Nonlife Actuarial Models

Chapter 9

Empirical Implementation of Credibility

Learning Objectives

1. Empirical Bayes method
2. Nonparametric estimation
3. Semiparametric estimation
4. Parametric estimation

9.1 Empirical Bayes Method

- In the Bühlmann and Bühlmann-Straub framework, the key quantities of interest are the expected value of the process variance, μ_{PV}, and the variance of the hypothetical means, σ^2_{HM}.

- These quantities can be derived from the Bayesian framework and depend on both the prior distribution and the likelihood.

- In a strictly Bayesian approach, the prior distribution is given and inference is drawn based on the given prior.

- For practical applications when researchers are not in a position to state the prior, empirical methods may be applied to estimate the hyperparameters.

- This is called the **empirical Bayes method**. Depending on the assumptions about the prior distribution and the likelihood, empirical Bayes estimation may adopt one of the following approaches

1. **Nonparametric approach:** No assumptions are made about the prior density $f_\Theta(\theta)$ and the conditional density $f_{X\,|\,\Theta}(x\,|\,\theta)$. The method is very general and applies to a wide range of models.

2. **Semiparametric approach:** Parametric assumptions concerning $f_{X\,|\,\Theta}(x\,|\,\theta)$ is made, while the prior distribution of the risk parameters $f_\Theta(\theta)$ remains unspecified.

3. **Parametric approach:** When the researcher makes specific assumptions about $f_{X\,|\,\Theta}(x\,|\,\theta)$ and $f_\Theta(\theta)$, the estimation of the parameters in the model may be carried out using the maximum likelihood estimation (MLE) method.

9.2 Nonparametric Estimation

- No specific assumption is made about the likelihood of the loss random variables and the prior distribution of the risk parameters.

- The key quantities required are the expected value of the process variance, $\mu_{\rm PV}$, and the variance of the hypothetical means, $\sigma^2_{\rm HM}$, which together determine the Bühlmann credibility parameter k.

- We extend the Bühlmann-Straub set-up to consider multiple risk groups, each with multiple samples of loss observations over possibly different periods.

- We formally state the assumptions of the extended set-up as follows

1. Let X_{ij} denote the loss per unit of exposure and m_{ij} denote the amount of exposure. The index i denotes the ith risk group, for

$i = 1, \cdots, r$, with $r > 1$. Given i, the index j denotes the jth loss observation in the ith group, for $j = 1, \cdots, n_i$, where $n_i > 1$ for $i = 1, \cdots, r$.

2. X_{ij} are assumed to be independently distributed. The risk parameter of the ith group is denoted by θ_i, which is a realization of the random variable Θ_i. We assume Θ_i to be independently and identically distributed as Θ.

3. The following assumptions are made for the hypothetical means and the process variance

$$\mathrm{E}(X_{ij} \,|\, \Theta = \theta_i) = \mu_X(\theta_i), \qquad \text{for } i = 1, \cdots, r;\; j = 1, \cdots, n_i, \quad (9.2)$$

and

$$\mathrm{Var}(X_{ij} \,|\, \theta_i) = \frac{\sigma_X^2(\theta_i)}{m_{ij}}, \qquad \text{for } i = 1, \cdots, r;\; j = 1, \cdots, n_i. \quad (9.3)$$

- We define the overall mean of the loss variable as

$$\mu_X = \mathrm{E}[\mu_X(\Theta_i)] = \mathrm{E}[\mu_X(\Theta)], \tag{9.4}$$

the mean of the process variance as

$$\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_X^2(\Theta_i)] = \mathrm{E}[\sigma_X^2(\Theta)], \tag{9.5}$$

and the variance of the hypothetical means as

$$\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mu_X(\Theta_i)] = \mathrm{Var}[\mu_X(\Theta)]. \tag{9.6}$$

- For future reference, we also define the following quantities

$$m_i = \sum_{j=1}^{n_i} m_{ij}, \qquad \text{for } i = 1, \cdots, r, \tag{9.7}$$

which is the total exposure for the ith risk group; and

$$m = \sum_{i=1}^{r} m_i, \tag{9.8}$$

which is the total exposure over all risk groups.

