(\hat{\gamma}_{jk})$ in Equation (C.27). Such improvement, valid for sufficiently many claims and using only strict mathematics with no other conditions than the Compound Poisson assumption and multiplicativity of claim frequency and mean claim, would be very desirable.

Acknowledgement

Thanks are due to the referee for many valuable suggestions, which helped to improve the paper.

References

Grigelionis, B. (1963). On the convergence of sums of random step processes to a Poisson process. *Theory of Probability and its Application* **8** (2), 177–182.

Mildenhall, S. J. (1999). A Systematic Relationship Between Minimum Bias and Generalized Linear Models. *Proceedings of Casualty Actuary Society* **86**, 393–487.

Ohlsson, E. & Johansson, B. (2010). *Non-Life Insurance Pricing with Generalized Linear Models.* Springer, Berlin.

Rosenlund, S. (2010). Dispersion estimates for Poisson and Tweedie models. *ASTIN Bulletin* **40** (1), 271–279.

Appendix A: Concepts and notation

As in Ohlsson & Johansson (2010), CV is the coefficient of variation of a random variable.

A.1. Methods

A.1.1. Point estimate methods

The following notation will be used for three methods of obtaining point estimates in multiplicative pricing, namely MMT, S-GLM and Tweedie.

MMT. Method of Marginal Totals. Solves a system of equations defined by prescribing that the sum of multiplicatively computed estimated claim costs over any argument class be equal to the empirical claim cost of the argument class.

S-GLM. Standard GLM method. A model that claim numbers are Poisson and mean claim severities are gamma distributed with constant CV yields a set of GLM Poisson log link equations for claim frequencies and a set of GLM gamma log link equations for mean claim severities. The estimated multiplicative claim frequency and claim severity per tariff cell are multiplied to give a point estimate for risk premium.

Tweedie. An exponent $p \in [1, 2]$ is attached to the method. A GLM model that claim cost has variance function $v(\mu) = \mu^p$ gives point estimates for risk premium. The case $p = 1$, Overdispersed Poisson, gives the same point estimates as MMT. The case $p = 2$ was historically denoted the "direct" method.

A.1.2. Variance estimate methods

The GLM models giving point estimates by S-GLM and Tweedie also give variance estimates, so we use the notation S-GLM and Tweedie for these variance estimate methods as well. For MMT the Overdispersed Poisson model for claim cost is however not realistic. Instead we attach the following variance estimate methods to MMT.

1984. A Compound Poisson model for the claim cost of a tariff cell is assumed. Simple estimated CVs for marginal (univariate) claim numbers, mean claims and claim costs are taken as valid for factor estimates. Entails a normally slight negative bias. 1984 was the first year the method was used at Länsförsäkringar Alliance. We describe it in Appendix C.

MVW. Same Compound Poisson model as for 1984. W is for weak assumptions. The GLM Poisson log link model for claim numbers follows from that model, so we use the S-GLM claim frequency variance estimates. The mean claim variance estimates of the 1984 method are adjusted upwards by factors resembling the ratios of S-GLM claim frequency variance estimates to 1984 claim frequency variance estimates. Description in Appendix C.

A.2. Basic assumptions and notation for parameters and random variables

We deal with a finite number of arguments, each with a finite number of classes. Let

s = number of arguments,
m_j = number of classes for argument j, $j = 1, 2, \dots, s$.
A combination of classes, where risk premium is constant, is called a tariff cell. Let

$n = m_1 m_2 \dots m_s$ = number of tariff cells.

We will index the tariff cells with $u = 1, \dots, n$. The classes in argument n:o 1 will vary slowest with u and the classes in argument n:o s will vary fastest. Let

e_u = exposure in tariff cell u,
N_u = observed number of claims in tariff cell u,
Y_u = observed total claim cost in tariff cell u,

and

$\nu_u = \mathrm{E}[N_u/e_u]$ = claim frequency,
$\theta_u = \mathrm{E}[Y_u/N_u \mid N_u > 0]$ = mean claim,
$\tau_u = \mathrm{E}[Y_u/e_u] = \nu_u\theta_u$ = risk premium.

We assume throughout a Poisson distribution for N_u and a Compound Poisson distribution for Y_u. Variables for different u are assumed to be independent.

Define the following parameters and random variables for argument j and class k. The factors have meaning if multiplicativity holds for ν_u, θ_u and τ_u, respectively. If two are multiplicative, then the third one is also multiplicative and it holds $\gamma_{jk} = \psi_{jk}\rho_{jk}$. We assume multiplicativity for our method for confidence intervals and for the simulations on these. For the simulations of MSEs of point estimates we assume some moderate deviations from multiplicativity.

