CREDIBILITY PROCEDURES

HANS BÜHLMANN
SWISS FEDERAL INSTITUTE OF TECHNOLOGY

1. Historical review

The concept of credibility is used by actuaries to estimate expected values (net premiums) from statistical data. The first papers on the subject were written by Whitney [8] and Perryman [7]. In the 1950's Arthur Bailey [1] in two special parametric cases gave a mathematical model from which the credibility procedures could be justified. Only in the last few years a nonparametric credibility theory was developed [3] which is now being further refined [4], [5]. The technique derived from the theory is becoming a major actuarial tool in non-life insurance.

As will be apparent from the formulation below, the method of estimation which actuaries mean when referring to credibility procedures is of quite general interest and can easily be transcribed to other fields of application where it may be used for forecasting. For reasons of intuitive appeal I shall, however, restrict the terminology to the actuarial application. The presentation here given follows in many respects that in [4] (in German).

2. The problem

For $i = 1, \cdots, n; j = 1, \cdots, N$, we consider random variables $X_{i,j}$, nonnegative real numbers $P_{i,j}$, and maps $\rho_{i,j}$ from R^∞ to R^1, where we think of j as indicating the risk (or risk group) within the collective of risks and i as indicating the period (say year) over which these risks (risk groups) can be observed. The above introduced abstract concepts have the following intuitive meaning:

$X_{i,j}$ is the observable risk performance (year i, risk j),

$P_{i,j}$ is the measure of exposure (year i, risk j), and

$\rho_{i,j}$ is the map assigning the risk performance $X_{i,j}$ to the doubly stochastic sequence of individual claims (year i, risk j). (We call $\rho_{i,j}$ the insurance conditions.)

Finally we introduce a parameter $\vartheta_{i,j}$ taking values in an abstract θ, $i = 1, \cdots, n$; $j = 1, \cdots, N$, and we think of it as characterizing the "quality" of the risk j in the year i. Using all symbols just introduced we write

$$F_\vartheta(x|P, \rho) = P[X_{i,j} \leqq x|\vartheta_{i,j} = \vartheta, P_{i,j} = P, \rho_{i,j} = \rho] \tag{2.1}$$

for the probability distribution function of $X_{i,j}$ given ϑ, P, ρ.

We now introduce the year $i = 0$. It is the one (just beginning now) for which a forecast (rating) has to be made. Actuaries would call it the rating year (rating period). On the other hand the positive values of the index i are thought of as referring to the past years (observation period) in the "reversed natural time order." We also reserve the index k for the one risk which we want to rate.

Mathematically we are faced with the following estimation problem: estimate $\mu(\vartheta_{0,k}, P_{0,k}, \rho_{0,k}) = E[X_{0,k}|\vartheta_{0,k}, P_{0,k}, \rho_{0,k}]$, where $P_{0,k}$ and $\rho_{0,k}$ (exposure and insurance conditions) are supposed to be *known* and $\vartheta_{0,k}$ (quality) is *unknown*. The data available for this estimation are:

the observations on $\mathscr{X} = \{(X_{i,j}, i = 1, \cdots, n; j = 1, \cdots, N)\}$,
the past exposures $\mathscr{P} = \{(P_{i,j}, i = 1, \cdots, n; j = 1, \cdots, N)\}$,
and the past insurance conditions $\mathscr{R} = \{(\rho_{i,j}, i = 1, \cdots, n; j = 1, \cdots, N)\}$.

Observe that we are treating here only the case of estimating the mean. In [5] this method is extended to estimating $\mu(\vartheta_{0,k}, P_{0,k}, \rho_{0,k}) + \beta\sigma^2(\vartheta_{0,k}, P_{0,k}, \rho_{0,k})$ and in principle one might try to estimate any functional of $F_{\vartheta_{0,k}}(x|P_{0,k}, \rho_{0,k})$ by a method similar to the one described here.

3. Assumptions

The following properties are assumed to hold throughout this paper.

(i) *Independence of risk performances.* Given the values of the parameter $\vartheta_{i,j}$ the risk performances $X_{i,j}$ shall be independent for all i and j.

