2004

Credibility Theory and Geometry

Elias S.W. Shiu
University of Iowa, eshiu@stat.uiowa.edu

Fuk Yum Sing
Hong Kong Polytechnic University

Credibility Theory and Geometry

Elias S.W. Shiu* and Fuk Yum Sing†

Abstract‡

We present a geometric approach to studying greatest accuracy credibility theory. Our main tool is the concept of orthogonal projections. We show, for example, that to determine the Bühlmann credibility premium is to find the coefficients of the minimum-norm vector in an affine space spanned by certain orthogonal random variables. Our approach is illustrated by deriving various common credibility formulas. Several equivalent forms of the credibility factor Z are derived by means of similar triangles.

Key words and phrases: *greatest accuracy credibility theory, Bühlmann credibility premium, credibility factor, affine space, inner product, orthogonal projection, Bühlmann-Straub model*

*Elias S. W. Shiu, Ph.D., A.S.A., is Principal Financial Group Foundation Professor of actuarial science at the University of Iowa, and visiting Chair Professor of actuarial science in the Department of Applied Mathematics, Hong Kong Polytechnic University, China. He received a Ph.D. in mathematics from the California Institute of Technology in 1975. From 1976 to 1991, he was a professor of actuarial science at the University of Manitoba and a consultant for the Great-West Life Assurance Company in Winnipeg, Canada. He is an editor of *Insurance: Mathematics and Economics* and a co-editor of the *North American Actuarial Journal.*

Dr. Shiu's address is: Department of Statistics and Actuarial Science, University of Iowa, Iowa City, Iowa 52242-1409, U.S.A. E-mail: *eshiu@stat.uiowa.edu*

†Fuk Yum Sing, M. Phil., is a visiting lecturer in the Department of Applied Mathematics, The Hong Kong Polytechnic University. He received his master's degree in mathematics from the University of Hong Kong in 1977. He joined the then Hong Kong Polytechnic (now The Hong Kong Polytechnic University) in 1976.

Mr. Sing's address is: Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, CHINA.

‡The authors thank the anonymous referees for their insightful comments. Elias Shiu gratefully acknowledges the generous support from the Principal Financial Group Foundation and Robert J. Myers, F.C.A., F.C.A.S., F.S.A.

1 Introduction

Credibility theory, which is called a cornerstone of actuarial science by some authors (Longley-Cook 1962, page 194; Hickman and Heacox 1999, page 1), is a required part of education syllabi of major international professional organizations including the Society of Actuaries, the Institute and Faculty of Actuaries, and the Casualty Actuarial Society. One of the texts recommended by the Society of Actuaries for studying credibility theory is Klugman, Panjer, and Willmot (1998). This text uses a traditional probability/statistics approach to derive credibility formulas. The main purpose of this paper is to present a geometric approach to derive and extend some of the results in Klugman, Panjer, and Willmot (1998, Sections 5.4.2, 5.4.3 and 5.4.4).

The main tool used in this paper is the concept of orthogonal projections. Background materials on the inner product, affine space, and inner product space of square-integrable random variables are presented in Section 2. The assumption of a risk parameter Θ, conditional on which the claims $\{X_j\}$ are independent, implies that the random variables $\{X_j - \mathbb{E}\left[X_j|\Theta\right]\}$ can be viewed as orthogonal vectors. Section 3 shows that to determine the credibility premium is to find the coefficients of the vector with the smallest length in an affine space containing these orthogonal vectors. With the expressions for the optimal coefficients, Section 4 derives various credibility formulas in the Klugman, Panjer, and Willmot textbook. For some readers, Section 5 may be the most intriguing section in this paper. By means of similar triangles, it derives various equivalent forms of the credibility factor Z. Section 6 presents several more interesting formulas.

There are many books and survey articles on credibility theory including: Bühlmann (1970), Kahn (1975), Goovaerts and Hoogstad (1987), Heilmann (1988), Straub (1988), Goovaerts et al. (1990), Venter (1990), Sundt (1993), Waters (1993), Goulet (1998), Klugman, Panjer, and Willmot (1998), Herzog (1999), Kaas et al. (2001), and Mahler and Dean (2001). These authors use probability theory and other tools to develop and explain credibility formulas and concepts. This paper's approach, which de-emphasizes probability theory, may be more appealing to some actuarial practitioners and students.

2 Some Mathematical Preliminaries

2.1 Inner Product Space and Orthogonal Projections

An inner product space is a vector space $\mathbf{V}$ (over the real numbers) together with an inner product (also called scalar product or dot product) defined on $\mathbf{V} \times \mathbf{V}$. Corresponding to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{V}$, the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is a real number. The inner product satisfies the following axioms:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$;
2. $\langle c\mathbf{u}, \mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$ for each real number c;
3. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$;
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$, the zero vector.