ψ_{jk} = claim frequency factor,
ρ_{jk} = mean claim factor,
γ_{jk} = risk premium factor,

e_{jk} = marginal exposure = sum of exposure for tariff cells in class k of argument j,

N_{jk} = marginal claim number = total claim number for tariff cells in class k of argument j,

Z_{jki} = claim amount n:o i, $i = 1, \ldots, N_{jk}$,

$S_{jk} = \sum_{i=1}^{N_{jk}} Z_{jki}$ = marginal claim cost,

$\overline{Z}_{jk} = S_{jk}/N_{jk}$ = marginal mean claim.

Let $T_j(u)$ = class number $\in \{1, 2, \ldots, m_j\}$ for tariff cell u in argument n:o j. Then the multiplicative assumptions are that there are constants ψ_0, ρ_0 and $\gamma_0 = \psi_0\rho_0$ such that

$$\nu_u = \psi_0 \prod_{j=1}^{s} \psi_{jT_j(u)}, \qquad \theta_u = \rho_0 \prod_{j=1}^{s} \rho_{jT_j(u)}, \qquad \tau_u = \gamma_0 \prod_{j=1}^{s} \gamma_{jT_j(u)}.$$

The estimates of ν_u and τ_u by the GLM Poisson log link (= MMT) method will be denoted $\hat{\nu}_u$ and $\tilde{\tau}_u$, respectively. A deduced estimate of θ_u is then $\tilde{\theta}_u = \tilde{\tau}_u/\hat{\nu}_u$.

The estimate of θ_u by the GLM gamma log link method will be denoted $\hat{\theta}_u$. A deduced estimate of τ_u in S-GLM is then $\hat{\tau}_u = \hat{\nu}_u\hat{\theta}_u$.

Estimates of the factors by the MMT method will be denoted $\hat{\psi}_{jk}$, $\hat{\rho}_{jk}$ and $\hat{\gamma}_{jk}$. We will not need notation for factor estimates by other methods.

A.2.1. Insurance exposure weighted MSE measure

Consider a method X with estimates $\dot{\tau}_u$ of τ_u. The MSEs per tariff cell have to be summed using weights of some kind, giving a total MSE measure. It is natural to use exposures as weights to obtain the average over the business of the MSEs. Thus we define the following exposure weighted MSE, which we use as a measure of goodness-of-fit. To suit its use in Section 3, we let the dependence on exposure be explicit in the notation, while other properties are implicit.

$$\mathcal{M}(\mathrm{X}, \{e_u\}) = \mathrm{E}\left[\sum_{u=1}^{n} e_u(\dot{\tau}_u - \tau_u)^2\right] \Big/ \sum_{u=1}^{n} e_u. \tag{A.1}$$

Appendix B: Equations for MMT and S-GLM point estimates

Denote the number of free parameters by

$$r = 1 + \sum_{j=1}^{s}(m_j - 1).$$

We follow Ohlsson & Johansson (2010), Chapter 2, in defining the design matrix

$$\mathbf{X} = \{x_{uj}\}\ (u = 1, \ldots, n;\ j = 1, \ldots, r).$$

The equations for the MMT estimates $\hat{\nu}_u$ and $\tilde{\tau}_u$ are, for $j = 1, \ldots, r$,

$$\sum_{u=1}^{n} x_{uj}\left[N_u - e_u\hat{\nu}_u\right] = 0, \tag{B.1}$$

$$\sum_{u=1}^{n} x_{uj}\left[Y_u - e_u\tilde{\tau}_u\right] = 0. \tag{B.2}$$

For S-GLM equations (B.1) are used together with mean claim equations, namely

$$\sum_{u=1}^{n} e_u\hat{\nu}_u x_{uj} = \sum_{u=1}^{n} N_u x_{uj} \quad \text{and} \quad \sum_{u=1}^{n} Y_u \frac{1}{\hat{\theta}_u} x_{uj} = \sum_{u=1}^{n} N_u x_{uj}. \tag{B.3}$$

See Ohlsson & Johansson (2010), Equation (2.30) with $p = 1$ and $p = 2$ respectively. The right hand side is the same. Replace in (B.3) the right hand side of the first set of Equations with the left hand side of the second set. Then

$$\sum_{u=1}^{n} x_{uj}\left[\frac{Y_u - e_u\hat{\nu}_u\hat{\theta}_u}{\hat{\theta}_u}\right] = 0, \tag{B.4}$$

or equivalently, since $\hat{\tau}_u = \hat{\nu}_u\hat{\theta}_u$,

$$\sum_{u=1}^{n} x_{uj}\left[\frac{Y_u - e_u\hat{\tau}_u}{\hat{\theta}_u}\right] = 0. \tag{B.5}$$

The S-GLM equations (B.5) are similar to the MMT equations (B.2) but put more weight on observations where estimated mean claim is small and less weight where it is large. This indicates that the variances of risk premium estimates from S-GLM typically will be smaller than the variances of MMT risk premium estimates, since it is realistic to assume that the variance of a claim amount is larger the larger its mean is.

