Iterative credibility

B.
Wissenschaftliche Mitteilungen

Iterative Credibility

By Fl. De Vijlder

Abstract

We consider a credibility theory model defined by random variables ${}_iX_s$ ($i = 1, 2, \ldots, k$; $s = 1, 2, \ldots, t$). We interpret i as the class-index and s as the time-index. We write ${}_iX$ for the column $({}_iX_1, {}_iX_2, \ldots, {}_iX_t)'$ and we call this column the i-th class. To ${}_iX$ is associated the structure variable ${}_i\Theta$. We set $\mu({}_i\Theta) = E({}_iX_s/{}_i\Theta)$, supposed to be independent of s (time homogeneous model).

Let us denote by $K = \{1, 2, \ldots, k\}$ the set of class-indices. For each non-void subset $I \subset K$, we may restrict our classes to those with index in I. The subset of classes $\{{}_iX/i \in I\}$ is a credibility theory model on its own. We can consider in it a credibility estimator, say ${}_{iI}M$ for $\mu({}_i\Theta)$, $(i \in I)$. Then, if we return to the whole set of classes, we may consider the best linear approximation

$${}_{iI}N = \alpha + \beta\, {}_{iI}M$$

to $\mu({}_i\Theta)$. In order to find the coefficients α, β in ${}_{iI}N$, we again use credibility methods. This makes our title "Iterative credibility" clear. Thus, by iterated credibility methods, we may force ${}_{iI}M$ to be used in an estimator for $\mu({}_i\Theta)$.

These considerations extend if we use more than one subset of classes. For example, we can construct the best linear approximation

$${}_{iIJ}N = \alpha + \beta\, {}_{iI}M + \gamma\, {}_{iJ}M, \quad (i \in I, J; I, J \subset K)$$

to $\mu({}_i\Theta)$. This gives a solution to a problem by *Straub* (1975) that we quote here: "If the collective is structured not only by one but by two criteria simultaneously (e.g. horsepower and driver's age), we would also like to have a nice and explicit rule for the calculation of the pure risk premium according to which one should take say.

γ% of the individual loss ratio

$+\delta$% of the loss ratio of all risks with the same horse-power class (1st criterion)

$+\varepsilon$% of the loss ratio of all risks with the same age (2nd criterion)

$+\phi$% of some overall loss ratio."

Indeed, nothing has to be changed if the classes ${}_{ij}X$ are doubly indexed. Then we may take, in the formula (4) of this paper,

$$I = \{(i, 1), (i, 2), \ldots\}; \qquad J = \{(1, j), (2, j), \ldots\}.$$

In this paper, we present the basic ideas in a simple case. Several extensions seem possible (n-way classifications, repeated iterations, regression assumption, multidimensional theory, ...).

Theorically, an estimator obtained iteratively can never be closer to $\mu({}_i\Theta)$ than the usual inhomogeneous credibility estimator. However, when we make this statement, we implicitly assume that the parameters involved are expressed exactly. Practically, these parameters are replaced by estimates. This may change the quality of the estimator. An illustration of this situation is given by the homogeneous and the inhomogeneous credibility estimators for $\mu({}_i\Theta)$. Theorically, the inhomogeneous estimator is best. But, practically, the homogeneous estimator gives automatically the optimal estimation for $m = E(\mu({}_i\Theta))$ (*De Vylder*, to be published). It is not excluded that iteratively constructed estimators, based on intuitively evident classifications, have similar properties. A wrong estimation of the parameters might be corrected automatically, in some way. Of course, this question should be examined more closely.

1. Model. Notations. Previous Results

1.1 Our model is that considered in *Bühlmann* and *Straub* (1970) in the case of uniform as-if statistics. It is defined by the observable random variables ${}_iX_s$ $(i = 1, 2, \ldots, k;\ s = 1, 2, \ldots, t)$ and the structure variables ${}_i\Theta$, subject to the conditions

(i) $({}_1X, {}_1\Theta), ({}_2X, {}_2\Theta), \ldots, ({}_kX, {}_k\Theta)$ are independent,

where $${}_iX = ({}_iX_1, {}_iX_2, \ldots, {}_iX_t)'.$$

(ii) ${}_1\Theta, {}_2\Theta, \ldots, {}_k\Theta$ have the same distribution. (These variables might be multidimensional.)