The norm (or length) of a vector $\mathbf{u}$ is $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. For each pair of nonzero vectors $\mathbf{u}$ and $\mathbf{v}$, the quantity $\langle \mathbf{u}, \mathbf{v} \rangle / (\|\mathbf{u}\| \, \|\mathbf{v}\|)$ can be interpreted as the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$. If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, we say that the vectors are *orthogonal* and we write $\mathbf{u} \perp \mathbf{v}$. Because

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2 \langle \mathbf{u}, \mathbf{v} \rangle,$$

the vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if the Pythagorean equation holds:

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Let $\mathbf{U}$ be a subspace of an inner product space $\mathbf{V}$ and $\mathbf{v}$ be an arbitrary vector in $\mathbf{V}$. We are interested in finding the vector $\mathbf{u}$ in $\mathbf{U}$ closest to $\mathbf{v}$ in the sense that it minimizes the norm $\|\mathbf{v} - \mathbf{u}\|$. It is not difficult to show (Luenberger 1969, page 50, Theorem 1) that, if there is $\mathbf{u}_0 \in \mathbf{U}$ such that

$$\|\mathbf{v} - \mathbf{u}_0\| \leq \|\mathbf{v} - \mathbf{u}\| \quad \text{for all} \quad \mathbf{u} \in \mathbf{U},$$

then $\mathbf{u}_0$ is unique. Furthermore, a necessary and sufficient condition that $\mathbf{u}_0 \in \mathbf{U}$ is a unique minimizing vector in $\mathbf{U}$ is the following:

$$(\mathbf{v} - \mathbf{u}_0) \perp \mathbf{u} \quad \text{for all} \quad \mathbf{u} \in \mathbf{U}. \tag{1}$$

It is easy to see that two conditions, each of which is equivalent to condition (1), are

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}_0, \mathbf{u} \rangle \quad \text{for all} \quad \mathbf{u} \in \mathbf{U} \tag{2}$$

and

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v} - \mathbf{u}_0\|^2 + \|\mathbf{u}_0 - \mathbf{u}\|^2 \quad \text{for all} \quad \mathbf{u} \in \mathbf{U}. \tag{3}$$

The vector $\mathbf{u}_0$ is called the *orthogonal projection of* $\mathbf{v}$ onto $\mathbf{U}$.

Consider the special case where $\mathbf{U}$ is a one-dimensional subspace spanned by a nonzero vector $\mathbf{u}^*$. Then it follows from equation (2) that the vector $\mathbf{u}_0$ is

$$\frac{\langle \mathbf{v}, \mathbf{u}^* \rangle}{\langle \mathbf{u}^*, \mathbf{u}^* \rangle}\mathbf{u}^* = \left\langle \frac{\mathbf{v}}{\|\mathbf{v}\|}, \frac{\mathbf{u}^*}{\|\mathbf{u}^*\|} \right\rangle \|\mathbf{v}\| \frac{\mathbf{u}^*}{\|\mathbf{u}^*\|}. \tag{4}$$

With the inner product on the right side of equation (4) being interpreted as the cosine of the angle between the vectors $\mathbf{v}$ and $\mathbf{u}^*$, the geometric explanation of the left side of equation (4) is obvious.

2.2 Vector with Minimal Norm in an Affine Space

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be m vectors in a vector space $\mathbf{V}$. The affine space (also called affine set or linear variety) spanned by these vectors is the set of vectors of the form $\sum_{j=1}^{m} c_j \mathbf{v}_j$ with real coefficients $c_1, c_2, \ldots, c_m$ satisfying

$$\sum_{j=1}^{m} c_j = 1. \tag{5}$$

There is no restriction on the sign of the coefficients. Assuming $\mathbf{V}$ is an inner product space and the m vectors are nonzero and mutually orthogonal, we claim the vector

$$\mathbf{w} = \sum_{j=1}^{m} \hat{c}_j \mathbf{v}_j, \tag{6}$$

with

$$\hat{c}_j = \frac{\dfrac{1}{\left\|\mathbf{v}_j\right\|^2}}{\displaystyle\sum_{k=1}^{m} \frac{1}{\|\mathbf{v}_k\|^2}} \quad j = 1, 2, \ldots, m, \tag{7}$$

is the vector having the minimal norm in the affine space spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. To see this, we use the assumption that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are mutually orthogonal to obtain

$$\|\sum_{j=1}^{m} c_j \mathbf{v}_j\|^2 = \sum_{j=1}^{m} c_j^2 \|\mathbf{v}_j\|^2, \tag{8}$$

which is called Parseval's identity. The optimal coefficients $\{\hat{c}_j\}$ are then determined by minimizing the right side of equation (8) subject to the constraint of equation (5). This optimization problem can be readily solved using the method of Lagrange multipliers, and the solution is the system of equations (7).

It follows from equations (6), (7), and (8) that

$$\|\mathbf{w}\|^2 = \frac{1}{\sum_{k=1}^{m} \frac{1}{\|\mathbf{v}_k\|^2}}. \tag{9}$$

Equation (9) shows that $\|\mathbf{w}\|^2$ is $1/m$ of the harmonic mean of $\|\mathbf{v}_1\|^2, \|\mathbf{v}_2\|^2, \ldots, \|\mathbf{v}_m\|^2$.

An alternative approach to deriving the system of equations (7) is to show that $\mathbf{w}$ is the vector of minimal norm in an affine space iff

$$\mathbf{w} \perp (\mathbf{v} - \mathbf{w}) \tag{10}$$

for all vectors $\mathbf{v}$ in the affine space. For further discussion, see Luenberger (1969, page 64).

2.3 Inner Product Space of Random Variables

For a given sample space, the set of square-integrable random variables (random variables with finite variance) forms an inner product space (Luenberger, 1969; Small and McLeish, 1994). For each pair of square-integrable random variables X and Y, the inner product is defined to be $\langle X, Y\rangle = \mathbb{E}[XY]$.

