Bühlmann credibility model with correlated risk parameters

Angelica, S Nurrohmah and I Fithriani
Department of Mathematics, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia

Corresponding author’s email: snurrohmah@sci.ui.ac.id

Abstract. Credibility theory is one of the tools to predict the amount of future claims by combining the experience of the claims in the past of a particular policyholder and external information, which is called manual rate, obtained from the experience of a large group of policyholders. One of the credibility theory that is widely used is Bühlmann credibility, which accommodates the heterogeneity of risk exposures, noted by unique risk parameters for each individual. However, Bühlmann credibility requires an assumption that the risk parameters are independent, which almost surely cannot be fulfilled by individuals living in the same area. Therefore, the Bühlmann credibility estimator with correlated risk parameters is formed by utilizing the orthogonal projection in Hilbert space. Also, the parameters included in the model are estimated. In addition, the standard Bühlmann credibility estimator and the Bühlmann credibility estimator assuming correlated risk parameters are compared to predict the amount of future claims based on data from a life insurance company. Comparing the root mean square error, the credibility estimator with correlated risk parameters is better in predicting the amount of future claim, meaning that the predicted amount of claim is closer to the original amount of claim. Also, as the correlation increases, the root mean square error becomes smaller. Moreover, the credibility estimator applied to the data that is partitioned based on the amount of past claims shows better performance than when applied to unpartitioned data.

Keywords: Bühlmann credibility, correlation, Hilbert space, orthogonal projection, risk parameter

1. Introduction
In daily activities, individuals face various risks. Risk is defined as danger, a result that could happen as a consequence of a process that is taking place or future events [1]. Insurance came to answer the worry of risk exposure by bearing individuals’ own risks. Thus, when an individual experiences loss, insurance company will compensate it. In return, individuals, often referred to as policyholders, must pay a certain amount of money regularly, which is called a premium.

In order to maintain the operation of the insurance company, it is useful to be able to predict the amount of claims in the future period. Prediction of claim amount is benefitting the company to be able to set the amount of premium precisely so that it will be able to pay out the benefits for the policyholders.

One of the tools to quantify the amount of premium that is suitable for each policyholder is credibility theory [2]. The idea of this theory is to combine the experience of past claims of a particular policyholder with manual rate, which is obtained from others’ claim experience.

Currently, one of the credibility theory that is frequently used is Bühlmann credibility. This theory covers the heterogeneity of risks for each individual, denoted by risk parameters. This is suitable for a
real life case where each individual is most likely to have different exposure of risks. However, Bühlmann credibility states that the risk parameters are independent, meaning that the risk that one individual is exposed to is not related to other individuals’ risks. This assumption sometimes cannot be fulfilled, for example for people living in the same area. It is shown in research held by Knuiman that spouses have positive correlated risks regarding the probability of getting cardiovascular illness [3].

In this paper, we propose Bühlmann credibility estimator assuming correlated risk parameters. By applying this estimator to a group of individuals living nearby, an insurance company can more accurately predict claims. Accordingly, more realistic premiums and reserves can be chosen.

2. Materials and method

2.1. Bühlmann credibility

Bühlmann credibility is a newer development of credibility premium which predicts claims by using the linear functions of past claims. First, we assume that for a particular individual (i), the claim \(X_{i(j)}\) has the same conditional mean for each period and also the same conditional variance for each period. This assumption can be written as:

\[
\mu(\theta_i) = E(X_{i(j)}|\theta_i), \quad \sigma^2(\theta_i) = Var(X_{i(j)}|\theta_i), \quad i = 1, 2, ..., k, j = 1, 2, ..., n
\]