If observed risk premiums Y_u/e_u are exactly multiplicative or if all observed mean claims Y_u/N_u are Equal, then it is seen that the risk premium solutions to Equations (B.2) and (B.5) are the same. (This implies that, if the true risk premiums are multiplicative, risk premium estimates converge almost surely to the true values for both S-GLM and MMT, as exposure tends to infinity with a factor c common to all tariff cells.) If not they are in general not the same. In many applications they are however very similar.

Appendix C: MVW variance estimates and confidence intervals

C.1. Overview of variance estimates built on the 1984 method

Let us start with taking simple estimated CVs for marginal (univariate) claim numbers, mean claims and claim costs as valid for factor estimates, given that the CVs of the factor estimates can be positive for all classes (see below). Only the Compound Poisson distribution is used. We call this the 1984 method, since it was first used that year at Länsförsäkringar Alliance.

This method gives approximate confidence intervals for factor estimates for risk premium, claim frequency and mean claim. They have been shown to be good approximations of the proper ones in theoretical and simulation studies, unless the exposure is extremely non-multiplicative (unevenly distributed) with respect to the arguments. Multiplicativity for exposure is defined in the same way as it is for risk premium, i. e. exposure in tariff cell n:o u is obtained by multiplying a constant ϵ_0 with factors ϵ_{jk} for $j = 1, 2, \ldots, s$, with k the class that the tariff cell belongs to in argument j. (If we consider the argument classes as random variables realized when a policy is chosen at random from the portfolio, then exposure multiplicativity means that these classes are independent for different arguments.) To show this approximation, first observe that risk premium factor estimates over an argument are proportional to univariate risk premium estimates, if exposure is multiplicative. This is a property of the marginal totals method, see below under Equation (C.7). Likewise for claim frequency, and thus also for mean claim. So in this case the 1984 method is completely justified. Secondly, analyze how deviations from exposure multiplicativity affect CVs of factor estimates. It can be shown that even rather large such deviations hardly affect these CVs. The bias in confidence interval width from using this approximation is negative, with size dependent on the degree of exposure non-multiplicativity.

We will try to eliminate in the MVW method, as far as possible, even the typically small bias of the 1984 method. Rewrite (B.2) as MMT mean claim equations by using $\tilde{\tau}_u = \hat{\nu}_u \tilde{\theta}_u$.

$$\sum_{u=1}^{n} x_{uj} \left[Y_u - e_u \hat{\nu}_u \tilde{\theta}_u \right] = 0 \quad (j = 1, \ldots, r). \tag{C.1}$$

Solving first $\hat{\nu}_u$ from (B.1) and then $\tilde{\theta}_u$ from (C.1) we get the same solution $\tilde{\tau}_u$ as from (B.2).

We will use the Fisher information matrix $\mathbf{I}$ in GLM Poisson log link theory and its estimate $\hat{\mathbf{I}}$. These matrices are indexed by F, M and R for frequency, mean claim and risk premium, respectively, and defined as follows. Let $\mathbf{diag}(d_u)$ be the diagonal $n \times n$ matrix with element d_u in row and column u. Then for the claim frequency equations (B.1) we have

$$\mathbf{I}_F = \mathbf{X}'\mathbf{diag}(e_u \nu_u)\mathbf{X}, \tag{C.2}$$

$$\hat{\mathbf{I}}_F = \mathbf{X}'\mathbf{diag}(e_u \hat{\nu}_u)\mathbf{X}, \tag{C.3}$$

and for the mean claim and risk premium Equations (C.1) and (B.2), respectively, it holds

$$\mathbf{I}_M = \mathbf{X}'\mathbf{diag}(e_u \nu_u \theta_u)\mathbf{X} = \mathbf{I}_R = \mathbf{X}'\mathbf{diag}(e_u \tau_u)\mathbf{X}, \tag{C.4}$$

$$\hat{\mathbf{I}}_M = \mathbf{X}'\mathbf{diag}(e_u \hat{\nu}_u \tilde{\theta}_u)\mathbf{X} = \hat{\mathbf{I}}_R = \mathbf{X}'\mathbf{diag}(e_u \tilde{\tau}_u)\mathbf{X}. \tag{C.5}$$

We will follow these three steps in the construction of the MVW method.

(1) Partition the square of the estimated CV for a risk premium factor estimate in the 1984 method in terms of claim frequency and mean claim.