- Also, we define

$$\bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{n_i} m_{ij} X_{ij}, \qquad \text{for } i = 1, \cdots, r, \tag{9.9}$$

as the exposure-weighted mean of the ith risk group; and

$$\bar{X} = \frac{1}{m} \sum_{i=1}^{r} m_i \bar{X}_i \tag{9.10}$$

as the overall weighted mean.

- The Bühlmann-Straub credibility predictor of the loss in the next period or a random individual of the ith risk group is

$$Z_i \bar{X}_i + (1 - Z_i)\mu_X, \tag{9.11}$$

where

$$Z_i = \frac{m_i}{m_i + k}, \tag{9.12}$$

with

$$k = \frac{\mu_{\text{PV}}}{\sigma^2_{\text{HM}}}. \tag{9.13}$$

- To implement the credibility prediction, we need to estimate μ_X, μ_{PV} and σ^2_{HM}.

- It is natural to estimate μ_X by $\bar{X}$. It can be shown that $\text{E}(\bar{X}) = \mu_X$, so that $\bar{X}$ is an unbiased estimator of μ_X.

Theorem 9.1: The following quantity is an unbiased estimator of μ_{PV}

$$\hat{\mu}_{\mathrm{PV}} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}. \tag{9.16}$$

Proof: We re-arrange the inner summation term in the numerator of equation (9.16) to obtain

$$\begin{aligned}
\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2 &= \sum_{j=1}^{n_i} m_{ij} \left\{[X_{ij} - \mu_X(\theta_i)] - [\bar{X}_i - \mu_X(\theta_i)]\right\}^2 \\
&= \sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\theta_i)]^2 + \sum_{j=1}^{n_i} m_{ij}[\bar{X}_i - \mu_X(\theta_i)]^2 \\
&\quad -2\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\theta_i)][\bar{X}_i - \mu_X(\theta_i)]. \qquad (9.17)
\end{aligned}$$

Simplifying the last two terms on the right-hand side of the above equation, we have

$$\begin{aligned}
&\sum_{j=1}^{n_i} m_{ij}[\bar{X}_i - \mu_X(\theta_i)]^2 - 2\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\theta_i)][\bar{X}_i - \mu_X(\theta_i)] \\
&= m_i[\bar{X}_i - \mu_X(\theta_i)]^2 - 2[\bar{X}_i - \mu_X(\theta_i)]\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\theta_i)] \\
&= -m_i[\bar{X}_i - \mu_X(\theta_i)]^2. \qquad (9.18)
\end{aligned}$$

Combining equations (9.17) and (9.18), we obtain

$$\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2 = \left[\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\theta_i)]^2\right] - m_i[\bar{X}_i - \mu_X(\theta_i)]^2. \quad (9.19)$$

We now take expectations of the two terms on the right-hand side of the above. First, we have

$$\mathrm{E}\left[\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\Theta_i)]^2\right] = \mathrm{E}\left[\mathrm{E}\left(\sum_{j=1}^{n_i} m_{ij}[X_{ij} - \mu_X(\Theta_i)]^2 \,|\, \Theta_i\right)\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left[\sum_{j=1}^{n_i} m_{ij} \mathrm{Var}(X_{ij} \,|\, \Theta_i)\right] \\
&= \mathrm{E}\left[\sum_{j=1}^{n_i} m_{ij} \left[\frac{\sigma_X^2(\Theta_i)}{m_{ij}}\right]\right] \\
&= \sum_{j=1}^{n_i} \mathrm{E}[\sigma_X^2(\Theta_i)] \\
&= n_i \, \mu_{\mathrm{PV}}, \qquad (9.20)
\end{aligned}
$$

and, noting that $\mathrm{E}(\bar{X}_i \,|\, \Theta_i) = \mu_X(\Theta_i)$, we have

