EXACT CREDIBILITY AND TWEEDIE MODELS

BY

ESBJÖRN OHLSSON AND BJÖRN JOHANSSON

ABSTRACT

Kaas, Dannenburg & Goovaerts (1997) generalized Jewell's theorem on exact credibility, from the classical Bühlmann model to the (weighted) Bühlmann-Straub model. We extend this result further to the "Bühlmann-Straub model with a priori differences" (Bühlmann & Gisler, 2005). It turns out that exact credibility holds for a class of Tweedie models, including the Poisson, gamma and compound Poisson distribution – the most important distributions for insurance applications of generalized linear models (GLMs). Our results can also be viewed as an alternative to the HGLM approach for combining credibility and GLMs, see Nelder and Verrall (1997).

KEYWORDS

Credibility theory, a priori differences, Jewell's theorem, generalized linear models, Tweedie models.

1. INTRODUCTION

Credibility estimators are often derived under the assumption of linearity in the observations. Exact credibility, on the other hand, refers to the situation where we can find distributional assumptions under which this linear credibility estimator is optimal (in mean square error) in the unrestricted class of all functions of the observations.

A famous theorem by Jewell (1974) states that exact credibility occurs when observations are drawn from a one-parameter exponential distribution with natural conjugate prior for the risk parameter. Another way to put this is that Jewell gives distributional assumptions under which the Bühlmann (1967) linear credibility estimator is the unrestricted best estimator.

This result was generalized by Kaas, Dannenburg & Goovaerts (1997) to the weighted Bühlmann-Straub model. It may be noted that they assume the observations to follow an exponential dispersion model, the class of distributions used in Generalized Linear Models (GLMs).

Astin Bulletin 36(1), 121-133. doi: 10.2143/AST.36.1.2014146

The Bühlmann-Straub model can be extended to the case known as "the Bühlmann-Straub model with a priori differences", where different contracts may have different a priori means, see Bühlmann & Gisler (2005, Chapter 4.13). We agree with these authors that this extension "enormously increases the applicability of the Bühlmann-Straub model in practice."

In practice this corresponds to the important situation where one has a number of ordinary rating factors alongside the factor estimated by credibility. An example is private motor car insurance, where we perform a GLM analysis with a number of rating factors such as *Sex* and *Age of driver*, *Age of car*, *Mileage per year*, or *Power of engine*, whereas a credibility approach is appropriate for the factor *Car model*. The reason for this is that *Car model* is a rating factor with far too many levels for accurate estimation in a GLM analysis. For more information on this and other practical applications, see Ohlsson & Johansson (2004).

The object of this paper is to prove an extension of Jewells theorem (in its generalized form given by Kaas et al.) to "the Bühlmann-Straub model with a priori differences". The placement of our result in credibility theory is shown schematically in Figure 1.

FIGURE 1:

SCHEMATIC RELATION OF OUR WORK TO OTHER PAPERS ON CREDIBILITY.

	Unweighted case	**Weighted case**	**Weighted with a priori differences**
Linear credibility	Bühlmann	Bühlmann-Straub	Bühlmann & Gisler
Exact credibility	Jewell	Kaas, Dannenburg & Goovaerts	**This paper**

It should be noted that our result suggests a different weighting than the one given in Bühlmann & Gisler, except in the Poisson case where they coincide. In Ohlsson (2004), we derive the linear credibility estimators that have precisely the same weights as our exact credibility estimators here.

As will be indicated below, our work here also has some relation to the work by Nelder and Verrall (1997) on the combination of credibility and GLMs using the theory of *hierarchical generalized linear models*, HGLM.

2. Tweedie models and random effects

In non-life insurance pricing with GLMs, one studies the effect of rating factors on some key ratio Y_i, typically the risk premium, claims frequency or average claim amount. By far, the most commonly used models are GLMs with a variance function of the form $v(\mu) = \mu^p$ for some p. We start out by repeating some basic facts about these so called Tweedie models, before deriving our extension of Jewell's theorem.

