Credibility Rating in a Multiplicative Tariff

Esbjörn Ohlsson*

Mathematical Statistics
Stockholm University
SE-106 91 Stockholm, Sweden
Phone +46 8 16 45 58
Fax +46 8 612 67 17
E-mail: esbj@math.su.se

Abstract

Credibility theory is a useful tool for the rating of *multi-level factors* (MLFs), i.e. categorical rating factors with a large number of levels that can not be grouped in a natural way. In practice we often have a number of ordinary rating factors besides the MLF. We extend the Bühlmann-Straub estimator to the situation with both an MLF and ordinary rating factors in a multiplicative tariff. An application to private motor car insurance is presented.

Keywords

Rating, credibility theory, Bühlmann-Straub, Jewell's theorem, multi-level factor, multiplicative model.

*The author is grateful to Björn Johansson, Länsförsäkringar insurance group, for valuable help in the preparation of this manuscript.

1 Multi-level factors and credibility

In non-life insurance, a customary rating technique is to estimate price relativities of a number of rating factors in a multiplicative model. The estimation has traditionally been done by some heuristic technique like *the method of marginal totals*, but is nowadays often carried out by the aid of Generalized Linear Models (GLMs). Rating factors are often categorical with a few levels (e.g. 'Sex', 'Number of persons in the household') or a grouping of a continuous variable (e.g. 'Age group', 'Mileage class'). If data are to sparse to produce reliable estimates for some groups, merging of neighbouring groups is often a remedy. However, for a nominal categorical rating factor with a large number of levels, a *multi-level factor* (MLF), there is usually no simple way to form groups that are (sufficiently) risk homogenous. Here are some examples of MLFs.

Example 1 (Car model) In private motor car insurance it is well known that the model of the car is an important rating factor, both for third-party liability, hull and theft. Even after some initial grouping of very similar models, there are thousands of different car models, some of which are popular so that sufficient data is available, but for most models we have moderate or few data. There is no sensible way to group the models *a priori* (and if there was enough data for posterior grouping, the grouping would not be necessary!). Hence, car model is a typical MLF. □

Example 2 (Geographic area) In order to get risk homogenous geographic areas one often has to use a very fine subdivision of the country, based on for instance ZIP codes. Neighbouring areas can have quite different risk profiles and hence a prior grouping can be very hard to achieve and we are again left with an MLF. □

Example 3 (Experience rating) Using the *customer* as a rating factor is another important example of an MLF. In the commercial lines it is important to base the rating to some extent on the individual claims experience, even though there is often not sufficient data for separate rating of each company. This is the classical case for employing credibility theory. In the private lines, Lemaire (1995) and others use credibility estimators for the construction of optimal bonus/malus systems, with the customer as an MLF (in our terminology). □

As indicated by Example 3, MLFs can be rated using *credibility theory*. However, classical (Bühlmann-Straub) credibility theory does not treat the important case where we have ordinary rating factors (covariates) besides the MLF. In private motor car insurance, e.g., the MLFs 'Car model' and 'Geographic area' are just two out of a large number of rating factors. In Ohlsson and Johansson (2003) we presented credibility estimators for this case by deriving an extension of the famous theorem by Jewell (1974), under distributional assumptions. In the present text we will use a non-parametric approach, extending the Bühlmann-Straub estimator, to derive the same estimator. The advantage of the present paper is that no distributional assumptions have to be made. However, the main aim of the paper is to show the power of combining credibility theory with a multiplicative tariff, see e.g. the application in the last section.

2 The Bühlmann-Straub estimator with covariates

Suppose, then, that we have a number of ordinary rating factors, dividing the portfolio into I tariff cells. For simplicity, we discuss the case with just one MLF, with K levels (groups). We observe some *key ratio* Y_{ikt}, where i denotes tariff cell and k the MLF level. Y_{ikt} is the ratio of some observation X_{ikt} to a measure of exposure w_{ikt}. In classical credibility, Y is the risk premium, for which the observation X is claim cost and w is policy years. With GLMs it is customary to carry out separate analyses with Y in turn as claim frequency and average claim cost, respectively. Our setting covers all these cases. The repeated observations for the combination (i, k) are indexed by t, since they are often presented as repetition over time, which is relevant in experience rating. They could, of course, be any repeated observations of group k — in Example 1, e.g., they could represent the individual policies for car model k.