where index \(i\) shows the individual and \(j\) shows the period. Henceforth, \(\mu(\theta_i)\) will be called the hypothetical mean and \(\sigma^2(\theta_i)\) is the process variance. In this standard Bühlmann credibility, it is also assumed that the risk parameters, denoted by \(\theta_i\) are independent across individuals. Here, we wish to predict the individual’s hypothetical mean, denoted with \(\mu_{n+1}(\theta_i)\). Index \(n + 1\) shows the next period in the future. Prediction of the hypothetical mean can be written as

\[
\mu_{n+1}(\theta_i) = \tilde{\alpha}_i + \sum_{j=1}^{n} \tilde{\alpha}_{ij}X_{i(j)} \tag{1}
\]

where \(\tilde{\alpha}_i\) dan \(\tilde{\alpha}_{ij}\) are real numbers, obtained by minimizing the mean square error of the hypothetical mean. From [2] we can derive equation 1 to

\[
\mu_{n+1}(\theta_i) = Z_iX_i + (1 - Z_i)\mu \tag{2}
\]

where

\[
\mu = E[\mu(\theta_j)] = E[X_{i(j)}], \forall j \tag{3}
\]

and

\[
Z_i = \frac{\sum_{j=1}^{n}E[Var(X_{i(j)}|\theta_i)]}{n + \frac{E[Var(X_{i(j)}|\theta_i)]}{Var(E(X_{i(j)}|\theta_i))}} \tag{4}
\]

2.2. \(L^2\) Space

In this paper, the formation of the credibility estimator requires the use of orthogonal projection on one of the Hilbert spaces, which is \(L^2\) [4]. This space’s elements are random variables, defined as

\[L^2 = \{X: X \text{ is random variable, } E(X^2) < \infty\}\]

with inner product [5]

\[\langle X, Y \rangle = E(XY)\]
and the distance between two elements are defined as
\[ X, Y \in \mathcal{L}^2, ||X - Y|| = \sqrt{E(X - Y)^2} \]
If \( X \) is assumed to be the estimator of \( Y \), then it is clear that the distance between two elements in \( \mathcal{L}^2 \) is in the form of a root mean square error. This is the main reason why \( \mathcal{L}^2 \) is used to form the credibility estimator, because root mean square error can measure the bias contained in the estimator.

### 2.2.1. Orthogonal projection in \( \mathcal{L}^2 \)
Here, we will explain the orthogonality and orthogonal projection in the context of \( \mathcal{L}^2 \). Two arbitrary elements in \( \mathcal{L}^2 \), denoted by \( X \) dan \( Y \) are orthogonal if
\[ (X, Y) = E(XY) = 0 \]
Furthermore, let \( M \subset \mathcal{L}^2 \), \( Y \in \mathcal{L}^2 \) and \( Y^* \in M \) is orthogonal projection of \( Y \). Then,
\[ (Y - Y^*, Z_1 - Z_2) = E[(Y - Y^*)(Z_1 - Z_2)] = 0, \forall Z_1, Z_2 \in M \]
In addition, we will mention a theorem related to the orthogonal projection. These three statements are equivalent to:
(i) \( Y^* = \text{proj}(Y|M) \), or in other words \( Y^* \) is orthogonal projection of \( Y \) on \( M \)
(ii) \( Y^* \in M \) and \( (Y - Y^*, Z - Y^*) = 0, \forall Z \in M \)
(iii) \( Y^* \in M \) and \( ||Y - Y^*|| \leq ||Y - Z||, \forall Z \in M \)
It is easy to prove that the orthogonal projection \( Y^* \) meets those three statements. As we can see from the third statement, it is clear that \( Y^* \) is the best estimator because it has the shortest distance from the object that is estimated. For this reason, a credibility estimator will be built by utilizing the orthogonal projection.

