**Zeitschrift:** Mitteilungen / Schweizerische Vereinigung der Versicherungsmathematiker = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries

**Herausgeber:** Schweizerische Vereinigung der Versicherungsmathematiker

**Band:** - (1994)

**Heft:** 2

**Artikel:** Dependent contracts in Bühlmann's credibility model

**Autor:** De Vijlder, F. / Cossette, H.

**DOI:** https://doi.org/10.5169/seals-967202

# B. Wissenschaftliche Mitteilungen

F. DE VIJLDER, H. COSSETTE*, Louvain-la-Neuve

## Dependent contracts in Bühlmann's credibility model

### 1. Introduction: credibility on a roulette, dependent contracts

We consider a roulette with holes numbered $1, 2, \ldots, k$. We denote by $\theta_j$ $(j = 1, \ldots, k)$ the probability that the ball falls in the hole $j$, in any play. In a perfect roulette,

$$\theta_1 = \theta_2 = \cdots = \theta_k = \frac{1}{k}\,.$$

We do not assume the roulette to be perfect here. We have

$$\theta_1 + \theta_2 + \cdots + \theta_k = 1.$$

Let us observe the roulette during $t$ plays, numbered $1, 2, \ldots, t$. We define $X_{js}$ $(j = 1, \ldots, k;\ s = 1, \ldots, t)$ to be equal to 1 if the ball falls in the hole $j$ at the play $s$, and to be equal to 0 if not. We want to find credibility estimators for $\theta_1, \ldots, \theta_k$. Bühlmann's model does not apply directly, because it treats portfolio's with independent contracts. The independence assumption is not satisfied on the roulette, because

$$X_{1s} + X_{2s} + \cdots + X_{ks} = 1 \qquad (s = 1, \ldots, t).$$

But Bühlmann's estimator is constructed on a fixed contract, and the portfolio with several contracts is used only for the estimation of the involved parameters. Starting with this idea, we were convinced that only a small adaptation of Bühlmann's model would solve our roulette problem.

We like to believe that the contracts of insurance portfolios can be considered as being independent, mostly because this is very convenient for the construction of a theory. But chain car crashes are not exceptional, and on misty days all cars of a region have higher probabilities to be involved in an accident. During dry hot summers, all wooden cottages are more exposed to fire. In such situations, the stochastic dependence can hardly be neglected. Next credibility model shows that it is not so difficult to take it into account.

*This research is supported by l'Université Laval and Fonds pour la Formation de Chercheurs et l'Aide à la Recherche.

The considerations of this note also apply, mutatis mutandis, to the Bühlmann-Straub model with weighted observations.

## 2. Bühlmann columns

A *Bühlmann column with characteristics* $\mu(.), m, a, s^2$ is a column

$$(\Theta, X_1, \dots, X_t)' \tag{1}$$

of random variables such that

$$\begin{aligned} E(X_s/\Theta) &= \mu(\Theta) \qquad (s = 1, \dots, t), \\ E\mu(\Theta) &= m, \\ \operatorname{Var}\mu(\Theta) &= a, \qquad\qquad\qquad\qquad (2) \\ E\operatorname{Cov}(X_r, X_s/\Theta) &= \delta_{rs}s^2 \qquad (r, s = 1, \dots, t). \qquad (3) \end{aligned}$$

(No confusion is possible between the time subscript $s$ and the $s$ occurring in the parameter $s^2$). Using the indicated covariances, we implicitly assume that

$$EX_s^2 < \infty \qquad (s = 1, \dots, t).$$

Any column (1), such that for fixed $\Theta$, the variables $X_1, \dots, X_t$ are conditionally i.i.d. is a Bühlmann column with characteristics defined by (2) and (3).
In the paper Bühlmann (1967), at the origin of all the developments of modern credibility theory, the author argued on a single column (1), and he only used the relations (2) and (3). He derived the *credibility estimator* $\widehat{\mu}$ for $\mu(\Theta)$:

$$\widehat{\mu} = z\frac{1}{t}(X_1 + \dots + X_t) + (1 - z)m, \qquad \text{where} \quad z = \frac{at}{s^2 + at}.$$

