Nonlife Actuarial Models

Chapter 7

Bühlmann Credibility

Learning Objectives

1. Basic framework of Bühlmann credibility

2. Variance decomposition

3. Expected value of the process variance

4. Variance of the hypothetical mean

5. Bühlmann credibility

6. Bühlmann-Straub credibility

7.1 Framework and Notations

- Consider a risk group or block of insurance policies with loss measure denoted by X, which may be claim frequency, claim severity, aggregate loss or pure premium.

- Assume that the risk profiles of the group are characterized by a parameter θ, which determines the distribution of the loss measure X.

- Denote the conditional mean and variance of X given θ by

$$\mathrm{E}(X \mid \theta) = \mu_X(\theta), \tag{7.1}$$

and

$$\mathrm{Var}(X \mid \theta) = \sigma_X^2(\theta). \tag{7.2}$$

- Assume that the insurance company has similar blocks of policies with different risk profiles.

- Thus, the parameter θ varies with different risk groups.

- We treat θ as the realization of a random variable Θ, the distribution of which is called the **prior distribution**.

- When θ varies over the support of Θ, the conditional mean and variance of X become random variables in Θ, and are denoted by $\mu_X(\Theta) = \mathrm{E}(X \mid \Theta)$ and $\sigma_X^2(\Theta) = \mathrm{Var}(X \mid \Theta)$, respectively.

Example 7.1: An insurance company has blocks of workers compensation policies. The claim frequency is known to be Poisson with parameter λ, where λ is 20 for low-risk group and 50 for high-risk group. Suppose

30% of the risk groups are low risk and 70% are high risk. What are the conditional mean and variance of the claim frequency?

Solution: The parameter determining the claim frequency X is λ, which we assume is a realization of the random variable Λ. As X is Poisson, the conditional mean and conditional variance of X are equal to λ. Thus, we have the following results

λ	$\Pr(\Lambda = \lambda)$	$\mathrm{E}(X \mid \lambda)$	$\mathrm{Var}(X \mid \lambda)$
20	0.3	20	20
50	0.7	50	50

so that

$$\mu_X(\Lambda) = \mathrm{E}(X \mid \Lambda) = \begin{cases} 20, & \text{with probability } 0.30, \\ \\ 50, & \text{with probability } 0.70. \end{cases}$$

Likewise, we have

$$\sigma_X^2(\Lambda) = \mathrm{Var}(X \,|\, \Lambda) = \begin{cases} 20, & \text{with probability } 0.30, \\ 50, & \text{with probability } 0.70. \end{cases}$$

Example 7.2: The claim severity X of a block of health insurance policies is normally distributed with mean θ and variance 10. If θ takes values within the interval [100, 200] and follows a uniform distribution, what are the conditional mean and conditional variance of X?

Solution: The conditional variance of X is 10, irrespective of θ. Hence, we have $\sigma_X^2(\Theta) = \mathrm{Var}(X \,|\, \Theta) = 10$ with probability 1. The conditional mean of X is Θ, i.e., $\mu_X(\Theta) = \mathrm{E}(X \,|\, \Theta) = \Theta$, which is uniformly distributed in [100, 200] with pdf

$$f_\Theta(\theta) = \begin{cases} 0.01, & \text{for } \theta \in [100, 200], \\ 0, & \text{otherwise.} \end{cases}$$

- The Bühlmann model assumes that there are n observations of losses, denoted by $\{X_1, \cdots, X_n\}$.

- The observations may be losses recorded in n periods and they are assumed to be iid as X, which depends on the parameter θ.

- The task is to update the prediction of X for the next period, i.e., X_{n+1}, based on $\{X_1, \cdots, X_n\}$.

- In the Bühlmann approach the solution depends on the variation between the conditional means as well as the average of the conditional variances of the risk groups.

7.2 Variance Components

- The variation of the loss measure X consists of two components: the **variation between risk groups** and the **variation within risk groups**.

- The first component, variation between risk groups, is due to the randomness of the risk profiles of each group and is captured by the parameter Θ.

- The second component, variation within risk group, is measured by the conditional variance of the risk group.