(ii) *Homogeneity in time.* The quality of each risk shall not vary in time, that is, $\vartheta_{i,j} = \vartheta_j$ independent of i for all $j = 1, \cdots, N$.

(iii) *Independence of parameter values.* The parameters ϑ_j, $j = 1, \cdots, N$, are independent random variables all obeying the same distribution ν with $\nu(M) = P[\vartheta_j \in M]$ for all measurable subsets of $M \subset \theta$.

(iv) *Existence of as-if-statistics per risk.* This means that it shall be possible to reconstruct "artificial statistics" as if the insurance conditions of each risk had always been the same, that is, $\rho_{i,j} = \rho_j$ independent of i for all $j = 1, \cdots, N$ and in particular $\rho_{0,k} = \rho_k$ for the risk to be rated.

The reader is referred to [4] for a discussion of these assumptions from an actuarial viewpoint. In particular the distribution ν is thought of as an idealized version of frequencies of risk qualities in the *collective* of risks from which the *portfolio* containing the risks j, where $j = 1, \cdots, N$, has been drawn. We therefore call $\nu(M)$ the *structural distribution* of the collective. Classical actuarial estimating methods implicitly or explicitly always assume this structural distribution to degenerate, which then leads to the fiction of a "homogeneous collective."

4. Equilibrium

Denote by X an element of $\mathscr{X}$, P an element of $\mathscr{P}$, and R an element of $\mathscr{R}$ and let $\bar{P} = (P, P_{0,k})$ and $\bar{R} = (\rho_1, \cdots, \rho_N)$. We then write $\hat{\mu}_k(X; \bar{P}; \bar{R})$ for the estimate of $\mu(\vartheta_k, P_{0,k}, \rho_k) = \mu_k$ and call μ_k the *correct premium* and $\hat{\mu}_k$ the *rating*. For our further investigations it is then important to distinguish the two cases and whether the structural distribution is *known* or *unknown*.

4.1. *Structural distribution known.* Given the structural distribution, we call a rating $\hat{\mu}_k$ *in equilibrium over a* (measurable) *set* $S \subset \mathscr{X} \times \theta^N$ if

$$\int_S \hat{\mu}_k(X; \bar{P}; \bar{R})\, dP = \int_S \mu(\vartheta_k, P_{0,k}, \rho_k)\, dP \tag{4.1}$$

where, with the usual abuse of notation,

$$dP = \prod_{j=1}^{N} \left(\prod_{i=1}^{n} dF_{\vartheta_j} \left[x_{i,j} | P_{i,j}, \rho_j \right] \right) dv(\vartheta_j). \tag{4.2}$$

From this form of the probability P it follows in particular, that(ϑ_1, X_1), $(\vartheta_2, X_2), \cdots, (\vartheta_N, X_N)$ are independent random variables.

Observe that we write here X_k for $(X_{1,k}, X_{2,k}, \cdots, X_{n,k})$. Later on we shall also write $\bar{P}_k$ for $(P_{0,k}, P_{1,k}, \cdots, P_{n,k})$.

We now ask that the rating $\hat{\mu}_k$ be in equilibrium for all cylinders $S \subset \mathscr{X} \times \theta^N$ with basis in $\mathscr{X}$. This means intuitively that the difference between correct premium and rating should average out to zero over any part of the collective which is characterized by experience alone. If we require this, we know from the Radon-Nikodym theorem that we must have

$$\hat{\mu}_k(X; \bar{P}; \bar{R}) = E[\mu(\vartheta_k, P_{0,k}, \rho_k) | X_1, \cdots, X_N] \tag{4.3}$$

and by the independence of $\{(\vartheta_k, X_k), k = 1, \cdots, N\}$ we have

$$\hat{\mu}(X; \bar{P}; \bar{R}) = E[\mu(\vartheta_k, P_{0,k}, \rho_k) | X_k]. \tag{4.4}$$

To express that the right side only depends on the observations made on the risk k alone, we write for the rating $\hat{\mu}_k$

$$\hat{\mu}_k(X_k; \bar{P}_k; \rho_k) = E[\mu(\vartheta, P_{0,k}, \rho_k) | X_k] \tag{4.5}$$

where we further drop the index k from ϑ_k to express that the structural distribution v is the same for all $\vartheta_j, j = 1, \cdots, N$. Observe that the rating formula (4.5) could also be derived from the postulate

$$\int_{\mathscr{X} \times \theta^N} [\hat{\mu}_k(X; \bar{P}; \bar{R}) - \mu(\vartheta_k, P_{0,k}, \rho_k)]^2\, dP = \text{minimum}, \tag{4.6}$$

but we prefer the equilibrium argument which corresponds to the requirements for a rating in practical applications.