(iii) $$\mu({}_i\Theta) = E({}_iX_s/{}_i\Theta) \text{ does not depend on } s.$$

(iiii) $$\mathrm{COV}({}_iX_r, {}_iX_s/{}_i\Theta) = \delta_{rs}\,\sigma^2({}_i\Theta)\,\frac{1}{{}_ip_s}.$$

1.2. We use the notations

$$m = E(\mu({}_i\Theta)) = E({}_iX_s); \qquad s^2 = E(\sigma^2({}_i\Theta)); \qquad a = \mathrm{VAR}(\mu({}_i\Theta));$$

$${}_ip_\Sigma = \sum_{s=1}^{t} {}_ip_s; \qquad {}_iX_E = \frac{1}{{}_ip_\Sigma} \sum_{s=1}^{t} {}_ip_s\, {}_iX_s$$

$${}_iz = a\left(a + \frac{s^2}{{}_ip_\Sigma}\right)^{-1}; \quad {}_\Sigma z = \sum_{i=1}^{k} {}_iz; \quad M = \frac{1}{{}_\Sigma z} \sum_{i=1}^{k} {}_iz\, {}_iX_E.$$

1.3. The following results are known.

$${}_iM = (1 - {}_iz)\, M + {}_iz\, {}_iX_E$$

is the best homogeneous linear unbiased approximation to $\mu({}_i\Theta)$ in terms of the observable variables ${}_jX_s$ (*Bühlmann* and *Straub*, 1970). The estimator M is an optimal estimator, in the sense of minimum-variance, for m (*De Vylder*, to be published). We call M the *credibility mean* over the whole set of classes.

2. The Restricted Model

We denote by $K = \{1, 2, \ldots, k\}$, the set of class-indices. For any non-void subset $I \subset K$, we may restrict our classes to the set $\{{}_iX / i \in I\}$. In this restricted model,

$$m, s^2, a, {}_ip_\Sigma, {}_iX_E, {}_iz \qquad (i \in I)$$

are the same as in the initial model, but the optimal estimator for m becomes

$${}_IM = \frac{1}{{}_Iz} \sum_{i \in I} {}_iz\, {}_iX_E, \qquad \text{(credibility mean over the classes with index in } I)$$

where, for each $I \subset K$, we write

$${}_Iz = \sum_{i \in I} {}_iz. \qquad \text{(In particular, } {}_\Sigma z = {}_Kz.)$$

For $i \in I$, the homogeneous credibility estimator for $\mu({}_i\Theta)$, in the restricted model, is

$${}_{iI}M = (1 - {}_iz)\, {}_IM + {}_iz\, {}_iX_E.$$

3. The Problem and its Solution

3.1. Let $i \in K$ be fixed. Let I, J be subsets of K having only i in common: $I \cap J = \{i\}$. We search the best approximation to $\mu({}_i\Theta)$ of the particular form

$$ {}_{IJ}N = \alpha\, m + \beta\, {}_{iI}M + \gamma\, {}_{iJ}M . $$

3.2. By least-squares theory, α, β, γ are solutions to the system

$$ 1 = \alpha + \beta + \gamma , $$

$$ \mathrm{COV}(\mu({}_i\Theta), {}_{iI}M) = \beta\ \mathrm{VAR}({}_{iI}M) + \gamma\ \mathrm{COV}({}_{iI}M, {}_{iJ}M), $$

$$ \mathrm{COV}(\mu({}_i\Theta), {}_{iJ}M) = \beta\ \mathrm{COV}({}_{iI}M, {}_{iJ}M) + \gamma\ \mathrm{VAR}({}_{iJ}M). $$

3.3. We denote:

$$ \dot{I} = \mathrm{I} - \{i\}, \qquad \dot{J} = J - \{i\}. $$

Then

$$ {}_{iI}M = (1 - {}_iz)\frac{1}{{}_Iz} \sum_{n \in \dot{I}} {}_nz\ {}_nX_E + \frac{{}_iz}{{}_Iz}(1 + {}_{\dot{I}}z)\, {}_iX_E . $$