We will now generalize the Bühlmann-Straub estimator, which was derived for the situation without covariates, i.e. without the index i above. Let μ_i be the expected value in tariff cell i, given by the ordinary rating factors (covariates). The relative deviation of group k from this value is considered as a random effect U_k with $E(U_k) = 1$ and our basic multiplicative model becomes

$$E(Y_{ikt}|U_k) = \mu_i U_k \tag{1}$$

where μ_i in turn is the product of a number of price relativities for the ordinary rating factors. Note that $E[Y_{ikt}] = \mu_i$. Next we give some basic assumptions, extending the classical ones. All random variables are supposed to have finite second moments.

Assumption 1 *(a) The U_k; $k = 1, 2, \ldots, K$ are independent random variables, all with $E[U_k] = 1$ and $\mathrm{Var}[U_k] = a$, for some $a > 0$.*

(b) The MLF groups are independent, i.e. (Y_{ikt}, U_k) and $(Y_{i'k't'}, U_k')$ are independent as soon as $k \neq k'$.

(c) For any k, the collection of Y_{ikt}, for all relevant i and t, is a sequence of conditionally independent variables, given U_k.

(d) For some function $\sigma^2(\cdot)$ we have

$$\mathrm{Var}(Y_{ikt}|U_k) = \frac{\sigma_i^2\,\sigma^2(U_k)}{w_{ikt}} \tag{2}$$

where w_{ikt} is the measure of exposure and σ_i^2 is another weight depending on i (possibly through μ_i).

In (c), 'all relevant i and t' means that we go through all the repetitions t in all the combinations of tariff cell i and MLF level k where there are observations (often, level k is not represented in all cells i).

Note. In the important case where Y_{ikt}, conditionally on U_k, follows a GLM with *variance function* $v(\mu) = \mu^p$ — a so called Tweedie model — well-known results on GLMs combined with (1) give that $\sigma_i^2 = \mu_i^p$ och $\sigma^2(U_k) = \phi U_k^p$, where ϕ is the dispersion parameter. This includes the important Poisson case $p = 1$ and Gamma case $p = 2$. □

We will now transform the random variables so that we can bring back this situation to the Bühlmann-Straub estimator, as presented in, e.g., Goovaerts & Hoogstad (1987, p. 43ff). Put

$$\tilde{Y}_{ikt} = \frac{Y_{ikt}}{\mu_i} \qquad \tilde{w}_{ikt} = \frac{w_{ikt}\mu_i^2}{\sigma_i^2} \tag{3}$$

Note that

$$E(\tilde{Y}_{ikt}|U_k) = U_k \qquad E(\tilde{Y}_{ikt}) = 1 \tag{4}$$

and, with $\sigma^2 = E[\sigma^2(U_k)]$,

$$\mathrm{Var}(\tilde{Y}_{ikt}|U_k) = \frac{\sigma^2(U_k)}{\tilde{w}_{ikt}} \qquad E[\mathrm{Var}(\tilde{Y}_{ikt}|U_k)] = \frac{\sigma^2}{\tilde{w}_{ikt}} \tag{5}$$

This transformation requires that the values of μ_i and σ_i are known. In practice we already have estimates of the μ_i, usually from a Tweedie type GLM, like the Poisson or Gamma model, or from the method of marginal totals. In the Tweedie GLM case $\sigma_i^2 = \mu_i^p$, and so we automatically get an estimate of σ_i^2 by plugging in $\hat{\mu}_i$. As usual in credibility theory, we now look for the best linear predictor $\hat{U}_k$ for each k, i.e. the one minimizing the expected quadratic error.

Note. Out of tradition one speaks of *credibility estimators*, though *credibility predictors* is more accurate, since we are predicting a random variable, not estimating a parameter. We use the terms interchangeably here. □

Theorem 1 *The best linear predictor of U_k is*

$$\hat{U}_k = z_k \overline{U}_k + (1 - z_k) \tag{6}$$

where

$$z_k = \frac{\tilde{w}_{\cdot k \cdot}}{\tilde{w}_{\cdot k \cdot} + \sigma^2 / a} \tag{7}$$

and

$$\overline{U}_k = \frac{\sum_{i,t} \tilde{w}_{ikt} \tilde{Y}_{ikt}}{\tilde{w}_{\cdot k \cdot}} = \frac{\sum_{i,t} \tilde{w}_{ikt} Y_{ikt} / \mu_i}{\sum_{i,t} \tilde{w}_{ikt}} \tag{8}$$

with

$$\tilde{w}_{ikt} = \frac{w_{ikt} \mu_i^2}{\sigma_i^2} \qquad \tilde{w}_{\cdot k \cdot} = \sum_{i,t} \tilde{w}_{ikt} \tag{9}$$

We will return to the estimation of σ^2 and a in Section 2.1 below.