The *classical Bühlmann model,* with several contracts, can be defined as being a finite number of independent Bühlmann columns with the same characteristics. The parameters $m$, $a$, $s^2$ can be estimated from observations in this model.
As we shall see in next sections, the independence assumption can be dropped if some bias is allowed in the estimation of $a$. In fact, already in the classical Bühlmann and Bühlmann-Straub models, the usual estimators for $a$ are biased. Indeed, pseudo-estimators are biased because they do not use the exact $z$'s. If the classical estimator is used, say $\widehat{a}$, it may take negative values and it is replaced by $\max(\widehat{a}, 0)$, introducing a bias again.

### 3. Bühlmann's generalized model

We consider the model with random variables

$$\begin{array}{ccccc} \Theta_1 & \dots & \Theta_i & \dots & \Theta_k \\ X_{11} & \dots & X_{i1} & \dots & X_{k1} \\ \dots & \dots & \dots & \dots & \dots \\ X_{1s} & \dots & X_{is} & \dots & X_{ks} \\ \dots & \dots & \dots & \dots & \dots \\ X_{1t} & \dots & X_{it} & \dots & X_{kt} \end{array}$$

We only assume that the columns of this tableau are Bühlmann columns with the same characteristics $\mu(.)$, $m$, $a$, $s^2$:

$$\begin{aligned} E(X_{is}/\Theta_i) &= \mu(\Theta_i), \\ E\mu(\Theta_i) &= m, \\ \operatorname{Var}\mu(\Theta_i) &= a \qquad (i = 1, \dots, k;\ \ s = 1, \dots, t) \\ E\operatorname{Cov}(X_{ir}, X_{is}/\Theta_i) &= \delta_{rs}s^2 \qquad (i = 1, \dots, k;\ \ r, s = 1, \dots, t). \end{aligned}$$

The credibility estimator for $\mu(\Theta_i)$ is

$$\widehat{\mu}_i = zX_{iM} + (1 - z)m,$$

$$\text{where} \quad z = \frac{at}{s^2 + at} \qquad \text{and} \quad X_{iM} = \frac{1}{t}(X_{i1} + \cdots + X_{it}).$$

Our problem is to estimate $m$, $s^2$, $a$.
Whereas the columns of the tableau are independent in the classical model, we later (section 6) make a conditional independence assumption on the lines of the tableau.

### 4. Parameter estimation

*Estimation of $m$.*

The obvious estimator of $m$ is

$$\widehat{m} = \frac{1}{kt}\sum_{is} X_{is}\,.$$

But $m$ might also be estimated in a larger portfolio than the one under consideration. Hereafter, $m'$ denotes any estimate (a constant, not a random variable), of $m$. In the roulette problem, $m$ is known exactly, and then we take of course $m' = m$.

*Estimation of* $s^2$.

As in the classical Bühlmann model, the right estimator of $s^2$ is

$$\widehat{s^2} = \frac{1}{k(t-1)} \sum_i \sum_s \left(X_{is} - X_{iM}\right)^2 \tag{4}$$

It is unbiased. Indeed, the unbiasedness of the estimator

$$\widehat{s_i^2} = \frac{1}{t-1} \sum_s \left(X_{is} - X_{iM}\right)^2,$$

based on the observations of the $i$-th contract only, results from the fact that the $i$-th column of the tableau of section 3 is a Bühlmann column. Moreover,

$$\widehat{s^2} = \frac{1}{k} \sum_i \widehat{s_i^2}\,.$$

*Estimation of* $a$.

We suggest the estimator

$$\widehat{a} = \frac{1}{kt(t-1)} \sum_i \sum_{r \neq s} \left(X_{ir} - m'\right)\left(X_{is} - m'\right). \tag{5}$$

