# Credibility for additive and multiplicative models

by Alois Gisler and Petra Müller
Winterthur Insurance Company

Abstract:
Additive and multiplicative models are common in modeling multivariate tariffs. Mostly classical multivariate statistical techniques and in particular generalized linear models are used to determine the premiums of such tariffs. However, often the number of risks in the different rating cells are rather small. In these cases a credibility approach would be more appropriate. In this paper we consider beside the classical additive and multiplicative model a Bayesian additive and multiplicative model and derive the corresponding credibility estimators. Moreover, estimators of the structural parameters are given and the methodology is applied to real data from the insurance practice.

Key Words: insurance pricing, additive and multiplicative models, generalized linear models, Tweedie models, credibility theory.

## 1 Introduction

Multivariate tariffs depending on several rating factors are common in most lines of business. Especially in personal lines like motor insurance there was a tendency in recent years to use more and more rating factors to construct tariffs reflecting as much as possible the riskiness of the risks insured. Mostly multivariate statistical techniques and in particular generalized linear models (GLM) are used to calculate such multivariate tariffs. It seems to have become a standard in insurance to use GLM with Poisson for claim frequencies and GLM with Gamma for claim severities.

However, there are many situations where the number of risks are rather scarce for many cells of the tariff. In general, the more tariff factors you use, the more cells you will have with little data. Another situation with little data is the modeling of big claims. Very often two or three percent of the biggest claims cause half or even more of the total claim amount. Therefore it is advisable to consider and to model the big claims separately. But then the data basis and the number of observations in the different cells are small.

If the data basis in the individual cells becomes small, then the point estimates resulting from generalized linear modeling may not be very accurate, i.e. the corresponding confidence intervals may become rather big. In such situations it is questionable, whether it is appropriate to simply use GLM.

Let us assume for the moment, that there is only one rating factor with $k$ levels dividing the risks into $k$ groups and that we want to estimate the expected value of the claim severity for each of these groups. The estimates resulting from the GLM machinery are simply the observed average claim sizes in the $k$ groups. However if the number of claims per risk group is small, then the random fluctuations of the observed claims averages will be very big. Hence taking them as an estimator of the expected value would probably not be a very good idea. A credibility approach by using for instance the Bühlmann and Straub model would be much more adequate in such a situation.

In the multivariate case it is exactly the same. But to use credibility in the case of several rating factors we have first to develop suitable Bayesian models and then we have to derive the corresponding credibility estimators. This is what is done in this paper.

In Section 2 we define what we mean by an additive and a multiplicative tariff structure. In Section 3 we introduce the classical additive model and summarize some known results regarding estimators of the pure risk premiums and the properties of these estimators. In Section 4 the Bayesian additive model is presented. We then derive the credibility estimators in this model. We also show, how we can estimate the structural parameters. Finally the results are applied to observed claims averages in the line third party liability.

Section 5 is similar to Section 3, but now for the multiplicative model, i.e. the classical multiplicative model and estimators of the premiums in this model are considered. In Section 6 we introduce the Bayesian multiplicative model. A pure credibility approach would not be suited here because of the multiplicative nature of the true pure risk premium. However, we show how we can define a credibility based estimator in this case and we derive the corresponding estimators. At the end of Section 6, we again apply the methodology to real data, namely on observed claim frequencies in third party liability.

# 2 Additive and Multiplicative Tariff Structure

Assume that in a given line of business there are $K$ rating factors to be used in a tariff and that rating factor $k$ has $I_k$ levels. Examples of rating factors in motor insurance are car model, cm$^3$-class, horse-power divided by weight, geographic area, sex, age-group, mileage-class, Bonus-Malus, etc. The rating factors subdivide the portfolio into

$$\prod_{k=1}^{K} I_k \text{ risk groups,}$$

and we call these risk-groups cells, where $cell_{i_1,i_2,\ldots,i_K}$ denotes the group of risks with level $i_1$ for rating factor 1, level $i_2$ for rating factor 2,etc..

Denote by $X_{i_1,i_2,\ldots,i_K}$ the observable random variable in $cell_{i_1,i_2,\ldots,i_K}$ and by $w_{i_1,i_2,\ldots,i_K}$ an associated measure of exposure. Usually $X_{i_1,i_2,\ldots,i_K}$ is either the average claim frequency, the average claim size or the average total claim amount divided by $w_{i_1,i_2,\ldots,i_K}$, where the "average" can also be a multi-year average. The exposure $w_{i_1,i_2,\ldots,i_K}$ is very often the number of year-risks in $cell_{i_1,i_2,\ldots,i_K}$, but it could also be something else like for instance the total sum insured in fire insurance.

Finally we denote by $P_{i_1 i_2 \ldots i_K}$ the "premium" of a risk in $cell_{i_1,i_2,\ldots,i_K}$. Here and in the following, "premium" is always to be understood as the expected value of the random variables $X_{i_1,i_2,\ldots,i_K}$ considered .

**Definition 2.1** *An additive tariff structure is defined by the property that*

$$P_{i_1 i_2 \dots i_K} = \mu + \sum_{k=1}^{K} \alpha_{k,i_k}, \tag{2.1}$$

*where* $\mu$ *and* $\alpha_{k,i_k}$, $k = 1, \dots, K$, $i_k = 1 \dots, I_k$, *are real numbers.*

We call $\mu$ and $\alpha_{k,i_k}$ the parameters of the tariff. The parameter $\mu$ might be the average premium (averaged over all cells) or it might refer to the premium of a reference cell, for instance of $cell_{1,1,,\dots,1}$. In the latter case $\mu = P_{1,1,..,1}$ and hence $\alpha_{k,1} = 0$ for $k = 1, \dots, K$. This shows that the tariff contains $I_\bullet - K + 1$ free parameters where $I_\bullet = \sum_k I_k$. Here and in the following a dot in the index means summation over the corresponding index.

**Definition 2.2** *A multiplicative tariff structure is defined by the property that*

$$P_{i_1 i_2 \dots i_K} = \mu \cdot \prod_{k=1}^{K} \alpha_{k,i_k}, \tag{2.2}$$

*where* $\mu$ *and* $\alpha_{k,i_k}$, $k = 1, \dots, K$, $i_k = 1 \dots, I_k$, *are real numbers.*

Again, the parameter $\mu$ might be the average premium or it might refer to the premium of $cell_{1,1,,\dots,1}$, in which case $\mu = P_{1,1,..,1}$ and $\alpha_{k,1} = 1$ for $k = 1, \dots, K$. This shows that the tariff contains $I_\bullet - K + 1$ free parameters.

Already at the birth of ASTIN, i.e. in the late 50-ies and in the 60-ies, several methods for estimating the tariff parameters have been suggested in the actuarial literature. We will encounter two of them later in this paper, namely the method of least squares and the method of marginal totals. A good survey of the different "classical methods" can be found in van Eghen and alias [12]. Nowadays, it is fairly common in the insurance industry to calculate such tariffs by using multivariate statistical methods and in particular generalized linear models.

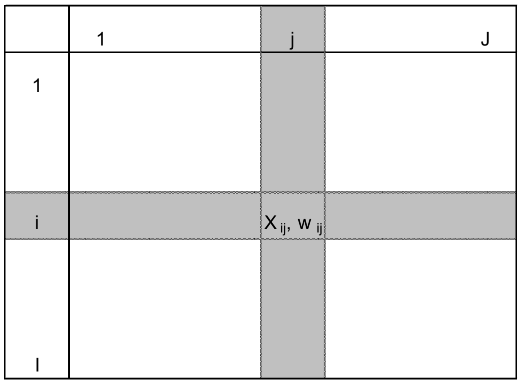
For simplicity and for didactical reasons, in the following, we will only consider the case of two rating factors. However all results can be extended in an obvious way to any number of rating factors. We will denote the two rating factors by $A$ and $B$, the corresponding levels by $i = 1, \dots, I$ and $j = 1, \dots, J$, the observable random variables by $X_{ij}$, the associated exposure measures by $w_{ij}$ and the parameters of the tariff by $\mu, \psi_i$ and $\varphi_j$. The additive tariff is then given by

$$P_{ij} = \mu + \psi_i + \varphi_j$$

and the multiplicative one by

$$P_{ij} = \mu \cdot \psi_i \cdot \varphi_j.$$

With two rating factors, we can visualize the situation by a $2 \times 2$ contingency table.




2 x 2 contingency table

# 3 The Classical Additive Model

## 3.1 Model Assumptions

Based on classical statistics, the underlying assumptions behind an additive tariff structure is the following fixed effect analysis of variance model.

**Model-Assumptions 3.1 (classical additive model)** *The observable random variables* $X_{ij}$, $i = 1, 2, \ldots, I$ *and* $j = 1, 2, \ldots, J$, *satisfy*

$$X_{ij} = \mu_0 + \psi_i + \varphi_j + \varepsilon_{ij},$$

*where* $\varepsilon_{ij}$ *are independent random variables with*

$$\begin{aligned} E[\varepsilon_{ij}] &= 0, &(3.1)\\ \mathrm{Var}\,(\varepsilon_{ij}) &= \frac{\sigma^2}{w_{ij}}, &(3.2)\end{aligned}$$

*where* $\mu_0, \psi_i, \varphi_j$ *are real numbers.*

The aim is to estimate for each $cell_{ij}$ the corresponding premium

$$P_{ij} = E\left[X_{ij}\right] = \mu_0 + \psi_i + \varphi_j. \tag{3.3}$$

**Remark:**

- The parameters $\mu_0, \psi_i, \varphi_j$ are determined only up to an additive constant. They are uniquely defined if we fix $\mu_0$. In the following $\mu_0$ is referred to as the average premium over the portfolio. This is equivalent to the side constraint

$$\sum_{i=1}^{I} w_{i\bullet}\psi_i = \sum_{j=1}^{J} w_{\bullet j}\varphi_j = 0. \tag{3.4}$$

## 3.2 Estimators of the tariff parameters

In the insurance practice, the tariff parameters $\mu_0, \psi_i, \varphi_j$ are unknown and have to be estimated from the observations $X_{ij}$.

