Rapp

Integrating ordinary GLM with credibility in a Compound Poisson model

Part 2 mean claim

Lecture in Indonesia november 2014
Stig Rosenlund

Summary of the introduction of Part 1 on claim frequency

The setup is that we have a number of arguments, each with a finite number of classes, the same as in my lecture on tariff analysis with Rapp, But one of the arguments has so many classes, that the uncertainty for risk premium is very large for many of the classes. And there is no natural ordering of the classes, such that risk premium can be assumed to be monotone, or have at most one minimum or maximum, in the ordering of the classes. So credibility must be used.

The focus here will be mainly on the mean claim variance components.

The argument with many classes will be called the credibility argument. The non-credibility arguments are of two kinds; those that depend functionally on the credibility argument and those that do not. The former are called Auxiliaries. For example, if the credibility argument is geographical parishes, median income and population density are Auxiliaries. Non-auxiliaries are e.g. policy holder age and object age, since a parish has policies in several classes of these.

Like Ohlsson & Johansson I treat non-auxiliaries by letting exposure be normalized duration, defined for an individual object as a number proportional to duration multiplied with estimated multiplicative risk premium for non-auxiliaries. After having stated that, I do not further treat non-auxiliaries, except in the section on iterations.

My claim is that my analysis gives better results than previous authors have given. For one Auxiliary my results are exact, while for several Auxiliaries they are approximations.

In order to describe the essentials of my work more clearly, I will here assume that there are no non-auxiliary ordinary arguments and that there is one single Auxiliary. You can think of geographical parishes as the credibility argument and population density as the Auxiliary. All mathematical intricacies are present in this setup. I treat claim freqency and mean claim separately in this lecture, concentrating on mean claim. I will also specialize some assumptions and omit some estimators in the submitted paper.

Model

The Compound Poisson distribution is assumed for total claim cost for group j = 1, 2, ... J, conditional on stochastic variables

Θ_{Fj} for claim frequency

Θ_{Mj} for mean claim

Namely, conditional on the Θs, the claim numbers are Poisson and the claim severities are IID and independent of the claim numbers.

Furthermore, all Θs are assumed to be independent with expectation 1. Their standard deviations are τ_F for claim frequency and τ_M for mean claim.

Notation for observables

We define the following observables, where J, N_j, J_1 and N_{j1} are considered deterministic, since in mean claim analysis we condition with respect to claim numbers.

$$J = \text{number of groups,} \tag{1}$$

$$N_j = \text{number of claims group } j, \tag{2}$$

$$J_1 = \sum_{j=1}^{J} \mathbf{1}_{\{N_j>0\}} = \text{number of groups with claims,} \tag{3}$$

$$N_0 = \sum_{j=1}^{J} N_j = \text{total number of claims,} \tag{4}$$

$$k_{j1} = \text{the class of group } j \text{ in its Auxiliary, } j \in \{1,\dots,J\}, \tag{5}$$

$$N_{j1} = \sum_{i:k_{i1}=k_{j1}} N_i = \text{total claim number over group } j\text{'s value in its Auxiliary,} \tag{6}$$

$$Z_{ji} = \text{individual claim amounts in group } j \in \{1,\dots,J\},\ \ i \in \{1,\dots,N_j\}, \tag{7}$$

$$X_j = \sum_{i=1}^{N_j} Z_{ji} \qquad X_{j1} = \sum_{i:k_{i1}=k_{j1}} X_i \tag{8}$$

$$Q_j = \sum_{i=1}^{N_j} Z_{ji}^2 \tag{9}$$

$$Y_{\mathsf{M}j} = X_j/N_j = \sum_{i=1}^{N_j} Z_{ji}/N_j = \text{observed mean claim group } j, \tag{10}$$

$$\hat{\mu}_{\mathsf{M}j} = X_{j1}/N_{j1} = \text{estimated collective mean claim group } j \text{ for its Auxiliary.} \tag{11}$$

Assumptions

We formulate four assumptions. In **A4M** we specialize the exponent p_{M} of the paper to 2, since this is the only value used in practice.

A1M. Conditional on stochastic variables $\Theta_{\mathsf{M}j}$ $(j = 1, \ldots, J)$, with expectation $\mathrm{E}[\Theta_{\mathsf{M}j}] = 1$ and variance $\mathrm{Var}[\Theta_{\mathsf{M}j}] = \tau_{\mathsf{M}}^2$, for any specific j the Z_{ji} are IID with expectation $\mu_{\mathsf{M}j}\Theta_{\mathsf{M}j}$, where $\mu_{\mathsf{M}j}$ is determined by the Auxiliary of j.