### 3. Results and discussion

#### 3.1. Bühlmann credibility estimator with correlated risk parameters
First, we will explain the notations that are used in the process of forming the credibility estimator, which are
(i) \( X_i = (X_{i1}, X_{i2}, ..., X_{in_i})' \) which is a vector containing an individual’s claim amount of each period
(ii) \( X = (X_1', X_2', ..., X_k')' \) which is a vector containing all individuals’ claim amount
Next, we will explain the assumptions required [6], which are
(i) \( X_{i(j)}|\theta_j \) with \( j = 1, 2, ..., n_i \), are i.i.d and have the same hypothetical mean and process variance:
\[ \mu(\theta_j) = E(X_{i(j)}|\theta_j) \]
\[ \sigma^2(\theta_j) = Var(X_{i(j)}|\theta_j) \]
(ii) \( E(\mu(\theta_j)) = \mu, \ Var(\mu(\theta_j)) = \tau^2, \ E(\sigma^2(\theta_j)) = \sigma^2 \)
(iii) \( \theta_1, \theta_2, ..., \theta_k \) are equally correlated
(iv) \( Corr(\mu(\theta_j), \mu(\theta_s)) = \rho, \ j \neq s \)
Next, let us define a closed subset of \( \mathcal{L}^2 \), which is \( \mathcal{L}(X, 1) \), whereas the elements are in the form of random variable as well, denoted
This closed subset contains all possible estimators of hypothetical mean, denoted \( \hat{\mu}(\theta) \).

Next, we define one of the possible estimators as the orthogonal projection of the hypothetical mean, denoted by \( \hat{\mu}(\theta) = \text{pro}(\mu(\theta) \mid L(\mathbf{X}, 1)) \in L(\mathbf{X}, 1) \).

\( \hat{\mu}(\theta) \) fulfills these statements:

(i) \( (\mu(\theta) - \hat{\mu}(\theta), 1) = 0 \)

(ii) \( (\mu(\theta) - \hat{\mu}(\theta), X_{i(j)}) = 0 \), for \( i = 1, \ldots, k \) and \( j = 1, \ldots, n_l \)

As a consequence, \( \hat{\mu}(\theta) \) also fulfills:

(i) \( E(\mu(\theta) - \hat{\mu}(\theta)) = 0 \)

(ii) \( \text{Cov}(\mu(\theta), X_{i(j)}) = \text{Cov}(\hat{\mu}(\theta), X_{i(j)}) \)

Since \( \hat{\mu}(\theta) \) also an element of \( L(\mathbf{X}, 1) \), then \( \hat{\mu}(\theta) \) can be represented as

\[
\hat{\mu}(\theta) = \hat{a}_0 + \sum_{i=1}^{k} \sum_{j=1}^{n_l} \hat{a}_{i(j)} X_{i(j)}
\]

(5)

In the form of matrix, \( \hat{\mu}(\theta) \) can be rewritten as

\[
\hat{\mu}(\theta) = \hat{a}_0 + \hat{\alpha} \mathbf{X}
\]

(6)

where \( \hat{\alpha} \) is a vector containing constants with size of \( N \times 1 \). \( N \) is the total number of claim periods observed for all individuals. It can be written as

\[
N = \sum_{i=1}^{k} n_l
\]

(7)

Now, we want to find the form of \( \hat{a}_0 \) and \( \hat{\alpha} \). First, let

\[
c' = \left( \text{Cov}(\mu(\theta_1), X_{1(1)}), \text{Cov}(\mu(\theta_1), X_{1(2)}), \ldots, \text{Cov}(\mu(\theta_1), X_{k(n_k)}) \right)
\]

(8)

and

\[
\Sigma_X = \begin{pmatrix}
\text{Var}(X_{1(1)}) & \text{Cov}(X_{1(1)}, X_{1(2)}) & \ldots & \text{Cov}(X_{1(1)}, X_{k(n_k)}) \\
\text{Cov}(X_{1(2)}, X_{1(1)}) & \text{Var}(X_{1(2)}) & \ldots & \text{Cov}(X_{1(2)}, X_{k(n_k)}) \\
\vdots & \vdots & \ddots & \vdots \\
\text{Cov}(X_{k(n_k)}, X_{1(1)}) & \text{Cov}(X_{k(n_k)}, X_{1(2)}) & \ldots & \text{Var}(X_{k(n_k)})
\end{pmatrix}
\]