$$
\begin{aligned}
\mathrm{E}\left\{m_i[\bar{X}_i - \mu_X(\Theta_i)]^2\right\} &= m_i \, \mathrm{E}\left[\mathrm{E}\left\{[\bar{X}_i - \mu_X(\Theta_i)]^2 \,|\, \Theta_i\right\}\right] \\
&= m_i \, \mathrm{E}\left[\mathrm{Var}(\bar{X}_i \,|\, \Theta_i)\right] \\
&= m_i \, \mathrm{E}\left[\mathrm{Var}\left(\frac{1}{m_i} \sum_{j=1}^{n_i} m_{ij} X_{ij} \,|\, \Theta_i\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= m_i \, \mathrm{E}\left[\frac{1}{m_i^2}\sum_{j=1}^{n_i} m_{ij}^2 \, \mathrm{Var}(X_{ij} \mid \Theta_i)\right] \\
&= m_i \, \mathrm{E}\left[\frac{1}{m_i^2}\sum_{j=1}^{n_i} m_{ij}^2 \, \left(\frac{\sigma_X^2(\Theta_i)}{m_{ij}}\right)\right] \\
&= \frac{1}{m_i}\sum_{j=1}^{n_i} m_{ij} \, \mathrm{E}[\sigma_X^2(\Theta_i)] \\
&= \mathrm{E}\left[\sigma_X^2(\Theta_i)\right] \\
&= \mu_{\mathrm{PV}}. \qquad (9.21)
\end{aligned}
$$

Combining equations (9.19), (9.20) and (9.21), we conclude that

$$
\mathrm{E}\left[\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2\right] = n_i\mu_{\mathrm{PV}} - \mu_{\mathrm{PV}} = (n_i - 1)\mu_{\mathrm{PV}}. \qquad (9.22)
$$

Thus, taking expectation of equation (9.16), we have

$$\begin{aligned} \mathrm{E}\left(\hat{\mu}_{\mathrm{PV}}\right) &= \frac{\sum_{i=1}^{r} \mathrm{E}\left[\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2\right]}{\sum_{i=1}^{r}(n_i - 1)} \\ &= \frac{\sum_{i=1}^{r}(n_i - 1)\mu_{\mathrm{PV}}}{\sum_{i=1}^{r}(n_i - 1)} \\ &= \mu_{\mathrm{PV}}, \end{aligned} \tag{9.23}$$

so that $\hat{\mu}_{\mathrm{PV}}$ is an unbiased estimator of μ_{PV}. □

Theorem 9.2: The following quantity is an unbiased estimator of σ^2_{HM}

$$\hat{\sigma}^2_{\text{HM}} = \frac{\left[\sum_{i=1}^r m_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\hat{\mu}_{\text{PV}}}{m - \dfrac{1}{m}\sum_{i=1}^r m_i^2}, \tag{9.27}$$

where $\hat{\mu}_{\text{PV}}$ is defined in equation (9.16).

Proof: We begin our proof by expanding the term $\sum_{i=1}^r m_i(\bar{X}_i - \bar{X})^2$ in the numerator of equation (9.27) as follows

$$\begin{aligned}
\sum_{i=1}^r m_i(\bar{X}_i - \bar{X})^2 &= \sum_{i=1}^r m_i\left[(\bar{X}_i - \mu_X) - (\bar{X} - \mu_X)\right]^2 \\
&= \sum_{i=1}^r m_i(\bar{X}_i - \mu_X)^2 + \sum_{i=1}^r m_i(\bar{X} - \mu_X)^2 - 2\sum_{i=1}^r m_i(\bar{X}_i - \mu_X)(\bar{X} - \mu_X) \\
&= \left[\sum_{i=1}^r m_i(\bar{X}_i - \mu_X)^2\right] - m(\bar{X} - \mu_X)^2. \qquad (9.28)
\end{aligned}$$

We then take expectations on both sides of equation (9.28) to obtain

$$\begin{aligned}
\mathrm{E}\left[\sum_{i=1}^{r} m_i(\bar{X}_i - \bar{X})^2\right] &= \left[\sum_{i=1}^{r} m_i \,\mathrm{E}\left[(\bar{X}_i - \mu_X)^2\right]\right] - m\,\mathrm{E}\left[(\bar{X} - \mu_X)^2\right] \\
&= \left[\sum_{i=1}^{r} m_i \mathrm{Var}(\bar{X}_i)\right] - m\mathrm{Var}(\bar{X}). \qquad (9.29)
\end{aligned}$$

As

$$\mathrm{Var}(\bar{X}_i) = \mathrm{Var}\left[\mathrm{E}(\bar{X}_i \,|\, \Theta_i)\right] + \mathrm{E}\left[\mathrm{Var}(\bar{X}_i \,|\, \Theta_i)\right], \qquad (9.30)$$

with

$$\mathrm{Var}(\bar{X}_i \,|\, \Theta_i) = \frac{\sigma_X^2(\Theta_i)}{m_i}. \qquad (9.31)$$

and $\mathrm{E}(\bar{X}_i \,|\, \Theta_i) = \mu_X(\Theta_i)$, equation (9.30) becomes

$$\mathrm{Var}(\bar{X}_i) = \mathrm{Var}\left[\mu_X(\Theta_i)\right] + \frac{\mathrm{E}\left[\sigma_X^2(\Theta_i)\right]}{m_i} = \sigma_{\mathrm{HM}}^2 + \frac{\mu_{\mathrm{PV}}}{m_i}. \qquad (9.32)$$