We see that $\hat{U}_k$ in (6) is a credibility weighted *adjustment factor* to the rating given the other factors. High credibility gives large weight to $\overline{U}_k$, which we call an *experience value* since it is a measure of the outcome Y as compared to the expectation μ, aggregated over the tariff cells where k occur.

Note that the original Bühlmann-Straub estimator is recovered from our result as the case without covariates, for which $\mu_i = \mu$, independent of i, if we simply multiply equation (6) by μ and choose $\sigma_i^2 = \mu^2$.

It can be shown, see Ohlsson and Johansson (2003, Section 2.3.2), that if $z_k = 1$ and $\sigma_i^2 = \mu_i^p$ then (6) represents the estimating equation we would have if U_k was treated just like another covariate in a GLM. (When $p = 1$ we get the equation for the method of marginal totals.) The satisfying conclusion is that if we have sufficient data, our method reduces to standard rating methods. On the other hand, when credibility is not that high, we put less emphasis on the experience value $\overline{U}_k$, and the adjustment factor will be closer to 1 than it would have been if treated like another covariate. With very little data, we will completely rely on the ordinary rating factors, i.e. $\hat{U} \approx 1$.

Note. In the GLM Tweedie case $\tilde{w}_{ikt} = w_{ikt}\mu_i^{2-p}$ and $\sigma^2 = \phi E[U^p]$ and we get the same result as was obtained by an extension of Jewell's theorem in (2.18)–(2.20) of Ohlsson and Johansson (2003) under distributional assumptions on both Y and U.□

Proof. The result follows from the Bühlmann-Straub estimator as presented in Goovaerts & Hoogstad (1987, p. 43ff) as follows. Change their X to Y and identify our index k as their i and let their repetitions ('periods') s be obtained by, for each k, running through all the tariff cells i where k occurs *and* all the repetitions t there within. Further, replace all occurances of 'the risk parameter' Θ_i by our random effect U_k and note that the conditional expectation $\mu(\Theta_i)$ is simply U_k. Then obtain the result directly from the equations on page 47 of Goovaerts & Hoogstad (1987). Note though, that since $E[U_k] = 1$ we do not have to use a so called 'collective estimator' — in our case the number '1' stands for the 'collective' and is given the weight $(1 - z_k)$.

A direct proof of the theorem is excluded here for the sake of brevity. □

2.1 Estimation of variance parameters

It remains to estimate the variance parameters σ^2 and a. Unbiased estimators are obtained directly from the corresponding estimators in Goovaerts & Hoogstad (1987, p. 48). Let n_k be the number of observations for group k and put

$$\hat{\sigma}_k^2 = \frac{1}{n_k - 1} \sum_{i,t} \tilde{w}_{ikt} (\tilde{Y}_{ikt} - \overline{U}_k)^2 \tag{10}$$

Then the unbiased estimators $\hat{s}^2$ and $\hat{a}$ in Goovaerts & Hoogstad become

$$\hat{\sigma}^2 = \frac{\sum_k (n_k - 1) \hat{\sigma}_k^2}{\sum_k (n_k - 1)} \tag{11}$$

and

$$\hat{a} = \frac{\sum_k \tilde{w}_{\cdot k \cdot} (\overline{U}_k - \overline{U}_\cdot)^2 - (K-1)\hat{\sigma}^2}{\tilde{w}_{\cdot\cdot} - \sum_i \tilde{w}_{\cdot k \cdot}^2 / \tilde{w}_{\cdot\cdot\cdot}} \tag{12}$$

where $\overline{U}_\cdot$ is the $\tilde{w}_{\cdot k \cdot}$-wighted average of the $\overline{U}_k$'s. A direct proof of the unbiasedness of these estimators can easily be derived along the lines of Ohlsson and Johansson (2003, Section 3.1). Note, though, that in our case these estimators are strictly unbiased only if μ_i is known, while in practice we plug in an estimate $\hat{\mu}_i$.

3 An algorithm

When rating ordinary factors one iterates between, in turn, the estimating equations for each of the rating factors — whether one uses the method of marginal totals or GLM. In many applications it is reasonable to extend the iteration to the prediction of MLF random effects U_k in (6), since there may be collinearity between the MLF and the ordinary rating factors. (An exception is bonus/malus systems, in which we might want the bonus factor to be an adjustment to a *given* tariff.) We get the following algorithm for simultaneous rating of ordinary factors and MLFs.

