Non-Life Insurance: Mathematics and Statistics

Solution sheet 11

Solution 11.1 (Inhomogeneous) Credibility Estimators for Claim Counts

We define

$$X_{i,1} = \frac{N_{i,1}}{v_{i,1}},$$

for all $i \in \{1, \dots, 5\}$. Then we have

$$\mathbb{E}[X_{i,1} \,|\, \Theta_i] = \frac{1}{v_{i,1}}\, \mathbb{E}[N_{i,1} \,|\, \Theta_i] = \frac{1}{v_{i,1}}\, \mu(\Theta_i)\, v_{i,1} = \mu(\Theta_i)$$

and

$$\mathrm{Var}(X_{i,1} \,|\, \Theta_i) = \frac{1}{v_{i,1}^2}\, \mathrm{Var}(N_{i,1} \,|\, \Theta_i) = \frac{1}{v_{i,1}^2}\, \mu(\Theta_i)\, v_{i,1} = \frac{\mu(\Theta_i)}{v_{i,1}} = \frac{\sigma^2(\Theta_i)}{v_{i,1}},$$

with $\sigma^2(\Theta_i) = \mu(\Theta_i) = \Theta_i \lambda_0$, for all $i \in \{1, \dots, 5\}$. Moreover, since

$$\mathbb{E}[\mu(\Theta_i)^2] = \mathrm{Var}(\mu(\Theta_i)) + \mathbb{E}[\mu(\Theta_i)]^2 = \tau^2 + \lambda_0^2 < \infty$$

and

$$\mathbb{E}[X_{i,1}^2 \,|\, \Theta_i] = \mathrm{Var}(X_{i,1} \,|\, \Theta_i) + \mathbb{E}[X_{i,1} \,|\, \Theta_i]^2 = \frac{\mu(\Theta_i)}{v_{i,1}} + \mu(\Theta_i)^2,$$

we get

$$\mathbb{E}[X_{i,1}^2] = \mathbb{E}\left[\mathbb{E}[X_{i,1}^2 \,|\, \Theta_i]\right] = \mathbb{E}\left[\frac{\mu(\Theta_i)}{v_{i,1}} + \mu(\Theta_i)^2\right] = \frac{\lambda_0}{v_{i,1}} + \tau^2 + \lambda_0^2 < \infty,$$

for all $i \in \{1, \dots, 5\}$. In particular, the Model Assumptions 8.13 of the lecture notes for the Bühlmann-Straub model are satisfied. The (expected) volatility σ^2 within the regions defined in formula (8.5) of the lecture notes is given by

$$\sigma^2 = \mathbb{E}[\sigma^2(\Theta_i)] = \mathbb{E}[\mu(\Theta_i)] = \lambda_0 = 0.088.$$

(a) Let $i \in \{1, \dots, 5\}$. Then, according to Theorem 8.17 of the lecture notes, the inhomogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}$ is given by

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha_{i,T}\, \widehat{X}_{i,1:T} + (1 - \alpha_{i,T})\, \mu_0,$$

with credibility weight $\alpha_{i,T}$ and observation based estimator $\widehat{X}_{i,1:T}$

$$\alpha_{i,T} = \frac{v_{i,1}}{v_{i,1} + \frac{\sigma^2}{\tau^2}} \qquad \text{and} \qquad \widehat{X}_{i,1:T} = \frac{1}{v_{i,1}}\, v_{i,1}\, X_{i,1} = X_{i,1}.$$

Hence, we get

$$\widehat{\widehat{\mu(\Theta_i)}} = \frac{v_{i,1}}{v_{i,1} + \frac{\sigma^2}{\tau^2}}\, X_{i,1} + \frac{\frac{\sigma^2}{\tau^2}}{v_{i,1} + \frac{\sigma^2}{\tau^2}}\, \mu_0 = \frac{v_{i,1}}{v_{i,1} + \frac{0.088}{0.00024}}\, X_{i,1} + \frac{\frac{0.088}{0.00024}}{v_{i,1} + \frac{0.088}{0.00024}}\, 0.088.$$