For $\text{Var}(\bar{X})$ in equation (9.29), we have

$$\begin{aligned}
\text{Var}(\bar{X}) &= \text{Var}\left(\frac{1}{m}\sum_{i=1}^{r} m_i \bar{X}_i\right) \\
&= \frac{1}{m^2}\sum_{i=1}^{r} m_i^2 \text{Var}(\bar{X}_i) \\
&= \frac{1}{m^2}\sum_{i=1}^{r} m_i^2 \left(\sigma^2_{\text{HM}} + \frac{\mu_{\text{PV}}}{m_i}\right) \\
&= \left[\sum_{i=1}^{r} \frac{m_i^2}{m^2}\right] \sigma^2_{\text{HM}} + \frac{\mu_{\text{PV}}}{m}. \qquad (9.33)
\end{aligned}$$

Substituting equations (9.32) and (9.33) into (9.29), we obtain

$$\begin{aligned}
\text{E}\left[\sum_{i=1}^{r} m_i (\bar{X}_i - \bar{X})^2\right] &= \left[\sum_{i=1}^{r} m_i \left(\sigma^2_{\text{HM}} + \frac{\mu_{\text{PV}}}{m_i}\right)\right] - \left[\left(\sum_{i=1}^{r} \frac{m_i^2}{m}\right) \sigma^2_{\text{HM}} + \mu_{\text{PV}}\right] \\
&= \left[m - \frac{1}{m}\sum_{i=1}^{r} m_i^2\right] \sigma^2_{\text{HM}} + (r-1)\mu_{\text{PV}}. \qquad (9.34)
\end{aligned}$$

Thus, taking expectation of $\hat{\sigma}^2_{\mathrm{HM}}$, we can see that

$$\begin{aligned} \mathrm{E}\left(\hat{\sigma}^2_{\mathrm{HM}}\right) &= \frac{\mathrm{E}\left[\sum_{i=1}^r m_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\mathrm{E}(\hat{\mu}_{\mathrm{PV}})}{m - \dfrac{1}{m}\sum_{i=1}^r m_i^2} \\ &= \sigma^2_{\mathrm{HM}}. \end{aligned} \tag{9.35}$$

□

- With estimated values of the model parameters, the Bühlmann-Straub credibility predictor of the ith risk group can be calculated as

$$\hat{Z}_i \bar{X}_i + (1 - \hat{Z}_i)\bar{X}, \tag{9.40}$$

where

$$\hat{Z}_i = \frac{m_i}{m_i + \hat{k}}, \tag{9.41}$$

with

$$\hat{k} = \frac{\hat{\mu}_{\mathrm{PV}}}{\hat{\sigma}^2_{\mathrm{HM}}}. \tag{9.42}$$

- While $\hat{\mu}_{\mathrm{PV}}$ and $\hat{\sigma}^2_{\mathrm{HM}}$ are unbiased estimators of μ_{PV} and σ^2_{HM}, respectively, $\hat{k}$ is not unbiased for k, due to the fact that k is a nonlinear function of μ_{PV} and σ^2_{HM}.

- Note that $\hat{\sigma}^2_{\mathrm{HM}}$ and $\tilde{\sigma}^2_{\mathrm{HM}}$ may be negative in empirical applications. In such circumstances, they may be set to zero, which implies that $\hat{k}$ and $\tilde{k}$ will be infinite, and that $\hat{Z}_i$ and $\tilde{Z}_i$ will be zero for all risk groups.

- The total loss experienced is $m\bar{X} = \sum_{i=1}^{r} m_i\bar{X}_i$. Now if future losses are predicted according to equation (9.40), the total loss predicted will in general be different from the total loss experienced.