We have

$$ \mathrm{COV}(\mu({}_i\Theta), {}_iX_s) = \mathrm{VAR}(\mu({}_i\Theta)) = a $$

and then

$$ \begin{aligned} \mathrm{COV}(\mu({}_i\Theta), {}_jX_E) &= \delta_{ij}\ \mathrm{COV}(\mu({}_i\Theta), {}_iX_E) \\ &= \delta_{ij} \frac{1}{{}_ip_\Sigma} \sum_s {}_ip_s\ \mathrm{COV}(\mu({}_i\Theta), {}_iX_s) \\ &= \delta_{ij} \frac{1}{{}_ip_\Sigma} \sum_s {}_ip_s\, a = \delta_{ij}\, a . \end{aligned} $$

Therefore

$$ \mathrm{COV}(\mu({}_i\Theta), {}_{iI}M) = a\,\frac{{}_iz}{{}_Iz}(1 + {}_{\dot{I}}z). $$

Similarly,

$$ \mathrm{COV}({}_iX_r, {}_iX_s) = \mathrm{COV}(E({}_iX_r/{}_i\Theta), E({}_iX_s/{}_i\Theta)) + E\ \mathrm{COV}({}_iX_r, {}_iX_s/{}_i\Theta) $$

$$ = \mathrm{VAR}(\mu({}_i\Theta)) + \delta_{rs}\frac{s^2}{{}_ip_s} = a + \delta_{rs}\frac{s^2}{{}_ip_s} $$

and then

$$\mathrm{COV}({}_iX_E, {}_iX_E) = \frac{1}{{}_ip_\Sigma^2} \sum_{r,s} {}_ip_r\, {}_ip_s\, \mathrm{COV}({}_iX_r, {}_iX_s)$$

$$= \frac{1}{{}_ip_\Sigma^2} \sum_{r,s} {}_ip_r\, {}_ip_s \left(a + \delta_{rs} \frac{s^2}{{}_ip_s} \right)$$

$$= a + \frac{1}{{}_ip_\Sigma^2} \sum_r {}_ip_r\, s^2 = a + \frac{s^2}{{}_ip_\Sigma} = \frac{a}{{}_iz}$$

Then

$$\mathrm{COV}({}_{iI}M, {}_{iJ}M) = \frac{{}_iz^2}{{}_Iz\, {}_Jz} (1 + {}_Iz)(1 + {}_Jz)\, \mathrm{COV}({}_iX_E, {}_iX_E)$$

$$= \frac{a\, {}_iz}{{}_Iz\, {}_Jz} (1 + {}_Iz)(1 + {}_Jz).$$

Finally,

$$\mathrm{VAR}({}_{iI}M) = \mathrm{COV}({}_{iI}M, {}_{iI}M)$$

$$= \frac{(1 - {}_Iz)^2}{{}_Iz^2} \sum_{i \in I} {}_nz^2\, \mathrm{VAR}({}_nX_E) + \frac{{}_iz^2}{{}_Iz^2} (1 + {}_Iz)^2\, \mathrm{VAR}({}_iX_E)$$

$$= \frac{a}{{}_Iz} (1 + {}_iz\, {}_Iz),$$

after obvious transformations.

3.4. If we use these results in the last equations of 3.2, we have the system

$${}_iz(1 + {}_Iz) = \beta (1 + {}_iz\, {}_Iz) + \gamma\, \frac{{}_iz}{{}_Jz} (1 + {}_Iz)(1 + {}_Jz),$$

$${}_iz(1 + {}_Jz) = \beta\, \frac{{}_iz}{{}_Iz} (1 + {}_Iz)(1 + {}_Jz) + \gamma\, (1 + {}_iz\, {}_Jz).$$

Its solution $(\beta, \gamma) = ({}_{IJ}z, {}_{JI}z)$ is

$${}_{IJ}z = {}_{iI}p / {}_{IJ}q, \quad {}_{JI}z = {}_{iJ}p / {}_{IJ}q \tag{1}$$

where

$$ {}_{iI}p = {}_iz\ {}_Iz(1+{}_Iz^{-1}),\ {}_{iJ}p = {}_iz\ {}_Jz(1+{}_Jz^{-1}), \tag{2} $$

$$ {}_{IJ}q = {}_iz({}_Iz+{}_Jz)\,(1+{}_Iz^{-1}{}_Jz^{-1})+(1+{}_iz)^2. \tag{3} $$