2.1. Tweedie models

The rating factors divide the portfolio into *tariff cells*, and the key ratio Y_i is computed over the policies in cell i. In GLMs, Y_i is assumed to have a frequency function of the form

$$f_{Y_i}(y_i;\theta_i,\phi) = \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{\phi / w_i} + c(y_i,\phi,w_i)\right\}, \tag{2.1}$$

where ϕ is the *dispersion parameter* and w_i is the known exposure weight, i.e. the denominator of the key ratio (number of policy years for claim frequencies, number of claims for average claim amount, etc.). The function $b(\theta)$ is twice differentiable with a unique inverse for the first derivative $b'(\theta)$. With $\phi = 1$ and all $w_i = 1$, (2.1) is the exponential family with canonical parameter considered by Jewell (1974). Following Jørgensen (1997), the models defined by (2.1) are called *exponential dispersion models* (EDMs).

From standard GLM theory we know that $\mu_i \doteq E(Y_i) = b'(\theta_i)$. If we further denote the inverse of b' by $h(\mu)$, we can express the variance as

$$\mathrm{Var}(Y_i) = \phi b''(\theta_i)/w_i = \phi v(\mu_i)/w_i, \tag{2.2}$$

where $v(\mu) = b''(h(\mu))$ is known as the *variance function*. In this paper, we will only consider the subclass of EDMs where, for some p,

$$v(\mu) = \mu^p. \tag{2.3}$$

In the terminology of Jørgensen, these are called the *Tweedie models*. In the rest of this section we recapitulate some of their theory – for proofs, see Jørgensen (1997). The Tweedie models are defined only for p outside the interval $0 < p < 1$. Renshaw (1994) concludes that models with $p \leq 0$ "are of no practical concequence" in non-life insurance rating – one reason being that they have support on the whole real line, while our key ratios are non-negative. We thus restrict ourselves to the class with $p \geq 1$. Our calculations below carry through for $p = 0$ (Gaussian distribution), but that model is not appropriate for the non-negative key ratios with multiplicative rating factors that we consider here and is hence omitted.

The following Tweedie models are of special interest in non-life insurance pricing:

- $p = 1$: (Weighted) Poisson distribution.
- $1 < p < 2$: Compound Poisson distribution with gamma distributed summands.
- $p = 2$: Gamma distribution.

The case $1 < p < 2$ is applicable to risk premiums Y_i with a Poisson distributed number of claims and gamma distributed claim sizes. Here $p = (2 + \gamma)/(1 + \gamma)$, where γ is the shape parameter of the gamma distribution, see Jørgensen & Paes de Souza (1994), and hence $\gamma^{-1/2}$ is the coefficient of variation of that distribution.

From the relation between $v(\mu)$ and $b''(\theta)$, Jørgensen derives the functional form of the $b(\theta)$ corresponding to a variance function as in (2.3). The result is (for $p \geq 1$)

$$b(\theta) = \begin{cases} e^{\theta} & \text{for } p = 1; \\ -\log(-\theta) & \text{for } p = 2; \\ -\frac{1}{p-2}\left[-(p-1)\theta\right]^{(p-2)/(p-1)} & \text{for } 1 < p < 2 \text{ and } p > 2. \end{cases} \tag{2.4}$$

The canonical (maximal) parameter space M is

$$M = \begin{cases} -\infty < \theta < \infty & \text{for } p = 1; \\ -\infty < \theta < 0 & \text{for } 1 < p \leq 2; \\ -\infty < \theta \leq 0 & \text{for } p > 2. \end{cases} \tag{2.5}$$

We will also need expressions for the derivative $b'(\theta)$,

$$b'(\theta) = \begin{cases} e^{\theta} & \text{for } p = 1; \\ \left[-(p-1)\theta\right]^{-1/(p-1)} & \text{for } p > 1; \end{cases} \tag{2.6}$$

and its inverse $h(\mu)$,

$$h(\mu) = \begin{cases} \log(\mu) & p = 1; \\ -\frac{1}{p-1}\mu^{-(p-1)} & p > 1. \end{cases} \tag{2.7}$$

In GLM theory, $h(\mu)$ is called the canonical link function. Note, however, that we do not assume the use of canonical link, but instead use a log-link (multiplicative model) throughout. Multiplicative models are standard in insurance practice and usually a very reasonable choice.