(9)

Because \( \text{Cov}(\mu(\theta_i), X_{i(j)}) = \text{Cov}(\hat{\mu}(\theta_i), X_{i(j)}) \), which is proven to be fulfilled, \( c' \) can be written as

\[
c' = \hat{\alpha} \Sigma_X
\]
If $\Sigma_X$ is a non-singular matrix, $\bar{a}$ can be expressed as

$$\bar{a} = c' \Sigma_X^{-1}$$  \hspace{1cm} (10)

Substituting equation 10 to the equation 5, we obtain

$$\hat{\mu}(\theta_i) = \bar{a}_0 + c' \Sigma_X^{-1} \cdot \hat{X}$$  \hspace{1cm} (11)

Because $E\left(\mu(\theta_i) - \hat{\mu}(\theta_i)\right) = 0$, then we also obtain

$$\bar{a}_0 = \mu_0 - c' \Sigma_X^{-1} \cdot \mu_X$$  \hspace{1cm} (12)

Finally, we can get the complete credibility estimator, which can be expressed as

$$\hat{\mu}(\theta_i) = \mu_0 + c' \Sigma_X^{-1} \cdot (X - \mu_X)$$  \hspace{1cm} (13)

Equation 13 is equivalent to

$$X_{i(n_1+1)} = E(X_{i(n_1+1)}) + \text{Cov}(X_{i(n_1+1)}, X) \cdot \text{Var}(X)^{-1} \cdot (X - E(X))$$  \hspace{1cm} (14)

We wish to find the formula for each component, which are $E(X_{i(n_1+1)})$, $\text{Cov}(X_{i(n_1+1)}, X)$, $\text{Var}(X)^{-1}$ and $(X - E(X))$. First, we wish to derive $E(X_{i(n_1+1)})$. Based on the assumptions that has been mentioned before,

$$E(X_{i(n_1+1)}) = \mu$$  \hspace{1cm} (15)

and

$$E(X) = \mu_1_N$$  \hspace{1cm} (16)

where $1_N$ is a vector with the size of $N \times 1$ and its elements are 1.

Next, we wish to find the form of $\text{Cov}(X_{i(n_1+1)}, X)$. The elements of this vector are

$$\text{Cov}(X_{i(n_1+1)}, X) = \left\{ \text{Cov}(X_{i(n_1+1)}, X_{1(i)}), ..., \text{Cov}(X_{i(n_1+1)}, X_{k(n_k)}) \right\}$$  \hspace{1cm} (17)

From the double expectation of conditional covariance, we have that for different individuals, such as in the first element, the covariance is $\tau^2$. As for the same individuals, the covariance is $\rho \tau^2$.

Lastly, we derive the formula of $\text{Var}(X)^{-1}$. $\text{Var}(X)$ has the same elements as in the equation 9, and the formula for each element can also be obtained with the double expectation of conditional covariance,

$$\text{Var}(X) = \begin{pmatrix} \sigma^2 + \tau^2 & \tau^2 & \cdots & \rho \tau^2 \\ \tau^2 & \sigma^2 + \tau^2 & \cdots & \rho \tau^2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho \tau^2 & \rho \tau^2 & \cdots & \sigma^2 + \tau^2 \end{pmatrix}$$  \hspace{1cm} (18)