3.5. Finally, if we replace m by its optimal estimator M, we obtain the following estimator for $\mu({}_i\Theta)$:

$$ {}_{IJ}M = (1-\ {}_{IJ}z-{}_{JI}z)\,M+(1-{}_iz)\ {}_{IJ}z\ {}_IM+(1-{}_iz)\ {}_{JI}z\ {}_JM+{}_iz\ ({}_{IJ}z+{}_{JI}z)\ {}_iX_E\,. \tag{4} $$

Here,
M is the credibility mean over the whole set of classes,
${}_IM$ is the credibility mean over the classes with index in I,
${}_JM$ is the credibility mean over the classes with index in J,
${}_iX_E$ is a mean over the variables in the i-th class.
Thus, (4) completed by (1), (2), (3) is a solution to *Straub*'s problem on double classification.
It is verified that the coefficients of M, ${}_IM$, ${}_JM$, ${}_iX_E$ in (4) are never negative and that their sum is 1.

4. Mean Quadratic Errors

4.1. The mean quadratic error in the approximation of $\mu({}_i\Theta)$ by the homogeneous form

$$ Y = \sum_{n=1}^{k} {}_nc\ {}_nX_E\,, \qquad \left(\sum_{n=1}^{k} {}_nc = 1\right) $$

equals, by results in 3.3:

$$ \begin{aligned} \|Y-\mu({}_i\Theta)\|^2 &= E(Y-\mu({}_i\Theta))^2 = E\left[(Y-m)-(\mu({}_i\Theta)-m)\right]^2 \\ &= \mathrm{VAR}\ Y-2\ \mathrm{COV}(Y,\mu({}_i\Theta))+\mathrm{VAR}\ \mu({}_i\Theta) \\ &= a\left(\sum_{n=1}^{k} {}_nc^2/{}_nz-2\ {}_ic+1\right). \end{aligned} $$

4.2. This mean quadratic error is minimum if Y is the credibility approximation

$$ \begin{aligned} {}_iM &= (1-{}_iz)M+{}_iz\ {}_iX_E \\ &= (1-{}_iz)\frac{1}{{}_Kz}\sum_{n\in\dot{K}} {}_nz\ {}_nX_E+\frac{{}_iz}{{}_Kz}(1+{}_{\dot{K}}z)\ {}_iX_E\,, \qquad (\dot{K} = K-\{i\}). \end{aligned} $$

Then

$$\frac{1}{a}\,\|{}_iM-\mu({}_i\Theta)\|^2 = (1-{}_iz)^2\,\frac{{}_{\dot K}z}{{}_Kz^2}+\frac{{}_iz}{{}_Kz^2}\,(1+{}_{\dot K}z)^2-2\,\frac{{}_iz}{{}_Kz}\,(1+{}_{\dot K}z)+1$$

$$= (1-{}_iz)\,(1+{}_{\dot K}z)/{}_Kz\,.$$

4.3. With the obvious meaning of ${}_Kc, {}_Ic, {}_Jc, {}_ic$, the iterative credibility approximation ${}_{IJ}M$ (4) can be written[1]

$${}_{IJ}M = {}_Kc\;M+{}_Ic\;{}_IM+{}_Jc\;{}_JM+{}_ic\;{}_iX_E\,.$$

Then, if we set $\dot K = K-(I\cup J)$, we have

$${}_{IJ}M = \frac{{}_Kc}{{}_Kz}\sum_{n\,\in\,\ddot K}{}_nz\;{}_nX_E+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Ic}{{}_Iz}\right)\sum_{n\,\in\,\dot I}{}_nz\;{}_nX_E$$

$$+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Jc}{{}_Jz}\right)\sum_{n\,\in\,\dot J}{}_nz\;{}_nX_E+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Ic}{{}_Iz}+\frac{{}_Jc}{{}_Jz}+\frac{{}_ic}{{}_iz}\right){}_iz\;{}_iX_E\,.$$