(2) Replace the claim frequency term with an expression derived from the GLM theory for the Poisson distribution with log link, using $\hat{\mathbf{I}}_F$ of (C.3).

(3) Multiply the original mean claim term with an enlargement factor resembling the ratio between the new and the original claim frequency term, but using $\hat{\mathbf{I}}_M$ of (C.5). We let this factor go only halfway. We surmise that the enlargement factor will work about as well as the corresponding one for claim frequency. This is supported by simulations. The method is, however, heuristic and a strict method is desirable.

The new estimated CV of step 2 is used for claim frequency, and the one of step 3 is used for mean claim. The square root of the sum of their squares is our new estimated CV for a risk premium factor estimate.

For claim frequency the step 2 replacement by GLM theory and adjustment to confidence intervals with positive widths for all argument classes is well justified. It removes the bias of the 1984 method as far as claim frequency is concerned.

In step 3 we will not require homoscedasticity in some form. Like the 1984 method the MVW method will give correct confidence intervals under the Compound Poisson assumption with enough claims, when exposure is multiplicative. When exposure is markedly non-multiplicative, the MVW method is intended to adjust the mean claim part upwards relative to the 1984 method in order to compensate the negative bias of the latter method.

No base class with factor 1 is specified for $\gamma_{jk}, \psi_{jk}, \rho_{jk}$.

CVs are here formulated as functionals $v_x()$ with estimates $\hat{v}_x()$, where $x = {\scriptstyle 1984}$ denotes the 1984 estimate, $x = {\scriptstyle \text{MVW}}$ the new estimate and $x = {\scriptstyle \text{GLM}}$ an estimate from GLM Poisson log link.

We will let $\hat{v}_{1984}()$ and $\hat{v}_{\text{MVW}}()$ take positive values for all estimated factors (unless, for mean claim, all Z_{jki} as defined above happen to be equal for some j and k). This is in contrast to GLM methods with zero confidence interval width for a base class with factor 1.

These all-positive CVs can be used for confidence intervals for factor estimates computed so that their mean value weighted by exposure over an argument j is 1. This can be done if the total number of claims for all classes is so large that the total claim cost has a sufficiently small CV. See C.3.

If one still prefers to use CVs and confidence intervals with values and widths zero for the base class and larger than the all-positive variants for other classes, then these are obtained from the all-positive ones by Equation (C.30).

When risk premium is not multiplicative the confidence intervals should be interpreted as pertaining to the almost sure limiting values of the factor estimates as exposure $\to \infty$ with a factor c equal for all tariff cells. Those limiting values are in general different for the MMT and S-GLM methods and depend also on the distribution of exposure. Likewise for claim frequency and mean claim.

C.2. Properties of MMT point estimates related to the 1984 method

Let

$$e(t_1, t_2, \dots, t_s) = \begin{cases} \text{exposure } e_u \text{ for tariff cell } u \text{ that is the combination of the} \\ \text{classes } t_1, t_2, \dots, t_s \text{ for arguments } 1, 2, \dots, s, \text{ respectively,} \end{cases}$$

and

$$E_{jk} = \sum_{t_1=1}^{m_1} \dots \sum_{t_{j-1}=1}^{m_{j-1}} \sum_{t_{j+1}=1}^{m_{j+1}} \dots \sum_{t_s=1}^{m_s} e(t_1, \dots, t_{j-1}, k, t_{j+1}, \dots, t_s) \hat{\gamma}_{1t_1} \dots \hat{\gamma}_{j-1,t_{j-1}} \hat{\gamma}_{j+1,t_{j+1}} \dots \hat{\gamma}_{st_s}. \quad \text{(C.6)}$$

With $a_j = \sum_{k=1}^{m_j} e_{jk} / \sum_{k=1}^{m_j} E_{jk}$ we call $a_j E_{jk}$ normed exposure for argument j and class k. The norming is with respect to the estimated risk premium factors for other arguments than j. The formula for E_{jk} seems unwieldy, but it has long been used by actuaries working with multiplicative tariffs. Here a_j is set to make total normed exposure equal to total exposure.

It follows from the marginal totals equations in the classical non-GLM form that

$$\hat{\gamma}_{jk} = S_{jk} c_j E_{j1} / E_{jk} \text{ for some } c_j \text{ not depending on } k. \quad \text{(C.7)}$$

Here we could have written just c_j rather than $c_j E_{j1}$ for Equation (C.7) to be true, but the formulation given serves to explain relations and approximations in the sequel.