2.2. Random effects in Tweedie models

Suppose we have a number of ordinary rating factors, dividing the portfolio into I tariff cells – by an ordinary rating factor we mean one that is not treated as a random effect. As before, μ_i is the expected value of our key ratio, and since we use a multiplicative model μ_i is in effect the product of a number of price relativities for the rating factors.

To this model we add a random effect with K levels, where level k is given by the random variable U_k. Let w_{ik} denote the exposure weight in the ith tariff cell with respect to the ordinary rating factors and for the kth level of the random effect. Let Y_{ik} denote the corresponding observed key ratio.

With u_k denoting the outcome of U_k, the multiplicative model is now extended to

$$E(Y_{ik} \mid U_k = u_k) = \mu_i u_k. \tag{2.8}$$

Since the systematic effects are captured by μ_i, we can let the U_k's be purely random, so that we have $E(U_k) = 1$, and hence $E(Y_{ik}) = \mu_i$.

Note. To explicitly write down the entire multiplicative model, let $\gamma_{j(i)}^{(r)}$ denote the price relativity for rating factor number r at the level $j(i)$, the level that is attained in the i:th tariff cell. Then, if there are R ordinary rating factors

$$\mu_i = \gamma_0 \gamma_{j(i)}^{(1)} \gamma_{j(i)}^{(2)} \cdots \gamma_{j(i)}^{(R)},$$

where γ_0 is the base value. (For some base cell, say $i = 1$, all $\gamma_{j(1)}^{(r)}$ are set to 1 for unambiguous parametrisation.) The model with an added random effect is now

$$E(Y_{ik} \mid U_k = u_k) = \gamma_0 \gamma_{j(i)}^{(1)} \gamma_{j(i)}^{(2)} \cdots \gamma_{j(i)}^{(R)} u_k.$$

The $\gamma_{j(i)}^{(r)}$'s may be estimated by the standard GLM procedure. For simplicity, we will not write out the entire multiplicative model below. □

Conditionally on $U_k = u_k$ we assume that Y_{ik} follows a Tweedie model with expectation $\mu_i u_k$. Now, the frequency function in (2.1) is defined in terms of the canonic parameter $\theta = h(\mu)$, rather than the expectation μ, and in our case this parameter becomes $\tilde{\theta}_{ik} = h(\mu_i u_k)$. We make the corresponding transformation of the random effect and introduce the random variable $\Theta_k = h(U_k)$, which corresponds to the *risk parameter* in Jewell (1974) and other sources on credibility theory, taking values $\theta_k = h(u_k)$.

Note that by (2.7)

$$\tilde{\theta}_{ik} = h(\mu_i u_k) = \begin{cases} \log(\mu_i) + h(u_k) & p = 1; \\ \mu_i^{-(p-1)} h(u_k) & p > 1; \end{cases} \tag{2.9}$$

and then by (2.4),

$$
\begin{aligned}
b(\tilde{\theta}_{ik}) &= b\big(h(\mu_i u_k)\big) \\
&= \begin{cases} \log(\mu_i) + b(\theta_k) & p = 2; \\ \mu_i^{2-p} b(\theta_k) & 1 \le p < 2 \text{ and } p > 2. \end{cases}
\end{aligned}
\tag{2.10}
$$

Now for all the p we consider, we can write

$$
\begin{aligned}
f_{Y_{ik}|\Theta_k}(y_{ik}|\theta_k) &= \exp\left\{ \frac{y_{ik}\tilde{\theta}_{ik} - b\left(\tilde{\theta}_{ik}\right)}{\phi / w_{ik}} + c_1 \right\} \\
&= \exp\left\{ \frac{w_{ik}}{\phi} \left[\frac{y_{ik}}{\mu_i^{p-1}} \theta_k - \frac{1}{\mu_i^{p-2}} b(\theta_k) \right] + c_2 \right\},
\end{aligned}
\tag{2.11}
$$

where c_1 and c_2 are constants that do not depend on θ_k, and into c_2 we have incorporated the terms $\log(\mu_i)$ appearing in (2.9) and (2.10).