- If it is desired to equate the total loss predicted to the total loss experienced, some re-adjustment is needed. This may be done by replacing $\bar{X}$ with $\hat{\mu}_X$ given by

$$\hat{\mu}_X = \frac{\sum_{i=1}^{r} \hat{Z}_i \bar{X}_i}{\sum_{i=1}^{r} \hat{Z}_i}, \tag{9.47}$$

and the loss predicted for the ith group is $\hat{Z}_i \bar{X}_i + (1 - \hat{Z}_i)\hat{\mu}_X$.

Example 9.1: An analyst has data of the claim frequencies of workers compensations of 3 insured companies. Table 9.1 gives the data of company A in the last 3 years and companies B and C in the last four years. The numbers of workers (in hundreds) and the numbers of claims each year per hundred workers are given.

Table 9.1: Data for Example 9.1

Company		Years 1	2	3	4
A	Claims per hundred workers	–	1.2	0.9	1.8
	Workers (in hundreds)	–	10	11	12
B	Claims per hundred workers	0.6	0.8	1.2	1.0
	Workers (in hundreds)	5	5	6	6
C	Claims per hundred workers	0.7	0.9	1.3	1.1
	Workers (in hundreds)	8	8	9	10

Calculate the Bühlmann-Straub credibility predictions of the numbers of claim per hundred workers for the three companies next year, without and with corrections for balancing the total loss with the predicted loss.

Solution: The total exposures of each company are

$$m_A = 10 + 11 + 12 = 33,$$

$$m_B = 5 + 5 + 6 + 6 = 22,$$

and

$$m_C = 8 + 8 + 9 + 10 = 35,$$

which give the total exposures of all companies as $m = 33 + 22 + 35 = 90$. The exposure-weighted means of the claim frequency of the companies are

$$\bar{X}_A = \frac{(10)(1.2) + (11)(0.9) + (12)(1.8)}{33} = 1.3182,$$

$$\bar{X}_B = \frac{(5)(0.6) + (5)(0.8) + (6)(1.2) + (6)(1.0)}{22} = 0.9182,$$

and

$$\bar{X}_C = \frac{(8)(0.7) + (8)(0.9) + (9)(1.3) + (10)(1.1)}{35} = 1.0143.$$

The numerator of $\hat{\mu}_{\rm PV}$ in equation (9.16) is

$$
\begin{aligned}
&(10)(1.2-1.3182)^2+(11)(0.9-1.3182)^2+(12)(1.8-1.3182)^2\\
&+(5)(0.6-0.9182)^2+(5)(0.8-0.9182)^2+(6)(1.2-0.9182)^2+(6)(1.0-0.9182)^2\\
&+(8)(0.7-1.0143)^2+(8)(0.9-1.0143)^2+(9)(1.3-1.0143)^2+(10)(1.1-1.0143)^2\\
&=7.6448.
\end{aligned}
$$

Hence, we have

$$\hat{\mu}_{\rm PV}=\frac{7.6448}{2+3+3}=0.9556.$$

The overall mean is

$$\bar{X}=\frac{(1.3182)(33)+(0.9182)(22)+(1.0143)(35)}{90}=1.1022.$$

The first term in the numerator of $\hat{\sigma}^2_{\rm HM}$ in equation (9.27) is

$$(33)(1.3182-1.1022)^2+(22)(0.9182-1.1022)^2+(35)(1.0143-1.1022)^2=2.5549,$$

and the denominator is

$$90 - \frac{1}{90}[(33)^2 + (22)^2 + (35)^2] = 58.9111,$$

so that

$$\hat{\sigma}^2_{\rm HM} = \frac{2.5549 - (2)(0.9556)}{58.9111} = \frac{0.6437}{58.9111} = 0.0109.$$

Thus, the Bühlmann-Straub credibility parameter estimate is

$$\hat{k} = \frac{\hat{\mu}_{\rm PV}}{\hat{\sigma}^2_{\rm HM}} = \frac{0.9556}{0.0109} = 87.6697,$$

and the Bühlmann-Straub credibility factor estimates of the companies are

$$\hat{Z}_A = \frac{33}{33 + 87.6697} = 0.2735,$$

$$\hat{Z}_B = \frac{22}{22 + 87.6697} = 0.2006,$$

and

$$\hat{Z}_C = \frac{35}{35 + 87.6697} = 0.2853.$$

We then compute the Bühlmann-Straub credibility predictors of the claim frequencies per hundred workers for company A as