(0) Initially, let $\hat{U}_k = 1$ for all k.

(1) Estimate the parameters for the ordinary rating factors as usual, with $\hat{U}_k$ treated as a known part of the expectation (in the GLM case with log-link, this means that $\log(\hat{U}_k)$ is an *offset*-variable). This yields $\hat{\mu}_i$.

(2) Compute $\hat{\sigma}^2$ and $\hat{a}$ from (11) and (12), using $\hat{\mu}_i$ from Step 1.

(3) Use (6) to compute $\hat{U}_k$ for $k = 1, 2, \ldots, K$, using the estimates from Step 1 and 2.

(4) Return to Step 1 with the new $\hat{U}_k$ from Step 3.

Repeat Step 1-4 until convergence.

Note. The algorithm can be extended to the case with more than one MLF in a straight-forward fashion. However, we often find it easier to rate one MLF at a time, forming posterior groups of the MLF classes by $\hat{U}_k$ before going on to the next MLF. □

4 Application (Car Model Rating)

We present some results from a study on car hull insurance, using data from the Swedish insurance group Länsförsäkringar. The same study was earlier reported in Ohlsson and Johansson (2003). There are a number of ordinary rating factors, e.g., the car age group — the details are left out here for confidentiality reasons. The MLF is car model, having roughly 2 500 levels. The analysis was made separately for claim frequency and average claim amount. Here we present the results from the claim frequency part of the study only. In step 1 of the algorithm above standard GLM software was used.

The idea is to describe the car models as far as possible by using auxiliary rating factors like weight and weight/power ratio; then the residual variation is taken care of by the U-predictors. The introduction of auxiliary factors, as an aid in the risk classification of car models, decreases the (residual) variation of the car models, with the effect that the $\overline{U}_k$-values become more concentrated around 1, as seen in the bar chart in Figure 1.

Examples of the 2 500 credibility estimates are given in Table 1, both with and without auxiliary rating factors. The product of the estimates for the latter is called $\hat{\mu}_k$ here, since they are unique to each car model.

Note especially, that for car models with high credibility, the rating is hardly affected by the introduction of auxiliaries, as seen by comparing the "No auxiliaries" $\hat{U}_k$ to the "With auxiliaries" column $\hat{\mu}_k\hat{U}_k$. Of course, the introduction of auxiliaries makes the $\overline{U}_k$:s themselves move towards 1.

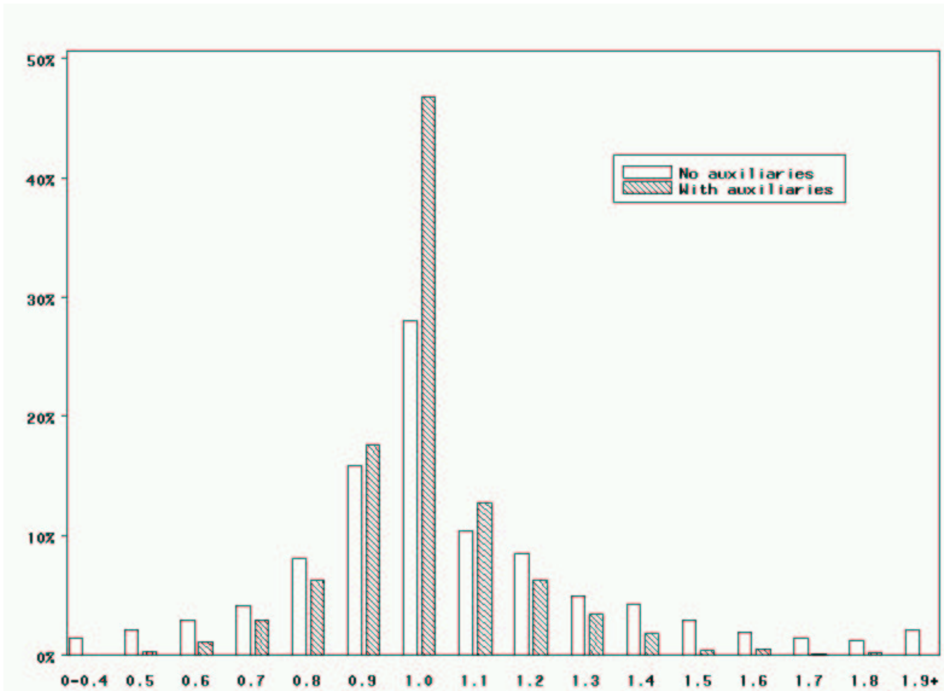

Figure 1: *Histogram of credibility predictors $\hat{u}_k$ with and without auxiliary rating factors.*

Sparse data gives low credibility, so for the car models at the end of the table one has to rely mostly on the auxiliary car model rating factors, while the credibility estimators $\hat{U}_k$ are close to 1.