Conditional on $\Theta_k = \theta_k$, or equivalently on $U_k = u_k$, the Y_{ik}'s are assumed independent. We can then perform a standard GLM analysis of the ordinary rating factors, using $\log(u_k)$ as an *offset* variable. Now, the U_k's are of course non-observable and must be estimated. We will follow Jewell (1974) and assume that the density function of $\Theta = h(U)$ is the *natural conjugate prior* (or associate conjugate) to the family in (2.1), which is given by

$$
f_\Theta(\theta) = \frac{1}{c(\delta,\alpha)} \exp\left\{ \frac{\theta\delta - b(\theta)}{1/\alpha} \right\},
\tag{2.12}
$$

for $\theta \in M$ (the canonical parameter space of (2.1)). Here δ and α are so called *hyperparameters* and $c(\delta,\alpha)$ is a normalizing constant. For all $p \ge 1$, this is a proper distribution if $\alpha > 0$ and $\delta > 0$, which we assume in the following.

Lemma 2.1. *Let* $U = b'(\Theta)$, *where* Θ *follows the distribution in (2.12) and let* $\inf M$ *and* $\sup M$ *denote the lower and upper bound of the interval M in (2.5).*

(a) Suppose that $f_\Theta(\inf M) = f_\Theta(\sup M) = 0$. *Then*

$$
\delta = E(U)
$$

(b) In addition to the assumption in (a), suppose that $f'_\Theta(\inf M) = f'_\Theta(\sup M) = 0$. *Then*

$$
\alpha = \frac{E\left[U^p\right]}{\mathrm{Var}(U)}.
\tag{2.13}
$$

Proof. We have

$$f'_\Theta(\theta) = \alpha(\delta - b'(\theta))f_\Theta(\theta),$$
$$f''_\Theta(\theta) = \alpha^2(\delta - b'(\theta))^2 f_\Theta(\theta) - \alpha b''(\theta) f_\Theta(\theta).$$

Upon integrating these equations we get, under the assumptions of the limiting behavior of $f_\Theta(\theta)$ and $f'_\Theta(\theta)$, respectively,

$$\begin{aligned} 0 &= \alpha\int_M (\delta - b'(\theta))\, f_\Theta(\theta)d\theta = \alpha(\delta - E[b'(\Theta)]) \\ 0 &= \alpha^2\int_M (b'(\theta) - \delta)^2 f_\Theta(\theta)d\theta - \alpha\int_M b''(\theta)\, f_\Theta(\theta)d\theta \\ &= \alpha^2 \mathrm{Var}(b'(\Theta)) - \alpha E[b''(\Theta)]. \end{aligned}$$

Now the fact that $u = b'(\theta)$ and that $b''(\theta) = b''(h(u)) = v(u) = u^p$ completes the proof. □

We next investigate to what extent the assumptions of the lemma are satisfied for Tweedie models with $p \geq 1$.

Lemma 2.2. *(a) The assumptions of Lemma 2.1(a) are satisfied for* $1 \leq p \leq 2$. *They are not valid for* $p > 2$.

(b) The assumptions of Lemma 2.1(b) are satisfied for $1 \leq p < 2$. *For* $p = 2$ *they are satisfied if* $\alpha > 1$, *but not for* $\alpha \leq 1$. *For* $p > 2$ *they are invalid.*

The proof of this lemma is straightforward. For $p = 2$ one may add that when $\alpha \leq 1$ the variance $\mathrm{Var}(U)$ does not exist.