The results for the 5 regions are summarized in the following table:

	region 1	region 2	region 3	region 4	region 5
$\alpha_{i,T}$	99.3%	96.5%	99.7%	99.0%	92.0%
$\widehat{X}_{i,1:T}$	7.8%	7.8%	7.4%	9.8%	7.5%
$\widehat{\widehat{\mu(\Theta_i)}}$	7.8%	7.9%	7.4%	9.8%	7.6%

Note that since the credibility coefficient $\kappa = \sigma^2/\tau^2 \approx 367$ is rather small compared to the volumes $v_{1,1}, \ldots, v_{5,1}$, the credibility weights $\alpha_{1,T}, \ldots, \alpha_{5,T}$ are fairly high. Moreover, the observation based estimators are almost the same for the regions $1, 2, 3$ and 5 and only $\widehat{X}_{4,1:T}$ is roughly 2% higher. As a result, only for the smallest two credibility weights $\alpha_{2,T}$ and $\alpha_{5,T}$ we see a slight upwards deviation of the corresponding inhomogeneous credibility estimators $\widehat{\widehat{\mu(\Theta_2)}}$ and $\widehat{\widehat{\mu(\Theta_5)}}$ from the observation based estimators $\widehat{X}_{2,1:T}$ and $\widehat{X}_{5,1:T}$ towards μ_0.

(b) Since the number of policies grows 5% in each region, next year's numbers of policies $v_{1,2}, \ldots, v_{5,2}$ are given by

	region 1	region 2	region 3	region 4	region 5
$v_{i,2}$	52'564	10'642	127'376	36'797	4'402

Similarly to part (a), we define

$$X_{i,2} = \frac{N_{i,2}}{v_{i,2}},$$

for all $i \in \{1, \ldots, 5\}$. Then, according to formula (8.17) of the lecture notes, the mean square error of prediction is given by

$$\mathbb{E}\left[\left(\frac{N_{i,2}}{v_{i,2}} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right] = \mathbb{E}\left[\left(X_{i,2} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right] = \frac{\sigma^2}{v_{i,2}} + (1 - \alpha_{i,T})\,\tau^2,$$

for all $i \in \{1, \ldots, 5\}$. We get the following square-rooted mean square errors of prediction for the five regions:

	region 1	region 2	region 3	region 4	region 5
$\sqrt{\text{mse of prediction}}$	0.185%	0.408%	0.119%	0.221%	0.627%
in % of the credibility estimators	2.4%	5.2%	1.6%	2.2%	8.3%

Note that we get the highest (square-rooted) mean square errors of prediction for the regions 2 and 5, i.e. exactly for those regions for which we also have the lowest volumes and, consequently, the lowest credibility weights. Of course, this is due to the formula for the mean square error of prediction given above.

Solution 11.2 (Homogeneous) Credibility Estimators for Claim Sizes

We define

$$X_{i,t} = \frac{Y_{i,t}}{v_{i,t}},$$

for all $i \in \{1, 2, 3, 4\}$, $t \in \{1, 2\}$. Then we have

$$\mathbb{E}[X_{i,t} \,|\, \Theta_i] = \frac{1}{v_{i,t}}\,\mathbb{E}[Y_{i,t} \,|\, \Theta_i] = \frac{1}{v_{i,t}}\,\frac{\mu(\Theta_i) c v_{i,t}}{c} = \mu(\Theta_i)$$

and

$$\mathrm{Var}(X_{i,t} \,|\, \Theta_i) = \frac{1}{v_{i,t}^2}\,\mathrm{Var}(Y_{i,t} \,|\, \Theta_i) = \frac{1}{v_{i,t}^2}\,\frac{\mu(\Theta_i) c v_{i,t}}{c^2} = \frac{\mu(\Theta_i)}{c v_{i,t}} = \frac{\sigma^2(\Theta_i)}{v_{i,t}},$$