$$(0.2735)(1.3182) + (1 - 0.2735)(1.1022) = 1.1613,$$

for company B as

$$(0.2006)(0.9182) + (1 - 0.2006)(1.1022) = 1.0653,$$

and for company C as

$$(0.2853)(1.0143) + (1 - 0.2853)(1.1022) = 1.0771.$$

Note that the total claim frequency predicted based on the historical exposure is

$$(33)(1.1613) + (22)(1.0653) + (35)(1.0771) = 99.4580,$$

which is not equal to the total recorded claim frequency of $(90)(1.1022) = 99.20$. To balance the two figures, we use equation (9.47) to obtain

$$\hat{\mu}_X = \frac{(0.2735)(1.3182) + (0.2006)(0.9182) + (0.2853)(1.0143)}{0.2735 + 0.2006 + 0.2853} = 1.0984.$$

Using this as the credibility complement, we obtain the updated predictors as

$$\begin{aligned}
\text{A} \quad &: \quad (0.2735)(1.3182) + (1 - 0.2735)(1.0984) = 1.1585, \\
\text{B} \quad &: \quad (0.2006)(0.9182) + (1 - 0.2006)(1.0984) = 1.0623, \\
\text{C} \quad &: \quad (0.2853)(1.0143) + (1 - 0.2853)(1.0984) = 1.0744.
\end{aligned}$$

It can be checked that the total claim frequency predicted based on the historical exposure is

$$(33)(1.1585) + (22)(1.0623) + (35)(1.0744) = 99.20,$$

which balances with the total claim frequency recorded. $\square$

9.3 Semiparametric Estimation

- In some applications, researchers may have information about the possible conditional distribution $f_{X_{ij}\,|\,\Theta_i}(x\,|\,\theta_i)$ of the loss variables. For example, claim frequency per exposure may be assumed to be Poisson distributed.

- In contrast, the prior distribution of the risk parameters, which are not observable, are usually best assumed to be unknown.

- Under such circumstances, estimates of the parameters of the Bühlmann-Straub model can be estimated using the semiparametric method.

- Suppose X_{ij} are the claim frequencies per exposure and $X_{ij} \sim \mathcal{P}(\lambda_i)$, for $i = 1, \cdots, r$ and $j = 1, \cdots, n_i$. As $\sigma_X^2(\lambda_i) = \lambda_i$, we have

$$\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_X^2(\Lambda_i)] = \mathrm{E}(\Lambda_i) = \mathrm{E}[\mathrm{E}(X\,|\,\Lambda_i)] = \mathrm{E}(X). \qquad (9.48)$$

Thus, μ_{PV} can be estimated using the overall sample mean of X, $\bar{X}$.

- From (9.27) an alternative estimate of σ^2_{HM} can then be obtained by substituting $\hat{\mu}_{\mathrm{PV}}$ with $\bar{X}$.

9.4 Parametric Estimation

- If the prior distribution of Θ and the conditional distribution of X_{ij} given Θ_i, for $i = 1, \cdots, r$ and $j = 1, \cdots, n_i$ are of known functional forms, then the hyperparameter of Θ, γ, can be estimated using the maximum likelihood estimation (MLE) method.

- The quantities μ_{PV} and σ^2_{HM} are functions of γ, and we denote them by $\mu_{\text{PV}} = \mu_{\text{PV}}(\gamma)$ and $\sigma^2_{\text{HM}} = \sigma^2_{\text{HM}}(\gamma)$.

- As k is a function of μ_{PV} and σ^2_{HM}, the MLE of k can be obtained by replacing γ in $\mu_{\text{PV}} = \mu_{\text{PV}}(\gamma)$ and $\sigma^2_{\text{HM}} = \sigma^2_{\text{HM}}(\gamma)$ by the MLE of γ, $\hat{\gamma}$.