Lemma 2.1 is fundamental to the proof of our main result and since it's conclusion is – by Lemma 2.2 – valid only when $1 \leq p \leq 2$, we restrict the discussion to that case in what follows. Under this restriction, the two lemmas show that $\delta = E(U) = 1$ so that, in effect, we have just one parameter in the conjugate distribution, $\alpha > 0$. Fortunately, $1 \leq p \leq 2$ contains the most commonly used distributions in insurance applications of GLMs, namely the Poisson, gamma and compound Poisson-gamma distributions. (Note that the assumptions of Lemma 2.1(a) also appear in the original theorem by Jewell (1974) and so his results are not valid for Tweedie models with $p > 2$).

3. Main results

In his classical result, Jewell (1974) assumed an exponential family of distributions for Y, conditionally on the so called *risk parameter* (our Θ_k). Kaas et al (1997) generalized Jewell's theorem to the exponential dispersion models used in GLMs, including weights w. Before presenting our extension of these results, we make some basic assumptions that are more or less standard in credibility theory.

Assumption 1. *(a) Θ_k; $k = 1, 2, \ldots, K$ are independant and identically distributed random variables.*

(b) For $k = 1, 2, \ldots, K$, the pairs (Y_{ik}, Θ_k) are independent.

(c) Conditioned on Θ_k the random variables $Y_{1k}, Y_{2k}, \ldots, Y_{I_k,k}$ are independent.

By (2.8) we have $E(Y_{ik} \mid U_k) = \mu_i U_k$, where μ_i is the mean given by the ordinary rating factors, which can be estimated by standard GLM methods once we have the u_k. Hence, in our case the search for credibility estimators amounts to finding an estimator of U_k, for every k. In exact credibility, this means that we look for the function g of our data vector $\boldsymbol{Y}$ that minimizes

$$E\left[(U_k - g(\boldsymbol{Y}))^2\right]. \tag{3.1}$$

It is well known that the solution to this minimization problem is $g(\boldsymbol{Y}) = E[U_k \mid \boldsymbol{Y}]$. By assumption 1(b) we can restrict the conditioning to $\boldsymbol{Y}_k = \{Y_{ik}; i = 1, 2, \ldots, I_k\}$ and our optimal estimator is $g(\boldsymbol{Y}) = E[U_k \mid \boldsymbol{Y}_k] = E[b'(\Theta_k) \mid \boldsymbol{Y}_k]$. An expression for this *posterior mean* is given in the following extension of Jewell's theorem, which is our main result.

Theorem 3.1. *Let Assumption 1 be satisfied. Suppose that conditionally on U_k we have a Tweedie model for Y_{ik} with $1 \leq p \leq 2$ and that $\Theta_k = h(U_k)$ follows the natural conjugate distribution given by (2.12), where $\alpha > 0$ and $\delta > 0$. Then the optimal estimator $E[U_k \mid \boldsymbol{Y}_k]$ of the random effect U_k can be written as*

$$\hat{u}_k = \frac{\sum_i w_{ik}\, y_{ik} / \mu_i^{p-1} + \phi\alpha}{\sum_i w_{ik}\, \mu_i^{2-p} + \phi\alpha}. \tag{3.2}$$

Proof. To compute the posterior expectation we need the posterior distribution of Θ_k, which we get from Bayes theorem, plus (2.11) and (2.12).

$$\begin{aligned}
f_{\Theta_k|\boldsymbol{Y}_k}(\theta|\boldsymbol{y}_k) &\propto f_{\Theta_k}(\theta) f_{\boldsymbol{Y}_k|\Theta_k}(\boldsymbol{y}_k|\theta) = f_{\Theta_k}(\theta) \prod_i f_{Y_{ik}|\Theta_k}(y_{ik}|\theta) \\
&\propto \exp\{\alpha(\theta - b(\theta))\} \prod_i \exp\left\{\frac{w_{ik}}{\phi}\left[\frac{y_{ik}}{\mu_i^{p-1}}\theta - \mu_i^{2-p} b(\theta)\right]\right\} \qquad (3.3) \\
&= \exp\left\{\theta\left(\alpha + \frac{1}{\phi}\sum_i w_{ik}\frac{y_{ik}}{\mu_i^{p-1}}\right) - b(\theta)\left(\alpha + \frac{1}{\phi}\sum_i w_{ik}\,\mu_i^{2-p}\right)\right\}.
\end{aligned}$$