A natural candidate is the *weighted least square estimator,* i.e. the parameters are determined in such a way that

$$
\begin{aligned}
Q &= \sum_{i,j} w_{ij}\left(X_{ij} - \widehat{P}_{ij}\right)^2, && (3.5)\\
\text{where } \widehat{P}_{ij} &= \widehat{\mu}_0 + \widehat{\psi}_i + \widehat{\varphi}_j, && (3.6)
\end{aligned}
$$

is minimized subject to the constraint

$$
\sum_{i=1}^{I} w_{i\bullet}\widehat{\psi}_i = \sum_{j=1}^{J} w_{\bullet j}\widehat{\varphi}_j = 0. \tag{3.7}
$$

Another well known method for determining the tariff parameters is the so called *method of marginal totals.* The parameters are fixed in such a way that there is equality between premiums and observed losses for the marginal totals. The method was first presented by Bailey (1963) and later by Jung (1968). The basic idea is that for large groups of insured the premium should be equal to the observed losses (for the past observation period). In the additive model these marginal total conditions are

$$
\begin{aligned}
\sum_{j} w_{ij}(\widehat{\mu}_0 + \widehat{\psi}_i + \widehat{\varphi}_j) &= \sum_{j} w_{ij}X_{ij}, && (3.8)\\
\sum_{i} w_{ij}(\widehat{\mu}_0 + \widehat{\psi}_i + \widehat{\varphi}_j) &= \sum_{i} w_{ij}X_{ij}. && (3.9)
\end{aligned}
$$

**Theorem 3.2** *Under the Model Assumptions 3.1 the weighted least square estimators of $\mu_0, \psi_i, \varphi_j$ minimizing* (3.5) *as well as the estimators by the method of marginal totals satisfying* (3.8) *and* (3.9) *are given by the equations*

$$
\begin{aligned}
\widehat{\mu}_0 &= \overline{X}_{\bullet\bullet}, && (3.10)\\
\widehat{\psi}_i &= \left(\overline{X}_{i\bullet} - \overline{X}_{\bullet\bullet}\right) - \sum_{j=1}^{J} \frac{w_{ij}}{w_{i\bullet}}\widehat{\varphi}_j \qquad i = 1, \ldots, I, && (3.11)\\
\widehat{\varphi}_j &= \left(\overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet}\right) - \sum_{i=1}^{I} \frac{w_{ij}}{w_{\bullet j}}\widehat{\psi}_i \qquad j = 1, \ldots, J, && (3.12)
\end{aligned}
$$

*where*

$$
\begin{aligned}
\overline{X}_{i\bullet} &= \sum_{j} \frac{w_{ij}}{w_{i\bullet}} X_{ij}, && (3.13)\\
\overline{X}_{\bullet j} &= \sum_{i} \frac{w_{ij}}{w_{\bullet j}} X_{ij}, && (3.14)\\
\overline{X}_{\bullet\bullet} &= \sum_{i,j} \frac{w_{ij}}{w_{\bullet\bullet}} X_{ij}. && (3.15)
\end{aligned}
$$

**Remark:**

- The equation system (3.11) and (3.12) can be solved iteratively by starting for instance with $\widehat{\varphi}_j = \overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet}$ in (3.11) and then inserting successively $\widehat{\psi}_i$ and $\widehat{\varphi}_j$ in (3.11) resp. (3.12). Usually the iteration process converges very quickly.

**Proof of Theorem 3.2:**
The results are easily obtained by equating on the right hand side of (3.5) the partial derivatives with respect to $\widehat{\mu}_0, \widehat{\psi}_i, \widehat{\varphi}_j$ to 0 and taking the constraint (3.7) into account. The validity of the marginal total conditions (3.8) and (3.9) can be verified by inserting (3.10) – (3.12) into these equations. □

The method of weighted least square as well as the method of marginal totals are natural but pragmatic procedures for estimating the tariff parameters. An interesting actuarial question is what properties do these estimator have? A desirable property would be that they minimize the mean square error

$$mse = E\left[\sum_{i,j} w_{ij}\left(X_{ij} - \widehat{P}_{ij}\right)^2\right]. \tag{3.16}$$

As the following theorems shows, this is indeed the case if we restrict on linear estimators. This result follows from the Gauss-Markov Theorem from classical statistics (see for instance [11]).

**Theorem 3.3** *Under the Model Assumptions 3.1, the weighted least square estimators of $\mu_0, \psi_i, \varphi_j$ given by Theorem 3.2 minimize the mean square error* (3.16) *among all linear unbiased estimators of $\mu_0, \psi_i, \varphi_j$.*

In the case of a normal distribution, the following results are also well known from classical statistics:

**Theorem 3.4** *Under the Model Assumptions 3.1 and if the random variables $X_{ij}$ are normally distributed then*

*i) The weighted least square estimators of $\mu_0, \psi_i, \varphi_j$ given by Theorem 3.2 minimize the mean square error* (3.16) *among all unbiased estimators of $\mu_0, \psi_i, \varphi_j$.*

*ii) The weighted least square estimators of $\mu_0, \psi_i, \varphi_j$ given by Theorem 3.2 are also the maximum likelihood estimators of $\mu_0, \psi_i, \varphi_j$.*

**Remark:**

- In the normal case, because of property ii) of the above theorem, the weighted least square estimators are also the GLM estimators (= estimators resulting from generalized linear modeling).

# 4 Bayesian Additive Model

## 4.1 Model Assumptions

In the Bayesian model the risk parameters $\psi_i$ and $\varphi_j$ are assumed to be realizations of random variables $\Psi_i$ and $\Phi_j$, and that conditionally, given $\Psi_i$ and $\Phi_j$, the assumptions of the classical linear model (Model Assumptions 3.1) are fulfilled.

**Model-Assumptions 4.1 (Bayes Additive Model)**

*i) The r.v. $X_{ij}$, $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$, are conditionally, given $\Theta_{ij} = (\Psi_i, \Phi_j)$, independent with*

$$\begin{aligned} E[X_{ij}|\,\Theta_{ij}] &= \mu_0 + \Psi_i + \Phi_j, & (4.1)\\ \mathrm{Var}(X_{ij}|\,\Theta_{ij}) &= \frac{\sigma^2(\Theta_{ij})}{w_{ij}}. & (4.2) \end{aligned}$$

*ii) The r.v. $\Psi_i$, $i = 1, \ldots, I$, are independent and identically distributed (i.i.d.) with*

$$\begin{aligned} E[\Psi_i] &= \mu_\Psi = 0, & (4.3)\\ \mathrm{Var}(\Psi_i) &= \tau_\Psi^2. & (4.4) \end{aligned}$$

*iii) The r.v. $\Phi_j$, $j = 1, 2, \ldots, J$, are i.i.d. with*

$$\begin{aligned} E[\Phi_j] &= \mu_\Phi = 0, & (4.5)\\ \mathrm{Var}(\Phi_j) &= \tau_\Phi^2. & (4.6) \end{aligned}$$

*iv) $\Psi_i, \Phi_j$ are independent.*

## 4.2 Credibility Estimators

The aim is to find for each $cell_{ij}$ an estimator of the corresponding true individual premium

$$P_{ij} = E[X_{ij}|\,\Theta_{ij}] = \mu(\Theta_{ij}) = \mu_0 + \Psi_i + \Phi_j. \tag{4.7}$$

Note the difference between (4.7) and (3.3): in the Bayesian model $P_{ij}$ defined by (4.7) is a random variable.

**Definition 4.2** *An estimator $\widehat{P}_{ij}$ is said to be a better or equal than another estimator $\widehat{P}^*_{ij}$, if*

$$E\left[\left(P_{ij} - \widehat{P}_{ij}\right)^2\right] \leq E\left[\left(P_{ij} - \widehat{P}^*_{ij}\right)^2\right]. \tag{4.8}$$

The above definition means that we use the expected quadratic loss

$$E\left[\left(P_{ij}-\widehat{P}_{ij}\right)^2\right] \tag{4.9}$$

as optimality criterion.

By definition the credibility estimator $P_{ij}^{Cred}$ is the best estimator out of the class of linear estimators minimizing the expected quadratic loss (4.9).

**Theorem 4.3**

*i) Under the Model Assumptions 4.1, the credibility estimator of $P_{ij}$ is*

$$P_{ij}^{Cred}=\widehat{\widehat{\mu(\Theta_{ij})}}=\mu_0+\Psi_i^{Cred}+\Phi_j^{Cred}, \tag{4.10}$$

*where $\Psi_i^{Cred}$ and $\Phi_j^{Cred}$ are the credibility estimators of $\Psi_i$ and $\Phi_j$.*

*ii) $\Psi_i^{Cred}$ and $\Phi_j^{Cred}$ are given by the following system of equations:*

$$\Psi_i^{Cred} = \alpha_i(\overline{X}_{i\bullet}-\mu_0)-\alpha_i\sum_{j=1}^{J}\frac{w_{ij}}{w_{i\bullet}}\Phi_j^{Cred}, \tag{4.11}$$

$$\Phi_j^{Cred} = \beta_j(\overline{X}_{\bullet j}-\mu_0)-\beta_j\sum_{i=1}^{I}\frac{w_{ij}}{w_{\bullet j}}\Psi_i^{Cred}, \tag{4.12}$$

*where*

$$\begin{aligned}
\alpha_i &= \frac{w_{i\bullet}}{w_{i\bullet}+\frac{\sigma^2}{\tau_\Psi^2}},\\
\beta_j &= \frac{w_{\bullet j}}{w_{\bullet j}+\frac{\sigma^2}{\tau_\Phi^2}},\\
\sigma^2 &= E\left[\sigma^2(\Theta_{ij})\right].
\end{aligned}$$

**Remarks:**

- $\overline{X}_{i\bullet}$ and $\overline{X}_{\bullet j}$ are the weighted averages defined in (3.13) and (3.14).
- As the in the classical additive model the (4.11) and (4.12) can again be solved iteratively.
- By comparing (4.11) and (4.12) with (3.11) and (3.12) we see that (4.11) and (4.12) are the credibility pendant to the classical estimators and hence also the credibility pendant to the GLM in the case of normally distributed $X_{ij}$.