- We first consider the calculation of the overall mean of the loss measure X. The **unconditional mean** (or **overall mean**) of X is

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X \mid \Theta)] = \mathrm{E}[\mu_X(\Theta)]. \tag{7.3}$$

- For the **unconditional variance** (or **total variance**), we have

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X \mid \Theta)] + \mathrm{Var}[\mathrm{E}(X \mid \Theta)]. \tag{7.4}$$

- $\mathrm{Var}(X \mid \Theta)$ measures the variance of a given risk group. It is a function of the random variable Θ and we call this the **process variance**.

- Thus, $\mathrm{E}[\mathrm{Var}(X \mid \Theta)]$ is the **expected value of the process variance** (**EPV**).

- On the other hand, $\mathrm{E}(X \mid \Theta)$ is the mean of a given risk group. We call this conditional mean the **hypothetical mean**.

- Thus, $\mathrm{Var}[\mathrm{E}(X \mid \Theta)]$ is the **variance of the hypothetical means** (**VHM**), as it measures the variations in the *means* of the risk groups.

- Verbally, equation (7.4) can be written as

$$\begin{aligned}\text{Total variance} &= \text{Expected value of process variance} \\ &\quad + \text{Variance of hypothetical means},\end{aligned} \tag{7.5}$$

or

$$\text{Total variance} = \text{EPV} + \text{VHM}. \tag{7.6}$$

- It can also be stated alternatively as

$$\begin{aligned}\text{Total variance} &= \text{Mean of conditional variance} \\ &\quad +\text{Variance of conditional mean}.\end{aligned} \tag{7.7}$$

- Symbolically, we use the following notations

$$\mathrm{E}[\mathrm{Var}(X \,|\, \Theta)] = \mathrm{E}[\sigma_X^2(\Theta)] = \mu_{\mathrm{PV}}, \tag{7.8}$$

and

$$\mathrm{Var}[\mathrm{E}(X \,|\, \Theta)] = \mathrm{Var}[\mu_X(\Theta)] = \sigma^2_{\mathrm{HM}}, \tag{7.9}$$

so that equation (7.4) can be written as

$$\mathrm{Var}(X) = \mu_{\mathrm{PV}} + \sigma^2_{\mathrm{HM}}. \tag{7.10}$$

Example 7.3: For Examples 7.1 and 7.2, calculate the unconditional mean, the expected value of the process variance, the variance of the hypothetical means and the total variance.

Solution: For Example 7.1, the unconditional mean is

$$\begin{aligned}
\mathrm{E}(X) &= \Pr(\Lambda = 20)\mathrm{E}(X \,|\, \Lambda = 20) + \Pr(\Lambda = 50)\mathrm{E}(X \,|\, \Lambda = 50) \\
&= (0.3)(20) + (0.7)(50) \\
&= 41.
\end{aligned}$$

The expected value of the process variance, EPV, is

$$\begin{aligned} \mathrm{E}[\mathrm{Var}(X \,|\, \Lambda)] &= \Pr(\Lambda = 20)\mathrm{Var}(X \,|\, \Lambda = 20) + \Pr(\Lambda = 50)\mathrm{Var}(X \,|\, \Lambda = 50) \\ &= (0.3)(20) + (0.7)(50) \\ &= 41. \end{aligned}$$

As the mean of the hypothetical means (i.e., the unconditional mean) is 41, the variance of the hypothetical means, VHM, is

$$\mathrm{Var}[\mathrm{E}(X \,|\, \Lambda)] = (0.3)(20 - 41)^2 + (0.7)(50 - 41)^2 = 189.$$

Thus, the total variance of X is

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X \,|\, \Lambda)] + \mathrm{Var}[\mathrm{E}(X \,|\, \Lambda)] = 41 + 189 = 230.$$

For Example 7.2, as Θ is uniformly distributed in [100, 200], the unconditional mean of X is

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X \,|\, \Theta)] = \mathrm{E}(\Theta) = 150.$$