Since we are using a conjugate prior, it is no surprise that the posterior distribution is a member of the same family, with new "updated" parameters

$$\alpha' = \alpha + \frac{1}{\phi}\sum_i w_{ik}\,\mu_i^{2-p}, \quad \delta' = \hat{u}_k, \tag{3.4}$$

where $\hat{u}_k$ is given by (3.2). Finally, from Lemma 2.1(a) and 2.2(a) the expectation of U_k in the posterior distribution is just δ' and the proof is complete. □

We shall rewrite (3.2) in the form of a credibility estimator. First note that since $Y_{ik} \mid U_k$ follows a Tweedie model, (2.2) and (2.3) give

$$\mathrm{Var}(Y_{ik} \mid U_k = u_k) = \phi \frac{v(\mu_i u_k)}{w_{ik}} = \phi \frac{\mu_i^p u_k^p}{w_{ik}}.$$

Introducing the new weights

$$\tilde{w}_{ik} = w_{ik}\,\mu_i^{2-p}, \tag{3.5}$$

we may write, with $\sigma^2 = \phi E[U_k^p]$,

$$\mathrm{Var}\left(\frac{Y_{ik}}{\mu_i}\middle| U_k\right) = \phi \frac{U_k^p}{\tilde{w}_{ik}}; \quad E\left[\mathrm{Var}\left(\frac{Y_{ik}}{\mu_i}\middle| U_k\right)\right] = \frac{\sigma^2}{\tilde{w}_{ik}}.$$

If we write $\tau^2 \doteq \mathrm{Var}(U_k)$, we have by Lemma 2.1 and 2.2,

$$\phi\alpha = \frac{\phi E\left[U_k^p\right]}{\mathrm{Var}(U_k)} = \frac{\sigma^2}{\tau^2}.$$

This result is valid under the assumptions of Theorem 3.1, except for the case $p = 2$ with $\alpha \leq 1$, where $\mathrm{Var}(U_k)$ is not finite. Next, introduce the weighted average

$$\bar{u}_k = \frac{\sum_i \tilde{w}_{ik}\, y_{ik}/\mu_i}{\sum_i \tilde{w}_{ik}}. \tag{3.6}$$

We call $\bar{u}_k$ the *experience factor*, indicating how one might adjust the expected values μ_i to take into account the observed multiplicative deviances y_{ik}/μ_i. The unconventional notation $\bar{u}_k$ is motivated by the fact that this is an average that estimates U_k, with $E[\bar{U}_k \mid U_k] = U_k$.

Corollary 3.1. *Under the assumptions of Theorem 3.1, except for the case $p = 2$ with $\alpha \leq 1$, the exact credibility estimator of U_k is*

$$\hat{u}_k = z_k \bar{u}_k + (1 - z_k) \cdot 1, \tag{3.7}$$

where the credibility factor z_k is defined by

$$z_k \doteq \frac{\sum_i \tilde{w}_{ik}}{\sum_i \tilde{w}_{ik} + \sigma^2 / \tau^2}. \tag{3.8}$$

Thus the estimator $\hat{u}_k$ is a credibility weighted adjustment factor to the rating by the ordinary factors, μ_i. The rating for insurances in cell i having level k on the random effect is then obtained from multiplying by μ_i,

$$\mu_i \hat{u}_k = z_k (\mu_i \bar{u}_k) + (1 - z_k) \cdot \mu_i.$$

The unknown μ_i's can be estimated by GLMs, treating $\hat{u}_k$ as a known *offset* variable. Since μ_i also appears in the definition of $\tilde{w}_{ik}$, this calls for an iterative procedure over the estimation of μ_i by a GLM, treating $\hat{u}_k$ as an offset, and the estimation of u_k, treating $\hat{\mu}_i$ as known. See Ohlsson (2004) for a detailed algorithm and an application.