**Proof:**

i) Equation (4.10) follows immediately from the linearity property of credibility estimators.

ii) By the iterativity property of projections (Theorem A.4) we have

$$\begin{aligned}\Psi_i^{Cred} &= \mathrm{Pro}\left(\Psi_i \middle| L(1,\mathbf{X})\right) \\ &= \mathrm{Pro}(\underbrace{\mathrm{Pro}(\Psi_i| L(1,\mathbf{X},\mathbf{\Phi}))}_{\Psi_i^*} | L(1,\mathbf{X})). \end{aligned} \tag{4.13}$$

iii)

$$\Psi_i^* = \alpha_i(\overline{X}_{i\bullet} - \mu_0) - \alpha_i \sum_{j=1}^{J} \frac{w_{ij}}{w_{i\bullet}} \Phi_j \tag{4.14}$$

To show (4.14) we check that the normal equations $(A.6)$ and $(A.7)$ are satisfied. Thus we have to verify that

$$E\left[\Psi_i^*\right] = E\left[\Psi_i\right] = 0, \tag{4.15}$$

$$\mathrm{Cov}\left(\Psi_i^*, X_{kl}\right) = \mathrm{Cov}\left(\Psi_i, X_{kl}\right) = \tau_\Psi^2 \delta_{ik}, \tag{4.16}$$

$$\mathrm{Cov}\left(\Psi_i^*, \Phi_l\right) = \mathrm{Cov}\left(\Psi_i, \Phi_l\right) = 0. \tag{4.17}$$

$$\begin{aligned} E\left[\Psi_i^*\right] &= \alpha_i \underbrace{E\left[(\overline{X}_{i\bullet} - \mu_0)\right]}_{=0} - \alpha_i \sum_{j=1}^{J} \frac{w_{ij}}{w_{i\bullet}} \underbrace{E\left[\Phi_j\right]}_{=0} \\ &= 0. \end{aligned}$$

$$\begin{aligned} Cov(\Psi_i^*, X_{kl}) &= \underbrace{\alpha_i \sum_j \frac{w_{ij}}{w_{i\bullet}} Cov(X_{ij}, X_{kl})}_{=\frac{\alpha_i}{w_{i\bullet}}\sigma^2\delta_{ik} + \alpha_i\tau_\Psi^2\delta_{ik} + \alpha_i\frac{w_{il}}{w_{i\bullet}}\tau_\Phi^2} - \alpha_i \sum_j \frac{w_{ij}}{w_{i\bullet}} \underbrace{Cov(\Phi_j, X_{kl})}_{=\tau_\Phi^2\delta_{jl}} \\ &= \frac{\alpha_i}{w_{i\bullet}}\sigma^2\delta_{ik} + \alpha_i\tau_\Psi^2\delta_{ik} + \alpha_i\frac{w_{il}}{w_{i\bullet}}\tau_\Phi^2 - \alpha_i\frac{w_{il}}{w_{i\bullet}}\tau_\Phi^2 \\ &= \alpha_i(\frac{\sigma^2}{w_{i\bullet}} + \tau_\Psi^2)\delta_{ik} \\ &= \tau_\Psi^2\delta_{ik}. \end{aligned}$$

$$\begin{aligned} Cov(\Psi_i^*, \Phi_l) &= \alpha_i \sum_j \frac{w_{ij}}{w_{i\bullet}} \underbrace{Cov(X_{ij}, \Phi_l)}_{=\tau_\Phi^2\delta_{jl}} - \alpha_i \sum_j \frac{w_{ij}}{w_{i\bullet}} \underbrace{Cov(\Phi_j, \Phi_l)}_{=\tau_\Phi^2\delta_{jl}} \\ &= \alpha_i\frac{w_{il}}{w_{i\bullet}}\tau_\Phi^2 - \alpha_i\frac{w_{il}}{w_{i\bullet}}\tau_\Phi^2 \\ &= 0. \end{aligned}$$

Hence equations (4.15) – (4.17) are fulfilled.

iv) (4.14) plugged into (4.13) gives

$$\begin{aligned}
\Psi_i^{Cred} &= \alpha_i(\overline{X}_{i\bullet} - \mu_0) - \alpha_i \sum_{j=1}^{J} \frac{w_{ij}}{w_{i\bullet}} \mathrm{Pro}\left(\Phi_j | \, L(1, \mathbf{X})\right) \\
&= \alpha_i(\overline{X}_{i\bullet} - \mu_0) - \alpha_i \sum_{j=1}^{J} \frac{w_{ij}}{w_{i\bullet}} \Phi_j^{Cred}. \qquad (4.18)
\end{aligned}$$

Thus we have proved the validity of (4.11). (4.12) can be proved analogously. □

To get some more insight into the structure of the resulting credibility estimators, it is worthwhile to look at the case of identical weights. In this case an explicit solution can be found.

**Corollary 4.4** *Under the Model Assumptions 4.1 with equal weights ($w_{ij} = 1$ for all $i$ and $j$) the credibility estimator is given by*

$$P_{ij}^{Cred} = \widehat{\widehat{\mu(\Theta_{ij})}} = \mu_0 + \alpha(\overline{X}_{i\bullet} - \mu_0) + \beta(\overline{X}_{\bullet j} - \mu_0) - \gamma(\overline{X}_{\bullet\bullet} - \mu_0), \qquad (4.19)$$

*where*

$$\begin{aligned}
\alpha &= \frac{J}{J + \frac{\sigma^2}{\tau_\Psi^2}}, \\
\beta &= \frac{I}{I + \frac{\sigma^2}{\tau_\Phi^2}}, \\
\gamma &= \frac{\alpha\beta}{1 - \alpha\beta}(2 - \alpha - \beta).
\end{aligned}$$

**Remark:**

- The first summand of the right hand side is the overall expected mean, the second summand a correction term for the level $i$ of rating factor $A$ and the third summand a correction term for the level $j$ of rating factor $B$. But surprisingly, there is also a further last summand which is a correction term based on the overall observed mean $\overline{X}_{\bullet\bullet}$.

**Proof of the Corollary:**

Equation (4.19) can also be written as

$$P_{ij}^{Cred} = \mu(\Theta_{ij})^{Cred} = \mu_0 + \Psi_i^{Cred} + \Phi_j^{Cred},$$

where

$$\begin{aligned}
\Psi_i^{Cred} &= \alpha(\overline{X}_{i\bullet} - \mu_0) - \frac{\alpha\beta}{1 - \alpha\beta}(1 - \alpha)(\overline{X}_{\bullet\bullet} - \mu_0), \qquad (4.20) \\
\Phi_j^{Cred} &= \beta(\overline{X}_{\bullet j} - \mu_0) - \frac{\alpha\beta}{1 - \alpha\beta}(1 - \beta)(\overline{X}_{\bullet\bullet} - \mu_0). \qquad (4.21)
\end{aligned}$$

It is easy to check that (4.20) and (4.21) fulfill the equations (4.11) and (4.12) of Theorem 4.3. □

## 4.3 Estimation of Structural Parameters

The (inhomogeneous) credibility estimator of Theorem 4.3 involves the four structural parameters $\mu_0$, $\sigma^2$, $\tau_\Psi^2$, $\tau_\Phi^2$. In practice, these four parameters are unknown and have to be estimated from the data of the collective. We suggest the following estimators of the variance components:

$$\widehat{\sigma^2} = \frac{1}{(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} w_{ij} (X_{ij} - \widehat{\mu_{ij}})^2, \tag{4.22}$$

where $\widehat{\mu_{ij}}$ is the weighted least square estimator from classical statistics;.

$$\begin{aligned} \widehat{\tau_\Psi^2} &= \max(\widehat{\widehat{\tau_\Psi^2}}, 0), & (4.23)\\ \widehat{\tau_\Phi^2} &= \max(\widehat{\widehat{\tau_\Phi^2}}, 0), & (4.24) \end{aligned}$$

where

$$\begin{aligned} \widehat{\widehat{\tau_\Psi^2}} &= c_1 \left\{ \frac{I}{I-1} \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \left(\overline{X}_{i\bullet} - \overline{X}_{\bullet\bullet}\right)^2 \right. & (4.25)\\ &\quad \left. - \frac{I}{I-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \left( \frac{w_{ij}}{w_{i\bullet}} - \frac{w_{\bullet j}}{w_{\bullet\bullet}} \right)^2 \widehat{\widehat{\tau_\Phi^2}} - \frac{I}{w_{\bullet\bullet}} \widehat{\sigma^2} \right\}, & \\ &\text{where } c_1 = \frac{I-1}{I} \left\{ \sum_i \frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}} \left(1 - \frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}}\right) \right\}^{-1}, & (4.26)\\ \widehat{\widehat{\tau_\Phi^2}} &= c_2 \left\{ \frac{J}{J-1} \sum_{j=1}^{J} \frac{w_{\bullet j}}{w_{\bullet\bullet}} \left(\overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet}\right)^2 \right. & (4.27)\\ &\quad \left. - \frac{J}{J-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{w_{\bullet j}}{w_{\bullet\bullet}} \left( \frac{w_{ij}}{w_{\bullet j}} - \frac{w_{i\bullet}}{w_{\bullet\bullet}} \right)^2 \widehat{\widehat{\tau_\Psi^2}} - \frac{J}{w_{\bullet\bullet}} \widehat{\sigma^2} \right\}, & \\ &\text{where } c_2 = \frac{J-1}{J} \left\{ \sum_j \frac{\omega_{\bullet j}}{\omega_{\bullet\bullet}} \left(1 - \frac{\omega_{\bullet j}}{\omega_{\bullet\bullet}}\right) \right\}^{-1}. & (4.28) \end{aligned}$$

It is well known from classical statistics that $\widehat{\sigma^2}$ is conditionally, given $\Psi_i$ and $\Phi_j$, an unbiased estimator. A proof can be found for instance in [10].

The estimators (4.25) and (4.27) are similar to the ”standard estimators” of the structural parameters of the Bühlmann and Straub model presented in the literature (see for instance [2], Section 4.8). They are more complicated because of the middle term. By straightforward but tedious calculations one can show that $\widehat{\widehat{\tau_\Psi^2}}$ and $\widehat{\widehat{\tau_\Phi^2}}$ are unbiased. Because of the middle term, (4.25) and (4.27) are not any more explicit equations but rather a system of equations, which can be solved iteratively. But nevertheless it makes things more complicated. However, these middle terms were very small and not of practical relevance in the examples considered. This might be the case in many insurance applications.

Therefore an alternative is to neglect the middle term and to replace (4.25) and (4.27) by

$$\widehat{\widehat{\tau_\Psi^2}} = c_1 \left\{ \frac{I}{I-1} \sum_{i=1}^{I} \frac{w_{i\bullet}}{w_{\bullet\bullet}} \left( \overline{X}_{i\bullet} - \overline{X}_{\bullet\bullet} \right)^2 - \frac{I}{w_{\bullet\bullet}} \widehat{\sigma^2} \right\}, \tag{4.29}$$

$$\widehat{\widehat{\tau_\Phi^2}} = c_2 \left\{ \frac{J}{J-1} \sum_{j=1}^{J} \frac{w_{\bullet j}}{w_{\bullet\bullet}} \left( \overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet} \right)^2 - \frac{J}{w_{\bullet\bullet}} \widehat{\sigma^2} \right\}. \tag{4.30}$$

In the numerical example of subsection 4.4 we have used these estimators of the structural parameters.