with

$$\sigma^2(\Theta_i) = \frac{\mu(\Theta_i)}{c} = \frac{\Theta_i}{c},$$

for all $i \in \{1,2,3,4\}$, $t \in \{1,2\}$. Moreover, using that

$$\mathbb{E}[X_{i,t}^2 \,|\, \Theta_i] = \mathrm{Var}(X_{i,t} \,|\, \Theta_i) + \mathbb{E}[X_{i,t} \,|\, \Theta_i]^2 = \frac{\mu(\Theta_i)}{cv_{i,t}} + \mu(\Theta_i)^2 = \frac{\Theta_i}{cv_{i,t}} + \Theta_i^2$$

we get

$$\mathbb{E}[X_{i,t}^2] = \mathbb{E}\left[\mathbb{E}[X_{i,t}^2 \,|\, \Theta_i]\right] = \mathbb{E}\left[\frac{\Theta_i}{cv_{i,t}} + \Theta_i^2\right] < \infty$$

by assumption, for all $i \in \{1,2,3,4\}$, $t \in \{1,2\}$. In particular, the Model Assumptions 8.13 of the lecture notes for the Bühlmann-Straub model are satisfied.

(a) In order to calculate the homogeneous credibility estimators, we need to estimate the structural parameters $\sigma^2 = \mathbb{E}[\sigma^2(\Theta_1)]$ and $\tau^2 = \mathrm{Var}(\mu(\Theta_1))$. First, following Theorem 8.17 of the lecture notes, we define the observation based estimator $\widehat{X}_{i,1:T}$ as

$$\widehat{X}_{i,1:T} = \frac{1}{\sum_{t=1}^T v_{i,t}} \sum_{t=1}^T v_{i,t} X_{i,t} = \frac{v_{i,1}X_{i,1} + v_{i,2}X_{i,2}}{v_{i,1} + v_{i,2}} = \frac{Y_{i,1} + Y_{i,2}}{v_{i,1} + v_{i,2}},$$

for all $i \in \{1,2,3,4\}$. According to formula (8.15) of the lecture notes, σ^2 can be estimated by

$$\widehat{\sigma}_T^2 = \frac{1}{I} \sum_{i=1}^I \frac{1}{T-1} \sum_{t=1}^T v_{i,t} \left(X_{i,t} - \widehat{X}_{i,1:T}\right)^2 \approx 1.3 \cdot 10^{10}.$$

For the estimator $\widehat{\tau}_T^2$ of τ^2, we define first the weighted sample mean $\bar{X}$ over all observations by

$$\bar{X} = \frac{1}{\sum_{i=1}^I \sum_{t=1}^T v_{i,t}} \sum_{i=1}^I \sum_{t=1}^T v_{i,t} X_{i,t} = \frac{\sum_{i=1}^I Y_{i,1} + Y_{i,2}}{\sum_{i=1}^I v_{i,1} + v_{i,2}} \approx 7004.$$

Then, as on page 219 of the lecture notes, we define $\widehat{v}_T^2$, c_w and $\widehat{t}_T^2$ as

$$\widehat{v}_T^2 = \frac{I}{I-1} \sum_{i=1}^4 \frac{v_{i,1} + v_{i,2}}{\sum_{j=1}^I v_{j,1} + v_{j,2}} \left(\widehat{X}_{i,1:T} - \bar{X}\right)^2 \approx 9.3 \cdot 10^7,$$

$$c_w = \frac{I-1}{I} \left[\sum_{i=1}^I \frac{v_{i,1} + v_{i,2}}{\sum_{j=1}^I v_{j,1} + v_{j,2}} \left(1 - \frac{v_{i,1} + v_{i,2}}{\sum_{j=1}^I v_{j,1} + v_{j,2}}\right)\right]^{-1} \approx 1.425$$

and

$$\widehat{t}_T^2 = c_w \left(\widehat{v}_T^2 - \frac{I\,\widehat{\sigma}_T^2}{\sum_{i=1}^I v_{i,1} + v_{i,2}}\right) \approx 1.25 \cdot 10^8.$$