3.1. Comparison with linear credibility

Here we specialise to the case where all exposure units at random effect level k have the same expectation, denoted here by μ_k. Then the risk premium (or whichever key ratio we are considering) is estimated by

$$\mu_k \hat{u}_k = z_k \bar{y}_k + (1 - z_k) \cdot \mu_k, \tag{3.9}$$

where

$$\bar{y}_k = \frac{\sum_i w_{ik} y_{ik}}{\sum_i w_{ik}},$$

since in this case

$$\mu_k \bar{u}_k = \frac{\sum_i \tilde{w}_{ik} y_{ik}}{\sum_i \tilde{w}_{ik}} = \frac{\sum_i w_{ik} y_{ik}}{\sum_i w_{ik}} = \bar{y}_k.$$

The credibility factor is now

$$z_k \doteq \frac{\mu_k^{2-p} \sum_i w_{ik}}{\mu_k^{2-p} \sum_i w_{ik} + \sigma^2 / \tau^2}.$$

In the Poisson case $p = 1$ we then have exactly the same credibility estimator as that given in Corollary 4.15 of Bühlmann & Gisler (2005). Indeed, they motivate their variance assumption with the Poisson case. The conclusion is that our result is the exact credibility counterpart of the Bühlmann & Gisler (2005) estimator in this case (as claimed in Figure 1).

In the gamma ($p = 2$) and compound Poisson ($1 < p < 2$) cases, our factor μ_k^{2-p} differs from the μ_k that is recommend Bühlmann & Gisler. Since they motivate the choice μ_k only in the Poisson case, our estimator is in this respect a "fine tuning" of theirs.

On the other hand, the strength of Bühlmann & Gisler (2005) is that the approach is distribution-free, i.e. only requires assumptions on expectations and variances, as is generally the case with linear credibility as opposed to exact credibility. In this distribution-free setting, our estimator can be derived directly from the Bühlmann-Straub estimator, see Ohlsson (2004), where the variance assumption is

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

for some constants σ_i^2 and some function $\sigma^2(U_k)$. The result of the present paper corresponds to the case $\sigma_i^2 = \mu_i^{2-p}$, while the assumption (4.83) of Bühlmann & Gisler is obtained by letting $\sigma_i^2 = \mu_i$. The advantage of exact credibility is that we get the proper variance directly from the GLM model used, while in the non-parametrical case we have to guess the value of σ_i^2.

We also work in a slightly different setting than Bühlmann & Gisler, by allowing the possibility that level k occurs with different μ_i for different contracts (a certain Car model k may be owned by drivers of different age driving different mileage, etc.). However, the Bühlmann & Gisler estimator could be extended to this case without difficulty.

3.2. Special cases

Finally, it is of interest to specialize our estimators to the important special cases $p = 1$ (Poisson) and $p = 2$ (gamma), by looking at the corresponding experience factors.

$$p = 1 \Rightarrow \bar{u}_k = \frac{\sum_i w_{ik} y_{ik}}{\sum_i w_{ik} \mu_i}. \tag{3.10}$$

$$p = 2 \Rightarrow \bar{u}_k = \frac{\sum_i w_{ik} y_{ik} / \mu_i}{\sum_i w_{ik}}. \tag{3.11}$$

It can be noted that equation (3.10) corresponds to an estimating equation in the so called *method of marginal totals* and that (3.11) corresponds to the so called *direct method* (see, e.g., Kaas et al., 2001, pp. 179-181).

In these cases – and in general – equation (3.6) is easily verified to be the estimating equation for u_k if considered a *fixed* effect in a standard GLM analysis (remembering that we are using a log-link). This is appealing: in a case with very high credibility our estimating equations are the same as those resulting from treating the random effect as just another covariate in our GLM. Under the iterative procedure mentioned after Corollary 3.1, high credibility would then give the same estimates $\hat{u}_k$ as a GLM would have done, while with less credibility these estimates are "shrinked" towards 1.