An alternative for estimating the structural parameters would be to calculate them iteratively together with the iterative procedure of estimating the credibility estimators $\Psi_i^{Cred}$ and $\Phi_j^{Cred}$ (see Theorem 4.3). In each iterative step one estimates $\tau_\Psi^2$ and $\tau_\Phi^2$ respectively by using the standard estimators of the Bühlmann and Straub model.

## 4.4 Numerical Example

The following table and the corresponding graph show the observed average claim sizes of "normal" claims of a larger Swiss insurance company in the line third-party motor liability. To illustrate the method, only two rating factors were considered. The first factor $A$ has 4 levels and the second factor $B$ has 12 levels. The exposure measure is the number of claims. We can see from the table and the graph that the exposures for the cells with rating factors $B1$ as well as for the cells with rating factors $A2-A4$ are rather small and that, due to this small exposure, the observed average claim sizes fluctuate heavily between the cells. With more rating factors, the exposure in the individual cells would be even more scarce and the random fluctuations of the observed average claim size even bigger.

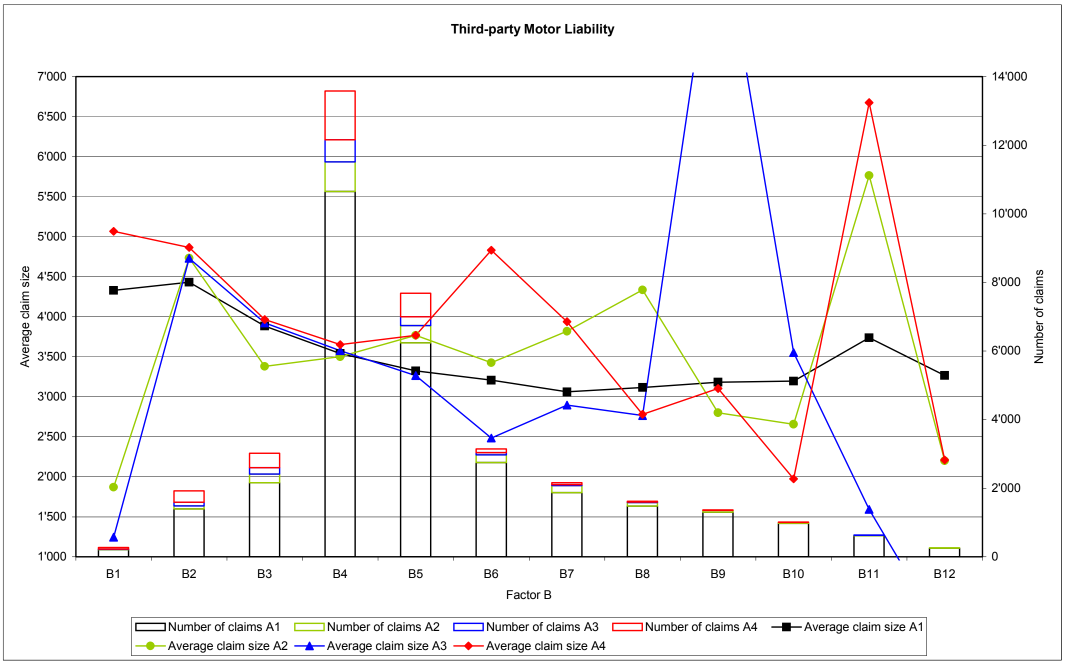
**Observed average claim size**

| | Factor B | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 |
| Factor A | A1 | 4'330 | 4'432 | 3'883 | 3'540 | 3'324 | 3'206 | 3'062 | 3'117 | 3'182 | 3'197 | 3'738 | 3'267 |
| | A2 | 1'871 | 4'733 | 3'379 | 3'501 | 3'769 | 3'426 | 3'818 | 4'335 | 2'800 | 2'657 | 5'765 | 2'200 |
| | A3 | 1'246 | 4'729 | 3'923 | 3'575 | 3'265 | 2'484 | 2'896 | 2'764 | 9'169 | 3'553 | 1'592 | 0 |
| | A4 | 5'066 | 4'866 | 3'966 | 3'652 | 3'769 | 4'830 | 3'939 | 2'780 | 3'103 | 1'974 | 6'674 | 2'211 |

**Number of claims**

| | Factor B | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 |
| Factor A | A1 | 204 | 1'397 | 2'161 | 10'650 | 6'239 | 2'746 | 1'870 | 1'478 | 1'306 | 974 | 620 | 252 |
| | A2 | 12 | 89 | 251 | 864 | 501 | 228 | 209 | 103 | 48 | 31 | 9 | 9 |
| | A3 | 9 | 108 | 184 | 644 | 261 | 64 | 23 | 15 | 3 | 2 | 3 | 0 |
| | A4 | 49 | 327 | 427 | 1'427 | 683 | 105 | 56 | 24 | 9 | 10 | 3 | 1 |

Observations




Observations

We have applied the techniques presented in sections 3 and 4 to these data. For the structural parameters we obtained

| $\widehat{\mu}_0$ | $\widehat{\sigma^2}$ | $\widehat{\tau^2_\Psi}$ | $\widehat{\tau^2_\Phi}$ |
|---|---|---|---|
| 3'512 | 32'647'932 | 41'434 | 112'348 |

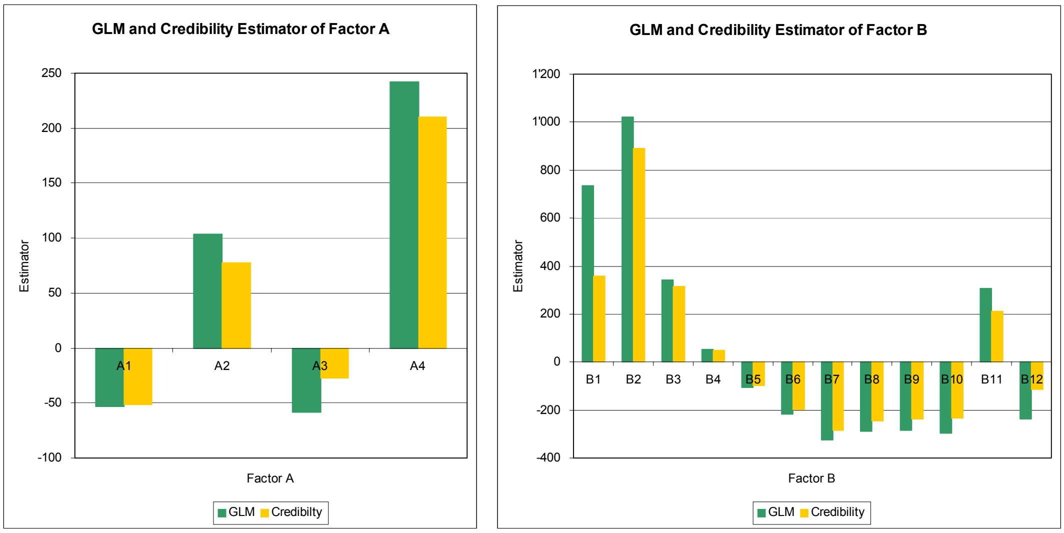
.

The following table and the corresponding graph show the credibility estimators as well as the GLM (weighted least square) estimators of the factor effects. From these results we can well see the "smoothing" effect of the credibility estimators: the absolute values resulting from the credibility estimators are all smaller than the corresponding ones resulting from GLM. The smaller the exposure (number of claims) in the underlying cells, the bigger is this smoothing effect (e.g. effect of factors $B1$ or $A3$).

**Estimators of Factor Effects**

| Factor | | Estimator GLM | Estimator Cred. |
|---|---|---|---|
| Factor A | $\psi_1$ | -54 | -52 |
| | $\psi_2$ | 103 | 77 |
| | $\psi_3$ | -58 | -27 |
| | $\psi_4$ | 242 | 210 |
| Factor B | $\varphi_1$ | 734 | 358 |
| | $\varphi_2$ | 1'020 | 889 |
| | $\varphi_3$ | 342 | 315 |
| | $\varphi_4$ | 52 | 49 |
| | $\varphi_5$ | -105 | -99 |
| | $\varphi_6$ | -218 | -199 |
| | $\varphi_7$ | -325 | -285 |
| | $\varphi_8$ | -287 | -243 |
| | $\varphi_9$ | -285 | -236 |
| | $\varphi_{10}$ | -297 | -232 |
| | $\varphi_{11}$ | 308 | 210 |
| | $\varphi_{12}$ | -239 | -114 |

Factor Effects




Factor Effects

Finally, the following table and the corresponding graph show the resulting premiums. One can see that the differences between the "Credibility premiums" and the "GLM premiums" are by no means negligible in a competitive environment.

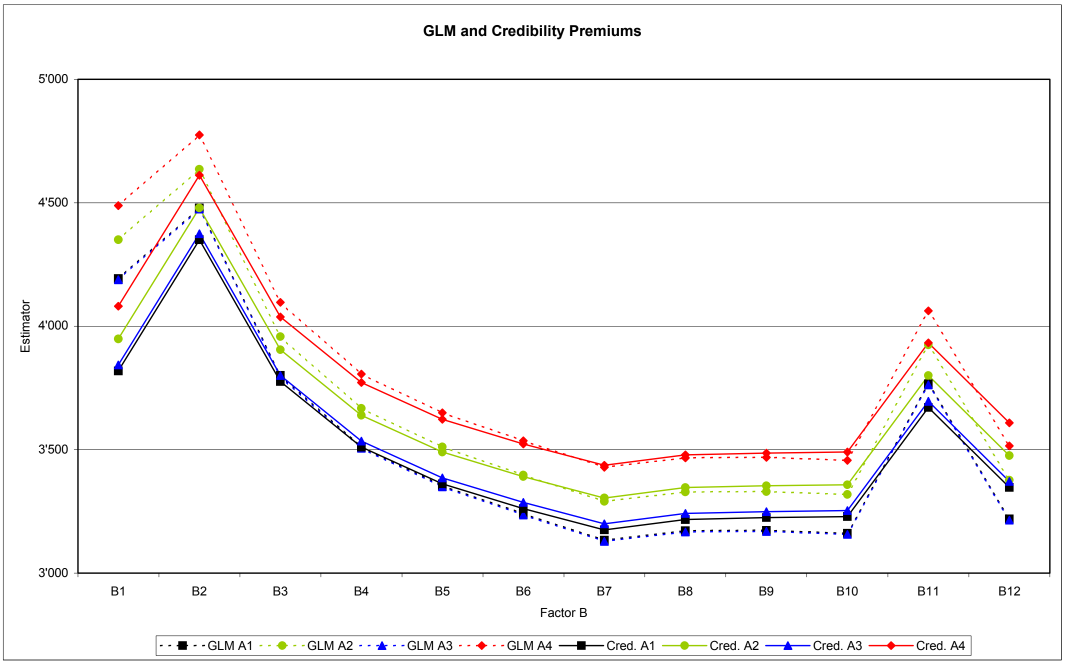
**GLM Premium**

| | Factor B | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 |
| Factor A | A1 | 4'193 | 4'479 | 3'801 | 3'510 | 3'354 | 3'241 | 3'133 | 3'172 | 3'173 | 3'161 | 3'767 | 3'220 |
| | A2 | 4'350 | 4'636 | 3'958 | 3'667 | 3'511 | 3'398 | 3'290 | 3'328 | 3'330 | 3'318 | 3'924 | 3'377 |
| | A3 | 4'188 | 4'474 | 3'796 | 3'506 | 3'349 | 3'236 | 3'129 | 3'167 | 3'169 | 3'157 | 3'762 | 3'215 |
| | A4 | 4'488 | 4'774 | 4'096 | 3'806 | 3'649 | 3'536 | 3'429 | 3'467 | 3'469 | 3'457 | 4'062 | 3'515 |