Let g be a function such that $g(Y)$ is a square-integrable random variable. Then, by the law of iterated expectations,

$$\begin{aligned}\langle X, g(Y)\rangle &= \mathbb{E}[Xg(Y)]\\ &= \mathbb{E}[\mathbb{E}[Xg(Y)|Y]]\\ &= \mathbb{E}[\mathbb{E}[X|Y]g(Y)]\\ &= \langle \mathbb{E}[X|Y], g(Y)\rangle.\end{aligned}$$

Hence, $(X - \mathbb{E}[X|Y]) \perp g(Y)$, and we have the Pythagorean equation:

$$\|X - g(Y)\|^2 = \|X - \mathbb{E}[X|Y]\|^2 + \|\mathbb{E}[X|Y] - g(Y)\|^2. \qquad (11)$$

The conditional expectation $\mathbb{E}[X|Y]$ is the orthogonal projection of X onto the subspace of square-integrable functions of Y. Note that, by the law of iterated expectations,

$$\|X - \mathbb{E}[X|Y]\|^2 = \mathbb{E}\left[\mathbb{E}\left[(X - \mathbb{E}[X|Y])^2|Y\right]\right] = \mathbb{E}[\mathbb{V}\mathrm{ar}(X|Y)]. \qquad (12)$$

If g(Y) is the constant random variable that takes the value $\mathbb{E}[X]$, i.e., if g(Y) $\equiv \mathbb{E}[X]$, then equation (11) is the well-known variance decomposition equation

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}[\mathbb{V}\mathrm{ar}(X|Y)] + \mathbb{V}\mathrm{ar}[\mathbb{E}[X|Y]]. \qquad (13)$$

The above can be generalized in various ways. In particular, we have Exercise 5.83(a) in Klugman, Panjer, and Willmot (1998):

$$\|X - g(\mathbf{X})\|^2 = \|X - \mathbb{E}[X|\mathbf{X}]\|^2 + \|\mathbb{E}[X|\mathbf{X}] - g(\mathbf{X})\|^2 \qquad (14)$$

where $\mathbf{X}$ denotes the random variables X_1, X_2, ..., X_n. Also, equations (11), (12), and (13) can be generalized as

$$\begin{aligned}\langle W - f(Y), X - g(Y)\rangle &= \langle W - \mathbb{E}[W|Y], X - \mathbb{E}[X|Y]\rangle \\ &\quad + \langle \mathbb{E}[W|Y] - f(Y), \mathbb{E}[X|Y] - g(Y)\rangle,\end{aligned}$$

$$\langle W - \mathbb{E}[W|Y], X - \mathbb{E}[X|Y]\rangle = \mathbb{E}[\mathbb{C}\mathrm{ov}[W, X|Y]] \qquad (15)$$

and

$$\mathbb{C}\mathrm{ov}[W, X] = \mathbb{E}[\mathbb{C}\mathrm{ov}[W, X|Y]] + \mathbb{C}\mathrm{ov}[\mathbb{E}[W|Y], \mathbb{E}[X|Y]],$$

respectively.

3 Greatest Accuracy Credibility Theory

Following Klugman, Panjer, and Willmot (1998, Chapter 5), let X_j denote the claim amount in the j^{th} period, $j = 1, 2, 3, \ldots$. In greatest accuracy credibility theory the objective is to determine the coefficients

α_0, α_1,..., α_n of the credibility premium for period $(n + 1)$ given the losses in the previous n periods,

$$P_{n+1} = \alpha_0 + \sum_{j=1}^{n} \alpha_j X_j \tag{16}$$

so that the norm

$$\|X_{n+1} - P_{n+1}\| \tag{17}$$

is minimized. Because P_{n+1} is a function of the random variables X_1, X_2,..., X_n, we have a special case of equation (14):

$$\begin{aligned}\|X_{n+1} - P_{n+1}\|^2 &= \|X_{n+1} - \mathbb{E}[X_{n+1}|X_1, X_2, \ldots, X_n]\|^2 \\ &\quad + \|\mathbb{E}[X_{n+1}|X_1, X_2, \ldots, X_n] - P_{n+1}\|^2.\end{aligned} \tag{18}$$

Hence, the credibility premium P_{n+1} can be determined by minimizing

$$\|\mathbb{E}[X_{n+1}|X_1, X_2, \ldots, X_n] - P_{n+1}\|, \tag{19}$$

which is not a surprising result.

As in Section 5.4 of Klugman, Panjer, and Willmot (1998), we assume the existence of a risk parameter random variable Θ, conditional on which the random variables X_1, X_2,..., X_j,... are independent. We write

$$\mu_j(\Theta) = \mathbb{E}\left[X_j|\Theta\right].$$

Thus,

$$\mu_{n+1}(\Theta) = \mathbb{E}[X_{n+1}|\Theta] = \mathbb{E}[X_{n+1}|\Theta, X_1, X_2, \ldots, X_n]$$

because of the conditional independence assumption. By the law of iterated expectations,

$$\begin{aligned}\mathbb{E}[\mu_{n+1}(\Theta)|X_1, X_2, \ldots, X_n] &= \mathbb{E}[\mathbb{E}[X_{n+1}|\Theta, X_1, X_2, \ldots, X_n]|X_1, X_2, \ldots, X_n] \\ &= \mathbb{E}[X_{n+1}|X_1, X_2, \ldots, X_n].\end{aligned}$$