A2M. $(\Theta_{\mathsf{M}1}, X_1, Q_1), \ \ldots \ , (\Theta_{\mathsf{M}J}, X_J, Q_J)$ are independent.

A3M. $\mathrm{E}[\hat{\mu}_{\mathsf{M}j}] = \mu_{\mathsf{M}j}$, always true with only one Auxiliary.

A4M. $\mathrm{Var}[Y_{\mathsf{M}j} \mid \Theta_{\mathsf{M}j}] = \phi_{\mathsf{M}}(\mu_{\mathsf{M}j}\Theta_{\mathsf{M}j})^2/N_j$ for $\phi_{\mathsf{M}} > 0$.

Objective. To predict $\mu_{\mathsf{M}j}\Theta_{\mathsf{M}j}$ as well as possible.

We use the words predict and predictor, by the convention to reserve the word estimator for guesses on non-stochastic parameters (the Bayesian formulation set aside). Our predictors are guesses on stochastic variable outcomes that occurred in the past and can never be observed.

Define

$$\mu_{\mathsf{M}j} = \mathrm{E}[Y_{\mathsf{M}j}] \quad \text{mean claim group } j, \tag{12}$$

$$\Lambda_{\mathsf{M}j} = \mu_{\mathsf{M}j}\Theta_{\mathsf{M}j} \quad \text{mean claim group } j \text{ conditional on } \Theta_{\mathsf{M}j}, \tag{13}$$

$$\sigma^2_{\mathsf{M}} = \phi_{\mathsf{M}}\mathrm{E}[\Theta^2_{\mathsf{M}j}] \text{ assumed independent of } j, \tag{14}$$

$$\sigma^2_{\mathsf{M}j} = \mathrm{E}[\mathrm{Var}[Y_{\mathsf{M}j} \mid \Theta_{\mathsf{M}j}]], \tag{15}$$

$$\tau^2_{\mathsf{M}} = \mathrm{Var}[\Theta_{\mathsf{M}j}] \text{ assumed independent of } j, \tag{16}$$

$$\nu^2_{\mathsf{M}j} = \frac{1}{{N_{j1}}^2} \sum_{i:k_{i1}=k_{j1}} N_i^2(\sigma^2_{\mathsf{M}i} + \mu^2_{\mathsf{M}j}\tau^2_{\mathsf{M}}) \text{ equal to } \mathrm{Var}[\hat{\mu}_{\mathsf{M}j}]. \tag{17}$$

The predictors and estimators are

$$\Lambda^*_{\mathsf{M}j} = z_{\mathsf{M}j} Y_{\mathsf{M}j} + (1 - z_{\mathsf{M}j}) \mu_{\mathsf{M}j} \quad \text{non-observable predictor of } \Lambda_{\mathsf{M}j}, \text{ see } (26), \tag{18}$$

$$\widehat{\Lambda}_{\mathsf{M}j} = \widehat{z}_{\mathsf{M}j} Y_{\mathsf{M}j} + (1 - \widehat{z}_{\mathsf{M}j}) \hat{\mu}_{\mathsf{M}j} \quad \text{estimated predictor of } \Lambda_{\mathsf{M}j}, \text{ see } (34), \tag{19}$$

$$\widehat{\sigma}^2_{\mathsf{M}} = \frac{\sum_{j=1}^{J} \hat{\mu}_{\mathsf{M}j}^{-2} \sum_{i=1}^{N_j} (Z_{ji} - Y_{\mathsf{M}j})^2}{\sum_{j=1}^{J} \mathbf{1}_{\{N_j > 0\}} (N_j - 1)} = \frac{\sum_{j=1}^{J} \hat{\mu}_{\mathsf{M}j}^{-2} \mathbf{1}_{\{N_j > 0\}} (Q_j - X_j^2 / N_j)}{\sum_{j=1}^{J} \mathbf{1}_{\{N_j > 0\}} (N_j - 1)}, \tag{20}$$

$$\widehat{\sigma}^2_{\mathsf{M}j} = \widehat{\sigma}^2_{\mathsf{M}} \hat{\mu}^2_{\mathsf{M}j} / N_j, \tag{21}$$

$$\widetilde{Y} = \frac{1}{N_0} \sum_{j=1}^{J} X_j / \hat{\mu}_{\mathsf{M}j} = \frac{1}{N_0} \sum_{j=1}^{J} N_j Y_{\mathsf{M}j} / \hat{\mu}_{\mathsf{M}j}, \tag{22}$$