**Credibility Premium**

| | Factor B | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 |
| Factor A | A1 | 3'819 | 4'350 | 3'776 | 3'510 | 3'361 | 3'262 | 3'175 | 3'217 | 3'225 | 3'229 | 3'671 | 3'347 |
| | A2 | 3'948 | 4'479 | 3'905 | 3'639 | 3'491 | 3'391 | 3'304 | 3'346 | 3'354 | 3'358 | 3'800 | 3'476 |
| | A3 | 3'844 | 4'374 | 3'800 | 3'535 | 3'386 | 3'286 | 3'200 | 3'242 | 3'249 | 3'253 | 3'696 | 3'371 |
| | A4 | 4'081 | 4'612 | 4'038 | 3'772 | 3'623 | 3'523 | 3'437 | 3'479 | 3'486 | 3'490 | 3'933 | 3'609 |

GLM and Credibility Premium




Comparison of GLM and Credibility Premiums

# 5 The Classical Multiplicative Model

## 5.1 Model Assumptions

In a classical statistical sense, the following assumptions are behind the multiplicative tariff structure:

**Model-Assumptions 5.1 (classical multiplicative model)** *The observable random variables $X_{ij}$, $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$, satisfy*

$$E[X_{ij}] = \mu_0 \cdot \psi_i \cdot \varphi_j, \tag{5.1}$$

*where $\psi_i$ and $\varphi_j$ are real numbers.*

The aim is again to estimate for each $cell_{ij}$ the corresponding premium

$$P_{ij} = E[X_{ij}] = \mu_0 \cdot \psi_i \cdot \varphi_j.$$

**Remark:**

- The tariff parameters $\mu_0, \psi_i$ and $\varphi_j$ are determined only up to a constant factor, i.e. if we multiply the $\psi_i$ with a constant factor $c_1$ and the $\varphi_j$ with a constant factor $c_2$ and then replace $\mu_0$ by $\mu_0^* = \mu_0 / c_1 c_2$, we get the same tariff.
- To make the tariff parameters uniquely defined we have to fix $\mu_0$. In the following, $\mu_0$ is referred to as the average premium, i.e.

$$\mu_0 = E\left[X_{\bullet\bullet}\right], \tag{5.2}$$

where

$$\overline{X}_{\bullet\bullet} = \sum_{ij} \frac{w_{ij}}{w_{\bullet\bullet}} X_{ij}.$$

## 5.2 Estimators of the tariff parameters

In the insurance practice, the tariff parameters $\mu_0, \psi_i, \varphi_j$ are unknown and have to be estimated from the observations $X_{ij}$. Several estimators have been suggested in the actuarial literature. Contrary to the additive model, the *weighted least square estimators* and the estimators obtained by the *method of marginal totals* are no more identical. We concentrate here on the method of marginal totals.

In the multiplicative model the marginal total conditions are

$$\sum_j w_{ij}(\widehat{\mu_0} \cdot \widehat{\psi_i} \cdot \widehat{\varphi_j}) = \sum_j w_{ij} X_{ij}, \tag{5.3}$$

$$\sum_i w_{ij}(\widehat{\mu_0} \cdot \widehat{\psi_i} \cdot \widehat{\varphi_j}) = \sum_i w_{ij} X_{ij}. \tag{5.4}$$

Again $\widehat{\mu_0}, \widehat{\psi_i}, \widehat{\varphi_j}$ are defined only up to a constant factor. Setting $\widehat{\mu_0}$ equal to the observed average then the *estimators of the method of marginal totals* are given by the following system of equations which can be solved iteratively by starting for instance with $\widehat{\varphi_j} = \overline{X}_{\bullet j}/\overline{X}_{\bullet\bullet}$.

$$\widehat{\mu_0} = \overline{X}_{\bullet\bullet}, \tag{5.5}$$

$$\widehat{\psi_i}\left(\sum_j w_{ij}\,\widehat{\varphi_j}\right) = \sum_j w_{ij}\frac{X_{ij}}{\overline{X}_{\bullet\bullet}}, \tag{5.6}$$

$$\widehat{\varphi_j}\left(\sum_i w_{ij}\widehat{\psi_i}\right) = \sum_i w_{ij}\frac{X_{ij}}{\overline{X}_{\bullet\bullet}}. \tag{5.7}$$

The following result is well known from the literature (see for instance [12]).

**Theorem 5.2** *Under the Model Assumptions 5.1 and if the random variables $X_{ij}$ are independent and Poisson distributed (or overdispersed Poisson distributed), then the method of marginal totals yields the Maximum Likelihood estimators.*

**Remark:**

- Since GLM estimators are maximum likelihood estimators in the case of a distribution of the exponential dispersion family, the estimators resulting from the method of marginal totals are the same as the GLM estimators in the case where the $X_{ij}$ are (overdispersed) Poisson distributed.

# 6 The Bayesian Multiplicative Model

## 6.1 Model Assumptions

In the Bayesian model the risk parameters $\psi_i$ and $\varphi_j$ are assumed to be realizations of random variables $\Psi_i$ and $\Phi_j$, and that conditionally, given $\Psi_i$ and $\Phi_j$, the assumptions of the classical multiplicative model (Model Assumptions 5.1) are fulfilled.

**Model-Assumptions 6.1 (Bayes Multiplicative Model)**

*i) The r.v. $X_{ij}$, $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$, are conditionally, given $\Theta_{ij} = (\Psi_i, \Phi_j)$, independent with*

$$E[X_{ij}|\,\Theta_{ij}] = \mu_0 \cdot \Psi_i \cdot \Phi_j, \tag{6.1}$$

$$\begin{aligned} \mathrm{Var}(X_{ij}|\,\Theta_{ij}) &= \frac{\sigma^2(\Theta_{ij})}{w_{ij}} \\ &= \frac{\eta \cdot (\mu_0 \cdot \Psi_i \cdot \Phi_j)^p}{w_{ij}}, \end{aligned} \tag{6.2}$$

*where $\eta$ and $p \in \mathbb{R}^+$ .*

*ii) The r.v. $\Psi_i$, $i = 1, \ldots, I$, are independent and identically distributed (i.i.d.) with*

$$E[\Psi_i] = \mu_\Psi = 1, \tag{6.3}$$

$$\mathrm{Var}(\Psi_i) = \tau_\Psi^2. \tag{6.4}$$

*iii) The r.v. $\Phi_j$, $j = 1, 2, \ldots, J$, are i.i.d. with*

$$E[\Phi_j] = \mu_\Phi = 1, \tag{6.5}$$

$$\mathrm{Var}(\Phi_j) = \tau_\Phi^2. \tag{6.6}$$

*iv) $\Psi_i, \Phi_j$ are independent.*

**Remark:**

- The variance condition (6.2) is fulfilled for the Tweedie family of distributions (see for instance in [9]). It includes the family of the normal distributions ($p = 1$), the family of the (overdispersed) Poisson distributions ($p = 1$), the family of the compound Poisson-distributions with Gamma distributed claim severities ($1 < p < 2$) and the family of the Gamma distributions ($p = 2$). The parameter $\eta$ is called dispersion parameter in the literature.

## 6.2 Credibility Estimator

The aim is to find for each $cell_{ij}$ an estimator of the corresponding true individual premium

$$P_{ij} = E[X_{ij} | \Theta_{ij}] = \mu(\Theta_{ij}) = \mu_0 \cdot \Psi_i \cdot \Phi_j. \tag{6.7}$$

By definition the credibility estimator $P_{ij}^{Cred}$ is an estimator which is a linear function of the observations $X_{ij}$. However, given the multiplicative structure in (6.7), such a linear estimator would not be meaningful. Therefore we have to find another procedure suited to the multiplicative structure.

**Definition 6.2** *We denote by $\Psi_i^*(\mathbf{\Phi})$ resp. $\Phi_j^*(\mathbf{\Psi})$ the credibility estimators of $\Psi_i$ resp. $\Phi_j$ on the condition given $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_J)'$ resp. $\mathbf{\Psi} = (\Psi_1, \ldots, \Psi_I)'$.*

Note that $\Psi_i^*(\mathbf{\Phi})$ and $\Phi_j^*(\mathbf{\Psi})$ depend on the hidden unknown random variables $\Psi_i$ and $\Phi_j$ respectively we want to estimate. Such estimators are often called "pseudo estimators" in the actuarial literature.