4.2. *Structural distribution unknown.* Under 4.1 we have derived the rating $\hat{\mu}_k$ and we have found that it is a function depending on the arguments $(X_k; \bar{P}_k; \rho_k)$.

The *form of this function* is however *undetermined* if the structural distribution is *not* given. In this case we follow the basic idea of Robbins [6] and use the full information $(X, \bar{P}, \bar{R})$ to *estimate* the form of the rating function. Instead of discussing this in full generality we will show how to proceed in the case of the "credibility rating" which we now want to present.

5. Credibility rating, generalities

In many actuarial applications it can at best be hoped that first and second moments relating to the structural distribution are known. We therefore want to develop suitable *approximations* to the rating $\hat{\mu}_k$, which itself can only be computed if the whole structural distribution is given. The easiest way to get such approximations is by restricting admissible predictors of the correct premium to the linear form. Following standard actuarial terminology we speak then of "credibility."

A *credibility rating* μ_k^* is defined as

$$\mu_k^*(X; \bar{P}; \bar{R}) = \sum_{i,j} \alpha_{i,j} X_{i,j} + \beta_k \tag{5.1}$$

where $\alpha_{i,j}$ and β_k are constants, and we say that μ_k^* approximates $\hat{\mu}_k$ best if

$$E[\mu_k^*(X; \bar{P}; \bar{R}) - \hat{\mu}_k(X_k; P_k; \rho_k)]^2 \tag{5.2}$$

is smallest among all credibility ratings μ_k^*. The expected value operation over $\mathscr{X} \times \theta^N$ is denoted by $E[\cdot]$ as in Section 4. Since

$$\begin{aligned} E[\mu_k^*(X; \bar{P}; \bar{R}) - \mu(\vartheta_k, P_{0,k}, \rho_k)]^2 &= E[\mu_k^*(X; \bar{P}; \bar{R}) - \hat{\mu}_k(X_k; \bar{P}_k; \rho_k)]^2 \\ &\quad + E[\hat{\mu}_k(X_k; \bar{P}_k; \rho_k) - \mu(\vartheta_k, P_{0,k}, \rho_k)]^2, \end{aligned} \tag{5.3}$$

the best credibility rating can also be derived from having

$$E[\mu_k^*(X; \bar{P}; \bar{R}) - \mu(\vartheta_k, P_{0,k}, \rho_k)]^2 \tag{5.4}$$

smallest among all μ_k^*.

In order to render our computations easier we are in the following also assuming some special properties of the distributions relating to $X_{i,j}$, $i = 1, \cdots, n; j = 1, \cdots, N$. (a) $E[X_{i,j} | \vartheta_{i,j} = \vartheta, P_{i,j} = P, \rho_{i,j} = \rho] = \mu(\vartheta, \rho)$ independent of P and (b) $\operatorname{Var}[X_{i,j} | \vartheta_{i,j} = \vartheta, P_{i,j} = P, \rho_{i,j} = \rho] = \sigma^2(\vartheta, \rho)/P$. These assumptions (expressing intuitively that $X_{i,j}$ is some "kind of an average") are discussed in [2]. Finally we introduce the observations $E[\mu(\vartheta, \rho_j)] = m_j$, $\operatorname{Var}[\mu(\vartheta, \rho_j)] = w_j$, and $E[\sigma^2(\vartheta, \rho_j)] = v_j$.