This shows that expression (19) is the same as

$$\|\mathbb{E}[\mu_{n+1}(\Theta)|X_1, X_2, \ldots, X_n] - P_{n+1}\|.$$

Similar to equation (18), we have

$$\|\mu_{n+1}(\Theta) - P_{n+1}\|^2 = \|\mu_{n+1}(\Theta) - \mathbb{E}\left[\mu_{n+1}(\Theta)|X_1, X_2, \ldots, X_n\right]\|^2 + \|\mathbb{E}\left[\mu_{n+1}(\Theta)|X_1, X_2, \ldots, X_n\right] - P_{n+1}\|^2. \quad (20)$$

Therefore, an alternative way to determine the credibility premium is to minimize

$$\|\mu_{n+1}(\Theta) - P_{n+1}\|. \quad (21)$$

By equation (15),

$$\left\langle X_j - \mu_j(\Theta), X_k - \mu_k(\Theta) \right\rangle = \mathbb{E}\left[\mathbb{C}\mathrm{ov}\left[X_j, X_k|\Theta\right]\right],$$

which is zero because of the conditional independence assumption. Hence, the random variables $\{X_j - \mu_j(\Theta)\}$ are mutually orthogonal. This fact will play a key role in determining the credibility premium.

We now follow Klugman, Panjer, and Willmot (1998, Section 5.4) and assume that $\mu_j(\Theta) = \mu(\Theta)$ for $j = 1, 2, 3, \ldots$, and write $\mathbb{E}\left[\mu(\Theta)\right] = \mu$. Thus, $\mathbb{E}\left[X_j\right] = \mu$ for $j = 1, 2, 3, \ldots$, and expression (21) becomes

$$\|\mu(\Theta) - P_{n+1}\|. \quad (22)$$

If we fix α_1, $\alpha_2, \ldots, \alpha_n$, which are the coefficients of $\{X_j\}$ in P_{n+1}, then the minimum of expression (22) is attained with

$$\alpha_0 = \mathbb{E}\left[\mu(\Theta) - \sum_{j=1}^{n} \alpha_j X_j\right] = \left(1 - \sum_{j=1}^{n} \alpha_j\right)\mu,$$

because the mean of a random variable is its orthogonal projection onto the subspace of constants. With the definition

$$c_0 = 1 - \sum_{j=1}^{n} \alpha_j, \quad (23)$$

equation (16) becomes

$$P_{n+1} = c_0\mu + \sum_{j=1}^{n} \alpha_j X_j$$

and, hence,

$$P_{n+1} - \mu(\Theta) = c_0[\mu - \mu(\Theta)] + \sum_{j=1}^{n} \alpha_j [X_j - \mu(\Theta)]. \quad (24)$$

It follows from equations (24) and (23) that $\hat{P}_{n+1}$ is the credibility premium minimizing expression (22) if and only if $\hat{P}_{n+1} - \mu(\Theta)$ is the minimum-norm vector in the affine space spanned by $\mu - \mu(\Theta)$ and $X_j - \mu(\Theta)$, $j = 1, 2, \ldots, n$.

We have pointed out earlier that the $\{X_j - \mu(\Theta)\}$ are mutually orthogonal. Also, $X_j - \mu(\Theta) = X_j - \mathbb{E}\left[X_j|\Theta\right]$ is orthogonal to $\mu - \mu(\Theta)$, because $\mu - \mu(\Theta)$ is a function of Θ. Therefore, we can apply the system of equations (7) to obtain the optimal coefficients:

$$\hat{c}_0 = \frac{\dfrac{1}{\|\mu - \mu(\Theta)\|^2}}{\dfrac{1}{\|\mu - \mu(\Theta)\|^2} + \sum\limits_{j=1}^{n} \dfrac{1}{\left\|X_j - \mu(\Theta)\right\|^2}}, \tag{25}$$

$$\hat{\alpha}_k = \frac{\dfrac{1}{\|X_k - \mu(\Theta)\|^2}}{\dfrac{1}{\|\mu - \mu(\Theta)\|^2} + \sum\limits_{j=1}^{n} \dfrac{1}{\left\|X_j - \mu(\Theta)\right\|^2}} \tag{26}$$

for $k = 1, 2, \ldots, n$.

To express the premium in the form $\hat{P}_{n+1} = (1 - Z)\mu + Z\bar{X}$, we set

$$Z = 1 - \hat{c}_0 = \frac{\sum\limits_{j=1}^{n} \dfrac{1}{\left\|X_j - \mu(\Theta)\right\|^2}}{\dfrac{1}{\|\mu - \mu(\Theta)\|^2} + \sum\limits_{j=1}^{n} \dfrac{1}{\left\|X_j - \mu(\Theta)\right\|^2}} \tag{27}$$

and

$$\bar{X} = \frac{\sum\limits_{j=1}^{n} \dfrac{X_j}{\left\|X_j - \mu(\Theta)\right\|^2}}{\sum\limits_{k=1}^{n} \dfrac{1}{\|X_k - \mu(\Theta)\|^2}}. \tag{28}$$

Thus, $\bar{X}$ is a weighted average of the X_js with the weight attached to X_j being inversely proportional to $\|X_j - \mu(\Theta)\|^2$. Also, note that

$$\|\mu - \mu(\Theta)\|^2 = \|\mu(\Theta) - \mu\|^2 = \mathbb{E}\left[(\mu(\Theta) - \mu)^2\right] = \mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right],$$