**Theorem 6.3** *Under the Model Assumptions 6.1 $\Psi_i^*(\mathbf{\Phi})$ and $\Phi_j^*(\mathbf{\Psi})$ are given by*

$$\Psi_i^*(\mathbf{\Phi}) = 1 + \alpha_i \left( \overline{X}_{i\bullet}^{(1)} - 1 \right), \tag{6.8}$$

$$\Phi_j^*(\mathbf{\Psi}) = 1 + \beta_j \left( \overline{X}_{\bullet j}^{(2)} - 1 \right), \tag{6.9}$$

*where*

$$\begin{aligned}
X_{ij}^{(1)} &= \frac{X_{ij}}{\Phi_j \mu_0}, \\
w_{ij}^{(1)} &= w_{ij} \left( \Phi_j \mu_0 \right)^{2-p}, \\
\overline{X}_{i\bullet}^{(1)} &= \sum_{j=1}^{J} \frac{w_{ij}^{(1)}}{w_{i\bullet}^{(1)}} X_{ij}^{(1)}, \\
\alpha_i &= \frac{w_{i\bullet}^{(1)}}{w_{i\bullet}^{(1)} + \frac{\sigma_\Psi^2}{\tau_\Psi^2}}, \\
\sigma_\Psi^2 &= \eta \cdot E[\Psi_i^p],
\end{aligned}$$

$$\begin{aligned}
X_{ij}^{(2)} &= \frac{X_{ij}}{\Psi_i \mu_0}, \\
w_{ij}^{(2)} &= w_{ij} \left( \Psi_i \mu_0 \right)^{2-p}, \\
\overline{X}_{\bullet j}^{(2)} &= \sum_{i=1}^{I} \frac{w_{ij}^{(2)}}{w_{\bullet j}^{(2)}} X_{ij}^{(2)}, \\
\beta_j &= \frac{w_{\bullet j}^{(2)}}{w_{\bullet j}^{(2)} + \frac{\sigma_\Phi^2}{\tau_\Phi^2}}, \\
\sigma_\Phi^2 &= \eta \cdot E[\Phi_j^p].
\end{aligned}$$

**Proof:**

On the condition given $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_J)'$, the random variables $X_{ij}^{(1)}$ together with the weights $w_{ij}^{(1)}$ fulfill the conditions of the Bühlmann and Straub model with

$$\begin{aligned} E\left[ X_{ij}^{(1)} \middle| \Theta_{ij} \right] &= \mu(\Psi_i) = \Psi_i, \\ \operatorname{Var}\left( X_{ij}^{(1)} \middle| \Theta_{ij} \right) &= \frac{\sigma^2(\Psi_i)}{w_{ij}^{(1)}} = \frac{\eta \cdot \Psi_i^p}{w_{ij}^{(1)}}. \end{aligned}$$

Therefore, the credibility estimator $\Psi_i^*(\mathbf{\Phi})$ based on $X_{ij}^{(1)}$ is given by (6.8). (6.9) can be proved analogously. □

$\Psi_i^*(\mathbf{\Phi})$ and $\Phi_j^*(\mathbf{\Psi})$ are pseudo estimators. However, we want to find estimators of $\Psi_i$ and $\Phi_j$ depending only on the observations. Given the iterative procedure for determining the credibility procedure in the additive model and given also the iterative procedure for determining the tariff parameters by the method of marginal totals in the classical multiplicative model, it looks natural, to proceed analogously in the multiplicative model. This means that we start with some meaningful initial estimate $\mathbf{\Phi}^{(1)}$. Inserting $\mathbf{\Phi}^{(1)}$ into (6.8) yields $\mathbf{\Psi}^{(1)}$, inserting $\mathbf{\Psi}^{(1)}$ into (6.9) gives $\mathbf{\Phi}^{(2)}$, and so on. In general, the iterative procedure converges rather quickly. Thus we suggest to use the following "credibility based" estimators:

**Estimator 6.4 (credibility based estimator )**

*i) The credibility based estimators of $\Psi_i$ and $\Phi_j$ in the multiplicative model are given by*

$$\Psi_i^{(Cred)} = \Psi_i^*\left(\mathbf{\Phi}^{(Cred)}\right), \tag{6.10}$$

$$\Phi_j^{(Cred)} = \Phi_j^*\left(\mathbf{\Psi}^{(Cred)}\right), \tag{6.11}$$

*where*

$$\begin{aligned} \mathbf{\Phi}^{(Cred)} &= (\Phi_1^{(Cred)}, \ldots, \Phi_J^{(Cred)})', \\ \mathbf{\Psi}^{(Cred)} &= (\Psi_1^{(Cred)}, \ldots, \Psi_I^{(Cred)})', \end{aligned}$$

*and where $\Psi_i^*(.)$ and $\Phi_j^*(.)$ are defined by* (6.8) *and* (6.9).

*ii) The credibility based premium in the multiplicative model is given by*

$$P_{ij}^{(Cred)} = \mu_0 \cdot \Psi_i^{(Cred)} \cdot \Phi_j^{(Cred)}. \tag{6.12}$$

**Remarks:**

- Note that we have written *credibility based* estimator and not credibility estimator and $\Psi_i^{(Cred)}$, $\Phi_j^{(Cred)}$ and not $\Psi_i^{Cred}$, $\Phi_j^{Cred}$. The reason is that these are not really credibility estimators, because they are not a linear function of the observations.
- (6.10) and (6.11) can be solved iteratively.

## 6.3 Estimation of structural parameters

The structural parameters can also be calculated iteratively by applying in each iterative step the standard estimators of the structural parameters in the Bühlmann and Straub model (the exact form of these estimators can for instance be found in [2], Section 4.8).

## 6.4 Numerical Example

The following table shows the observed claim frequencies of large claims of a larger Swiss insurance company in the line third-party motor liability. We make the usual assumption that the claim number is conditionally Poisson distributed. From this assumption follows, that the dispersion parameter $\eta$ is equal to one and that $p = 2$. To illustrate the method we again consider only two rating factors. The first factor $A$ has 4 levels and the second factor $B$ has 27 levels. The exposure measure is the number of year risks.

**Observed claim frequency**

| | Factor A | | | |
|---|---|---|---|---|
| Factor B | A1 | A2 | A3 | A4 |
| B1 | 0.1% | 0.1% | 0.0% | 0.2% |
| B2 | 0.1% | 0.0% | 0.0% | 2.4% |
| B3 | 0.1% | 0.0% | 0.0% | 0.2% |
| B4 | 0.1% | 0.1% | 0.1% | 0.2% |
| B5 | 0.1% | 0.2% | 0.0% | 0.1% |
| B6 | 0.0% | 0.2% | 0.2% | 0.1% |
| B7 | 0.1% | 0.0% | 0.2% | 0.1% |
| B8 | 0.1% | 0.1% | 0.1% | 0.0% |
| B9 | 0.0% | 0.0% | 0.0% | 0.0% |
| B10 | 0.0% | 0.1% | 0.3% | 0.0% |
| B11 | 0.1% | 0.0% | 0.0% | 0.0% |
| B12 | 0.1% | 0.1% | 0.0% | 0.2% |
| B13 | 0.0% | 0.1% | 0.2% | 0.5% |
| B14 | 0.0% | 0.0% | 0.0% | 0.0% |
| B15 | 0.0% | 0.0% | 0.0% | 0.0% |
| B16 | 0.1% | 0.1% | 0.3% | 0.2% |
| B17 | 0.0% | 0.0% | 0.0% | 0.5% |
| B18 | 0.1% | 0.0% | 0.0% | 0.0% |
| B19 | 0.0% | 0.0% | 0.0% | 0.0% |
| B20 | 0.1% | 0.0% | 0.0% | 0.3% |
| B21 | 0.0% | 0.1% | 0.0% | 0.0% |
| B22 | 0.0% | 0.0% | 0.0% | 0.0% |
| B23 | 0.0% | 0.1% | 0.0% | 0.1% |
| B24 | 0.1% | 0.0% | 0.3% | 0.1% |
| B25 | 0.1% | 0.0% | 0.0% | 0.2% |
| B26 | 0.1% | 0.1% | 0.1% | 0.2% |
| B27 | 0.1% | 0.0% | 0.0% | 0.0% |

**Number of year risks**

| | Factor A | | | |
|---|---|---|---|---|
| Factor B | A1 | A2 | A3 | A4 |
| B1 | 38'537 | 1'764 | 437 | 2'300 |
| B2 | 918 | 30 | 7 | 41 |
| B3 | 5'506 | 265 | 101 | 424 |
| B4 | 82'838 | 2'742 | 1'472 | 3'285 |
| B5 | 16'270 | 1'205 | 249 | 1'113 |
| B6 | 7'077 | 845 | 462 | 985 |
| B7 | 15'686 | 583 | 1'267 | 1'017 |
| B8 | 25'342 | 4'613 | 4'096 | 3'753 |
| B9 | 2'853 | 155 | 46 | 162 |
| B10 | 16'516 | 934 | 389 | 580 |
| B11 | 3'886 | 236 | 178 | 141 |
| B12 | 26'724 | 786 | 380 | 990 |
| B13 | 13'235 | 1'872 | 1'266 | 555 |
| B14 | 4'273 | 90 | 26 | 53 |
| B15 | 2'816 | 54 | 17 | 88 |
| B16 | 40'676 | 2'790 | 656 | 4'190 |
| B17 | 9'505 | 491 | 171 | 1'030 |
| B18 | 15'790 | 635 | 124 | 948 |
| B19 | 9'357 | 356 | 82 | 533 |
| B20 | 21'309 | 1'393 | 356 | 1'313 |
| B21 | 25'746 | 5'348 | 682 | 1'258 |
| B22 | 2'245 | 55 | 15 | 77 |
| B23 | 43'607 | 4'877 | 3'839 | 3'011 |
| B24 | 13'842 | 975 | 1'121 | 826 |
| B25 | 7'716 | 246 | 56 | 412 |
| B26 | 107'242 | 6'969 | 1'933 | 7'423 |
| B27 | 45'225 | 497 | 110 | 161 |

**Number of large claims**

| | Factor A | | | |
|---|---|---|---|---|
| Factor B | A1 | A2 | A3 | A4 |
| B1 | 27 | 1 | 0 | 4 |
| B2 | 1 | 0 | 0 | 1 |
| B3 | 3 | 0 | 0 | 1 |
| B4 | 62 | 3 | 1 | 5 |
| B5 | 12 | 3 | 0 | 1 |
| B6 | 2 | 2 | 1 | 1 |
| B7 | 9 | 0 | 3 | 1 |
| B8 | 17 | 5 | 6 | 1 |
| B9 | 0 | 0 | 0 | 0 |
| B10 | 7 | 1 | 1 | 0 |
| B11 | 3 | 0 | 0 | 0 |
| B12 | 22 | 1 | 0 | 2 |
| B13 | 4 | 2 | 2 | 3 |
| B14 | 1 | 0 | 0 | 0 |
| B15 | 0 | 0 | 0 | 0 |
| B16 | 29 | 2 | 2 | 9 |
| B17 | 4 | 0 | 0 | 5 |
| B18 | 12 | 0 | 0 | 0 |
| B19 | 2 | 0 | 0 | 0 |
| B20 | 12 | 0 | 0 | 4 |
| B21 | 11 | 7 | 0 | 0 |
| B22 | 1 | 0 | 0 | 0 |
| B23 | 11 | 4 | 1 | 2 |
| B24 | 8 | 0 | 3 | 1 |
| B25 | 5 | 0 | 0 | 1 |
| B26 | 102 | 7 | 2 | 14 |
| B27 | 33 | 0 | 0 | 0 |

Observations

We have applied the techniques presented in sections 5 and 6 to these data. Since the assumption that the claim number is conditionally Poisson distributed we get $\sigma_\Phi^2 = \sigma_\Psi^2 = 1$. For the remaining structural parameters we obtained

| $\widehat{\mu_0}$ | $\widehat{\tau_\Psi^2}$ | $\widehat{\tau_\Phi^2}$ |
|---|---|---|
| 0.07% | 0.28 | 0.06 |

The following table and the corresponding graph show the credibility based estimators as well as the GLM estimators of the factor effects. Note that in this case the GLM estimators are obtained by the method of marginal totals.