- Specifically, the MLE of k is

$$\hat{k} = \frac{\mu_{\rm PV}(\hat{\gamma})}{\sigma^2_{\rm HM}(\hat{\gamma})}. \tag{9.49}$$

- We now consider the estimation of γ. For simplicity, we assume $m_{ij} \equiv 1$. The marginal pdf of X_{ij} is given by

$$f_{X_{ij}}(x_{ij} \,|\, \gamma) = \int_{\theta_i \in \Omega_\Theta} f_{X_{ij} \,|\, \Theta_i}(x_{ij} \,|\, \theta_i) f_{\Theta_i}(\theta_i \,|\, \gamma) \, d\theta_i. \tag{9.50}$$

- Given the data X_{ij}, for $i = 1, \cdots, r$ and $j = 1, \cdots, n_i$, the likelihood function $L(\gamma)$ is

$$L(\gamma) = \prod_{i=1}^{r} \prod_{j=1}^{n_i} f_{X_{ij}}(x_{ij} \,|\, \gamma), \tag{9.51}$$

and the log-likelihood function is

$$\log[L(\gamma)] = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \log f_{X_{ij}}(x_{ij} \,|\, \gamma). \tag{9.52}$$

- The MLE of γ, $\hat{\gamma}$, is obtained by maximizing $L(\gamma)$ in equation (9.51) or $\log[L(\gamma)]$ in equation (9.52) with respect to γ.

Example 9.5: The claim frequencies X_{ij} are assumed to be Poisson distributed with parameter λ_i, i.e., $X_{ij} \sim \mathcal{PN}(\lambda_i)$. The prior distribution of Λ_i is gamma with hyperparameters α and β, where α is a known constant. Derive the MLE of β and k.

Solution: As α is a known constant, the only hyperparameter of the prior is β. The marginal pf of X_{ij} is

$$\begin{aligned} f_{X_{ij}}(x_{ij} \,|\, \beta) &= \int_0^\infty \left[\frac{\lambda_i^{x_{ij}} \exp(-\lambda_i)}{x_{ij}!}\right] \left[\frac{\lambda_i^{\alpha-1} \exp\left(-\frac{\lambda_i}{\beta}\right)}{\Gamma(\alpha)\beta^\alpha}\right] d\lambda_i \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha x_{ij}!} \int_0^\infty \lambda_i^{x_{ij}+\alpha-1} \exp\left[-\lambda_i\left(\frac{1}{\beta}+1\right)\right] d\lambda_i \end{aligned}$$

$$
\begin{aligned}
&= \frac{\Gamma(x_{ij}+\alpha)}{\Gamma(\alpha)\beta^{\alpha}x_{ij}!}\left(\frac{1}{\beta}+1\right)^{-(x_{ij}+\alpha)} \\
&= \frac{c_{ij}\beta^{x_{ij}}}{(1+\beta)^{x_{ij}+\alpha}},
\end{aligned}
$$

where c_{ij} does not involve β. Thus, the likelihood function is

$$L(\beta) = \prod_{i=1}^{r}\prod_{j=1}^{n_i} \frac{c_{ij}\beta^{x_{ij}}}{(1+\beta)^{x_{ij}+\alpha}},$$

and ignoring the term that does not involve β, the log-likelihood function is

$$\log[L(\beta)] = (\log\beta)\left(\sum_{i=1}^{r}\sum_{j=1}^{n_i} x_{ij}\right) - [\log(1+\beta)]\left[n\alpha + \sum_{i=1}^{r}\sum_{j=1}^{n_i} x_{ij}\right],$$

where $n = \sum_{i=1}^{r} n_i$. The derivative of $\log[L(\beta)]$ with respect to β is

$$\frac{n\bar{x}}{\beta} - \frac{n\,(\alpha+\bar{x})}{1+\beta},$$

where

$$\bar{x} = \frac{1}{n}\left(\sum_{i=1}^{r}\sum_{j=1}^{n_i} x_{ij}\right).$$

The MLE of β, $\hat{\beta}$, is obtained by solving for β when the first derivative of $\log[L(\beta)]$ is set to zero. Hence, we obtain

$$\hat{\beta} = \frac{\bar{x}}{\alpha}.$$

As $X_{ij} \sim \mathcal{PN}(\lambda_i)$ and $\Lambda_i \sim \mathcal{G}(\alpha, \beta)$, $\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_X^2(\Lambda_i)] = \mathrm{E}(\Lambda_i) = \alpha\beta$. Also, $\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mu_X(\Lambda_i)] = \mathrm{Var}(\Lambda_i) = \alpha\beta^2$, so that

$$k = \frac{\alpha\beta}{\alpha\beta^2} = \frac{1}{\beta}.$$

Thus, the MLE of k is

$$\hat{k} = \frac{1}{\hat{\beta}} = \frac{\alpha}{\bar{x}}.$$

□