It satisfies

$$E\widehat{a} = a + (m - m')^2, \tag{6}$$

i.e. it has the bias $(m - m')^2$.
Indeed, we have

$$\text{Cov}\left(X_{ir}, X_{is}\right) = a + \delta_{rs} s^2 \qquad (i = 1, \ldots, k;\; r, s = 1, \ldots, t)$$

because the $i$-th contract is a Bühlmann column. Also,

$$\begin{aligned}
&\left(X_{ir} - m'\right)\left(X_{is} - m'\right) \\
&\quad = \left(\left(X_{ir} - m\right) + \left(m - m'\right)\right)\left(\left(X_{is} - m\right) + \left(m - m'\right)\right) \\
&\quad = \left(X_{ir} - m\right)\left(X_{is} - m\right) + \left(m - m'\right)\left(X_{is} - m\right) \\
&\qquad + \left(X_{ir} - m\right)\left(m - m'\right) + \left(m - m'\right)^2 .
\end{aligned}$$

Hence, if $r \neq s$,

$$E\big((X_{ir}-m')(X_{is}-m')\big)$$
$$= E\big((X_{ir}-m)(X_{is}-m)\big) + (m-m')^2$$
$$= \mathrm{Cov}\,(X_{ir}, X_{is}) + (m-m')^2 = a + (m-m')^2.$$

This implies (6) because the double sum in the last member of (5) has $kt(t-1)$ terms with the same expectation

$$a + (m-m')^2.$$

We notice that

$$\widehat{a} = \frac{1}{k}\sum_i \widehat{a}_i\,,$$

where

$$\widehat{a}_i = \frac{1}{t(t-1)}\sum_{r\neq s}(X_{ir}-m')(X_{is}-m').$$

Because $a \geq 0$, we cannot accept negative estimates and we have to replace $\widehat{a}$ by $\max(0,\widehat{a})$, or, perhaps even better, by

$$\frac{1}{k}\sum_i \max(0,\widehat{a}_i).$$

## 5. Application to the roulette problem

In the roulette problem of the introduction, we interpret the unknown vector $(\theta_1,\ldots,\theta_k)$ as realization of some random vector $(\Theta_1,\ldots,\Theta_k)$. To other realizations of that random vector correspond other hypothetical roulettes.
At each play, the ball falls in exactly one hole. Hence

$$X_{1s} + \cdots + X_{ks} = 1.$$

Taking expectations, we have $km = 1$. Hence $m = 1/k$ and we can use the estimator (5) with $m' = 1/k$. Then (6) shows its unbiasedness.
We use the notations

$$X_{i\Sigma} = X_{i1} + \cdots + X_{it} \qquad (i=1,\ldots,k),$$
$$X_{\Sigma\Sigma} = X_{1\Sigma} + \cdots + X_{k\Sigma} = t,$$
$$Q = X_{1\Sigma}^2 + \cdots + X_{k\Sigma}^2\,.$$

We have

$$t \leq Q \leq t^2.$$

The case $Q = t$ occurs if the ball falls in different holes in the plays $s = 1, \ldots, t$. This is only possible if $t \leq k$. The case $Q = t^2$ occurs if the ball falls in the same hole in each of the plays $s = 1, \ldots, t$.

Here we have

$$\begin{aligned} \mu(\Theta_i) &= E\big(X_{is}/\Theta_i\big) \\ &= P\big(X_{is} = 1/\Theta_i\big) \\ &= \Theta_i \end{aligned}$$

and $\widehat{\mu}_i$ is the credibility estimator $\widehat{\theta}_i$ of $\theta_i$:

$$\widehat{\theta}_i = zX_{iM} + (1-z)\frac{1}{k} \qquad (i = 1, \ldots, k). \tag{7}$$

*Estimation of* $s^2$.

From (4) results

$$\begin{aligned} k(t-1)\widehat{s}^2 &= \sum_i \sum_s \left(X_{is} - \frac{1}{t}X_{i\Sigma}\right)^2 \\ &= \sum_i \sum_s \left(X_{is} - \frac{2}{t}X_{is}X_{i\Sigma} + \frac{1}{t^2}X_{i\Sigma}^2\right) \\ &= X_{\Sigma\Sigma} - \frac{2}{t}\sum_i X_{i\Sigma}^2 + \frac{1}{t^2}t\sum_i X_i^2 \\ &= t - \frac{Q}{t}\,. \end{aligned}$$