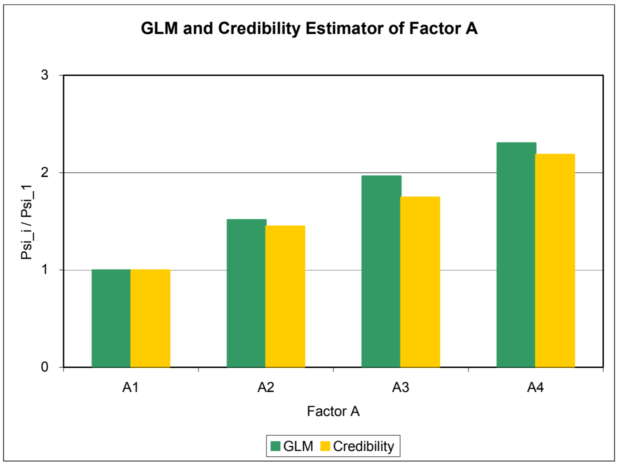
In the following table and graph, we can again well see the "smoothing" effect of the credibility based estimators, in particular for the factors with small exposure (number of year risks) in the underlying cells (e.g. effect of factors $B2$). A special situation occurs for $B9$ and $B15$, where the number of large claims is zero in all underlying cells. The GLM estimator yields a value of zero for $\Phi_9$ and $\Phi_{15}$, which is of course not meaningful. The credibility based estimates of $\Phi_9$ and $\Phi_{15}$, however, are smoothed towards one and are reasonable.

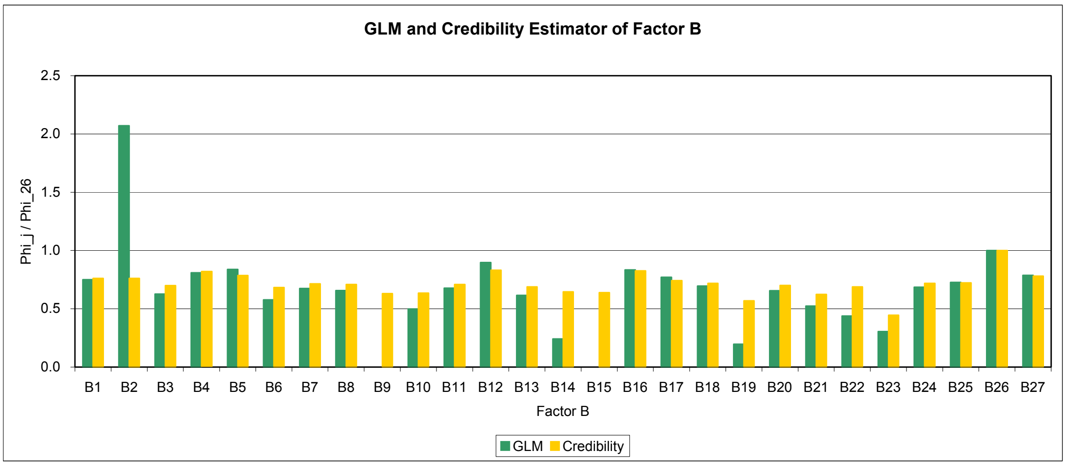
**Estimators of Factor Effects**

| Factor A | $\psi_1$ | $\psi_2$ | $\psi_3$ | $\psi_4$ |
|---|---|---|---|---|
| GLM | 0.10 | 0.15 | 0.19 | 0.23 |
| Credibility | 0.83 | 1.20 | 1.44 | 1.81 |

| Factor B | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\phi_4$ | $\phi_5$ | $\phi_6$ | $\phi_7$ | $\phi_8$ | $\phi_9$ | $\phi_{10}$ | $\phi_{11}$ | $\phi_{12}$ | $\phi_{13}$ | $\phi_{14}$ | $\phi_{15}$ | $\phi_{16}$ | $\phi_{17}$ | $\phi_{18}$ | $\phi_{19}$ | $\phi_{20}$ | $\phi_{21}$ | $\phi_{22}$ | $\phi_{23}$ | $\phi_{24}$ | $\phi_{25}$ | $\phi_{26}$ | $\phi_{27}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GLM | 9.34 | 25.80 | 7.81 | 10.07 | 10.46 | 7.19 | 8.40 | 8.18 | 0.00 | 6.21 | 8.43 | 11.17 | 7.67 | 3.02 | 0.00 | 10.40 | 9.60 | 8.65 | 2.45 | 8.15 | 6.53 | 5.45 | 3.80 | 8.54 | 9.07 | 12.46 | 9.80 |
| Credibility | 1.07 | 1.07 | 0.99 | 1.16 | 1.11 | 0.96 | 1.01 | 1.00 | 0.89 | 0.90 | 1.00 | 1.17 | 0.97 | 0.91 | 0.90 | 1.17 | 1.05 | 1.01 | 0.80 | 0.99 | 0.88 | 0.97 | 0.63 | 1.01 | 1.02 | 1.41 | 1.10 |

Factor Effects




Comparison of Factor Effects

Finally, the following table and the corresponding graph show the resulting estimators of the frequency. Note in particular the columns $B9$ and $B15$ in the table GLM frequency: the estimated frequencies are zero, which does not make sense. The estimated frequencies in the same columns of the table credibility based frequency do not have this deficiency.

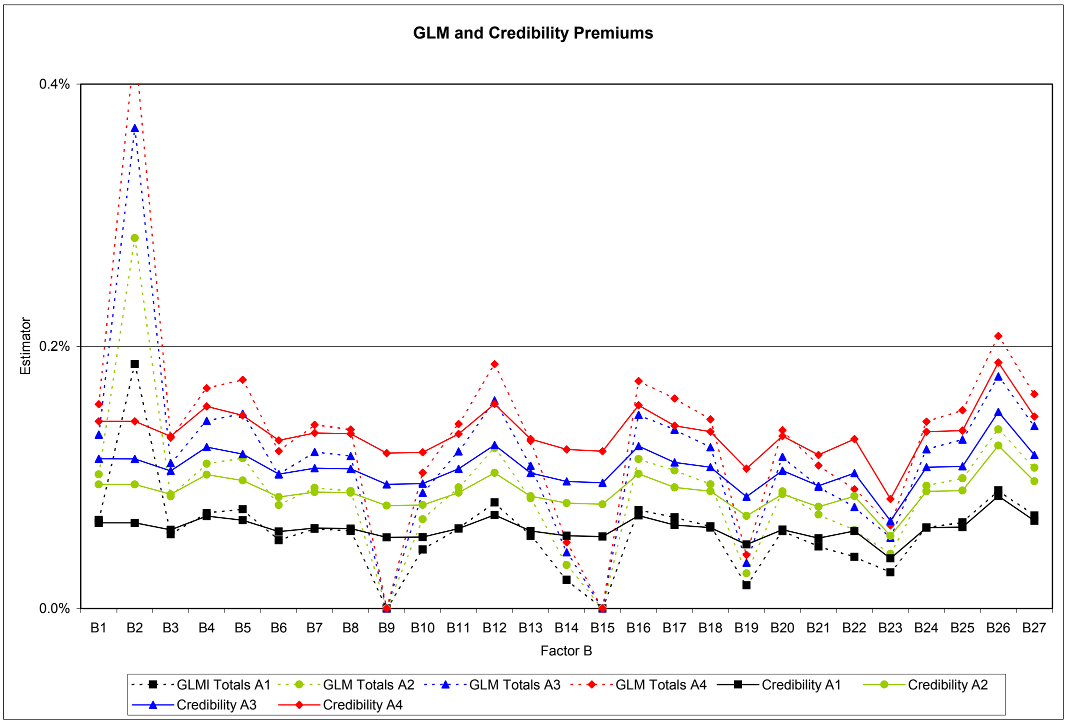
**GLM large claim frequency**

| | | Factor A | | | |
|---|---|---|---|---|---|
| | | A1 | A2 | A3 | A4 |
| Factor B | B1 | 0.07% | 0.10% | 0.13% | 0.16% |
| | B2 | 0.19% | 0.28% | 0.37% | 0.43% |
| | B3 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B4 | 0.07% | 0.11% | 0.14% | 0.17% |
| | B5 | 0.08% | 0.11% | 0.15% | 0.17% |
| | B6 | 0.05% | 0.08% | 0.10% | 0.12% |
| | B7 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B8 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B9 | 0.00% | 0.00% | 0.00% | 0.00% |
| | B10 | 0.04% | 0.07% | 0.09% | 0.10% |
| | B11 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B12 | 0.08% | 0.12% | 0.16% | 0.19% |
| | B13 | 0.06% | 0.08% | 0.11% | 0.13% |
| | B14 | 0.02% | 0.03% | 0.04% | 0.05% |
| | B15 | 0.00% | 0.00% | 0.00% | 0.00% |
| | B16 | 0.08% | 0.11% | 0.15% | 0.17% |
| | B17 | 0.07% | 0.11% | 0.14% | 0.16% |
| | B18 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B19 | 0.02% | 0.03% | 0.03% | 0.04% |
| | B20 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B21 | 0.05% | 0.07% | 0.09% | 0.11% |
| | B22 | 0.04% | 0.06% | 0.08% | 0.09% |
| | B23 | 0.03% | 0.04% | 0.05% | 0.06% |
| | B24 | 0.06% | 0.09% | 0.12% | 0.14% |
| | B25 | 0.07% | 0.10% | 0.13% | 0.15% |
| | B26 | 0.09% | 0.14% | 0.18% | 0.21% |
| | B27 | 0.07% | 0.11% | 0.14% | 0.16% |