As X has a constant variance of 10, the expected value of the process variance is

$$\mathrm{E}[\mathrm{Var}(X \mid \Theta)] = \mathrm{E}(10) = 10.$$

The variance of the hypothetical means is

$$\mathrm{Var}[\mathrm{E}(X \mid \Theta)] = \mathrm{Var}(\Theta) = \frac{(200 - 100)^2}{12} = 833.33,$$

and the total variance of X is

$$\mathrm{Var}(X) = 10 + 833.33 = 843.33.$$

□

Example 7.5: An insurance company sells workers compensation policies, each of which belongs to one of three possible risk groups. The risk groups have claim frequencies N that are Poisson distributed with parameter λ and claim severity X that are gamma distributed with parameters

α and β. Claim frequency and claim severity are independently distributed given a risk group, and the aggregate loss is S. The data of the risk groups are given in Table 7.2.

Table 7.2: Data for Example 7.5

Risk group	Relative frequency	Distribution of N: $\mathcal{PN}(\lambda)$	Distribution of X: $\mathcal{G}(\alpha, \beta)$
1	0.2	$\lambda = 20$	$\alpha = 5,\ \beta = 2$
2	0.4	$\lambda = 30$	$\alpha = 4,\ \beta = 3$
3	0.4	$\lambda = 40$	$\alpha = 3,\ \beta = 2$

For each of the following loss measures: (a) claim frequency N, (b) claim severity X, and (c) aggregate loss S, calculate EPV, VHM and the total variance.

Solution: **(a) Claim frequency** We first calculate the conditional mean and conditional variance of N given the risk group, which is characterized by the parameter Λ. As N is Poisson, the mean and variance are equal to Λ, so that we have the results in Table 7.3.

Table 7.3: Results for Example 7.5 (a)

Risk group	Probability	$\mathrm{E}(N \mid \Lambda) = \mu_N(\Lambda)$	$\mathrm{Var}(N \mid \Lambda) = \sigma_N^2(\Lambda)$
1	0.2	20	20
2	0.4	30	30
3	0.4	40	40

Thus, the EPV is

$$\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(N \mid \Lambda)] = (0.2)(20) + (0.4)(30) + (0.4)(40) = 32,$$

which is also equal to the unconditional mean $\mathrm{E}[\mu_N(\Lambda)]$. For VHM, we first calculate

$$\mathrm{E}\{[\mu_N(\Lambda)]^2\} = (0.2)(20)^2 + (0.4)(30)^2 + (0.4)(40)^2 = 1{,}080,$$

so that

$$\sigma_{\mathrm{HM}}^2 = \mathrm{Var}[\mu_N(\Lambda)] = \mathrm{E}\{[\mu_N(\Lambda)]^2\} - \{\mathrm{E}[\mu_N(\Lambda)]\}^2 = 1{,}080 - (32)^2 = 56.$$

Therefore, the total variance of N is

$$\mathrm{Var}(N) = \mu_{\mathrm{PV}} + \sigma_{\mathrm{HM}}^2 = 32 + 56 = 88.$$

(b) Claim severity There are three claim-severity distributions, which are specific to each risk group. Note that the relative frequencies of the risk groups as well as the claim frequencies in the risk groups jointly determine the relative occurrence of each claim-severity distribution. The probabilities of occurrence of the severity distributions, as well as their conditional means and variances are given in Table 7.4, in which Γ denotes the vector random variable representing α and β.

Table 7.4: Results for Example 7.5 (b)

Group	Group probability	λ	Col 2 $\times$ Col 3	Probability of severity X	$\mathrm{E}(X \mid \Gamma)$ $= \mu_X(\Gamma)$	$\mathrm{Var}(X \mid \Gamma)$ $= \sigma_X^2(\Gamma)$
1	0.2	20	4	0.125	10	20
2	0.4	30	12	0.375	12	36
3	0.4	40	16	0.500	6	12

Column 4 gives the expected number of claims in each group weighted by the group probability. Column 5 gives the probability of occurrence of each type of claim-severity distribution, which is obtained by dividing the corresponding figure in Column 4 by the sum of Column 4 (e.g., $0.125 = 4/(4+12+16)$). The last two columns give the conditional mean $\alpha\beta$ and conditional variance $\alpha\beta^2$ corresponding to the three different distributions of claim severity. Similar to the calculation in (a), we have