$$\overset{*}{\tau}{}^{2}_{\mathsf{M}} = \max\left(0, \frac{\sum_{j=1}^{J} N_j (Y_{\mathsf{M}j}/\hat{\mu}_{\mathsf{M}j} - \widetilde{Y})^2 - (J_1 - 1)\hat{\sigma}^2_{\mathsf{M}}}{N_0 - \sum_{j=1}^{J} N_j^2 / N_0}\right), \text{ estimator of } \tau^2_{\mathsf{M}}, \tag{23}$$

an adaptation of (4.27) in Ohlsson and Johansson (2010),

$$\hat{\tau}^2_{\mathsf{M}} = \text{pseudo-estimator of } \tau^2_{\mathsf{M}}, \text{ see (33)}, \tag{24}$$

$$\hat{\nu}^2_{\mathsf{M}j} = \frac{1}{{N_{j1}}^2} \sum_{i:k_{i1}=k_{j1}} N_i^2 (\widehat{\sigma}^2_{\mathsf{M}i} + \hat{\mu}^2_{\mathsf{M}j} \hat{\tau}^2_{\mathsf{M}}) \text{ estimator of } \mathrm{Var}[\hat{\mu}_{\mathsf{M}j}]. \tag{25}$$

The estimators of σ^2_{M} and $\sigma^2_{\mathsf{M}j}$ are as in Ohlsson and Johansson (2010).

Best linear predictor of the credibility factor

For pricing we wish to use a predictor of $\mu_{\mathsf{M}j}\Theta_{\mathsf{M}j}$ as a mean claim factor for the value j of the credibility argument. It consists of the collective mean claim $\mu_{\mathsf{M}j}$, the same for all groups with the same value of the Auxiliary, and the specific factor $\Theta_{\mathsf{M}j}$.

Theorem 1. *For the BLP in L^2-norm of $\mu_{\mathsf{M}j}\Theta_{\mathsf{M}j}$ it holds*

$$z_{\mathsf{M}j} = \frac{\mathrm{Var}[\hat{\mu}_{\mathsf{M}j}] + \mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2 - \mathrm{Cov}(\hat{\mu}_{\mathsf{M}j}, Y_{\mathsf{M}j}) - \mu_{\mathsf{M}j}\mathrm{Cov}(\hat{\mu}_{\mathsf{M}j}, \Theta_{\mathsf{M}j})}{\sigma_{\mathsf{M}j}^2 + \mathrm{Var}[\hat{\mu}_{\mathsf{M}j}] + \mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2 - 2\mathrm{Cov}(\hat{\mu}_{\mathsf{M}j}, Y_{\mathsf{M}j})}. \tag{26}$$

Here τ_{M}^2 is a variance component for variation between the groups. The variance component for variation within group j is $\sigma_{\mathsf{M}j}^2$. And $\mathrm{Var}[\hat{\mu}_{\mathsf{M}j}]$ is a component that gives the estimation variance. We shall give estimators of these variance components which can be inserted in the expression (26). With only one Auxiliary, where $\hat{\mu}_{\mathsf{M}j} = X_{j1}/N_{j1}$, we can show the following.

Corollary 1.

$$z_{\mathsf{M}j} = \frac{\nu_{\mathsf{M}j}^2 + \mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2 - \frac{N_j}{N_{j1}}\sigma_{\mathsf{M}j}^2 - \frac{2N_j}{N_{j1}}\mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2}{\sigma_{\mathsf{M}j}^2 + \nu_{\mathsf{M}j}^2 + \mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2 - \frac{2N_j}{N_{j1}}\sigma_{\mathsf{M}j}^2 - \frac{2N_j}{N_{j1}}\mu_{\mathsf{M}j}^2\tau_{\mathsf{M}}^2}. \tag{27}$$

Estimator of mean claim variance component within a group

Lemma 1. *Given* **A1M - A4M** *and $\widehat{\sigma}_{\mathsf{M}}^2$ by (20),* $\mathrm{E}[\widehat{\sigma}_{\mathsf{M}}^2] \approx \sigma_{\mathsf{M}}^2$.

Lemma 2. *Given* **A1M - A4M** *and* $\widehat{\sigma}^2_{\mathsf{M}j}$ *by (21),* $\mathrm{E}[\widehat{\sigma}^2_{\mathsf{M}j}] \approx \sigma^2_{\mathsf{M}j}$.

Pseudo-estimator of variance component between groups

Now assume that the claim amounts Z_{ji} are Γ-distributed, conditional on $\Theta_{\mathsf{M}j}$. Then they have the same CV conditional on $\Theta_{\mathsf{M}j}$. In other words we assume the mean claim part of Standard GLM. This enables a mathematically consistent optimization of the pseudo-estimator, which sometimes might perform well even if these assumptions are only partly true. The pseudo-estimator will anyway be approximatively unbiased.