If exposure is multiplicative, then $E_{j1}/E_{jk} = e_{j1}/e_{jk}$ i. e. non-stochastic. Because, if

$$e(t_1, t_2, \dots, t_s) = \epsilon_0 \prod_{j=1}^{s} \epsilon_{jt_j},$$

then

$$E_{jk} = \epsilon_0 \left(\sum_{t=1}^{m_1} \epsilon_{1t} \hat{\gamma}_{1t} \right) \cdots \left(\sum_{t=1}^{m_{j-1}} \epsilon_{j-1,t} \hat{\gamma}_{j-1,t} \right) \epsilon_{jk} \left(\sum_{t=1}^{m_{j+1}} \epsilon_{j+1,t} \hat{\gamma}_{j+1,t} \right) \cdots \left(\sum_{t=1}^{m_s} \epsilon_{st} \hat{\gamma}_{st} \right),$$

and so

$$E_{j1}/E_{jk} = \epsilon_{j1}/\epsilon_{jk}.$$

For multiplicative exposure it holds

$$e_{jk} = \epsilon_0 \left(\sum_{t=1}^{m_1} \epsilon_{1t} \right) \cdots \left(\sum_{t=1}^{m_{j-1}} \epsilon_{j-1,t} \right) \epsilon_{jk} \left(\sum_{t=1}^{m_{j+1}} \epsilon_{j+1,t} \right) \cdots \left(\sum_{t=1}^{m_s} \epsilon_{st} \right),$$

and this implies

$$e_{j1}/e_{jk} = \epsilon_{j1}/\epsilon_{jk} = E_{j1}/E_{jk}.$$

So risk premium factor estimates over an argument are proportional to univariate risk premium estimates, if exposure is multiplicative. Likewise for claim frequency and mean claim. This is actuarial knowledge from long before GLM theory was developed. Even if exposure deviates from multiplicativity rather much, the E_{jk} will be sums of approximately independent variables and furthermore E_{j1} and E_{jk} will be positively correlated. Hence, if we let c_j be non-stochastic, the CV of $\hat{\gamma}_{jk}$ will be dominated by the one of S_{jk}. Estimating the former CV with an estimate of the latter one gives a negative bias with ordinarily small absolute value, with size dependent on the degree of exposure non-multiplicativity. But if exposure deviates from multiplicativity very much this bias can be non-negligible. See Tables 7 and 9 in Section 4.3.

If we replace $\hat{\gamma}_{jk}$ in the left side of (C.7) with $\hat{\gamma}_{jk}^{(B)}$, denoting factors with base class 1, then $c_j = 1/S_{j1}$. Then ordinarily the CV of $\hat{\gamma}_{jk}^{(B)}$ will be dominated by the one of S_{jk}/S_{j1}.

C.3. Approximate confidence intervals with all-positive widths

We will now describe the use of all-positive CVs in confidence intervals for factor estimates with exposure-weighted mean 1. Let E_{jk} above be defined from some set of risk premium factor estimates, for example with value 1 for a base level. Define a new set of $\hat{\gamma}_{jk}$ such that its exposure-weighted mean over all k is 1. That is, let

$$\hat{\gamma}_{jk} = (S_{jk}/E_{jk})\left(\sum_{t=1}^{m_j} e_{jt}\right)\left(\sum_{t=1}^{m_j} e_{jt}S_{jt}/E_{jt}\right)^{-1}, \tag{C.8}$$

in order to obtain

$$\sum_{k=1}^{m_j} \frac{e_{jk}}{e_{j1} + \cdots + e_{jm_j}}\hat{\gamma}_{jk} = 1.$$

Then for likewise exposure-weighted true risk premium factors we have

$$\gamma_{jk} \approx \mathrm{E}[S_{jk}/E_{jk}]\left(\sum_{t=1}^{m_j} e_{jt}\right)\left(\sum_{t=1}^{m_j} e_{jt}\mathrm{E}[S_{jt}/E_{jt}]\right)^{-1}. \tag{C.9}$$

If the total number of claims is large enough and if exposure is not too strongly non-multiplicative, then all components except S_{jk} in (C.8) have negligible CVs. Hence $\mathrm{E}[\hat{\gamma}_{jk}] \approx \gamma_{jk}$ and an estimate $\hat{v}_{jk}$ of the CV of S_{jk} can be applied to $\hat{\gamma}_{jk}$ and used in confidence intervals, for example

$$\mathrm{P}(\hat{\gamma}_{jk}(1 - 1.96\hat{v}_{jk}) < \gamma_{jk} < \hat{\gamma}_{jk}(1 + 1.96\hat{v}_{jk})) \approx 0.95. \tag{C.10}$$

If exposure is multiplicative, then normed exposure is proportional to exposure. Then E_{jk} can be changed to e_{jk} in the expression for $\hat{\gamma}_{jk}$ and the approximate expression (C.9) will be exact.