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X \mid \Gamma)] = (0.125)(10) + (0.375)(12) + (0.5)(6) = 8.75,$$

and

$$\mu_{\mathrm{PV}} = (0.125)(20) + (0.375)(36) + (0.5)(12) = 22.$$

To calculate VHM, we first compute the raw second moment of the conditional mean of X, which is

$$\mathrm{E}\{[\mu_X(\Gamma)]^2\} = (0.125)(10)^2 + (0.375)(12)^2 + (0.5)(6)^2 = 84.50.$$

Hence,

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mu_X(\Gamma)] = \mathrm{E}\{[\mu_X(\Gamma)]^2\} - \{\mathrm{E}[\mu_X(\Gamma)]\}^2 = 84.50 - (8.75)^2 = 7.9375.$$

Therefore, the total variance of X is

$$\mathrm{Var}(X) = \mu_{\mathrm{PV}} + \sigma^2_{\mathrm{HM}} = 22 + 7.9375 = 29.9375.$$

(c) Aggregate loss The distribution of the aggregate loss S is determined jointly by Λ and Γ, which we shall denote as Θ. For the conditional mean of S, we have

$$\mathrm{E}(S \mid \Theta) = \mathrm{E}(N \mid \Theta)\mathrm{E}(X \mid \Theta) = \lambda\alpha\beta.$$

For the conditional variance of S, we use the result on compound distribution with Poisson claim frequency stated in equation (A.123), and make use of the assumption of gamma severity to obtain

$$\mathrm{Var}(S \mid \Theta) = \lambda[\sigma_X^2(\Gamma) + \mu_X^2(\Gamma)] = \lambda(\alpha\beta^2 + \alpha^2\beta^2).$$

The conditional means and conditional variances of S are summarized in Table 7.5.

Table 7.5: Results for Example 7.5 (c)

Group	Group probability	Parameters λ, α, β	$\mathrm{E}(S \mid \Theta)$ $= \mu_S(\Theta)$	$\mathrm{Var}(S \mid \Theta)$ $= \sigma_S^2(\Theta)$
1	0.2	20, 5, 2	200	2,400
2	0.4	30, 4, 3	360	5,400
3	0.4	40, 3, 2	240	1,920

The unconditional mean of S is

$$\mathrm{E}(S) = \mathrm{E}[\mathrm{E}(S \mid \Theta)] = (0.2)(200) + (0.4)(360) + (0.4)(240) = 280,$$

and the EPV is

$$\mu_{\mathrm{PV}} = (0.2)(2{,}400) + (0.4)(5{,}400) + (0.4)(1{,}920) = 3{,}408.$$

Also, the VHM is given by

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mu_S(\Theta)]$$

$$
\begin{aligned}
&= \mathrm{E}\{[\mu_S(\Theta)]^2\} - \{\mathrm{E}[\mu_S(\Theta)]\}^2 \\
&= \left[(0.2)(200)^2 + (0.4)(360)^2 + (0.4)(240)^2\right] - (280)^2 \\
&= 4{,}480.
\end{aligned}
$$

Therefore, the total variance of S is

$$
\mathrm{Var}(S) = 3{,}408 + 4{,}480 = 7{,}888.
$$

□

- EPV and VHM measure two different aspects of the total variance.
- When a risk group is homogeneous so that the loss claims are similar within the group, the conditional variance is small. If all risk groups have similar loss claims within the group, the expected value of the process variance is small.

- On the other hand, if the risk groups have very different risk profiles across groups, their hypothetical means will differ more and thus the variance of the hypothetical means will be large.

- Thus, it will be easier to distinguish between risk groups if the hypothetical means differ more and the average of the process variance is small.

- We define k as the ratio of EPV to VHM, i.e.,

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{\mathrm{EPV}}{\mathrm{VHM}}. \tag{7.11}$$

- A small EPV or large VHM will gives rise to a small k. The risk groups will be more *distinguishable* in the mean when k is smaller, in which case we may put more weight on the data in updating our revised prediction for future losses.