6. Credibility rating if m_j, v_j, w_j are given

Since

$$E[\mu_k^*(X; \bar{P}; \bar{R}) - \mu(\vartheta_k, \rho_k)]^2 \tag{6.1}$$

$$= E\left[\sum_i \alpha_{i,k}X_{i,k} + \beta_k + \sum_{i,j\neq k} \alpha_{i,j}E(X_{i,j}) + \sum_{i,j\neq k} \alpha_{i,j}[X_{i,j} - E(X_{i,j})] - \mu(\vartheta_k, \rho_k)\right]^2$$

$$= E\left[\sum_{i,j\neq k} \alpha_{i,j}[X_{i,j} - E(X_{i,j})]\right]^2 + E\left[\sum_i \alpha_{i,k}X_{i,k} + \beta_k' - \mu(\vartheta_k, \rho_k)\right]^2,$$

where for brevity $\beta_k' = \beta_k + \sum_{i,j\neq k} \alpha_{i,j}E(X_{i,j})$, the optimal μ_k^* can be searched for among those linear estimators for which $\alpha_{i,j} = 0$ for $j \neq k$. Hence we have

$$E\left[\sum_i \alpha_{i,k}X_{i,k} + \beta_k - \mu(\vartheta_k, \rho_k)\right]^2 = E\left[\sum_i \alpha_{i,k}[X_{i,k} - \mu(\vartheta_k, \rho_k)]\right]^2 + E\left[\left(\sum_i \alpha_{i,k} - 1\right)\mu(\vartheta_k, \rho_k) + \beta_k\right]^2 = \sum_i \alpha_{i,k}^2 \frac{v_k}{P_{i,k}} + \left(\sum_i \alpha_{i,k} - 1\right)^2 w_k. \tag{6.2}$$

Put $\beta_k = (1 - \Sigma_i \alpha_{i,k})m_k$, which minimizes the second term on the right of (6.2). The minimum is achieved for

$$\alpha_{i,k} = \frac{P_{i,k}w_k}{v_k + \sum_i P_{i,k}w_k}. \tag{6.3}$$

Hence we have for the credibility rating

$$\mu_k^* = \frac{\sum_i P_{i,k}X_{i,k}}{\sum_i P_{i,k}} \frac{\sum_i P_{i,k}w_k}{v_k + \sum_i P_{i,k}w_k} + m_k \frac{v_k}{v_k + \sum_i P_{i,k}w_k}. \tag{6.4}$$

Observe that we have obtained a *weighted average* between the individual average experience and the theoretical mean over the whole collective. It is customary to call the weight attached to the individual average experience the *credibility factor* which thus turns out to be $\Sigma_i P_{i,k}w_k/(v_k + \Sigma_i P_{i,k}w_k)$.

7. Credibility rating if m_j, v_j, w_j are not given

As indicated in Section 4 we will now use all the observations (including those on the "other risks" $j \neq k$) to estimate these quantities. Of course if such a procedure should be meaningful the risks must be "comparable" in some sense. This means that we must have some knowledge about the functional relationship between the $m_j, j = 1, \cdots, N$ (and similarly for the v_j and the w_j). The presentation of the basic ideas becomes clearest if we assume that $m_j = m$, $v_j = v$ and $w_j = w$ for all j. (This is equivalent to the postulate of equal insurance conditions $\rho_j = \rho$ for all risks in all years under consideration.) We then try to find an estimate by again starting with a linear credibility rating

$$\mu_k^* = \sum_{i,j} \alpha_{i,j}X_{i,j} \tag{7.1}$$

and we postulate

$$E\left[\sum_{i,j} \alpha_{i,j} X_{i,j} - \mu(\vartheta_k, \rho)\right]^2 \tag{7.2}$$

minimum under all credibility ratings of the form (7.1).