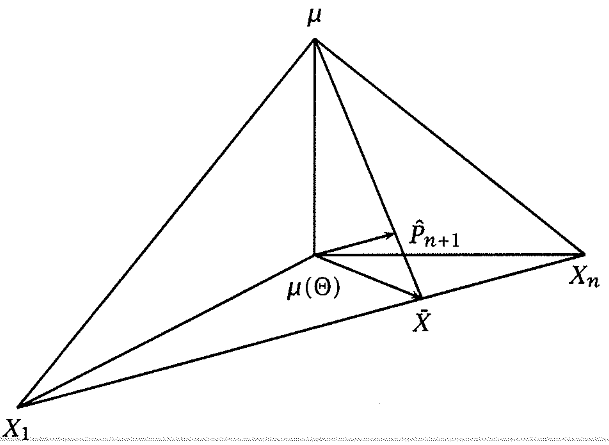

Figure 1: Credibility Premium as an Orthogonal Projection

and, by equation (12),

$$\|X_j - \mu(\Theta)\|^2 = \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[X_j|\Theta\right]\right].$$

An illustration of this geometric approach to credibility theory is shown in Figure 1. The affine space spanned by $\mu - \mu(\Theta)$ and $X_j - \mu(\Theta)$, $j = 1, 2, \ldots, n$, is the linear space spanned by μ and X_j, $j = 1, 2, \ldots, n$, translated by $-\mu(\Theta)$. The vector $\hat{P}_{n+1} - \mu(\Theta)$, being the minimum-norm vector in the affine space, is orthogonal to all vectors in the linear space spanned by μ and X_j, $j = 1, 2, \ldots, n$; see also condition (10).

4 Applications

The purpose of this section is to derive some of the results in Klugman, Panjer, and Willmot (1998, Chapter 5) using the results above.

(i) In the Bühlmann model as explained in Section 5.4.3 of Klugman, Panjer, and Willmot (1998),

$$\|\mu - \mu(\Theta)\|^2 = \mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right] = a$$

and

$$\|X_j - \mu(\Theta)\|^2 = \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[X_j|\Theta\right]\right] = \mathbb{E}\left[v(\Theta)\right] = v.$$

Hence, equation (27) becomes

$$Z = \frac{\sum_{j=1}^{n} \frac{1}{v}}{\frac{1}{a} + \sum_{j=1}^{n} \frac{1}{v}} = \frac{\frac{n}{v}}{\frac{1}{a} + \frac{n}{v}} = \frac{n}{\frac{v}{a} + n},$$

and equation (28) is

$$\bar{X} = \frac{\sum_{j=1}^{n} \frac{X_j}{v}}{\sum_{k=1}^{n} \frac{1}{v}} = \frac{\sum_{j=1}^{n} X_j}{n}.$$

As a check, we evaluate equation (26),

$$\hat{\alpha}_k = \frac{\frac{1}{v}}{\frac{1}{a} + \frac{n}{v}} = Z\frac{1}{n}, \quad k = 1, 2, \ldots, n.$$

(ii) In the Bühlmann-Straub model as explained in Section 5.4.4 of Klugman, Panjer, and Willmot (1998),

$$\|\mu - \mu(\Theta)\|^2 = \mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right] = a$$

and

$$\|X_j - \mu(\Theta)\|^2 = \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[X_j|\Theta\right]\right] = \mathbb{E}\left[v(\Theta)/m_j\right] = v/m_j.$$

Hence, with $m = \sum_{j=1}^{n} m_j$, we have from equation (27)

$$Z = \frac{\sum_{j=1}^{n} \frac{m_j}{v}}{\frac{1}{a} + \sum_{j=1}^{n} \frac{m_j}{v}} = \frac{\sum_{j=1}^{n} m_j}{\frac{v}{a} + \sum_{j=1}^{n} m_j} = \frac{m}{\frac{v}{a} + m},$$

and from equation (28)

$$\bar{X} = \frac{\sum_{j=1}^{n} \frac{m_j X_j}{v}}{\sum_{j=1}^{n} \frac{m_j}{v}} = \frac{\sum_{j=1}^{n} m_j X_j}{m}.$$

As a check, we evaluate equation (26),

$$\hat{\alpha}_k = \frac{\frac{m_k}{v}}{\frac{1}{a} + \sum_{j=1}^{n} \frac{m_j}{v}} = Z\frac{m_k}{m}, \quad k = 1, \ldots, n.$$

(iii) In Example 5.40 of Klugman, Panjer, and Willmot (1998),

$$\|\mu - \mu(\Theta)\|^2 = \mathbb{V}\mathrm{ar}\,[\mu(\Theta)] = a$$

and

$$\|X_j - \mu(\Theta)\|^2 = \mathbb{E}\left[\mathbb{V}\mathrm{ar}(X_j|\Theta)\right] = \mathbb{E}\left[w(\Theta) + v(\Theta)/m_j\right] = w + v/m_j.$$

Hence, with

$$m^* = \sum_{j=1}^{n} \frac{1}{\left\|X_j - \mu(\Theta)\right\|^2} = \sum_{j=1}^{n} \frac{m_j}{v + w m_j},$$

we have

$$Z = \frac{m^*}{\frac{1}{a} + m^*} = \frac{a m^*}{1 + a m^*},$$

and

$$\bar{X} = \frac{\sum_{j=1}^{n} \frac{m_j X_j}{v + w m_j}}{\sum_{k=1}^{n} \frac{m_k}{v + w m_k}} = \frac{\sum_{j=1}^{n} \frac{m_j X_j}{v + w m_j}}{m^*}.$$