Hence

$$\widehat{s}^2 = \frac{t^2 - Q}{kt(t-1)}\,, \tag{8}$$

$$0 \leq \widehat{s}^2 \leq \frac{1}{k}\,. \tag{9}$$

*Estimation of a.*

$$\begin{aligned}
&\sum_i \sum_{r \neq s} (X_{ir} - m)(X_{is} - m) \\
&= \sum_{irs} (X_{ir} - m)(X_{is} - m) - \sum_{is} (X_{is} - m)^2 \\
&= \sum_{irs} (X_{ir}X_{is} - mX_{is} - mX_{ir} + m^2) \\
&\quad - \sum_{is} (X_{is} - 2mX_{is} + m^2) \\
&= Q - mt^2 - mt^2 + m^2kt^2 - t + 2mt - ktm^2 \\
&= Q - \frac{1}{k}t^2 - \frac{1}{k}t^2 + \frac{1}{k^2}kt^2 - t + \frac{2}{k}t - kt\frac{1}{k^2} \\
&= (Q - t) + \frac{t}{k} - \frac{t^2}{k} = t(t-1) - \frac{1}{k}t(t-1) - (t^2 - Q).
\end{aligned}$$

Hence, by (5) and (8),

$$\widehat{a} = \left(\frac{1}{k} - \widehat{s}^2\right) - \frac{1}{k^2} \tag{10}$$

This is a relation between the unbiased estimators $\widehat{a}$, $\widehat{s}^2$. Taking expectations, we obtain that the corresponding relation holds between the exact values $a$, $s^2$. In fact, the latter relation is obvious because

$$\begin{aligned}
a + s^2 = \operatorname{Var} X_{is} &= E(X_{is}^2) - E^2(X_{is}) = E(X_{is}) - m^2 \\
&= m - m^2 = \frac{1}{k} - \frac{1}{k^2}\,.
\end{aligned}$$

*Discussion of the solution.*

*Case* 1: The last member of (10) is negative.

Then we can only adopt the estimate $a = 0$. Then $z = 0$ and from (7) results

$$\widehat{\theta}_i = \frac{1}{k} \qquad (i = 1, \ldots, k).$$

This means that we cannot contradict the perfectness of the roulette.
Case 1 certainly occurs if the balls fall in different holes in the considered plays. Indeed, then $Q = t$, $\widehat{s}^2 = 1/k$ and the last member of (10) equals $-1/k^2$.

*Case* 2: The last member of (10) is positive.

Then we adopt (8) and (10) as estimators for $s^2$, $a$ and the estimate of $\theta_i$ results from (7). Of course, all involved random variables $X_{is}$ are supposed to be replaced by their realizations $x_{is}$.
We tried different values of $k$, $t$, $Q$. In each case there was complete agreement with the intuition that we have of the roulette problem, or at least, there was no contradiction with this intuition. As extreme example, let us assume that in all the considered plays the ball falls in the hole 1. Then

$$\begin{aligned} &Q = t^2, \qquad &&\widehat{s}^2 = 0, \qquad &&z = 1, \\ &\widehat{\theta}_1 = 1, \qquad &&\widehat{\theta}_i = 0 \qquad &&(i = 2, \ldots, k). \end{aligned}$$

*Conclusion.*

This study implies the following advice to the gamblers: play the numbers that come out most frequently. Most strangely, a lot of gamblers do just the opposite. They commit a double error. First they believe that the roulette is perfect. This is certainly not the case in obscure gamblers clubs, and even not in big casinos, as one of us was told by a manager. Further, they know just enough of the law of large numbers to apply it incorrectly. If a number has not come out for some time, they believe that the law of large numbers sanctions the roulette by forcing its outcome in the next plays. Of course they ignore a statement by Borel, we believe, "La roulette n'a ni yeux, ni mémoire".