The excess of a random variable is defined in the following way.

$$\mathrm{e}(Z) = \frac{\mathrm{E}[(Z-\mu)^4]}{\mathrm{E}[(Z-\mu)^2]^2} - 3 \text{ for a random variable } Z \text{ with } \mathrm{E}[Z] = \mu.$$

This gives

$$\mathrm{Var}[(Z-\mu)^2] = \mathrm{E}[(Z-\mu)^4] - \mathrm{E}[(Z-\mu)^2]^2 = [\mathrm{e}(Z)+2]\mathrm{E}[(Z-\mu)^2]^2 = [\mathrm{e}(Z)+2]\mathrm{Var}[Z]^2. \tag{28}$$

We give a pseudo-estimator of τ^2_{M} under the conditions that the 3:rd central moment and the excess of $\Theta_{\mathsf{M}j}$ are 0. This is true if $\Theta_{\mathsf{M}j} \sim N(1, \tau^2_{\mathsf{M}})$. Now $\Theta_{\mathsf{M}j}$ is non-negative, but the normal distribution can hold approximatively with a τ^2_{M} giving a small probability to negative values in $N(1, \tau^2_{\mathsf{M}})$.

We have not been able to prove that there is exactly one solution $\hat{\tau}^2_{\mathsf{M}}$ of (33), only that there is at least one, which can be 0. For definiteness we thus define the solution as the largest solution of possibly more than one. The equation is easily solved by a simple binary search.

Let

$$
\begin{aligned}
a_4(x) &= x^{-3}(x+3)(x+2)(x+1), & (29)\\
a_3(x) &= x^{-2}(x+2)(x+1), & (30)\\
\mathrm{v}(x,y) &= (3x^2+6x+1)a_4\left(\frac{x+1}{y}\right) - 4(3x+1)a_3\left(\frac{x+1}{y}\right)\\
&\quad + 6(x+y) - (x+y)^2 + 3, & (31)
\end{aligned}
$$

and

$$
c_{\mathsf{M}j} = \mathbf{1}_{\{N_j>0\}}\left(\frac{\widehat{\sigma}^2_{\mathsf{M}j}}{\hat{\mu}^2_{\mathsf{M}j}} + \hat{\tau}^2_{\mathsf{M}}\right)^2 \mathrm{v}(\hat{\tau}^2_{\mathsf{M}}, \widehat{\sigma}^2_{\mathsf{M}}/N_j)^{-1}\left[\sum_{i=1}^{J}\mathbf{1}_{\{N_i>0\}}\left(\frac{\widehat{\sigma}^2_{\mathsf{M}i}}{\hat{\mu}^2_{\mathsf{M}i}} + \hat{\tau}^2_{\mathsf{M}}\right)^2 \mathrm{v}(\hat{\tau}^2_{\mathsf{M}}, \widehat{\sigma}_{\mathsf{M}}/N_i)^{-1}\right]^{-1}, \tag{32}
$$

$$
\hat{\tau}^2_{\mathsf{M}} = \sum_{j=1}^{J} c_{\mathsf{M}j}\hat{\tau}^2_{\mathsf{M}}\left(\frac{\widehat{\sigma}^2_{\mathsf{M}j}}{\hat{\mu}^2_{\mathsf{M}j}} + \hat{\tau}^2_{\mathsf{M}}\right)^{-1}\left(\frac{Y_{\mathsf{M}j}}{\hat{\mu}_{\mathsf{M}j}} - 1\right)^2. \tag{33}
$$

Lemma 3. *Assume* **A1M - A4M** *and that* $p_{\mathsf{M}} = 2$ *and* Z_{ji} *are Gamma distributed, conditional on* $\Theta_{\mathsf{M}j}$*. For the pseudo-estimator* $\hat{\tau}^2_{\mathsf{M}}$ *we have* $\mathrm{E}[\hat{\tau}^2_{\mathsf{M}}] \approx \tau^2_{\mathsf{M}}$*. If* $\mathrm{E}[(\Theta_{\mathsf{M}j} - 1)^3] = 0$ *and* $\mathrm{e}(\Theta_{\mathsf{M}j}) = 0$*, then* $\mathrm{Var}[(Y_{\mathsf{M}j}/\mu_{\mathsf{M}j} - 1)^2] = \mathrm{v}(\tau^2_{\mathsf{M}}, \sigma^2_{\mathsf{M}}/N_j)$ *and* $\hat{\tau}^2_{\mathsf{M}}$ *is approximately optimal in the sense of having the smallest variance of estimators in the form (33) with* $c_{\mathsf{M}1} + \ldots + c_{\mathsf{M}J} = 1$*.*

Credibility factor estimator for predictor and correction for bias

We can now put estimators into the right side of equation (27).