As it is by no means clear that the minimum solution of (7.2) is in equilibrium over $\mathscr{X} \times \theta^N$, the whole collective, (contrary to the case treated in Section 6 where this is automatically the case) we require this equilibrium property in addition to (7.2)

$$E\left[\sum_{i,j} \alpha_{i,j} X_{i,j}\right] = m \tag{7.3}$$

which is equivalent to $\Sigma_{i,j}\, \alpha_{i,j} = 1$. Then formulate the Lagrangian

$$\begin{aligned} \phi(\alpha_{1,1}, \cdots, \alpha_{n,1}, \alpha_{1,2}, \cdots, \alpha_{n,2}, \cdots, \alpha_{1,N}, \cdots, \alpha_{n,N}, \alpha) \\ = E\left[\sum_{i,j} \alpha_{i,j} X_{i,j} - \mu(\vartheta_k, \rho)\right]^2 - 2\alpha \sum_{i,j} \alpha_{i,j} E[X_{i,j}] \end{aligned} \tag{7.4}$$

from which the system of equations is obtained

$$\frac{\partial \phi}{\partial \alpha_{\ell,h}} = 2E\left[\left\{\sum_{i,j} \alpha_{i,j} X_{i,j} - \mu(\vartheta_k, \rho)\right\} X_{\ell,h}\right] - 2\alpha E[X_{\ell,h}] = 0, \tag{7.5}$$

$$\ell = 1, \cdots, n; h = 1, \cdots, N.$$

Hence using (7.3) and the conditional independence of all $X_{i,j}$

$$E\left[\sum_{i,j} \alpha_{i,j}\{\mu(\vartheta_h, \rho)[\mu(\vartheta_j, \rho) - \mu(\vartheta_k, \rho)]\} + \alpha_{\ell,h} \frac{\sigma^2(\vartheta_h, \rho)}{P_{\ell,h}}\right] = \alpha m \tag{7.6}$$

and as

$$E[\mu(\vartheta_h, \ell)\mu(\vartheta_j, \rho)] = m^2 + \delta_{h,j} w, \qquad \delta_{h,j} = \begin{matrix} 1, h = j \\ 0, h \neq j, \end{matrix} \tag{7.7}$$

we get

$$\sum_{i=1}^{n} \alpha_{i,h} w + \alpha_{\ell,h} \frac{v}{P_{\ell,h}} = \delta_{h,k} w + \alpha m. \tag{7.8}$$

Hence

$$\alpha_{\ell,h} = \frac{P_{\ell,h}}{v}\left[\left(\delta_{h,k} - \sum_{i=1}^{n} \alpha_{i,h}\right) w + \alpha m\right] = P_{\ell,h} \cdot C, \tag{7.9}$$

(where C is independent of ℓ). Summing over the first index we get

$$\sum_{i=1}^{n} \alpha_{i,h} = \frac{\sum_{i=1}^{n} P_{i,h}}{v}\left[\left(\delta_{h,k} - \sum_{i=1}^{n} \alpha_{i,h}\right) w + \alpha m\right] \tag{7.10}$$

or

$$\sum_{i=1}^{n} \alpha_{i,h}\left(1 + \frac{\sum_i P_{i,h}w}{v}\right) = \frac{\sum_{i=1}^{n} P_{i,h}w}{v}\,\delta_{h,k} + \frac{\sum_{i=1}^{n} P_{i,h}}{v}\,\alpha m \tag{7.11}$$

or equivalently

$$\sum_{i=1}^{n} \alpha_{i,h} = \frac{\sum_i P_{i,h}w}{v + \sum_i P_{i,h}w}\,\delta_{h,k} + \frac{\sum_i P_{i,h}}{v + \sum_i P_{i,h}w}\,\alpha m \tag{7.12}$$

from which

$$\alpha_{\ell,h} = \frac{P_{\ell,h}w}{v + \sum_{i=1}^{n} P_{i,h}w}\,\delta_{h,k} + \frac{P_{\ell,h}}{v + \sum_{i=1}^{n} P_{i,h}w}\,\alpha m, \qquad \ell = 1, \cdots, n;\ h = 1, \cdots, N. \tag{7.13}$$

Let us write $P_{\cdot h}$ for $\Sigma_{i=1}^{n} P_{i,h}$. Then $\Sigma_{i,h}\,\alpha_{i,h} = 1$ yields

$$1 = \frac{P_{\cdot k}w}{v + P_{\cdot k}w} + \sum_{h=1}^{N} \frac{P_{\cdot h}}{v + P_{\cdot h}w}\,\alpha m. \tag{7.14}$$