## 6. A restriction on Bühlmann's generalized model

Our next problem is the detection of the stochastic dependence of the contracts (i.e. the columns in the tableau of 3.) in Bühlmann's generalized model, restricted conveniently.
We shall assume that, in any fixed portfolio, corresponding to a realization

$$\theta = (\theta_1, \ldots, \theta_k)$$

of the vector

$$\Theta = (\Theta_1, \ldots, \Theta_k),$$

the observations in different years $s = 1, \ldots, t$, are independent. In fact we assume less. Precisely, we suppose that

$$E\big(X_{is}/\Theta\big) = E\big(X_{is}/\Theta_i\big) \qquad (i = 1, \ldots, k; \;\; s = 1, \ldots, t)$$

and that constants $b$, $c$ exist, satisfying the relations

$$\begin{aligned} \mathrm{Cov}(\mu(\Theta_i), \mu(\Theta_j)) &= b \qquad (i,j = 1,\dots,k;\ i \neq j), \\ E\,\mathrm{Cov}\left(X_{ir}, X_{js}/\Theta\right) &= \delta_{rs}c \quad (i,j = 1,\dots,k;\ i \neq j;\ r,s = 1,\dots,t). \end{aligned}$$

If the contracts are independent, then $b = c = 0$.

The values of $b$ and $c$ give an idea of the stochastic dependence, more precisely of the correlation, of the contracts. In next section, we indicate how they can be estimated.

## 7. Parameter estimation in the restricted model

The estimation of $b$ and $c$ is based on the relation

$$\mathrm{Cov}\left(X_{ir}, X_{js}\right) = b + \delta_{rs}c \qquad (i,j = 1,\dots,k;\ i \neq j;\ r,s = 1,\dots,t) \tag{11}$$

proved in the usual way from the assumptions on the restricted model. Indeed, for $i \neq j$,

$$\begin{aligned} \mathrm{Cov}\left(X_{ir}, X_{js}\right) &= \mathrm{Cov}\left(E(X_{ir}/\Theta), E(X_{js}/\Theta)\right) + E\,\mathrm{Cov}\left(X_{ir}, X_{js}/\Theta\right) \\ &= \mathrm{Cov}\left(E(X_{ir}/\Theta_i), E(X_{js}/\Theta_j)\right) + \delta_{rs}c \\ &= \mathrm{Cov}\left(\mu(\Theta_i), \mu(\Theta_j)\right) + \delta_{rs}c \\ &= b + \delta_{rs}c. \end{aligned}$$

*Estimation of b.*

For $b$ we suggest the estimator

$$\widehat{b} = \frac{1}{k(k-1)t(t-1)} \sum_{i\neq j, r\neq s} (X_{ir} - m')(X_{js} - m'), \tag{12}$$

where $m'$ is an approximation of $m$.

By a proof similar to that of (6), we obtain

$$E\widehat{b} = b + (m - m')^2. \tag{13}$$

In the roulette case, we can take $m' = m$ and then $\widehat{b}$ is unbiased.

*Estimation of c.*

For $c$ we suggest the estimator

$$\widehat{c} = \frac{1}{tk(k-1)} \sum_{s} \sum_{i \neq j} (X_{is} - m')(X_{js} - m') - \widehat{b}. \tag{14}$$

It is unbiased: $E\widehat{c} = c$.
Indeed, from (14) results that

$$E(\widehat{b} + \widehat{c}) = b + c + (m - m')^2,$$

and the difference with (13) proves the unbiasedness of $\widehat{c}$, whatever be $m'$.

*General expressions for the estimators.*

Let us abbreviate

$$\sum_{ijrs} = \sum_{ijrs} (X_{ir} - m')(X_{js} - m'),$$

$$\sum_{ijs} = \sum_{ijs} (X_{is} - m')(X_{js} - m'),$$

$$\sum_{irs} = \sum_{irs} (X_{ir} - m')(X_{is} - m'),$$

$$\sum_{is} = \sum_{is} (X_{is} - m')(X_{is} - m'),$$

where $i, j = 1, \ldots, k$ and $r, s = 1, \ldots, t$.
Then

$$\widehat{b} = \frac{1}{k(k-1)t(t-1)} \left( \sum_{ijrs} - \sum_{ijs} - \sum_{irs} + \sum_{is} \right), \tag{15}$$

$$\widehat{c} = \frac{1}{tk(k-1)} \left( \sum_{ijs} - \sum_{is} \right) - \widehat{b}, \tag{16}$$

$$\widehat{a} = \frac{1}{kt(t-1)} \left( \sum_{irs} - \sum_{is} \right) \tag{17}$$

From (15) and (17) results

$$\widehat{b} = \frac{1}{k(k-1)t(t-1)} \left( \sum_{ijrs} - \sum_{ijs} \right) - \frac{1}{k-1} \widehat{a} \tag{18}$$

## 8. Estimation of $b$ and $c$ in the roulette case

In the roulette case we take $m' = m = \frac{1}{k}$.