This matrix can be rewritten as

$$\text{Var}(X) = \text{diag}(\Lambda_1, \Lambda_2, ..., \Lambda_k) + \rho \tau^2 1_N 1_N'$$  \hspace{1cm} (19)

where

$$\Lambda_i = \sigma^2 I_{n_i} + \tau^2 (1 - \rho) 1_{n_i} 1_{n_i}'$$  \hspace{1cm} (20)
Using the formula of the inverse of matrix addition in [7], one can get

\[
Var(\mathbf{X})^{-1} = \text{diag}(\Lambda_1^{-1}, \Lambda_2^{-1}, ..., \Lambda_k^{-1})
\]

\[
- \frac{1}{\rho r^2 + \sum_{i=1}^{k} n_i (1 - \rho) r^2} \left( \frac{1}{\sigma^2} \mathbf{1} n_1 \cdots \frac{1}{\Lambda_{n_1}^{-1}} \mathbf{1} n_1 \cdots \frac{1}{\Lambda_{n_k}^{-1}} \mathbf{1} n_k \right)
\]

Furthermore, by the formula of matrix inverse

\[
(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\]

we can get

\[
\Lambda_i^{-1} = \frac{1}{\sigma^2} \left( I_{n_i} - \frac{1 - \rho r^2}{\sigma^2 + n_i (1 - \rho) r^2} \mathbf{1} n_i \mathbf{1} n_i \right)
\]

Substituting all of the components' formulas to the original credibility estimator formula in equation 14, one can get

\[
X_{i(n_i+1)} = Z_{i1} \overline{X}_i + Z_{i2} \overline{X}_\lambda + (1 - Z_{i1} - Z_{i2}) \mu
\]

with

\[
\overline{X}_\lambda = \frac{1}{\lambda} \sum_{i=1}^{k} \lambda_i \overline{X}_i, \quad \lambda = \sum_{i=1}^{k} \lambda_i, \quad \lambda_i = \frac{n_i}{\sigma^2 + n_i (1 - \rho) r^2}
\]

and the credibility estimators' formulas are

\[
Z_{i1} = (1 - \rho) r^2 \lambda_i, \quad Z_{i2} = \frac{\lambda \rho r^2 (1 - (1 - \rho) r^2 \lambda_i)}{\lambda \rho r^2 + 1}
\]

If \(n_1 = n_2 = \cdots = n_k = n\) and \(\rho = 0\), it is easy to prove that

\[
X_{i(n_i+1)} = Z_{i1} \overline{X}_i + (1 - Z_{i1}) \mu
\]

with \(Z_{i1} = Z_i\) in equation 4. This is the standard Bühlmann credibility estimator with no correlation between risk parameters.

3.2. Parameter estimation

Here, we wish to find the unbiased estimator of each parameter included in the credibility estimator formula, which are \(\mu, \sigma^2, r^2\). The unbiased estimator of \(\mu\) is \(\overline{X}_\lambda\), with proving that

\[
E(\overline{X}_\lambda) = \mu
\]

Next, the search of the unbiased estimator of \(\sigma^2\) is as follow. First, we find that

\[
\text{Cov}(X_{i(0)}, \overline{X}_i) = \frac{\sigma^2}{n_i} + r^2
\]

and

\[
E\left(\left(\overline{X}_i - \overline{X}_\lambda\right)^2\right) = \sigma^2 - \frac{\sigma^2}{n_i}
\]
Based on equation 29 and equation 30, one can easily get

\[ \hat{\sigma}^2 = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{i(k)} - \bar{X}_i)^2}{\sum_{i=1}^{k} n_i (1 - k)} \]  

(31)

Lastly, we wish to find the unbiased estimator of \( \tau^2 \). We find that

\[ E(\bar{X}) = \mu, \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{N} + \frac{\sum_{i=1}^{k} n_i^2 (1 - \rho) \tau^2}{N^2} + \rho \tau^2 \]  

(32)

and also

\[ \text{Cov}(\bar{X}_i, \bar{X}) = \frac{1}{N} (\sigma^2 + n_i (1 - \rho) \tau^2) + \rho \tau^2 \]  

(33)

Based on that, one can easily get

\[ \hat{\tau}^2 = \frac{\sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})^2 - \hat{\sigma}^2 (k - 1)}{N - \sum_{i=1}^{k} n_i^2 (1 - \rho)} \]  