Hence

$$\alpha m = \frac{\dfrac{v}{v + P_{\cdot k}w}}{\sum_{h=1}^{n} \dfrac{P_{\cdot h}}{v + P_{\cdot h}w}}. \tag{7.15}$$

We finally get

$$\alpha_{\ell,h} = \frac{P_{\ell,k}w}{v + P_{\cdot k}w}\,\delta_{h,k} + \frac{\dfrac{P_{\ell,h}}{v + P_{\cdot h}w}}{\sum_{h=1}^{N} \dfrac{P_{\cdot h}}{v + P_{\cdot h}w}}\,\frac{v}{v + P_{\cdot k}w} \tag{7.16}$$

and

$$\mu_k^* = \sum_{i,h} \alpha_{i,h}X_{i,h} = \frac{P_{\cdot k}w}{v + P_{\cdot k}w}\,\frac{\sum_{i=1}^{n} P_{i,k}X_{i,k}}{\sum_{i=1}^{n} P_{i,k}} + \frac{v}{v + P_{\cdot k}w}\,\frac{\sum_{h=1}^{N}\left(\dfrac{\sum_{i=1}^{n} P_{i,h}X_{i,h}}{\sum_i P_{i,h}}\right)\Pi_h}{\sum_{h=1}^{N} \Pi_h} \tag{7.17}$$

or in shorter form

$$\mu_k^* = \frac{P_{\cdot k}w}{v + P_{\cdot k}w}\,\bar{X}_{\cdot k} + \frac{v}{v + P_{\cdot k}w}\,\bar{X}_{\cdot\cdot}. \tag{7.18}$$

with

$$\bar{X}_{\cdot k} = \frac{\sum_i P_{i,k} X_{i,k}}{\sum_i P_{i,k}}$$

$$\bar{X}_{\cdot\cdot} = \frac{\sum_h \Pi_h \bar{X}_{\cdot h}}{\sum_h \Pi_h} \tag{7.19}$$

$$\Pi_h = \frac{P_{\cdot h}}{v + P_{\cdot h} w}.$$

Observe the very close relationship between (6.4) and (7.18). In the latter the theoretical mean over the collective (which we denoted by m) is replaced by the estimate $\bar{X}_{\cdot\cdot}$. Note, however, that our formula still involves the unknowns v and w. (Actually only the proportion of the two plays a role.) We want to show in the sequel how these quantities as well may be estimated from the observations on $\mathscr{X} = \{X_{i,j}, i = 1, \cdots, n; j = 1, \cdots, N\}$.

8. Estimating v and w

We recall that $v = E[\sigma^2(\vartheta, \rho)]$ and $w = \operatorname{Var}[\mu(\vartheta, \rho)]$. Intuitively v measures the variability within each risk and w the variation between different risks. It is hence natural to consider the following statistics V and W to estimate the two above quantities

$$\text{(a)} \qquad V = \frac{1}{N} \sum_{h=1}^{N} \frac{1}{n-1} \sum_{\ell=1}^{n} \frac{P_{\ell,h}}{P} (X_{\ell,h} - \bar{X}_{\cdot h})^2$$

$$= \frac{1}{N(n-1)} \sum_{\ell,h} \frac{P_{\ell,h}}{P} X_{\ell,h}^2 - \frac{1}{N(n-1)} \sum_{h=1}^{N} \frac{P_{\cdot h}}{P} \bar{X}_{\cdot h}^2,$$

with $P_{\cdot h} = \Sigma_{\ell=1}^n P_{\ell,h}$, $\Sigma_{h=1}^N P_{\cdot h} = P$, and $\bar{X}_{\cdot h} = \Sigma_{\ell=1}^n P_{\ell,h} X_{\ell,h}/P_{\cdot h}$,

and

$$\text{(b)} \quad W = \frac{1}{N(n-1)} \sum_{\ell,h} \frac{P_{\ell,h}}{P} (X_{\ell,h} - \tilde{X})^2 = \frac{1}{N(n-1)} \sum_{\ell,h} \frac{P_{\ell,h}}{P} X_{\ell,h}^2 - \frac{1}{N(n-1)} \tilde{X}^2$$

with $\tilde{X} = \Sigma_{\ell,h} P_{\ell,h} X_{\ell,h}/P = \Sigma_{n=1}^N P_{\cdot h} \bar{X}_{\cdot h}/P$.