Instead of a simple univariate CV-estimate $\hat{v}_{jk}$ we can use the CV $\hat{v}_{\mathrm{MVW}}(\hat{\gamma}_{jk})$ of (C.27) in (C.10). This will correct the negative bias mentioned above.

C.4. Partitioning the 1984 estimate

We assume that the marginal number of claims $N_{jk} > 0$ and claim cost $S_{jk} > 0$. (In cases where this is not so, the system of equations has to be reformulated by eliminating classes with no claims or claim cost, respectively. This is done in practice but is beside the point here.)

The original 1984 CV estimates, positive for all k, are equal to the ones of the corresponding marginal statistics. Here the squares are considered.

$$\hat{v}_{1984}(\hat{\psi}_{jk})^2 = \frac{1}{N_{jk}}, \tag{C.11}$$

$$\hat{v}_{1984}(\hat{\gamma}_{jk})^2 = \left[\sum_{i=1}^{N_{jk}} Z_{jki}^2\right] / S_{jk}^2, \tag{C.12}$$

and with a natural estimate for mean claim CV

$$\hat{v}_{1984}(\hat{\rho}_{jk})^2 = \left[\frac{1}{N_{jk}}\left(\frac{1}{N_{jk}}\sum_{i=1}^{N_{jk}}(Z_{jki} - \overline{Z}_{jk})^2\right)\right] / \overline{Z}_{jk}^2, \tag{C.13}$$

we have

$$\hat{v}_{1984}(\hat{\gamma}_{jk})^2 = \hat{v}_{1984}(\hat{\psi}_{jk})^2 + \hat{v}_{1984}(\hat{\rho}_{jk})^2. \tag{C.14}$$

The numerator in the expression (C.13) is equal to the square of the customary standard error for the arithmetic mean of N_{jk} IID random variables except that the latter has $\frac{1}{N_{jk}-1}$ rather than $\frac{1}{N_{jk}}$ within the parentheses (and), but this difference is negligible if N_{jk} is not too small.

A further illumination is the following. Using S-sufficiency and S-ancillarity and an independence condition in the prior distribution of the claim frequency and mean claim parameters in a Bayesian setup, we can treat $\hat{\psi}_{jk}$ and $\hat{\rho}_{jk}$ as if they were independent random variables. With $v()$ the ordinary CV and T and U independent random variables, we have

$$v(TU)^2 = v(T)^2 + v(U)^2 + v(T)^2 v(U)^2. \tag{C.15}$$

If both $v(T)$ and $v(U)$ are not too large the last term is negligible. If for example $v(T)$ and $v(U)$ are both less than 0.4, then $1 < v(TU)^2/[v(T)^2 + v(U)^2] < 1.08$.

For large N_{jk} the CVs of $\hat{\psi}_{jk}$ and $\hat{\rho}_{jk}$ are small, so that the identity (C.14) for CV estimates corresponds to an asymptotic identity for CVs.

C.5. Adjusted claim frequency CVs

Let superindex (B) denote claim frequency factors with base class 1, i. e.

$$\psi_{jk}^{(B)} = \psi_{jk}/\psi_{j1}. \tag{C.16}$$

In practice any class k can be used as a base class. Class 1 is here used for simple notation.

Let $c_{jk}^{(F)}$ be the diagonal element of $\hat{\mathbf{I}}_F^{-1}$ by (C.3) pertaining to argument j and class k for $k \geq 2$, setting $c_{j1}^{(F)} = 0$. Now $c_{jk}^{(F)}$ is in GLM theory an estimate of $\mathrm{Var}[\log \psi_{jk}^{(B)}]$ and since

$\mathrm{Var}[\log T] \approx v(T)^2$ if $v(T)$ is small,

we can form a GLM CV estimate from

$$\hat{v}_{\text{GLM}}(\hat{\psi}_{jk}^{(B)})^2 = c_{jk}^{(F)}. \tag{C.17}$$

We now proceed by writing an analogous expression (C.18) for estimated CVs of factor estimates with base class 1, using the original 1984 method. (C.18) can be transformed back to (C.11). The same transformation is applied to Equation (C.17), so that we get adjusted GLM CVs for claim frequency with positive value also for $k = 1$. Here subindex $_{\text{1984B}}$ for the CV estimate means that the 1984 method is modified so that class 1 is used as base class.

$$\hat{v}_{\text{1984B}}(\hat{\psi}_{jk}^{(B)})^2 = \begin{cases} 0, & k = 1, \\ \frac{1}{N_{jk}} + \frac{1}{N_{j1}}, & k \geq 2. \end{cases} \tag{C.18}$$