Example 7.6: Frequency of claim per year, N, is distributed as a Binomial random variable $\mathcal{BN}(10, \theta)$, and claim severity, X, is distributed as an exponential random variable with mean $c\theta$, where c is a known constant. Given θ, claim frequency and claim severity are independently distributed. Derive an expression of k for the aggregate loss per year, S, in terms of c and the moments of Θ, and show that it does not depend on c. If Θ is 0.3 or 0.7 with equal probabilities, calculate k.

Solution: We first calculate the conditional mean of S as a function of θ. Due to the independence assumption of N and X, the hypothetical mean of S is

$$\mathrm{E}(S \,|\, \Theta) = \mathrm{E}(N \,|\, \Theta)\mathrm{E}(X \,|\, \Theta) = (10\Theta)(c\Theta) = 10c\Theta^2.$$

The process variance is

$$\mathrm{Var}(S \,|\, \Theta) \quad = \quad \mu_N(\Theta)\sigma_X^2(\Theta) + \sigma_N^2(\Theta)\mu_X^2(\Theta)$$

$$
\begin{aligned}
&= (10\Theta)(c\Theta)^2 + [10\Theta(1-\Theta)](c\Theta)^2 \\
&= 10c^2\Theta^3 + 10c^2\Theta^3(1-\Theta) \\
&= 10c^2\Theta^3(2-\Theta).
\end{aligned}
$$

Hence, the unconditional mean of S is

$$
\mathrm{E}(S) = \mathrm{E}[\mathrm{E}(S \,|\, \Theta)] = \mathrm{E}(10c\Theta^2) = 10c\mathrm{E}(\Theta^2)
$$

and the variance of the hypothetical means is

$$
\begin{aligned}
\sigma^2_{\mathrm{HM}} &= \mathrm{Var}[\mathrm{E}(S \,|\, \Theta)] \\
&= \mathrm{Var}(10c\Theta^2) \\
&= 100c^2\mathrm{Var}(\Theta^2) \\
&= 100c^2\{\mathrm{E}(\Theta^4) - [\mathrm{E}(\Theta^2)]^2\}.
\end{aligned}
$$

The expected value of the process variance is

$$
\mu_{\mathrm{PV}} = \mathrm{E}[\mathrm{Var}(S \,|\, \Theta)]
$$

$$
\begin{aligned}
&= \mathrm{E}[10c^2\Theta^3(2-\Theta)] \\
&= 10c^2[2\mathrm{E}(\Theta^3) - \mathrm{E}(\Theta^4)].
\end{aligned}
$$

Combining the above results we conclude that

$$
k = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}} = \frac{10c^2[2\mathrm{E}(\Theta^3) - \mathrm{E}(\Theta^4)]}{100c^2\{\mathrm{E}(\Theta^4) - [\mathrm{E}(\Theta^2)]^2\}} = \frac{2\mathrm{E}(\Theta^3) - \mathrm{E}(\Theta^4)}{10\{\mathrm{E}(\Theta^4) - [\mathrm{E}(\Theta^2)]^2\}}.
$$

Thus, k does not depend on c. To compute its value for the given distribution of Θ, we present the calculations as follows:

θ	$\Pr(\Theta = \theta)$	θ^2	θ^3	θ^4
0.3	0.5	0.09	0.027	0.0081
0.7	0.5	0.49	0.343	0.2401

Thus, the required moments of Θ are

$$
\mathrm{E}(\Theta) = (0.5)(0.3) + (0.5)(0.7) = 0.5,
$$

$$\mathrm{E}(\Theta^2) = (0.5)(0.09) + (0.5)(0.49) = 0.29,$$

$$\mathrm{E}(\Theta^3) = (0.5)(0.027) + (0.5)(0.343) = 0.185$$

and

$$\mathrm{E}(\Theta^4) = (0.5)(0.0081) + (0.5)(0.2401) = 0.1241,$$

so that

$$k = \frac{2(0.185) - 0.1241}{10\left[0.1241 - (0.29)^2\right]} = 0.6148.$$

□

In this example, note that both EPV and VHM depend on c. However, as the effects of c on these components are the same, the ratio of EPV to

VHM is invariant to c. Also, though X and N are independent *given* θ, they are correlated *unconditionally* due to their common dependence on Θ.