(34)

### 3.3. Data analysis

Here, we apply the standard Bühlmann credibility estimator and Bühlmann credibility estimator with correlated risk parameters to predict the amount of claim for 50 policyholders living in the same area, claiming to suffer from dengue fever in the same period. This data is obtained from a multinational life insurance company operating in Indonesia. The data consists of 5 periods for each policyholder, from January 2018 until May 2018. To analyse which estimator is better, we will utilize cross validation to divide the data into groups consisting of testing and training data. To be specific, we will apply cross validation 4-fold, which forms 4 testing data consisting of 25 % of claim data at the 5th period (May 2018) and 4 train data which consists of the first four periods (January-April 2018) claim data, including the data for the 5th period which are not used in the test data. We then obtain the root mean square errors depicted in figure 1.

![Figure 1. Root mean square error](image-url)
Table 1. Root mean square errors for partitioned data

|                   | Optimum $\rho$ | 2 partitions | 5 partitions | Without partition |
|-------------------|----------------|--------------|--------------|------------------|
| Train             | 6013980.625    | 55446751.1   | 5823877      |
| Test              | 5597454.875    | 4683706.5    | 5823877      |

The dotted lines are root mean square errors from 4 training and testing data, which are determined by cross-validation. The solid lines are the average of the 4 sets’ of data root mean square errors. Based on figure 1 generally, when $\rho = 0$, the estimator generates a higher root mean square error, and as the correlation increases, it is shown that the root mean square error decreases.

This proves that for this particular data, it is more suitable for us to apply the Bühlmann credibility estimator with correlated risk parameters to predict the claim. Furthermore, we also split the data into two and five, based on the homogeneity of the claim amount’s mean. We obtained the root mean square errors shown in table 1.

As seen from table 1, data partition has the potential to produce smaller root mean square errors as the number of partitions increase, although there is always a chance that the partitioned data will yield a larger root mean square error. In table 1, this happens when the credibility estimator using the optimum $\rho$ in training data is applied to a two-partition data. It has a greater root mean square error compared to the non-partitioned data.

4. Conclusion

The formation of Bühlmann credibility estimator assuming equally correlated risk parameters is carried out by utilizing orthogonal projections in one of Hilbert’s subspaces, namely $L(\mathbf{X}, 1)$. Each component of the orthogonal projection is determined by its formula; then, the parameter estimation is performed. The credibility factors $(Z_{11}, Z_{12})$, and the parameters’ estimators are in the form of the function of correlation ($\rho$). Also, the credibility estimators are compared to predict the claims based on data and it is proven that the Bühlmann credibility estimator assuming correlated risk parameters is superior. Partition of data based on the homogeneity of claims shows that it helps reduce the root mean square error.

References

[1] Hanafi M M 2006 *Manajemen Risiko* (Yogyakarta: UPP STIM YKPN)
[2] Klugman S A, Panjer H H and Willmot G E 2012 *Loss Models: From Data to Decisions* 4th edition (New Jersey: John Wiley & Sons)
[3] Knuiman M W, Divitini M L, Bartholomew H C and Welborn T A 1996 *Am. J. Epidomol.* **143** 48-53
[4] Bühlmann H and Alois G 2005 *A Course in Credibility Theory and Its Application* (Berlin: Springer)
[5] Reschenhofer E 2019 *L2* (Austria: University of Vienna) available at https://homepage.univie.ac.at/erhard.reschenhofer/pdf/advanced/L2.pdf
[6] Wen L and Deng W 2011 *J. Syst. Sci. Complex.* **24** 532-9
[7] Sherman J and Morrison W J 1949 *Adjustment of an Inverse Matrix Corresponding to Changes in the Elements of a Given Column or a Given Row of the Original Matrix* available https://projecteuclid.org/download/pdf_1/euclid.aoms/1177729893