The rationale for this transformation is this approximation, valid if the Poisson distributed variables N_{jk} are large and their CVs hence are small. Here $\hat{v}()$ is a general CV estimate.

$$\hat{v}\left(\frac{1}{N_{j1}}\right)^2 \approx \hat{v}(N_{j1})^2 = \frac{1}{N_{j1}}. \tag{C.19}$$

By the expression (C.15) above

$$\hat{v}\left(\frac{N_{jk}}{N_{j1}}\right)^2 \approx \frac{1}{N_{jk}} + \frac{1}{N_{j1}}, \quad k \geq 2. \tag{C.20}$$

It is seen that Equation (C.11) is obtained from Equation (C.18) by adding $1/N_{j1}$ for $k = 1$ and subtracting $1/N_{j1}$ for $k \geq 2$.

Analogously, for a new type of CV estimate with positive values for all k, modify Equation (C.17) by adding $1/N_{j1}$ for $k = 1$ and subtracting $1/N_{j1}$ for $k \geq 2$. This applies to CV estimates for $\hat{\psi}_{jk}$ with no specified base class.

$$\hat{v}_{\text{MVW}}(\hat{\psi}_{jk})^2 = \begin{cases} \frac{1}{N_{j1}}, & k = 1, \\ c_{jk}^{(F)} - \frac{1}{N_{j1}}, & k \geq 2. \end{cases} \tag{C.21}$$

It is somewhat arbitrary why, when no base class is specified, class $k = 1$ or any specified class should have the same value in $\hat{v}_{1984}()$ and $\hat{v}_{\text{MVW}}()$ and the other values of k should get larger values in $\hat{v}_{\text{MVW}}()$. However, this facilitates a transformation of the CV estimates obtained to estimates relevant when a base class with factor 1 is specified.

We now define a variable $N_{jk}^{(G)}$ called Adjusted Claim Number. It is not an integer, but can be used instead of N_{jk} in new versions of Equations (C.11) and (C.14), to get better CV estimates for risk premium factor estimates.

$$N_{jk}^{(G)} = \begin{cases} N_{j1}, & k = 1, \\ \left(c_{jk}^{(F)} - \frac{1}{N_{j1}}\right)^{-1}, & k \geq 2. \end{cases} \tag{C.22}$$

It is immediate that

$$c_{jk}^{(F)} = \begin{cases} 0, & k = 1, \\ \frac{1}{N_{jk}^{(G)}} + \frac{1}{N_{j1}}, & k \geq 2, \end{cases} \tag{C.23}$$

$$\hat{v}_{\text{MVW}}(\hat{\psi}_{jk})^2 = \frac{1}{N_{jk}^{(G)}} = \frac{N_{jk}}{N_{jk}^{(G)}}\hat{v}_{1984}(\hat{\psi}_{jk})^2. \tag{C.24}$$

It can be shown that if exposure is multiplicative, then $N_{jk}^{(G)} = N_{jk}$. This is because, see above under Equation (C.7), with multiplicative exposure estimated claim frequency factors are proportional to estimated univariate claim frequencies, i. e.

$$\hat{\psi}_{jk}^{(B)} = \frac{N_{jk}/e_{jk}}{N_{j1}/e_{j1}}.$$

From this identity we can deduce that, if exposure is multiplicative,

$$c_{jk}^{(F)} = \hat{v}_{1984\text{B}}(\hat{\psi}_{jk}^{(B)})^2 = \hat{v}_{\text{GLM}}(\hat{\psi}_{jk}^{(B)})^2.$$

C.6. Adjusted mean claim CVs

In Equation (C.24) we can call $N_{jk}/N_{jk}^{(G)}$ an enlargement factor correcting the negative bias in the original CV estimate due to non-multiplicativity in exposure. An analogous enlargement factor for mean claim should be useful. This is heuristic and must be justified by simulations.

Let $c_{jk}^{(M)}$ be the diagonal element of $\hat{\mathbf{I}}_M^{-1}$ by (C.5) pertaining to argument j and class k for $k \geq 2$, with $c_{j1}^{(M)} = 0$. Superindex (M) denotes mean claim.

Define, analogously to Equation (C.22), a variable $S_{jk}^{(G)}$ called Adjusted Claim Cost.

$$S_{jk}^{(G)} = \begin{cases} S_{j1}, & k = 1, \\ \left(c_{jk}^{(M)} - \frac{1}{S_{j1}}\right)^{-1}, & k \geq 2. \end{cases} \tag{C.25}$$

We could, analogously to Equation (C.24), define $(S_{jk}/S_{jk}^{(G)})\hat{v}_{1984}(\hat{\rho}_{jk})^2$ as the square of an adjusted mean claim CV. However, in simulations we found a positive bias from this, while an expression halfway between this and the 1984 expression gave mostly approximately a zero bias. So we let

$$\hat{v}_{\text{MVW}}(\hat{\rho}_{jk})^2 = \frac{1}{2}\left(1 + \frac{S_{jk}}{S_{jk}^{(G)}}\right)\hat{v}_{1984}(\hat{\rho}_{jk})^2. \tag{C.26}$$

If exposure is multiplicative, then $S_{jk}^{(G)} = S_{jk}$.