7.3 Bühlmann Credibility

- Bühlmann's approach of updating the predicted loss measure is based on a linear predictor using past observations.

- It is also called the **greatest accuracy approach** or the **least squares approach**.

- For the classical credibility approach, the updated prediction U is given by (see equation (6.1))

$$U = ZD + (1 - Z)M. \tag{7.12}$$

- The Bühlmann credibility method has a similar basic equation, in which D is the sample mean of the data and M is the overall prior mean $\mathrm{E}(X)$.

- The Bühlmann credibility factor Z depends on the sample size n and the EPV to VHM ratio k. In particular, Z varies with n and k as follows:

1. Z increases with the sample size n of the data.

2. Z increases with the *distinctiveness* of the risk groups. As argued above, the risk groups are more distinguishable when k is small. Thus, Z increases as k decreases.

- We now state formally the assumptions of the Bühlmann model and derive the updating formula as the **least mean-squared-error (MSE) linear predictor**.

1. $\{X_1, \cdots, X_n\}$ are loss measures that are independently and identically distributed as the random variable X. The distribution of X depends on the parameter θ.

2. The parameter θ is a realization of a random variable Θ. Given θ, the conditional mean and variance of X are

$$\mathrm{E}(X \mid \theta) = \mu_X(\theta), \tag{7.13}$$

and

$$\mathrm{Var}(X \mid \theta) = \sigma_X^2(\theta). \tag{7.14}$$

3. The unconditional mean of X is $\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X \mid \Theta)] = \mu_X$. The mean of the conditional variance of X is

$$\begin{aligned} \mathrm{E}[\mathrm{Var}(X \mid \Theta)] &= \mathrm{E}[\sigma_X^2(\Theta)] \\ &= \mu_{\mathrm{PV}} \\ &= \text{Expected value of process variance} \\ &= \mathrm{EPV}, \end{aligned} \tag{7.15}$$

and the variance of the conditional mean is

$$\begin{aligned}
\mathrm{Var}[\mathrm{E}(X \mid \Theta)] &= \mathrm{Var}[\mu_X(\Theta)] \\
&= \sigma^2_{\mathrm{HM}} \\
&= \text{Variance of hypothetical means} \\
&= \mathrm{VHM}. \qquad (7.16)
\end{aligned}$$

The unconditional variance (or total variance) of X is

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}[\mathrm{Var}(X \mid \Theta)] + \mathrm{Var}[\mathrm{E}(X \mid \Theta)] \\
&= \mu_{\mathrm{PV}} + \sigma^2_{\mathrm{HM}} \\
&= \mathrm{EPV} + \mathrm{VHM}. \qquad (7.17)
\end{aligned}$$

4. The Bühlmann approach formulates a predictor of X_{n+1} based on a linear function of $\{X_1, \cdots, X_n\}$, where X_{n+1} is assumed to have the

same distribution as X. The predictor minimizes the mean squared error in predicting X_{n+1} over the joint distribution of Θ, X_{n+1} and $\{X_1, \cdots, X_n\}$. Specifically, the predictor is given by

$$\hat{X}_{n+1} = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n, \tag{7.18}$$

where $\beta_0, \beta_1, \cdots, \beta_n$ are chosen to minimize the mean squared error, MSE, defined as

$$\text{MSE} = \text{E}\left[\left(X_{n+1} - \hat{X}_{n+1}\right)^2\right]. \tag{7.19}$$

- We skip the proof and state the result here that

$$\hat{X}_{n+1} = Z\bar{X} + (1 - Z)\mu_X, \tag{7.34}$$

where

$$Z = \frac{n}{n + k}. \tag{7.35}$$

- Z defined in equation (7.35) is called the **Bühlmann credibility factor** or simply the **Bühlmann credibility**. It depends on the EPV to VHM ratio k, which is called the **Bühlmann credibility parameter**.

- The optimal linear forecast $\hat{X}_{n+1}$ given in equation (7.34) is also called the **Bühlmann premium**.

- Note that k depends only on the parameters of the model, while Z is a function of k and the size n of the *data*.