Then, by the result in 4.1, we have

$$\frac{1}{a}\|\,{}_{IJ}M-\mu({}_i\Theta)\,\|^2 = \left(\frac{{}_Kc}{{}_Kz}\right)^2{}_{\ddot K}z+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Ic}{{}_Iz}\right)^2{}_{\dot I}z+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Jc}{{}_Jc}\right)^2{}_{\dot J}z$$

$$+\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Ic}{{}_Iz}+\frac{{}_Jc}{{}_Jz}+\frac{{}_ic}{{}_iz}\right)^2{}_iz\;-2\left(\frac{{}_Kc}{{}_Kz}+\frac{{}_Ic}{{}_Iz}+\frac{{}_Jc}{{}_Jz}+\frac{{}_ic}{{}_iz}\right){}_iz+1\,.$$

This expression can perhaps be cast in a simpler form, but numerically, it can be used as it stands.

4.4. By the expressions in 4.2 and 4.3, we can calculate the relative difference of the mean quadratic errors in the approximation of $\mu({}_i\Theta)$ by ${}_iM$ and ${}_{IJ}M$ respectively, i. e.

$$r = \left[\|{}_{IJ}M-\mu({}_i\Theta)\|^2-\|{}_iM-\mu({}_i\Theta)\|^2\right]/\|{}_iM-\mu({}_i\Theta)\|^2\,.$$

In order to have an idea of the value of r, let we suppose that ${}_iz = z$, independent of i. Let m, n be the number of indices in I, J resp. Then the following table gives the value of r (independent of a) for different values of z, m, n, k (the number of indices in K).

z	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.1	0.1
m	10	10	10	5	5	5	2	2	2	2	2	500
n	10	10	10	10	10	10	10	10	10	100	2	500
k	10[illegible]	100	100	50	50	50	20	20	20	1000	1000	1000
r	0.0[illegible]	0.03	0.004	0.04	0.03	0.005	0.02	0.03	0.004	0.03	0.03	0.0007

We also examined some cases with different ${}_iz$ and we always found small values for r, of the same order of magnitude as in similar cases with equal ${}_iz$.

[1] The notaion ${}_Kc$, ${}_Ic$, ... is not very adequate if i, I, J may vary since ${}_Kc$, ... depends on i, I and J. Of course, here i, I, J are kept fixed.

5. Conclusion

Practicaly ${}_{IJ}M$ is as close to $\mu({}_i\Theta)$ as ${}_iM$. In consideration of its intuitive appeal, ${}_{IJ}M$ can be prefered to ${}_iM$ in some concrete cases, especially in cases of double classification (*Straub*'s problem).

Bibliography

Bühlmann H. and *Straub E.* (1970): Glaubwürdigkeit für Schadensätze, Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker.
De Vylder Fl. (to be published): Parameter estimation in credibility theory.
Straub E. (1975): Credibility in practice. Paper in the book "Credibility" edited by P.M. Kahn, Academic Press.

Floriaen De Vijlder
Chargé de Cours
à l'Université de Louvain
470, Steenweg op Aalst
B-9400 Ninove

Zusammenfassung

Der Autor löst und verallgemeinert erstmals das Problem der «zweidimensionalen Credibility» (d.h.: Ein Versicherungsbestand wird nach zwei statt nach einem einzigen Kriterium in Klassen aufgeteilt, vgl. Problemstellung in den «Mitteilungen» 1976, Heft 1, Seite 103).

Résumé

L'auteur résout et généralise pour la première fois le problème de la «crédibilité bidimensionelle» (un portefeuille d'assurance est réparti en différentes catégories non plus selon un, mais deux critères. Cf. problématique dans le cahier 1, p. 103 du «Bulletin» de 1976).

Riassunto

L'autore risolve e generalizza per la prima volta il problema della «credibilità bidimensionale» (cioè: un portafoglio viene suddiviso in classi secondo due criteri, invece di uno; cf. problematica in «Mitteilungen» 1976, fascicolo 1, pag. 103).

Summary

For the first time, the two-dimensional credibility problem is formularized and solved by the author (i.e. an insurance portfolio is classified according to two criteria instead of one. Cf. problem discussion in «Mitteilungen» 1976, No. 1, p. 103).