Compute then

$$\text{(a)} \qquad N(n-1)E[V|\vartheta_1, \vartheta_2, \cdots, \vartheta_N] = \sum_{\ell,h} \frac{P_{\ell,h}}{P} \left\{ \mu^2(\vartheta_h, \rho) + \frac{\sigma^2(\vartheta_h, \rho)}{P_{\ell,h}} \right\}$$

$$- \sum_{h=1}^{N} \frac{P_{\cdot h}}{P} \left\{ \mu^2(\vartheta_h, \rho) + \frac{\sigma^2(\vartheta_h, \rho)}{P_{\cdot h}} \right\}$$

$$= \frac{n-1}{P} \sum_{h=1}^{N} \sigma^2(\vartheta_h, \rho).$$

(Observe that we have used Var $[\bar{X}_{\cdot h}|\vartheta_h] = \sigma^2(\vartheta_h, \rho)/P_{\cdot h}$ which follows from our assumptions.)

Hence

$$E[V] = \frac{E[\sigma^2(\vartheta, \rho)]}{P} = \frac{v}{P}. \tag{8.1}$$

On the other hand

$$\begin{aligned}
\text{(b)}\quad (nN-1)E[W|\vartheta_1, \vartheta_2, \cdots, \vartheta_N] &= \sum_{\ell,h} \frac{P_{\ell,h}}{P}\left\{\mu^2(\vartheta_h, \rho) + \frac{\sigma^2(\vartheta_h, \rho)}{P_{\ell,h}}\right\} \\
&\quad - \left(\sum_{h=1}^{N} \frac{P_{\cdot h}}{P}\mu(\vartheta_h, \rho)\right)^2 - \sum_{h=1}^{N} \frac{P_{\cdot h}}{P^2}\sigma^2(\vartheta_h, \rho) \\
&= \sum_{h=1}^{N} \frac{P_{\cdot h}}{P}\mu^2(\vartheta_h, \rho) - \left(\sum_{h=1}^{N} \frac{P_{\cdot h}}{P}\mu(\vartheta_h, \rho)\right)^2 \\
&\quad + \sum_{h=1}^{N}\left(\frac{n}{P} - \frac{P_{\cdot h}}{P^2}\right)\sigma^2(\vartheta_h, \rho).
\end{aligned}$$

Hence

$$(nN-1)E[w] = \sum_{n=1}^{N} \frac{P_{\cdot h}}{P}(m^2 + w) - m^2 - \sum_{n=1}^{N} \frac{P_{\cdot h}^2}{P^2} w + (nN-1)\frac{v}{P} \tag{8.2}$$

or

$$E[W] = \frac{v}{P} + \frac{1}{nN-1}\sum_{h=1}^{N} \frac{P_{\cdot h}}{P}\left(1 - \frac{P_{\cdot h}}{P}\right) w \tag{8.3}$$

which we abbreviate

$$E[W] = \frac{v}{P} + \tilde{\Pi} w \qquad \text{with} \quad \tilde{\Pi} = \frac{1}{nN-1}\sum_{h=1}^{N} \frac{P_{\cdot h}}{P}\left(1 - \frac{P_{\cdot h}}{P}\right). \tag{8.4}$$

This leads to the linear equations

$$\begin{aligned}
\frac{\hat{v}}{P} &= V \\
\frac{\hat{v}}{P} + \tilde{\Pi}\hat{w} &= W
\end{aligned} \tag{8.5}$$

for the unbiased and consistent estimators $\hat{v}$ and $\hat{w}$.

There is however one difficulty in using these estimates. The estimator $\hat{w}$ turns out negative whenever $W < V$. This is indeed possible and we therefore modify our estimates as follows: (1) If $W > V$ choose the estimator as given by (8.5).

(2) If $W < V$ put $\hat{w} = 0$, that is, use the credibility rating as for a homogeneous collective

$$\mu_k^* = \tilde{X} = \sum_{\ell,h} \frac{P_{\ell,h}}{P} X_{\ell,h}. \tag{8.6}$$

We believe this procedure to be reasonable in spite of the fact that it destroys the unbiasedness of $\hat{w}$.