*Estimation of b.*

$$\begin{aligned}\sum_{ijrs} &= \sum_{ijrs} (X_{ir}X_{js} - mX_{js} - mX_{ir} + m^2)\\ &= t^2 - mkt^2 - mkt^2 + m^2k^2t^2\\ &= t^2 - t^2 - t^2 + t^2\\ &= 0\\ \sum_{ijs} &= \sum_{ijs} (X_{is}X_{js} - mX_{is} - mX_{js} + m^2)\\ &= t - mkt - mkt + m^2k^2t\\ &= t - t - t + t\\ &= 0.\end{aligned}$$

By (18),

$$\widehat{b} = -\frac{1}{k-1}\widehat{a}. \tag{19}$$

Taking expectations, we have

$$b = -\frac{1}{k-1}a \tag{20}$$

For a direct proof of (20), we first observe the obvious result:

$$\sum_i \Theta_i = \sum_i E(X_{is}/\Theta_i) = E\Big(\sum_i X_{is}/\Theta\Big) = E(1/\Theta) = 1.$$

Then

$$\begin{aligned}0 = \mathrm{Cov}\left(\sum_i \Theta_i, \sum_j \Theta_j\right) &= \sum_{ij} \mathrm{Cov}(\Theta_i, \Theta_j)\\ &= \sum_{i\neq j} \mathrm{Cov}(\Theta_i, \Theta_j) + \sum_i \mathrm{Var}(\Theta_i)\\ &= k(k-1)b + ka.\end{aligned}$$

*Estimation of c.*

$$
\begin{aligned}
\sum_{is} &= \sum_{is} (X_{is}^2 - 2mX_{is} + m^2) \\
&= \sum_{is} (X_{is} - 2mX_{is} + m^2) \\
&= t - 2mt + m^2 kt \\
&= t - \frac{t}{k} \\
&= \frac{1}{k}(k-1)t.
\end{aligned}
$$

Then, by (10) and (19)

$$
\widehat{c} = -\frac{1}{k^2} - \widehat{b} = -\frac{1}{k^2} + \frac{1}{k-1}\left(\frac{1}{k} - \widehat{s}^2 - \frac{1}{k^2}\right) = -\frac{1}{k-1}\widehat{s}^2 ,
$$

$$
\widehat{c} = -\frac{1}{k-1}\widehat{s}^2. \tag{21}
$$

Taking expectations, we obtain

$$
c = -\frac{1}{k-1}s^2 . \tag{22}
$$

For a direct proof of (22):

$$
\begin{aligned}
0 &= E\,\mathrm{Cov}(1, 1/\Theta) \\
&= E\,\mathrm{Cov}\left(\sum_i X_{is}, \sum_j X_{js}/\Theta\right) \\
&= \sum_{i \neq j} E\,\mathrm{Cov}\left(X_{is}, X_{js}/\Theta\right) + \sum_i E\,\mathrm{Var}\left(X_i/\Theta\right) \\
&= k(k-1)c + ks^2.
\end{aligned}
$$

*Remarks.*

In the roulette case, the dependence between the "contracts" is negative: $b \leq 0$, $c \leq 0$. In automobile insurance, fire insurance, ... positive dependencies must be expected.

In the roulette case, $\widehat{a}$, $\widehat{b}$, $\widehat{c}$ depend on $k$ and $\widehat{s}^2$ only, and $\widehat{s}^2$ only depends on $k$ and $Q$.

The quality of the estimator $\widehat{s}^2$ increases with $t$ and with $k$. For small values of $t$ and $k$, say 2 or 3, the number of observations $X_{is}$ is too small, and $\widehat{s}^2$ can only be a poor estimator of $s^2$.

