Non-Life Insurance: Mathematics and Statistics

Exercise sheet 11

Exercise 11.1 (Inhomogeneous) Credibility Estimators for Claim Counts
Suppose that we are given the following current year's claim counts data for 5 different regions, where $v_{i,1}$ = number of policies in region i and $N_{i,1}$ = number of claims in region i, for all $i \in \{1, \dots, 5\}$:

region i	$v_{i,1}$	$N_{i,1}$
1	50'061	3'880
2	10'135	794
3	121'310	8'941
4	35'045	3'448
5	4'192	314

Table 1: Observed numbers of policies $v_{i,1}$ and claim counts $N_{i,1}$ in the 5 regions.

We assume that we are in the Bühlmann-Straub model framework with $I = 5$, $T = 1$ and

$$N_{i,1}|\Theta_i \sim \text{Poi}(\mu(\Theta_i)v_{i,1}),$$

where $\mu(\Theta_i) = \Theta_i\lambda_0$ and $\lambda_0 = 0.088$, for all $i \in \{1, \dots, 5\}$. Moreover, we assume that the pairs $(\Theta_1, N_{1,1}), \dots, (\Theta_5, N_{5,1})$ are independent and $\Theta_1, \dots, \Theta_5$ are i.i.d. with $\Theta_1 > 0$ a.s. Finally, we set

$$\mu_0 = \mathbb{E}[\mu(\Theta_1)] = \lambda_0 \qquad \text{and} \qquad \tau^2 = \text{Var}(\mu(\Theta_1)) = 0.00024.$$

(a) Calculate the inhomogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}$ for each region $i \in \{1, \dots, 5\}$ and comment on the results. What would we observe if we decreased the volatility τ^2 between the risk classes?

(b) We denote next year's numbers of policies by $v_{1,2}, \dots, v_{5,2}$ and next year's numbers of claims by $N_{1,2}, \dots, N_{5,2}$. Suppose that $N_{i,1}$ and $N_{i,2}$ are independent, conditionally given Θ_i, for all $i \in \{1, \dots, 5\}$, and that the number of policies grows 5% in each region. For all $i \in \{1, \dots, 5\}$, under the assumption $N_{i,2}|\Theta_i \sim \text{Poi}(\mu(\Theta_i)v_{i,2})$, calculate the mean square error of prediction

$$\mathbb{E}\left[\left(\frac{N_{i,2}}{v_{i,2}} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right].$$

Exercise 11.2 (Homogeneous) Credibility Estimators for Claim Sizes
Suppose that we are given the following claim size data for two different years and four different risk classes, where $v_{i,t}$ = number of claims in risk class i and year t and $Y_{i,t}$ = total claim size in risk class i and year t, for all $i \in \{1, 2, 3, 4\}$ and $t \in \{1, 2\}$:

risk class i	$v_{i,1}$	$Y_{i,1}$	$v_{i,2}$	$Y_{i,2}$
1	1'058	8'885'738	1'111	13'872'665
2	3'146	7'902'445	3'303	4'397'183
3	238	2'959'517	250	6'007'351
4	434	10'355'286	456	15'629'998

Table 2: Observed numbers of claims $v_{i,1}$ and $v_{i,2}$ and total claim sizes $Y_{i,1}$ and $Y_{i,1}$ in the 4 risk classes.

We assume that we are in the Bühlmann-Straub model framework with $I = 4$, $T = 2$ and

$$Y_{i,t}|\Theta_i \sim \Gamma(\mu(\Theta_i)cv_{i,t}, c),$$

where $\mu(\Theta_i) = \Theta_i$ and $c > 0$, for all $i \in \{1, 2, 3, 4\}$ and $t \in \{1, 2\}$. Moreover, we assume that $(\Theta_1, Y_{1,1}, Y_{1,2}), \ldots, (\Theta_4, Y_{4,1}, Y_{4,2})$ are independent and that $\Theta_1, \ldots, \Theta_4$ are i.i.d. with $\mathbb{E}[\Theta_1^2] < \infty$ and $\Theta_1 > 0$ a.s. Finally, we also assume that $Y_{i,1}$ and $Y_{i,2}$ are independent, conditionally given Θ_i, for all $i \in \{1, 2, 3, 4\}$.

(a) Calculate the homogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}}$ for each risk class $i \in \{1, 2, 3, 4\}$ and comment on the results.

(b) We denote next year's numbers of claims by $v_{1,3}, \ldots, v_{4,3}$ and next year's total claim sizes by $Y_{1,3}, \ldots, Y_{4,3}$. Suppose that $Y_{i,1}, Y_{i,2}$ and $Y_{i,3}$ are independent, conditionally given Θ_i, for all $i \in \{1, 2, 3, 4\}$, and that the number of claims grows 5% in each risk cell. For all $i \in \{1, 2, 3, 4\}$, under the assumption $Y_{i,3} | \Theta_i \sim \Gamma(\mu(\Theta_i) c v_{i,3}, c)$, estimate the mean square error of prediction

$$\mathbb{E}\left[\left(\frac{Y_{i,3}}{v_{i,3}} - \widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}}\right)^2\right].$$

Exercise 11.3 Pareto-Gamma Model

Suppose that $\Lambda \sim \Gamma(\gamma, c)$ with prior shape parameter $\gamma > 0$ and prior scale parameter $c > 0$ and, conditionally given Λ, the components of $\mathbf{Y} = (Y_1, \ldots, Y_T)$ are independent with $Y_t \sim \text{Pareto}(\theta, \Lambda)$ for some threshold $\theta > 0$, for all $t \in \{1, \ldots, T\}$.

(a) Show that the posterior distribution of Λ, conditional on $\mathbf{Y}$, is given by

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

(b) For the estimation of the unknown parameter Λ, we define the prior estimator λ_0 and the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ as

$$\lambda_0 = \mathbb{E}[\Lambda] = \frac{\gamma}{c} \qquad \text{and} \qquad \widehat{\lambda}_T^{\text{post}} = \mathbb{E}[\Lambda | \mathbf{Y}] = \frac{\gamma + T}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}}.$$

Show that the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ has the following credibility form

$$\widehat{\lambda}_T^{\text{post}} = \alpha_T \, \widehat{\lambda}_T + (1 - \alpha_T) \, \lambda_0,$$

with credibility weight α_T and observation based estimator $\widehat{\lambda}_T$ given by

$$\alpha_T = \frac{\sum_{t=1}^{T} \log \frac{Y_t}{\theta}}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}} \qquad \text{and} \qquad \widehat{\lambda}_T = \frac{T}{\sum_{t=1}^{T} \log \frac{Y_t}{\theta}}.$$

(c) Show that the (conditional mean square error) uncertainty of the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ is given by

$$\mathbb{E}\left[\left(\Lambda - \widehat{\lambda}_T^{\text{post}}\right)^2 \middle| \mathbf{Y}\right] = (1 - \alpha_T) \, \frac{1}{c} \, \widehat{\lambda}_T^{\text{post}}.$$

(d) Let $\widehat{\lambda}_{T-1}^{\text{post}}$ denote the posterior estimator in the sub-model where we only have observed $(Y_1, \ldots, Y_{T-1})$. Show that the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ has the following recursive update structure

$$\widehat{\lambda}_T^{\text{post}} = \beta_T \, \frac{1}{\log \frac{Y_T}{\theta}} + (1 - \beta_T) \, \widehat{\lambda}_{T-1}^{\text{post}},$$

with credibility weight

$$\beta_T = \frac{\log \frac{Y_T}{\theta}}{c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}}.$$