- For predicting claim frequency N, the sample size n is the number of periods over which the number of claims is aggregated. For predicting claim severity X, the sample size n is the number of claims. As aggregate loss S refers to the total loss payout per period, the sample size is the number of periods of claim experience.

Example 7.7: Refer to Example 7.5. Suppose the claim experience last year was 26 claims with an average claim size of 12. Calculate the updated prediction of (a) the claim frequency, (b) the average claim size, and (c) the aggregate loss, for next year.

Solution: **(a) Claim frequency** From Example 7.5, we have $k = 0.5714$ and $M = \mathrm{E}(N) = 32$. Now we are given $n = 1$ and $D = 26$. Hence,

$$Z = \frac{1}{1 + 0.5714} = 0.6364,$$

so that the updated prediction of the claim frequency of this group is

$$U = (0.6364)(26) + (1 - 0.6364)(32) = 28.1816.$$

(b) Claim severity We have $k = 2.7717$ and $M = \mathrm{E}(X) = 8.75$, with

$n = 26$ and $D = 12$. Thus,

$$Z = \frac{26}{26 + 2.7717} = 0.9037,$$

so that the updated prediction of the claim severity of this group is

$$U = (0.9037)(12) + (1 - 0.9037)(8.75) = 11.6870.$$

(c) Aggregate loss With $k = 0.7607$, $M = \mathrm{E}(S) = 280$, $n = 1$ and $D = (26)(12) = 312$, we have

$$Z = \frac{1}{1 + 0.7607} = 0.5680,$$

so that the updated prediction of the aggregate loss of this group is

$$U = (0.5680)(312) + (1 - 0.5680)(280) = 298.1760.$$

7.4 Bühlmann-Straub Credibility

- An important limitation of the Bühlmann credibility theory is that the loss observations X_i are assumed to be *identically* distributed. This assumption is violated if the data are over different periods with different exposures (the definition of exposure will be explained below).

- The **Bühlmann-Straub credibility model** extends the **Bühlmann theory** to cases where the loss data X_i are not identically distributed. In particular, the process variance of the loss measure is assumed to depend on the exposure.

- We denote the exposure by m_i, and the *loss per unit of exposure* by X_i. Note that the exposure needs not be the number of insureds, although it may often be the case. We then assume the following for the conditional variance of X_i

$$\mathrm{Var}(X_i \,|\, \Theta) = \frac{\sigma_X^2(\Theta)}{m_i}, \tag{7.36}$$

for a suitably defined $\sigma_X^2(\Theta)$.

- The following are some examples.

1. X_i is the average number of claims per insured in year i, $\sigma_X^2(\Theta)$ is the variance of the claim frequency of an insured, and the exposure m_i is the number of insureds covered in year i.

2. X_i is the average aggregate loss per month of the ith block of policies, $\sigma_X^2(\Theta)$ is the variance of the aggregate loss of the block in a month, and the exposure m_i is the number of months of insurance claims for the ith block of policies.

3. X_i is the average loss per unit premium in year i, $\sigma_X^2(\Theta)$ is the variance of the claim amount of an insured per year divided by the premium per insured, and the exposure m_i is the amount of premiums received in year i.

- The parameter θ is a realization of a random variable Θ. Given θ, the conditional mean and variance of X_i are

$$\mathrm{E}(X_i \,|\, \theta) = \mu_X(\theta), \tag{7.38}$$

and

$$\mathrm{Var}(X_i \,|\, \theta) = \frac{\sigma_X^2(\theta)}{m_i}, \tag{7.39}$$

for $i \in \{1, \cdots, n\}$, where $\sigma_X^2(\theta)$ is suitably defined as in the examples above.