3.3. Comparison with HGLM

Besides the credibility interpretation, our models could also be described as GLMs with a random effect. Lee and Nelder (1996) and Nelder & Verall (1997) suggest a general likelihood-based approach to this type of models, using the concept of h-likelihood, under the name HGLM (Hierarchical Generalised Linear Models). For the Tweedie case with $p = 1$, our estimator (3.2) is a w-weighted version of the estimator on p. 623 in Lee & Nelder (1996). In the case $p = 2$ with weights $w_{ik} \equiv 1$, our (3.2) becomes, with n_k denoting the number of cells i for level k,

$$\hat{u}_k = \frac{\sum_i y_{ik} / \mu_i + \phi\alpha}{n_k + \phi\alpha},$$

while Lee & Nelder's (2.12) is, in our notation,

$$\hat{u}_k = \frac{\sum_i y_{ik} / \mu_i + \phi\alpha}{n_k + \phi\alpha + \phi},$$

the difference being the term $+\phi$ that appears in the denominator by which the latter is not a credibility estimator proper. The case $1 < p < 2$ is not explicitly treated in Lee & Nelder (1996) or Nelder & Verall (1997). Notwithstanding the rather similar results, note that our approach to these models is quite different from HGLM.

3.4. Estimation of variance parameters

In applications, one has to estimate $\phi\alpha$, or alternatively σ^2 and τ^2, whose ratio is $\phi\alpha$. There are several possibilities, the simplest perhaps being to use a counterpart of the unbiased estimators in the classical Bühlmann-Straub model. These estimators are given in Ohlsson (2004, Section 2.1).

Acknowledgement

The authors wish to thank an anonymous referee for many valuable comments leading to large improvements of the paper.

References

Bühlmann, H. (1967) Experience rating and credibility. *ASTIN Bulletin*, **4**, 199-207.

Bühlmann, H. and Gisler, A. (2005) *A Course in Credibility Theory and its Applications.* Springer-Verlag, Berlin.

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.

Jørgensen, B. (1997) The Theory of Dispersion Models. Chapman and Hall, London.

Jørgensen, B. and Paes de Souza, M.C. (1994) Fitting Tweedie's compound Poisson model to insurance claim data. *Scand. Actuarial Journal*, 69-93.

Kaas, R., Dannenburg, D. and Goovaerts, M.J. (1997) Exact credibility for weighted observations. *ASTIN Bulletin*, **27**, 287-295.

Kaas, R., Goovaerts, M.J., Dhaene, J. and Denuit, M. (2001) *Modern Actuarial Risk Theory.* Kluwer Academic Publishers, Dordrecht.

Lee, Y. and Nelder, J.A. (1996) Hierarchical Generalized linear models. *J.R. Statist. Soc. B*, **58(4)**, 619-678.

Nelder, J.A. and Verrall, R.J. (1997) Credibility Theory and Generalized linear models. *ASTIN Bulletin*, **27(1)**, 71-82.

Ohlsson, E. (2004) *Credibility Rating in a Multiplicative Tariff*. Paper presented at the 35th ASTIN Colloquium, http://www.astin2004.no/.

Ohlsson, E. and Johansson, B. (2004) *Combining Credibility and GLM for Rating of Multi-Level Factors*. CAS 2004 Discussion Paper Program, http://www.casact.org/pubs/dpp/dpp04/.

Renshaw, A.E. (1994) Modeling the claims process in the precense of covariates. *ASTIN Bulletin*, **24(2)**, 265-285.

Esbjörn Ohlsson
Mathematical Statistics
Department of Mathematics
Stockholm University
SE-106 91 Stockholm, Sweden
E-mail: esbjorn.ohlsson@math.su.se

Björn Johansson
Länsförsäkringar Alliance
SE-106 50 Stockholm, Sweden
E-mail: bjorn.johansson@lansforsakringar.se