## 9. Improved credibility estimators in the restricted model

The credibility estimator $\widehat{\mu}_i$ is the linear combination of the random variables $1, X_{i1}, \ldots, X_{it}$ closest to $\mu(\Theta_i)$ in the least squares sense. Let $\mu_i^*$ be the linear combination of 1 and all the observable random variables $X_{js}$ $(j = 1, \ldots, k;$ $s = 1, \ldots, t)$ closest to $\mu(\Theta_i)$ in the least squares sense. In the classical model $\widehat{\mu}_i = \mu_i^*$ because the contracts are independent. Here $\mu_i^*$ is a strictly better estimator than $\widehat{\mu}_i$ (except in special cases, where the two estimators may be equal). For symmetry reasons, $\mu_i^*$ can be displayed as

$$\mu_i^* = z_1 X_{iM} + z_2 X_{MM} + z_3 m, \tag{23}$$

where $X_{iM}$ is defined in section 3 and

$$X_{MM} = \frac{1}{k} \sum_i X_{iM} = \frac{1}{kt} \sum_{is} X_{is}$$

Then the unknown coefficients $z_1$, $z_2$, $z_3$ result from the relations

$$E\mu_i^* = E\mu(\Theta_i) \tag{24}$$

$$\text{Cov}\left(\mu_i^*, X_{js}\right) = \text{Cov}\left(\mu(\Theta_i), X_{js}\right) \qquad (j = 1, \ldots, k; \; s = 1, \ldots, t). \tag{25}$$

(For instance, see Theorem 5 of De Vylder e.a. (1992)).
In order to explicit (25), we need more

*Covariance relations.*

From the relation following (6) for $\text{Cov}\left(X_{ir}, X_{is}\right)$ and the relation (11) for $\text{Cov}\left(X_{ir}, X_{js}\right)$ $(i \neq j)$, results

$$\begin{aligned}\text{Cov}\left(X_{ir}, X_{js}\right) &= \text{Cov}\left(X_{ir}, X_{is}\right)\delta_{ij} + \text{Cov}\left(X_{ir}, X_{js}\right)(1 - \delta_{ij}) \\ &= b + c\delta_{rs} + (a - b)\delta_{ij} + (s^2 - c)\delta_{ij}\delta_{rs} \\ &\qquad (i, j = 1, \ldots, k; \; r, s = 1, \ldots, t).\end{aligned} \tag{26}$$

From (26) we obtain, for all $i, j = 1, \ldots, k$ and $r, s = 1, \ldots, t$:

$$\begin{aligned}\operatorname{Cov}\left(X_{ir}, X_{jM}\right) &= \operatorname{Cov}\left(X_{iM}, X_{jM}\right)\\ &= b + (a-b)\delta_{ij} + \frac{1}{t}\left(c + (s^2 - c)\delta_{ij}\right) \qquad (27)\end{aligned}$$

$$\begin{aligned}\operatorname{Cov}\left(X_{ir}, X_{iM}\right) &= \operatorname{Cov}\left(X_{iM}, X_{iM}\right)\\ &= a + \frac{1}{t}s^2 \qquad (28)\end{aligned}$$

$$\begin{aligned}\operatorname{Cov}\left(X_{ir}, X_{MM}\right) &= \operatorname{Cov}\left(X_{iM}, X_{MM}\right)\\ &= b + \frac{1}{k}(a-b) + \frac{1}{t}c + \frac{1}{kt}(b-c). \qquad (29)\end{aligned}$$

$$\begin{aligned}\operatorname{Cov}\left(\mu(\Theta_i), X_{js}\right) &= E\operatorname{Cov}\left(\mu(\Theta_i), X_{js}/\Theta\right)\\ &\quad + \operatorname{Cov}\left(E(\mu(\Theta_i)/\Theta), E(X_{js}/\Theta)\right)\\ &= 0 + \operatorname{Cov}\left(\mu(\Theta_i), \mu(\Theta_j)\right)\\ &= a\delta_{ij} + b(1-\delta_{ij}). \qquad (30)\end{aligned}$$

*Determination of* $\mu_i^*$.