For a numerical example where this method of estimation is applied the reader is referred to [4]. We give the results of the computations as presented there. The data are shown in Table I. In the last row we have tabulated $P_{\cdot h} = \sum_\ell P_{\ell,h}$ and $\bar{X}_{\cdot h} = \sum_\ell P_{\ell,h} X_{\ell,h}/P_{\cdot h}$. Using the estimates, formula (8.5), in Section 8 we get

$$\begin{aligned} \hat{v} &= 209.0 \\ \hat{w} &= 12.1. \end{aligned} \tag{8.7}$$

TABLE I

DATA FOR EXAMPLE OF THE METHOD OF ESTIMATION

	Risk h						
Year ℓ	1 P X	2 P X	3 P X	4 P X	5 P X	6 P X	7 P X
5	5 0·0	14 11.3	18 8.0	20 5.4	21 9.7	43 9.7	70 9.0
4	6 0.0	14 25.0	20 1.9	22 5.9	24 8.9	47 14.5	77 9.6
3	8 4.2	13 18.5	23 7.0	25 7.1	28 6.7	53 10.8	85 8.7
2	10 0.0	11 14.3	25 3.1	29 7.2	34 10.3	61 12.0	92 11.7
1	12 7.7	10 30.0	27 5.2	35 8.3	42 11.1	70 13.1	100 7.0
	41 3.1	62 19.5	113 5.0	131 7.0	149 9.5	274 12.1	424 9.2

On the basis of these values for $\hat{v}$ and $\hat{w}$ we tabulate

the credibility factors $\gamma_h = \dfrac{P_{\cdot h} w}{v + P_{\cdot h} w}$ using formula (6.4),

the average experience $\bar{X}_{\cdot\cdot} = \sum_h \dfrac{\prod_h}{\prod} \bar{X}_{\cdot h}$ using formula (7.18), and

the credibility rating $\mu_h^* = \gamma_h \bar{X}_{\cdot h} + (1 - \gamma_h)\bar{X}_{\cdot\cdot}$ using formula (7.18).

The resulting estimates are shown in Table II. It is instructive to compare the last two rows of the table, the credibility rating and the individual average experience.

TABLE II

RESULTING ESTIMATES

	Risk h						
	1	2	3	4	5	6	7
γ_h (in %)	70.4	78.2	86.7	88.4	89.6	94.1	96.1
$\bar{X}_{..}$	9.4	9.4	9.4	9.4	9.4	9.4	9.4
μ_h^*	5.0	17.3	5.6	7.3	9.5	11.9	9.2
$\bar{X}_{.h}$	3.1	19.5	5.0	7.0	9.5	12.1	9.2

REFERENCES

[1] A. L. BAILEY, "Credibility procedures: Laplace generalization of Bayes' rule and the combination of collateral knowledge with observed data," *Proceedings Casualty Actuarial Society*, Vol. 37 (1950), pp. 7–23.

[2] H. BÜHLMANN, "A distribution free method for general risk problems," *ASTIN-Bulletin*, Vol. 3 (1964), pp. 144–152.

[3] ———, "Experience rating and credibility," *ASTIN-Bulletin*, Vol. 4 (1967), pp. 199–227.

[4] H. BÜHLMANN and E. STRAUB, "Glaubwürdigkeit für Schadensätze," *Mitt. Verein. Schweiz. Versich.-math.*, Vol. 70 (1970), pp. 111–133.

[5] H. BÜHLMANN, *Mathematical Methods in Risk Theory*, Heidelberg, Springer Verlag, 1970.

[6] H. ROBBINS, "The empirical Bayes approach to statistical decision problems," *Ann. Math. Statist.*, Vol. 35 (1964), pp. 1–20.

[7] F. S. PERRYMAN, "Some notes on credibility," *Proceedings Casualty Actuarial Society*, Vol. 19 (1932), pp. 65–84.

[8] A. W. WHITNEY, "The theory of experience rating," *Proceedings Casualty Actuarial Society*, Vol. 4 (1918), pp. 274–292.