In Sweden, historically there was a tradition of 'car model grouping' on a national level, but nowadays each company has to handle car models from their own data. Our experience is that the method presented here is a useful tool in that process: e.g., we might form new car model groups from the $\hat{\mu}_k \hat{U}_k$ in Table 1.

5 Concluding remarks

Credibility theory offers a solution to the problem of rating multi-level factors, MLFs. With the extension of the Bühlmann-Straub estimator presented here, credibility can quite easily be used in the context of a multiplicative tariff, a familiar environment for non-life actuaries. In our practical experience, this opens up for many new applications of credibility in various actuarial areas.

In this paper we have presented our results as a way to introduce fixed effects into the credibility framework. Another approach was followed in Ohlsson and Johansson (2003) where the same credibility predictor was derived by introducing random effects into GLMs. This situation is analogous to the classical Bühlmann-Straub estimator, which can be derived either non-parametrically or with distributional assumptions as in Jewell's theorem.

6 References

Goovaerts, M.J. and Hoogstad, W.J. (1987): *Credibility Theory.* Survey of Actuarial Studies No. 4, Nationale-Nederlanden N.V.

Jewell, W.S: (1974): *Credible means are exact Bayesian for exponential families.* ASTIN Bulletin, 8, 77-90.

Lemaire, J. (1995): *Bonus-Malus Systems in Automobile In-*

surance. Kluwer Academic Publishers, Norwell, Massachusetts.

OHLSSON, E. and JOHANSSON, B. (2003): *Credibility theory and GLM revisited.* (Submitted to ASTIN bulletin.) Research Report 2003:15, Mathematical Statistics, Stockholm University, `http://www.math.su.se/matstat/reports/seriea/`.

k	$w_{\cdot k}$	*No auxiliaries*			*With auxiliaries*			
		$\overline{U}_k$	$\hat{U}_k$	z_k	$\overline{U}_k$	$\hat{U}_k$	z_k	$\hat{\mu}_k\hat{U}_k$
1	41275	0.74	0.74	1.00	0.98	0.98	0.99	0.75
2	39626	0.58	0.58	1.00	0.89	0.89	0.99	0.59
3	39188	0.59	0.59	1.00	0.86	0.86	0.99	0.60
4	31240	0.82	0.82	1.00	0.93	0.93	0.99	0.82
5	28159	0.49	0.50	1.00	0.74	0.75	0.98	0.50
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
401	803	2.08	1.95	0.88	1.43	1.35	0.82	1.99
402	802	0.97	0.97	0.86	1.11	1.08	0.70	0.95
403	801	1.77	1.66	0.86	1.54	1.40	0.74	1.62
404	799	0.74	0.78	0.86	0.83	0.88	0.69	0.79
405	798	1.32	1.27	0.86	0.73	0.78	0.82	1.41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
901	181	1.38	1.22	0.58	1.14	1.06	0.42	1.29
902	180	1.61	1.38	0.63	0.91	0.95	0.56	1.70
903	180	2.28	1.76	0.59	1.35	1.18	0.51	2.01
904	179	0.79	0.88	0.56	0.86	0.95	0.34	0.88
905	179	2.38	1.80	0.58	1.52	1.25	0.48	1.98
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1801	7	2.39	1.07	0.05	2.05	1.03	0.03	1.22
1802	7	4.63	1.19	0.05	3.86	1.08	0.03	1.31
1803	7	0.00	0.96	0.04	0.00	0.99	0.01	0.55
1804	7	0.00	0.95	0.05	0.00	0.98	0.02	0.87
1805	7	0.00	0.94	0.06	0.00	0.98	0.02	0.58
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 1: Selected car models k, sorted by number of policy years $w_{\cdot k \cdot}$, with experience factors $\overline{U}_k$, credibility predictors $\hat{U}_k$ and credibility factors z_k; without and with auxiliary rating factors, the product of the latter called $\hat{\mu}_k$ here.