From (23), (24):

$$z_1 + z_2 + z_3 = 1. \qquad (31)$$

From (23), (27), (29):

$$\begin{aligned}\operatorname{Cov}\left(\mu_i^*, X_{js}\right) &= z_1\operatorname{Cov}\left(X_{iM}, X_{js}\right) + z_2\operatorname{Cov}\left(X_{MM}, X_{js}\right)\\ &= z_1\left(b + (a-b)\delta_{ij} + \frac{1}{t}(c + (s^2 - c)\delta_{ij})\right) + z_2 d, \qquad (32)\end{aligned}$$

where $d$ is the last member of (29).
Hence, from (25) for $i = j$ and for $i \neq j$, using (30):

$$z_1\left(a + \frac{1}{t}s^2\right) + z_2 d = a \qquad (33)$$

$$z_1\left(b + \frac{1}{t}c\right) + z_2 d = b. \qquad (34)$$

Then $z_1, z_2, z_3$ result from (31), (33), (34).
If $m$, in (23), is estimated and replaced by $X_{MM}$, we obtain

$$\mu_i^* = z_1 X_{iM} + (1 - z_1)X_{MM} \qquad (i = 1, \ldots, k). \qquad (35)$$

Then we only need $z_1$. From (33) and (34) results, taking the difference of these relations:

$$z_1 = \frac{(a-b)t}{(s^2-c)+(a-b)t} \tag{36}$$

In the roulette case, (20) and (22) imply $z = z_1, \widehat{\mu}_i = \mu_i^*$.

*Best unbiased homogeneous linear estimator.*

The estimator $\mu_i^*$ defined by (35) is the best (least squares sense) unbiased approximation of $\mu(\Theta_i)$, linear homogeneous in the observable random variables $X_{js}$ $(j = 1, \ldots, k;\ s = 1, \ldots, t)$. This results from Theorem 6 of De Vylder e.a. (1992). By that theorem it is enough to verify that

$$\begin{aligned} E(\mu(\Theta_i) - \mu_i^*)X_{js} &= E((\mu(\Theta_i) - m) - (\mu_i^* - m))\big(X_{js} - m\big) \\ &= \operatorname{Cov}\big(\mu(\Theta_i), X_{js}\big) - \operatorname{Cov}\big(\mu_i^*, X_{js}\big) \end{aligned}$$

does not depend on the subscripts $j$, $s$. By (27), (29), (30), it is easily verified that this is so.

F. De Vijlder, H. Cossette,
U.C.L. Place des Doyens 1
B-1348 Louvain-la-Neuve

## References

1 Bühlmann H. (1967). Experience rating and credibility. The ASTIN Bulletin, vol IV, part II, 199–207.

2 De Vylder F., Goovaerts M. (1992). Optimal parameter estimation under zero excess assumptions in a classical model. IME 11, 1–6.

## Abstract

We show how Bühlmann's model in credibility theory can be adapted in order to cope with contracts that are not necessarily stochasticly independent. As illustration, outside the insurance world, we estimate the probability that the ball falls in a fixed hole of an imperfect casino roulette.
We also indicate how the dependence between the contracts could be detected in a distribution-free set-up.

## Zusammenfassung

Das Modell von Bühlmann in der Credibility-Theorie kann so angepasst werden, dass es auf nicht notwendig unabhängige Verträge angewendet werden kann. Zur Illustrierung ausserhalb des Versicherungswesens schätzen wir die Wahrscheinlichkeit, dass die Kugel bei einem unvollkommenen Roulette in ein festes Loch fällt.
Wir weisen ebenfalls darauf hin, wie die Abhängigkeit zwischen Verträgen ohne die Kenntnis von Verteilungen gefunden werden kann.

## Résumé

Nous montrons comment le modèle de Bühlmann de la théorie de la crédibilité peut être modifié afin qu'il puisse également être appliqué lorsque l'indépendance des contrats d'assurance n'est pas garantie. Comme illustration hors du domaine de l'assurance, nous estimons la probabilité que la boule d'une roulette de casino non parfaite s'arrête dans une case donnée.
Nous montrons également comment la dépendance des contrats peut être détectée lorsque les fonctions de distributions ne sont pas connues.