Then, using formula (8.16) of the lecture notes, τ^2 can be estimated by

$$\widehat{\tau}_T^2 = \max\left\{\widehat{t}_T^2, 0\right\} = \widehat{t}_T^2 \approx 1.25 \cdot 10^8.$$

Now we are ready to calculate the inhomogeneous credibility estimators. Let $i \in \{1,2,3,4\}$. Then, according to Theorem 8.17 of the lecture notes, the inhomogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}}$ is given by

$$\widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}} = \alpha_{i,T}\, \widehat{X}_{i,1:T} + (1 - \alpha_{i,T})\, \widehat{\mu}_T$$

with credibility weight $\alpha_{i,T}$ and estimate $\widehat{\mu}_T$ given by

$$\alpha_{i,T} = \frac{v_{i,1} + v_{i,2}}{v_{i,1} + v_{i,2} + \widehat{\sigma}_T^2/\widehat{\tau}_T^2} \qquad \text{and} \qquad \widehat{\mu}_T = \frac{1}{\sum_{i=1}^I \alpha_{i,T}} \sum_{i=1}^I \alpha_{i,T} \widehat{X}_{i,1:T}.$$

Hence, we get

$$\widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}} = \alpha_{i,T}\,\widehat{X}_{i,1:T} + \frac{(1-\alpha_{i,T})}{\sum_{i=1}^I \alpha_{i,T}} \sum_{i=1}^I \alpha_{i,T} \widehat{X}_{i,1:T}$$

The results for the 4 risk classes are summarized in the following table:

	risk class 1	risk class 2	risk class 3	risk class 4
$\alpha_{i,T}$	95.4%	98.4%	82.4%	89.6%
$\widehat{X}_{i,1:T}$	10'493	1'907	18'375	29'197
$\widehat{\widehat{\mu(\Theta_i)}}$	10'677	2'107	17'702	27'665

Moreover, we get $\hat{\mu}_T \approx$ 14'538. Looking at the credibility weights $\alpha_{1,1}, \alpha_{2,1}, \alpha_{3,1}$ and $\alpha_{4,1}$, we see that the estimated credibility coefficient $\hat{\kappa} = \hat{\sigma}_T^2/\hat{\tau}_T^2 \approx 104$ has the biggest impact on risk classes 3 and 4 where we have less volumes compared to risk classes 1 and 2. As a result, the smoothing of the observation based estimators $\widehat{X}_{1,1:T}, \widehat{X}_{2,1:T}, \widehat{X}_{3,1:T}$ and $\widehat{X}_{4,1:T}$ towards $\hat{\mu}_T$ is strongest for risk classes 1 and 2.

(b) Since the number of claims grows 5% in each region, next year's numbers of claims $v_{1,3}, \ldots, v_{4,3}$ are given by

	risk class 1	risk class 2	risk class 3	risk class 4
$v_{i,3}$	1'167	3'468	262	479

Similarly to part (a), we define

$$X_{i,3} = \frac{Y_{i,3}}{v_{i,3}},$$

for all $i \in \{1,2,3,4\}$. Then, according to formula (8.17) of the lecture notes, the mean square error of prediction can be estimated by

$$\widehat{\mathbb{E}}\left[\left(\frac{Y_{i,3}}{v_{i,3}} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right] = \widehat{\mathbb{E}}\left[\left(X_{i,3} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right] = \frac{\widehat{\sigma}_T^2}{v_{i,3}} + (1-\alpha_{i,T})\,\widehat{\tau}_T^2,$$

for all $i \in \{1,2,3,4\}$. We get the following square-rooted mean square errors of prediction for the four risk classes:

	risk class 1	risk class 2	risk class 3	risk class 4
$\sqrt{\text{estimated mse of prediction}}$	4'099	2'390	8'446	6'331
in % of the credibility estimators	38.4%	113.4%	47.7%	22.9%

According to the formula given above for the estimated mean square error of prediction, we observe that, the smaller the volumes of a particular risk class, the bigger the corresponding (square-rooted) estimated mean square error of prediction. Moreover, note that these square-rooted estimated mean square errors of prediction are rather high compared to the credibility estimators, which indicates a high variability within the individual risk classes.