- The unconditional mean of X_i is $\mathrm{E}(X_i) = \mathrm{E}[\mathrm{E}(X_i \,|\, \Theta)] = \mathrm{E}[\mu_X(\Theta)] = \mu_X$. The mean of the conditional variance of X_i is

$$\begin{aligned} \mathrm{E}[\mathrm{Var}(X_i \,|\, \Theta)] &= \mathrm{E}\left[\frac{\sigma_X^2(\Theta)}{m_i}\right] \\ &= \frac{\mu_{\mathrm{PV}}}{m_i}, \end{aligned} \tag{7.40}$$

for $i \in \{1, \cdots, n\}$, where $\mu_{\mathrm{PV}} = \mathrm{E}[\sigma_X^2(\Theta)]$, and the variance of its conditional mean is

$$\begin{aligned} \mathrm{Var}[\mathrm{E}(X_i \,|\, \Theta)] &= \mathrm{Var}[\mu_X(\Theta)] \\ &= \sigma_{\mathrm{HM}}^2. \end{aligned} \tag{7.41}$$

- Now we define

$$m = \sum_{i=1}^{n} m_i, \tag{7.48}$$

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{n} m_i X_i \tag{7.51}$$

and

$$k = \frac{\mu_{\mathrm{PV}}}{\sigma^2_{\mathrm{HM}}}. \tag{7.52}$$

- Denoting

$$Z = \frac{m}{m+k}. \tag{7.54}$$

we have

$$\hat{X}_{n+1} = Z\bar{X} + (1-Z)\mu_X. \tag{7.57}$$

Example 7.9: The number of accident claims incurred per year for each insured is distributed as a binomial random variable $\mathcal{BN}(2, \theta)$, and the claim incidences are independent across insureds. The probability θ of the binomial has a beta distribution with parameters $\alpha = 1$ and $\beta = 10$. The data in Table 7.7 are given for a block of policies.

Table 7.7: Data for Example 7.9

Year	Number of insureds	Number of claims
1	100	7
2	200	13
3	250	18
4	280	–

Calculate the Bühlmann-Straub credibility prediction of the number of claims in the fourth year.

Solution: Let m_i be the number of insureds in Year i, and X_i be the number of claims per insured in Year i. Define X_{ij} as the number of claims for the jth insured in Year i, which is distributed as $\mathcal{BN}(2, \theta)$. Thus, we have

$$\mathrm{E}(X_i \,|\, \Theta) = \frac{1}{m_i} \sum_{j=1}^{m_i} \mathrm{E}(X_{ij} \,|\, \Theta) = 2\Theta,$$

and

$$\sigma^2_{\mathrm{HM}} = \mathrm{Var}[\mathrm{E}(X_i \,|\, \Theta)] = \mathrm{Var}(2\Theta) = 4\mathrm{Var}(\Theta).$$

As Θ has a beta distribution with parameters $\alpha = 1$ and $\beta = 10$, we have

$$\mathrm{Var}(\Theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{10}{(11)^2(12)} = 0.006887.$$

For the conditional variance of X_i, we have

$$\mathrm{Var}(X_i \,|\, \Theta) = \frac{2\Theta(1-\Theta)}{m_i}.$$

Thus,

$$\mu_{\mathrm{PV}} = 2\mathrm{E}[\Theta(1-\Theta)].$$

As

$$\mathrm{E}(\Theta) = \frac{\alpha}{\alpha+\beta} = 0.0909,$$

we have

$$\begin{aligned}\mu_{\text{PV}} &= 2[\text{E}(\Theta) - \text{E}(\Theta^2)] \\ &= 2\left\{\text{E}(\Theta) - \left(\text{Var}(\Theta) + [\text{E}(\Theta)]^2\right)\right\} \\ &= 2\{0.0909 - [0.006887 + (0.0909)^2]\} = 0.1515.\end{aligned}$$

Thus,

$$k = \frac{\mu_{\text{PV}}}{\sigma^2_{\text{HM}}} = \frac{0.1515}{(4)(0.006887)} = 5.5.$$

As $m = 100 + 200 + 250 = 550$, we have

$$Z = \frac{550}{550 + 5.5} = 0.9901.$$

Now

$$\mu_X = \text{E}[\text{E}(X_i \,|\, \Theta)] = (2)(0.0909) = 0.1818$$

and

$$\bar{X} = \frac{7+13+18}{550} = 0.0691.$$

Thus, the predicted number of claims per insured is

$$(0.9901)(0.0691) + (1-0.9901)(0.1818) = 0.0702,$$

and the predicted number of claims in Year 4 is

$$(280)(0.0702) = 19.66.$$

□