As a check, we evaluate equation (26),

$$\hat{\alpha}_k = \frac{\dfrac{m_k}{v + w m_k}}{\dfrac{1}{a} + m^*} = Z \frac{1}{m^*} \cdot \frac{m_k}{v + w m_k}, \quad k = 1, \ldots, n.$$

(iv) In Example 5.41 of Klugman, Panjer, and Willmot (1998), $m = \sum_{j=1}^{n} m_j$, $\|\mu - \mu(\Theta)\|^2 = \mathbb{V}\mathrm{ar}[\mu(\Theta)] = a + b/m$ and $\left\|X_j - \mu(\Theta)\right\|^2 = \mathbb{E}\left[\mathbb{V}\mathrm{ar}(X_j|\Theta)\right] = w + v/m_j$.
Hence,

$$Z = \frac{m^*}{\dfrac{1}{a + b/m} + m^*} = \frac{(a + b/m)m^*}{1 + (a + b/m)m^*},$$

and

$$\bar{X} = \frac{\sum_{j=1}^{n} \dfrac{m_j X_j}{v + w m_j}}{m^*}.$$

As a check, we evaluate equation (26),

$$\hat{\alpha}_k = \frac{\dfrac{m_k}{v + w m_k}}{\dfrac{1}{a + b/m} + m^*} = Z \frac{1}{m^*} \cdot \frac{m_k}{v + w m_k}, \quad k = 1, \ldots, n.$$

(v) To solve Exercise 5.51 in Klugman, Panjer, and Willmot (1998), consider X_j/β_j in the exercise as X_j in Section 3 above.

(vi) To solve Exercise 5.56 in Klugman, Panjer, and Willmot (1998), consider X_j/τ^j in the exercise as X_j in Section 3 above.

5 Similar Triangles

Similar triangles are now used to derive several equivalent forms for the credibility factor, Z, and, hence, several equivalent forms for the credibility premium,

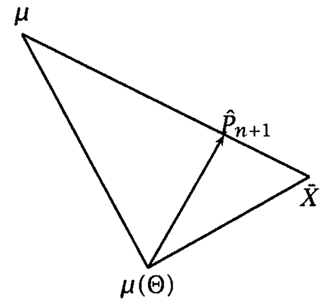

Figure 2: Three Similar Right-Angled Triangles

$$\hat{P}_{n+1} = Z\bar{X} + (1 - Z)\mu. \tag{29}$$

It follows from equation (29) that

$$Z = \frac{\|\hat{P}_{n+1} - \mu\|}{\|\bar{X} - \mu\|} \tag{30}$$

and

$$1 - Z = \frac{\|\bar{X} - \hat{P}_{n+1}\|}{\|\bar{X} - \mu\|}.$$

(Thus, Z is the ratio of the standard deviation of $\hat{P}_{n+1}$ to that of $\bar{X}$.) Now, equation (29) is equivalent to

$$\hat{P}_{n+1} - \mu(\Theta) = Z[\bar{X} - \mu(\Theta)] + (1 - Z)[\mu - \mu(\Theta)].$$

As $\bar{X}$ is an average of $\{X_j\}$, we have $\mathbb{E}[\bar{X}|\Theta] = \mu(\Theta)$, from which it follows that $[\bar{X} - \mu(\Theta)]$ and $[\mu - \mu(\Theta)]$ are orthogonal to each other. Figure 2 illustrates the geometric relationships among the random variables; note that Figure 2 is a slice in Figure 1.

There are three similar right-angled triangles in Figure 2. We shall show that each triangle gives a different form for Z (and for $1 - Z$). In each triangle, there are two acute angles complementary to each other. We shall also show that the square of the cosine of one of the acute angles gives the value of the credibility factor Z, while the square of the cosine of the other is $1 - Z$.

The three triangles yield three equivalent sets of ratios,

$$\begin{aligned}\|\bar{X} - \mu\| : \|\bar{X} - \mu(\Theta)\| : \|\mu(\Theta) - \mu\| \\ = \|\bar{X} - \mu(\Theta)\| : \|\bar{X} - \hat{P}_{n+1}\| : \|\mu(\Theta) - \hat{P}_{n+1}\| \\ = \|\mu(\Theta) - \mu\| : \|\mu(\Theta) - \hat{P}_{n+1}\| : \|\hat{P}_{n+1} - \mu\|.\end{aligned} \tag{31}$$

In particular, we have the equation

$$\frac{\|\bar{X} - \mu\|}{\|\mu(\Theta) - \mu\|} = \frac{\|\mu(\Theta) - \mu\|}{\|\hat{P}_{n+1} - \mu\|}, \tag{32}$$

which applied to equation (30) yields

$$Z = \frac{\|\mu(\Theta) - \mu\|^2}{\|\bar{X} - \mu\|^2} \tag{33}$$

and

$$Z = \frac{\|\hat{P}_{n+1} - \mu\|^2}{\|\mu(\Theta) - \mu\|^2}. \tag{34}$$

From (30) and (31), we also obtain

$$Z = \frac{\|\hat{P}_{n+1} - \mu(\Theta)\|^2}{\|\bar{X} - \mu(\Theta)\|^2}. \tag{35}$$

Corresponding to equations (33), (34), and (35), we have

$$1 - Z = \frac{\|\bar{X} - \mu(\Theta)\|^2}{\|\bar{X} - \mu\|^2}, \tag{36}$$

$$1 - Z = \frac{\|\mu(\Theta) - \hat{P}_{n+1}\|^2}{\|\mu(\Theta) - \mu\|^2} \tag{37}$$

and

$$1 - Z = \frac{\|\bar{X} - \hat{P}_{n+1}\|^2}{\|\bar{X} - \mu(\Theta)\|^2}, \tag{38}$$

respectively.