Solution 11.3 Pareto-Gamma Model

(a) Let $f_{\mathbf{Y}|\Lambda}$ and f_Λ denote the density of $\mathbf{Y}\,|\,\Lambda$ and f_Λ, respectively. Then we have

$$f_{\mathbf{Y}|\Lambda}(y_1,\ldots,y_T\,|\,\Lambda=\alpha) = \prod_{t=1}^T \frac{\alpha}{\theta}\left(\frac{y_t}{\theta}\right)^{-(\alpha+1)} \cdot 1_{\{y_t\geq\theta\}} = \alpha^T \left(\prod_{t=1}^T \frac{y_t}{\theta}\right)^{-\alpha} \prod_{t=1}^T \frac{y_t}{\theta} \cdot 1_{\{y_t\geq\theta\}}$$

and

$$f_\Lambda(\alpha) = \frac{c^\gamma}{\Gamma(\gamma)}\alpha^{\gamma-1}\exp\{-c\alpha\}\cdot 1_{\{\alpha>0\}}.$$

Let $f_{\Lambda|\mathbf{Y}}$ denote the density of $\Lambda|\mathbf{Y}$. Then, for all $\alpha>0$ and $y_1,\ldots,y_T\geq\theta$, we have

$$\begin{aligned}
f_{\Lambda|\mathbf{Y}}(\alpha\,|\,Y_1=y_1,\ldots,Y_T=y_T) &= \frac{f_{\mathbf{Y}|\Lambda}(y_1,\ldots,y_T\,|\,\Lambda=\alpha)\,f_\Lambda(\alpha)}{\int_0^\infty f_{\mathbf{Y}|\Lambda}(y_1,\ldots,y_T\,|\,\Lambda=x)\,f_\Lambda(x)\,dx}\\
&\propto \alpha^T\left(\prod_{t=1}^T\frac{y_t}{\theta}\right)^{-\alpha}\alpha^{\gamma-1}\exp\{-c\alpha\}\\
&= \alpha^{\gamma+T-1}\exp\left\{-\alpha\sum_{t=1}^T\log\frac{y_t}{\theta}\right\}\exp\{-c\alpha\}\\
&= \alpha^{\gamma+T-1}\exp\left\{-\alpha\left(\sum_{t=1}^T\log\frac{y_t}{\theta}+c\right)\right\},
\end{aligned}$$

i.e. we have shown that

$$\Lambda\,|\,\mathbf{Y}\sim\Gamma\left(\gamma+T, c+\sum_{t=1}^T\log\frac{Y_t}{\theta}\right).$$

(b) We calculate

$$\begin{aligned}
\alpha_T\,\widehat{\lambda}_T + (1-\alpha_T)\,\lambda_0 &= \frac{\sum_{t=1}^T\log\frac{Y_t}{\theta}}{c+\sum_{t=1}^T\log\frac{Y_t}{\theta}}\,\frac{T}{\sum_{t=1}^T\log\frac{Y_t}{\theta}} + \frac{c}{c+\sum_{t=1}^T\log\frac{Y_t}{\theta}}\,\frac{\gamma}{c}\\
&= \frac{\gamma+T}{c+\sum_{t=1}^T\log\frac{Y_t}{\theta}}\\
&= \widehat{\lambda}_T^{\text{post}}.
\end{aligned}$$