Theorem 2. *An estimated credibility factor is for* $N_j > 0$

$$\widehat{z}_{\mathsf{M}j} = \frac{\hat{\nu}^2_{\mathsf{M}j} + \hat{\mu}^2_{\mathsf{M}j}\left(1 - \dfrac{2N_j}{N_{j1}}\right)\hat{\tau}^2_{\mathsf{M}} - \dfrac{N_j}{N_{j1}}\widehat{\sigma}^2_{\mathsf{M}j}}{\widehat{\sigma}^2_{\mathsf{M}j} + \hat{\nu}^2_{\mathsf{M}j} + \hat{\mu}^2_{\mathsf{M}j}\left(1 - \dfrac{2N_j}{N_{j1}}\right)\hat{\tau}^2_{\mathsf{M}} - \dfrac{2N_j}{N_{j1}}\widehat{\sigma}^2_{\mathsf{M}j}}. \tag{34}$$

For $N_j = 0$ we set $\widehat{z}_{\mathsf{M}j} = 0$.

We have thus determined the predictor $\widehat{\Lambda}_{\mathsf{M}j} = \widehat{z}_{\mathsf{M}j} Y_{\mathsf{M}j} + (1 - \widehat{z}_{\mathsf{M}j})\hat{\mu}_{\mathsf{M}j}$ by (19) depending on variance estimates. For the non-observable predictors the following unbiasedness holds for the total.

$$\mathrm{E}[\sum_{j=1}^{J} \Lambda^*_{\mathsf{M}j} N_j] = \sum_{j=1}^{J} [z_{\mathsf{M}j}\mu_{\mathsf{M}j} + (1 - z_{\mathsf{M}j})\mu_{\mathsf{M}j}] N_j = \sum_{j=1}^{J} \mu_{\mathsf{M}j} N_j = \mathrm{E}[\sum_{j=1}^{J} Y_{\mathsf{M}j} N_j] = \mathrm{E}[X]. \tag{35}$$

The replacement of $z_{\mathsf{M}j}$ and $\mu_{\mathsf{M}j}$ with estimates $\widehat{z}_{\mathsf{M}j}$ and $\hat{\mu}_{\mathsf{M}j}$ can cause a bias. In examples the bias was found to be negligible. It can be corrected by a simple factor.

Combining claim frequency and mean claim

The claim frequency results are combined with corresponding mean claim results. if we wish to use these results for risk premium. We then define

$$\widehat{\Lambda}_{\mathsf{FM}j} = \widehat{\Lambda}_{\mathsf{F}j} \widehat{\Lambda}_{\mathsf{M}j} \tag{36}$$

and apply a correction factor

We have to assume that $\{\Theta_{\mathsf{F}j}\}_1^J$ and $\{\Theta_{\mathsf{M}j}\}_1^J$ are independent. Claim numbers are S-ancillary for the mean claim parameters in the Compound Poisson model, which motivates (36).

Simulation results

We have compared our pseudo-estimators with the classical type ones that are rendered in Ohlsson and Johansson (2010).

Claim frequency and mean claim were analyzed separately. The basic model assumptions were obeyed, but we did not let the distributions of $\Theta_{\mathsf{F}j}$ and $\Theta_{\mathsf{M}j}$ have zero excess or even mostly third central moment zero, as our pseudo-estimator theorems presuppose. From a practical viewpoint these assumptions are unrealistic, but they admit relatively simple and mathematically consistent estimators. The estimators have to be reasonably robust against departures from the assumptions in order to be useful, though. Instead we used some uniform and Γ distributions for $\Theta_{\mathsf{F}j}$ and $\Theta_{\mathsf{M}j}$.

We found that the pseudo-estimator $\hat{\tau}_{\mathsf{F}}^2$ is the best one, except for some cases with small τ_{F}^2.

The mean claim pseudo-estimator $\hat{\tau}_{\mathsf{M}}^2$ by (33) generally seemed to be best for light-tailed conditional claim distributions, not only Γ-distributions, while the non-pseudo-estimator by (23) seemed best for heavy-tailed ones.

Application

Credibility analysis according to this and the larger submitted paper can be done with Rapp, with a choice of choosing the classical estimators or the pseudo-estimators given in my paper.

A demo run of Taranmultiklass2.Rpp follows. It might take place in a separate session.