The usual credibility premium equation is obtained by applying equations (33) and (36),

$$\hat{P}_{n+1} = \frac{\|\mu(\Theta) - \mu\|^2}{\|\bar{X} - \mu\|^2}\bar{X} + \frac{\|\bar{X} - \mu(\Theta)\|^2}{\|\bar{X} - \mu\|^2}\mu \tag{39}$$

$$= \frac{\mathbb{V}\text{ar}\left[\mu(\Theta)\right]}{\mathbb{V}\text{ar}\left[\bar{X}\right]}\bar{X} + \frac{\mathbb{E}\left[\mathbb{V}\text{ar}\left[\bar{X}|\Theta\right]\right]}{\mathbb{V}\text{ar}\left[\bar{X}\right]}\mu. \tag{40}$$

The credibility premium can thus be viewed as a weighted average of $\bar{X}$ and μ, with weights distributed according to the Pythagorean equation

$$\|\bar{X} - \mu\|^2 = \|\bar{X} - \mu(\Theta)\|^2 + \|\mu(\Theta) - \mu\|^2,$$

or its equivalent variance-decomposition equation

$$\mathbb{V}\text{ar}\left[\bar{X}\right] = \mathbb{E}\left[\mathbb{V}\text{ar}\left[\bar{X}|\Theta\right]\right] + \mathbb{V}\text{ar}\left[E(\bar{X}|\Theta)\right].$$

Equation (39) follows from equations (6) and (7), with $m = 2$, $\mathbf{v}_1 = [\bar{X} - \mu(\Theta)]$, and $\mathbf{v}_2 = [\mu - \mu(\Theta)]$.

The cosine of the angle between $[\mu(\Theta) - \mu]$ and $[\bar{X} - \mu]$ is the correlation coefficient between $\mu(\Theta)$ and $\bar{X}$, which we call $\rho_{\bar{X},\mu(\Theta)}$. Hence, it follows from equation (33) that Z is the square of the correlation coefficient, i.e.,

$$Z = \rho^2_{\mu(\Theta),\bar{X}},$$

and the credibility premium is

$$\hat{P}_{n+1} = \rho^2_{\mu(\Theta),\bar{X}}\bar{X} + (1 - \rho^2_{\mu(\Theta),\bar{X}})\mu.$$

Also, it follows from equation (34) that the credibility factor Z is the square of the correlation coefficient between $\mu(\Theta)$ and $\hat{P}_{n+1}$,

$$Z = \rho^2_{\mu(\Theta),\hat{P}_{n+1}}.$$

We remark that

$$\mathbb{C}\text{ov}\left[\bar{X},\mu(\Theta)\right] = \mathbb{C}\text{ov}\left[\mathbb{E}\left[\bar{X}|\Theta\right],\mu(\Theta)\right] = \|\mu(\Theta) - \mu\|^2,$$

which may be viewed as a consequence of equation (2). Also,

$$\mathbb{C}\text{ov}\left[\hat{P}_{n+1},\mu(\Theta)\right] = \|\hat{P}_{n+1} - \mu\|^2.$$

6 Miscellaneous Equations and Remarks

We conclude this paper with some equations that readily follow from the discussion above. These equations provide further insights for understanding credibility theory.

From the ratios (31) we can obtain

$$\|\mu(\Theta) - \mu\| \|\bar{X} - \mu(\Theta)\| = \|\mu(\Theta) - \hat{P}_{n+1}\| \|\bar{X} - \mu\|. \tag{41}$$

If we divide both sides of equation (41) by 2, then the two sides of the equation represent two ways for finding the area of the largest triangle in Figure 2. Another consequence of the ratios (31) is

$$\frac{1}{\|\mu(\Theta) - \hat{P}_{n+1}\|^2} = \frac{1}{\|\mu(\Theta) - \mu\|^2} + \frac{1}{\|\bar{X} - \mu(\Theta)\|^2},$$

which also follows from equation (9).

From equation (32) we see that $\mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right]$ is the geometric mean of $\mathbb{V}\mathrm{ar}\left[\bar{X}\right]$ and $\mathbb{V}\mathrm{ar}\left[\hat{P}_{n+1}\right]$. Let us rewrite equations (33) and (34) as

$$\mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right] = Z\mathbb{V}\mathrm{ar}\left[\bar{X}\right] \tag{42}$$

and

$$\mathbb{V}\mathrm{ar}\left[\hat{P}_{n+1}\right] = Z\mathbb{V}\mathrm{ar}\left[\mu(\Theta)\right], \tag{43}$$

respectively. Applying equation (42) to (43) yields

$$\mathbb{V}\mathrm{ar}\left[\hat{P}_{n+1}\right] = Z^2\mathbb{V}\mathrm{ar}\left[\bar{X}\right],$$

which is also a consequence of equation (30).