C.7. Adjusted risk premium CVs

The partitioning identity applied to the new claim frequency and mean claim terms gives

$$\hat{v}_{\text{MVW}}(\hat{\gamma}_{jk})^2 = \hat{v}_{\text{MVW}}(\hat{\psi}_{jk})^2 + \hat{v}_{\text{MVW}}(\hat{\rho}_{jk})^2. \tag{C.27}$$

The claim frequency part $\hat{v}_{\text{MVW}}(\hat{\psi}_{jk})^2$ is rigorously justified, while the mean claim part $\hat{v}_{\text{MVW}}(\hat{\rho}_{jk})^2$ is heuristic. Unfortunately, not half of the risk premium CV is rigorously justified, since in typical real applications the claim frequency CV is small compared to the mean claim CV.

C.8. Confidence intervals

Three CV estimates by MVW are now available: $\begin{cases} \text{Claim frequency} & \hat{v}_{\text{MVW}}(\hat{\psi}_{jk}) \text{ in (C.24)} \\ \text{Mean claim} & \hat{v}_{\text{MVW}}(\hat{\rho}_{jk}) \text{ in (C.26)} \\ \text{Risk premium} & \hat{v}_{\text{MVW}}(\hat{\gamma}_{jk}) \text{ in (C.27)} \end{cases}$

Let η_{jk} be anyone of $\psi_{jk}, \rho_{jk}, \gamma_{jk}$ and let $\hat{v}_{jk}$ be the corresponding CV estimate.

The immediate confidence interval, with positive widths for all argument classes, is obtained using a normal approximation, if there are sufficiently many claims. With multiplicative exposure the central limit theorem can be applied for $\psi_{jk}, \rho_{jk}, \gamma_{jk}$. For claim frequency the GLM theory gives a normal approximation in any case. The same should hold also for mean claim and risk premium. See (C.10) in C.3 for a use of (C.28).

$$\eta_{jk} = \hat{\eta}_{jk}(1 \pm 1.96\hat{v}_{jk}) \ \ (95\%). \tag{C.28}$$

Since a lognormal approximation is better if the normal approximation will give negative lower confidence limits, we might also use

$$\eta_{jk} = \hat{\eta}_{jk} \exp\{\pm 1.96\hat{v}_{jk}\} \ \ (95\%). \tag{C.29}$$

In Länsförsäkringar Alliance practice the 95 % level is, somewhat ad hoc, reduced to 90 % in order to compensate for aberration from normality and the additional uncertainty introduced by substituting $\hat{v}_{jk}$ for the unknown v_{jk}.

A confidence interval strictly relating to factors for other classes than 1 when class 1 has factor 1 by definition, as in S-GLM, can be written for $k \neq 1$

$$\eta_{jk}^{(B)} = \hat{\eta}_{jk}^{(B)} \exp\{\pm 1.96\sqrt{\hat{v}_{j1}^2 + \hat{v}_{jk}^2}\} \ \ (95\%). \tag{C.30}$$

No interval for class 1 shall be given with this way of computing confidence intervals. Superindex (B) denotes factors with base class 1. See (C.16) in C.5 for claim frequency. Class 1 can be replaced with some other class taken as the base class. When $\eta_{jk}^{(B)} = \psi_{jk}^{(B)}$, the claim frequency factors, Equation (C.30) can be shown to be the S-GLM claim frequency confidence interval.

The disadvantage of (C.30) compared to (C.28) and (C.29) is that a large $\hat{v}_{j1}$ for class 1 will obscure the comparison between two other classes.

The reasoning leading to Equation (C.30) is similar to the reasoning leading to Equation (C.20), namely that for a ratio of two independent random variables we add the squares of their CVs to approximately obtain the square of the CV of the ratio, if the CVs are reasonably small.

The accuracy of these confidence intervals depend on the claims in the various classes. If there are few claims and/or unevenly distributed claim amounts in a certain class the confidence intervals for this class will be uncertain. Typically, however, these confidence intervals are then large anyway, so that no very useful estimate can be obtained in any case. An exception to this is if there are just a few claims in a class, who happen to be nearly equal. Then the confidence intervals for this class will be misleadingly small. It is thus advisable to check classes with few claims and small confidence intervals to see if this exception holds for these.