**Credibility large claim frequency**

| | | Factor A | | | |
|---|---|---|---|---|---|
| | | A1 | A2 | A3 | A4 |
| Factor B | B1 | 0.07% | 0.09% | 0.11% | 0.14% |
| | B2 | 0.07% | 0.09% | 0.11% | 0.14% |
| | B3 | 0.06% | 0.09% | 0.10% | 0.13% |
| | B4 | 0.07% | 0.10% | 0.12% | 0.15% |
| | B5 | 0.07% | 0.10% | 0.12% | 0.15% |
| | B6 | 0.06% | 0.08% | 0.10% | 0.13% |
| | B7 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B8 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B9 | 0.05% | 0.08% | 0.09% | 0.12% |
| | B10 | 0.05% | 0.08% | 0.10% | 0.12% |
| | B11 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B12 | 0.07% | 0.10% | 0.12% | 0.16% |
| | B13 | 0.06% | 0.09% | 0.10% | 0.13% |
| | B14 | 0.06% | 0.08% | 0.10% | 0.12% |
| | B15 | 0.05% | 0.08% | 0.10% | 0.12% |
| | B16 | 0.07% | 0.10% | 0.12% | 0.15% |
| | B17 | 0.06% | 0.09% | 0.11% | 0.14% |
| | B18 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B19 | 0.05% | 0.07% | 0.09% | 0.11% |
| | B20 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B21 | 0.05% | 0.08% | 0.09% | 0.12% |
| | B22 | 0.06% | 0.09% | 0.10% | 0.13% |
| | B23 | 0.04% | 0.06% | 0.07% | 0.08% |
| | B24 | 0.06% | 0.09% | 0.11% | 0.13% |
| | B25 | 0.06% | 0.09% | 0.11% | 0.14% |
| | B26 | 0.09% | 0.12% | 0.15% | 0.19% |
| | B27 | 0.07% | 0.10% | 0.12% | 0.15% |

GLM and credibility based frequency



Addresses of authors:


Alois Gisler
Chief Actuary Non-Life
Winterthur Insurance Company
P.O. Box 357
CH 8401 Winterthur

Email: alois.gisler@winterthur.ch

Petra Müller
Winterthur Insurance Company
P.O. Box 357
CH 8401 Winterthur

Email: petra.mueller@winterthur.ch

# Appendix

## A Summary of Some Results from Credibility Theory.

Assume that $\mathbf{X} = (X_1, \ldots, X_j)'$ is a vector of observable random variables and that we want to estimate another random variable $\mu(\Theta)$. Usually, the random variable $\Theta$ is the latent risk characteristic of an individual risk to be rated and $\mu(\Theta)$, its pure risk premium, is a function of $\Theta$.

An estimator $\widehat{\mu_1(\Theta)}$ of $\mu(\Theta)$ is said to be a better estimator than $\widehat{\mu_2(\Theta)}$ if

$$E\left[\left(\widehat{\mu_1(\Theta)} - \mu(\Theta)\right)^2\right] < E\left[\left(\widehat{\mu_2(\Theta)} - \mu(\Theta)\right)^2\right]. \tag{A.1}$$

**Definition A.1**

*i) The credibility estimator of $\mu(\Theta)$ (based on $\mathbf{X}$) is the best possible estimator in the class*

$$L(\mathbf{X}, 1) := \left\{\widehat{\mu(\Theta)} : \widehat{\mu(\Theta)} = a_0 + \sum a_j X_j, \quad a_0, a_1, \cdots \in \mathbb{R}\right\}. \tag{A.2}$$

*ii) The homogeneous credibility estimator of $\mu(\Theta)$ (based on $\mathbf{X}$) is the best possible estimator in the class*

$$L_e(\mathbf{X}) := \left\{\widehat{\mu(\Theta)} : \widehat{\mu(\Theta)} = \sum a_j X_j, \; a_1, a_2, \cdots \in \mathbb{R}, \; \sum a_j \mu_{X_j} = E[\mu(\Theta)]\right\} \tag{A.3}$$

*where $\mu_{X_j} = E[X_j]$.*

We denote the credibility estimator of $\mu(\Theta)$ by $\widehat{\widehat{\mu(\Theta)}}$ and the homogeneous credibility estimator by $\widehat{\widehat{\mu(\Theta)}}^{\text{hom}}$. These estimators defined as a solution to minimizing the mean square error within given classes of estimators are most elegantly understood as projections on the Hilbert space of all square integrable functions $\mathcal{L}^2$. The following definition is equivalent to the definition A.1.

**Definition A.2**

*i) The (inhomogeneous) credibility estimator is defined as*

$$\widehat{\widehat{\mu(\Theta)}} = \text{Pro}(\mu(\Theta) | \, L(\mathbf{X}, 1)). \tag{A.4}$$

*ii) The homogeneous credibility estimator is defined as*

$$\widehat{\widehat{\mu(\Theta)}}^{\text{hom}} = \text{Pro}(\mu(\Theta) | \, L_e(\mathbf{X})). \tag{A.5}$$

The following well known result characterizes the credibility estimators.

**Theorem A.3 (normal equations)** $\widehat{\widehat{\mu(\Theta)}} = \widehat{a}_0 + \sum_j \widehat{a}_j X_j$ *is the credibility estimator of* $\mu(\Theta)$, *if and only if the following normal equations are satisfied:*

$$i) \qquad \widehat{a}_0 = \mu_0 - \sum_j \widehat{a}_j \, \mu_{X_j}. \tag{A.6}$$

$$ii) \qquad \sum_j \widehat{a}_j \mathrm{Cov}\,(X_j, X_k) = \mathrm{Cov}(\mu(\Theta), X_k), \quad k = 1, \ldots, n, \tag{A.7}$$

The following result on iterative projections is often very useful for the derivation of credibility estimators.

**Theorem A.4 (Iterativity of projections)** *Let* $M$ *and* $M'$ *be closed subspaces (or affine spaces) of* $\mathcal{L}^2$ *with* $M \subset M'$, *then we have*

$$\mathrm{Pro}\,(Y \mid M) = \mathrm{Pro}\,(\mathrm{Pro}\,(Y \mid M') \mid M) \tag{A.8}$$

*and*

$$\|Y - \mathrm{Pro}\,(Y \mid M)\|^2 = \|Y - \mathrm{Pro}\,(Y \mid M')\|^2 + \|\mathrm{Pro}\,(Y \mid M') - \mathrm{Pro}\,(Y \mid M)\|^2 \tag{A.9}$$

The Bühlmann and Straub model (see [3]) is still by far the most used and the most important credibility model for the insurance practice. In this model a portfolio of risks $i = 1, \ldots, I$ is considered. Each risk is characterized by a latent risk characteristics $\Theta_i$ and for each of the risks there is is given a random vector

$$\mathbf{X}_i = (X_{i1,} \ldots, X_{in})',$$

where $X_{ij}$ is the observation of risk $i$ in year $j$ associated with a weight $w_{ij}$.

**Model-Assumptions A.5 (Bühlmann Straub)** *The risk* $i$ *is characterized by an individual risk profile* $\vartheta_i$, *which is itself the realization of a random variable* $\Theta_i$, *and we have that:*

*BS1 Conditionally, given* $\Theta_i$, *the* $\{X_{ij} : j = 1, 2, \ldots, n\}$ *are independent with*

$$E\,[X_{ij} \mid \Theta_i, w_{ij}] \;=\; \mu\,(\Theta_i)\ , \tag{A.10}$$

$$\mathrm{Var}\,[X_{ij} \mid \Theta_i, w_{ij}] \;=\; \frac{\sigma^2(\Theta_i)}{w_{ij}}\ . \tag{A.11}$$

*BS2 The pairs* $(\Theta_1, \mathbf{X}_1)$, $(\Theta_2, \mathbf{X}_2)\,, \ldots$ *are independent, and* $\Theta_1$, $\Theta_2, \ldots$ *are independent and identically distributed.*

**Theorem A.6** *The credibility estimator in the Bühlmann-Straub Model (Model Assumptions A.5) is given by*

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha_i \overline{X}_{i\bullet} + (1-\alpha_i)\,\mu_0 = \mu_0 + \alpha_i \left(\overline{X}_{i\bullet} - \mu_0\right) , \tag{A.12}$$

$$\textit{where} \quad \overline{X}_{i\bullet} = \sum\nolimits_j \frac{w_{ij}}{w_{i\bullet}} X_{ij} , \tag{A.13}$$

$$w_{i\bullet} = \sum\nolimits_j w_{ij} , \tag{A.14}$$

$$\alpha_i = = \frac{w_{i\bullet}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{w_{i\bullet}}{w_{i\bullet} + \kappa} . \tag{A.15}$$

# References


[1] Bailey R. A. (1963). Insurance rates with minimum bias. Proceedings of the Casualty Actuarial Society, Vol. L, 4-11.

[2] Bühlmann H. and Gisler A. (2005). A Course in Credibility Theory and its Applications. Universitext, Springer Verlag.

[3] Bühlmann H. and Straub A. (1970). Glaubwürdigkeit für Schadensätze. Bulletin of Swiss Association of Actuaries, 111-133.

[4] Jung J. (1968). On automobile insurance rate making. ASTIN Bulletin, 5/1, 41-48.

[5] Müller P. (2006). Ausgleichsverfahren, verallgemeinerte lineare Modelle und Credibility bei additiven und multiplikativen Tarifstrukturen. Diploma thesis, Departement of Mathematics, ETH Zürich.

[6] McCullagh P. and Nelder J.A. (1989). Generalized linear models. Chapman and Hall, Cambridge, 2nd edition.

[7] Nelder J.A. and Verrall R.J. (1997). Credibility theory and generalized linear model. ASTIN-Bulletin 27/1, 71-82.

[8] Norberg R. (1986). Hierarchical credibility: Analysis of a random effect linear model with nested classification. Scandinavian Actuarial Journal 3-4, 204-222.

[9] Ohlsson E. and Johannsen B. (2006). Exact credibility and Tweedie models. ASTIN-Bulletin 36/1, 121-133.

[10] Radhakrishna Rao C. and Toutenburg H. (1995). Linear Models:Least Squares and Alternatives. Springer, New York.

[11] Scheffé H. (1959). The Analysis of Variance. John Wiley & Sons.

[12] Van Eghen J., Greup E.K. and Nijssen J.A. (1983). Rate making. Surveys of Actuarial Studies 2.