Recall that $\hat{P}_{n+1}$ is the solution in minimizing (17). Thus, it follows from equation (3) that

$$\|X_{n+1} - \mu\|^2 = \|X_{n+1} - \hat{P}_{n+1}\|^2 + \|\hat{P}_{n+1} - \mu\|^2,$$

or

$$\mathbb{V}\mathrm{ar}(X_{n+1}) = \|X_{n+1} - \hat{P}_{n+1}\|^2 + \mathbb{V}\mathrm{ar}(\hat{P}_{n+1}).$$

Also, if we write expression (17) as

$$\|X_{n+1} - [Z\bar{X} + (1 - Z)\mu]\| = \|(X_{n+1} - \mu) - Z(\bar{X} - \mu)\|,$$

we see from the left side of equation (4) that the coefficient of $(\bar{X} - \mu)$ is

$$Z = \frac{\langle X_{n+1} - \mu, \bar{X} - \mu \rangle}{\langle \bar{X} - \mu, \bar{X} - \mu \rangle} = \frac{\mathbb{C}\mathrm{ov}\,[X_{n+1}, \bar{X}]}{\mathbb{V}\mathrm{ar}(\bar{X})}. \tag{44}$$

Equation (44) can be found in Fuhrer (1989, equation 1). Fuhrer (1989, page 84) derived the equation without assuming the existence of the risk parameter Θ; he also made some interesting remarks concerning the equation. A parameter-free approach to credibility theory can be found in Jones and Gerber (1975) and in Section 6.3 of Gerber (1979). Jones and Gerber (1975) also provided an appendix entitled "credibility theory ... in the light of functional analysis."

For further discussions on credibility and geometry, we refer the reader to De Vylder (1976a, 1976b, 1996), Gisler (1990), Hiss (1991), Jones and Gerber (1975), Norberg (1992), and Taylor (1977). We also recommend the book by Small and McLeish (1994).

References

Bühlmann, H. *Mathematical Methods in Risk Theory.* Heidelberg, Germany: Springer-Verlag, 1970.

De Vylder, F.E. "Optimal Semilinear Credibility." *Bulletin of the Swiss Association of Actuaries* (1976a): 27–40.

De Vylder, F.E. "Geometric Credibility." *Scandinavian Actuarial Journal* (1976b): 121–149.

De Vylder, F.E. *Advanced Risk Theory: A Self-Contained Introduction.* Brussels: Editions de l'Université de Bruxelles, 1996.

Fuhrer, C.S. Discussion of "Credibility: The Bayesian Model Versus Bühlmann's Model." *Transactions of the Society of Acturaries* 41 (1989): 83–86.

Gerber, H.U. *An Introduction to Mathematical Risk Theory.* S.S. Huebner Foundation Monograph Series No. 8. Homewood, IL: Irwin, 1979.

Gisler, A. "Credibility Theory Made Easy." *Bulletin of the Swiss Association of Actuaries* (1990): 75–100.

Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E., and Bauwelinckx, T. *Effective Actuarial Methods.* Amsterdam: North-Holland, 1990.

Goovaerts, M.J., and Hoogstad, W.J. *Credibility Theory.* Survey of Actuarial Studies No. 4. Rotterdam, The Netherlands: Nationale-Nederlanden N.V., 1987.

Goulet, V. "Principles and Application of Credibility Theory." *Journal of Actuarial Practice* 6 (1998): 5–66.

Heilmann, W.-R. *Fundamentals of Risk Theory.* Karlsruhe, Germany: VVW, 1988.

Herzog, T.N. *Introduction to Credibility Theory,* 3rd Edition. Winsted, CT: Actex, 1999.

Hickman, J.C., and Heacox, L. "Credibility Theory: The Cornerstone of Actuarial Science." *North American Actuarial Journal* 3 no. 2 (1999): 1–8.

Hiss, K. "Lineare Filtration und Kredibilitätstheorie." *Bulletin of the Swiss Association of Actuaries* (1991): 85–103.

Jones, D.A., and Gerber, H.U. "Credibility Formulas of the Updating Type." *Transactions of the Society of Acturaries* 27 (1975): 31–46; Discussion 47–52.

Kaas, R., Goovaerts, M.J., Dhaene, J., and Denuit, M. *Modern Actuarial Risk Theory.* Dordrecht, The Netherlands: Kluwer, 2001.

Kahn, P.M. (Editor) *Credibility: Theory and Applications.* New York: Academic Press, 1975.

Klugman, S.A., Panjer, H.H., and Willmot, G.E. *Loss Models: From Data to Decision.* New York: Wiley, 1998.

Longley-Cook, L.H. "An Introduction to Credibility Theory." *Proceedings of the Casualty of Actuarial Society* 49 (1962): 194–221.

Luenberger, D.G. *Optimization by Vector Space Methods.* New York: Wiley, 1969.

Mahler, H.C., and Dean, C.G. "Credibility." In *Foundations of Casualty Actuarial Science,* 4th Edition. Arlington, VA: Casualty Actuarial Society, 2001: 485–659.

Norberg, R. "Linear Estimation and Credibility in Continuous Time." *ASTIN Bulletin* 22 (1992): 149–165.

Small, C.G., and McLeish, D.L. *Hilbert Space Methods in Probability and Statistical Inference.* New York: Wiley, 1994.

Straub, E. *Non-Life Insurance Mathematics.* Heidelberg, Germany: Springer-Verlag, 1988.

Sundt, B. *An Introduction to Non-Life Insurance Mathematics.* Karlsruhe, Germany: VVW, 1993.

Taylor, G. "Abstract Credibility." *Scandinavian Actuarial Journal* (1977): 149–168.

Venter, G.G. "Credibility." In *Foundations of Casualty Actuarial Science.* New York: Casualty Actuarial Society, 1990: 375–483.

Waters, H.R. *Credibility Theory.* Edinburgh, Scotland: Department of Actuarial Mathematics and Statistics, Heriot-Watt University, 1993.