(c) For the (mean square error) uncertainty of the posterior estimator $\widehat{\lambda}_T^{\text{post}}=\mathbb{E}[\Lambda\,|\,\mathbf{Y}]$ we have

$$\begin{aligned}
\mathbb{E}\left[\left(\Lambda-\widehat{\lambda}_T^{\text{post}}\right)^2\,\middle|\,\mathbf{Y}\right] &= \mathbb{E}\left[(\Lambda-\mathbb{E}[\Lambda\,|\,\mathbf{Y}])^2\,\middle|\,\mathbf{Y}\right]\\
&= \operatorname{Var}(\Lambda\,|\,\mathbf{Y})\\
&= \frac{\gamma+T}{\left(c+\sum_{t=1}^T\log\frac{Y_t}{\theta}\right)^2}\\
&= \frac{1}{c+\sum_{t=1}^T\log\frac{Y_t}{\theta}}\,\widehat{\lambda}_T^{\text{post}}\\
&= (1-\alpha_T)\,\frac{1}{c}\,\widehat{\lambda}_T^{\text{post}}.
\end{aligned}$$

(d) Analogously to $\widehat{\lambda}_T^{\text{post}}$, the posterior estimator $\widehat{\lambda}_{T-1}^{\text{post}}$ in the sub-model where we only have observed $(Y_1, \ldots, Y_{T-1})$ is given by

$$\widehat{\lambda}_{T-1}^{\text{post}} = \frac{\gamma + T - 1}{c + \sum_{t=1}^{T-1} \log \frac{Y_t}{\theta}}.$$

Then we can calculate

$$\begin{aligned}
\beta_T \frac{1}{\log \frac{Y_T}{\theta}} + (1 - \beta_T)\, \widehat{\lambda}_{T-1}^{\text{post}} &= \frac{\log \frac{Y_T}{\theta}}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}} \frac{1}{\log \frac{Y_T}{\theta}} + \frac{c + \sum_{t=1}^{T-1} \log \frac{Y_t}{\theta}}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}} \frac{\gamma + T - 1}{c + \sum_{t=1}^{T-1} \log \frac{Y_t}{\theta}} \\
&= \frac{\gamma + T}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}} \\
&= \widehat{\lambda}_T^{\text{post}}.
\end{aligned}$$

<u>Remark:</u> Suppose we would like to use the observations $Y_1, \ldots, Y_{T-1}$ in order to estimate Y_T in a Bayesian sense. Then we have

$$\begin{aligned}
\mathbb{E}[Y_T \,|\, Y_1, \ldots, Y_{T-1}] &= \mathbb{E}\left[\mathbb{E}\left[Y_T \,|\, Y_1, \ldots, Y_{T-1}, \Lambda\right] \,|\, Y_1, \ldots, Y_{T-1}\right] \quad \text{a.s.} \\
&= \mathbb{E}\left[\mathbb{E}\left[Y_T \,|\, \Lambda\right] \,|\, Y_1, \ldots, Y_{T-1}\right] \quad \text{a.s.},
\end{aligned}$$

where in the second equality we used that, conditionally given Λ, $Y_1, \ldots, Y_T$ are independent. Now, by assumption,

$$Y_T \,|\, \Lambda \sim \text{Pareto}(\theta, \Lambda).$$

In particular, $\mathbb{E}[Y_T \,|\, \Lambda] < \infty$ if and only if $\Lambda > 1$. However, according to part (a), we have

$$\Lambda \,|\, (Y_1, \ldots, Y_{T-1}) \sim \Gamma\left(\gamma + T - 1, c + \sum_{t=1}^{T-1} \log \frac{Y_t}{\theta}\right).$$

Since the range of a gamma distribution is the whole positive real line, this implies that

$$0 < \mathbb{P}\left[\Lambda \leq 1 \,|\, Y_1, \ldots, Y_{T-1}\right] = \mathbb{P}\left[\mathbb{E}\left[Y_T \,|\, \Lambda\right] = \infty \,|\, Y_1, \ldots, Y_{T-1}\right] \quad \text{a.s.}$$

We conclude that

$$\mathbb{E}[Y_T \,|\, Y_1, \ldots, Y_{T-1}] = \infty \quad \text{a.s.}$$