TRANSACTIONS OF SOCIETY OF ACTUARIES
1975 VOL. 27

CREDIBILITY FORMULAS OF THE UPDATING TYPE*

DONALD A. JONES AND HANS U. GERBER

ABSTRACT

A credibility formula is said to be of the *updating type* if the premium for any period is a weighted average of the premium and the claim experience of the preceding period. A characterization is given in terms of the first- and second-order moments of the joint distribution of annual claims. The resulting family of covariance matrices can be developed by the concept of risk parameters.

The special case of constant weights from year to year, which leads to credibility premiums that are geometrically weighted averages of the claims for all prior years, is discussed in some detail. In particular, these premiums have stability properties which motivate their use even in cases where the underlying model does not justify it on statistical grounds. Also, in a continuous time model, these formulas may be considered as the underwriting analogue of the Ornstein-Uhlenbeck model for Brownian motion in physics.

I. INTRODUCTION

Let S^1, S^2, S^3, . . . be a sequence of random variables, for which we know the joint distribution. We assume that the first two moments of these random variables exist. In applications of this model, S_n is interpreted as the claims produced by a given policyholder in year n (although in many applications S_n could be the loss ratio in year n, or the time period could be other than a year).

The basic problem of credibility theory in prospective rate-making is this: Given the outcomes of $S_1, S_2, \ldots, S_n$ (the claim experience of the first n years), establish an appropriate premium P_{n+1} for year $n+1$.

The most popular solutions to this problem can be devised by the method of least squares [1, 5, 7, 11–15, 18], in which P_{n+1} is defined as the random variable X that solves the problem

$$\text{Minimize} \quad E[(X - S_{n+1})^2] \qquad (1)$$

* This paper was presented at the Berkeley Actuarial Research Conference on Credibility, which was sponsored in part by the Committee on Research of the Society of Actuaries and held at the University of California, September 19–21, 1974.

subject to the constraint that X be a member of a given class of random variables, say H (which usually is a linear space of square-integrable random variables that form a Hilbert space). The choice of H determines the resulting premium P_{n+1}, which is illustrated by the following three examples.

Example 1: Let H_1 be the set of all real numbers: that is, we demand that the X in expression (1) be a constant. By differentiating (1) with respect to X, we see that the resulting premium is

$$P_{n+1} = E[S_{n+1}] . \tag{2}$$

Example 2: Let H_2 be the set of all square-integrable functions of S_1, $S_2, \ldots, S_n$. The resulting premium is now the conditional expected value

$$P_{n+1} = E[S_{n+1} | S_1, \ldots, S_n] . \tag{3}$$

Example 3: Let H_3 be the set of all square-integrable functions of S_1, $S_2, \ldots, S_{n+1}$. The resulting premium is obviously

$$P_{n+1} = S_{n+1} . \tag{4}$$

The method of least squares dates back about two thousand years in Euclidean geometry, where it may be represented graphically. The following substitutions allow us to utilize this representation to summarize the results of the three examples graphically in Figure 1.

H_1	x-axis
H_2	xy-plane
H_3	xyz(3-dimensional)-space
$\sqrt{E[(X - S_{n+1})^2]}$	Euclidean distance between X and S_{n+1}
P_{n+1}	Orthogonal projection of S_{n+1} on H_i

II. LINEAR CREDIBILITY FORMULAS

Each of the preceding examples has its shortcomings. Example 1 ignores the available claim experience, while Example 3 requires the knowledge of S_{n+1}, which is of course unknown in advance! The premium of Example 2 is satisfactory in theory. However, it requires complete knowledge of the joint distribution of $(S_1, S_2, \ldots, S_{n+1})$, depends substantially on this distribution, and quite often is of an analytical form that has little appeal to the practitioner.

For these reasons most authors [7, 11–14] have assumed that H is the class of all linear functions of $S_1, \ldots, S_n$.

$$X = b_0 + \sum_{i=1}^{n} b_i S_i . \tag{5}$$

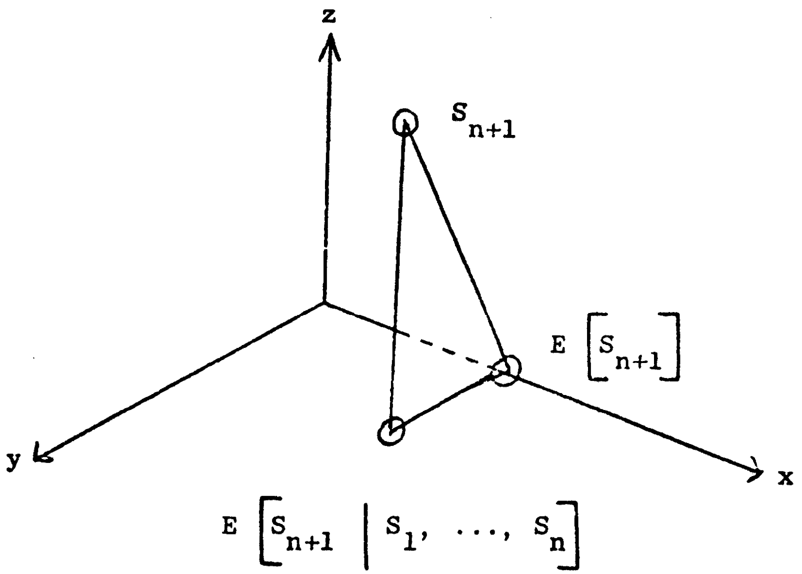

FIG. 1.—Graphical representation of least squares

In the following we shall adopt this assumption, so that problem (1) reduces to the minimization of

$$E\left[\left(b_0 + \sum_{i=1}^{n} b_i S_i - S_{n+1}\right)^2\right]. \tag{6}$$

By definition of the class H, the resulting premium will be of the form

$$P_{n+1} = {}_n a_0 + \sum_{i=1}^{n} {}_n a_i S_i \tag{7}$$

and therefore acceptable as far as simplicity is concerned.

A system of linear equations for the coefficients ${}_n a_0, {}_n a_1, \ldots, {}_n a_n$ may be obtained by setting the $n + 1$ partial derivatives of expression (6) equal to zero. From the partial derivative with respect to b_0 we obtain

$${}_n a_0 + \sum_{i=1}^{n} {}_n a_i E[S_i] = E[S_{n+1}], \tag{8}$$

and from the other partial derivatives we obtain the equations

$${}_n a_0 E[S_j] + \sum_{i=1}^{n} {}_n a_i E[S_i S_j] = E[S_j S_{n+1}] \tag{9}$$

for $j = 1, 2, \ldots, n$. Furthermore, if we multiply the members of equation (8) by $E[S_j]$ and subtract it term by term from the corresponding equation (9), we obtain the equations

$$\sum_{i=1}^{n} {}_n a_i \operatorname{Cov}[S_i, S_j] = \operatorname{Cov}[S_j, S_{n+1}] \tag{10}$$

for $j = 1, 2, \ldots, n$. This is a linear system of n equations for the n credibility coefficients ${}_na_1, {}_na_2, \ldots, {}_na_n$. In the sequel we shall assume that the matrix of this system, that is, the covariance matrix of $(S_1, \ldots, S_n)$ is nonsingular for every n (which guarantees a unique solution for these credibility coefficients). In turn, equation (8) may be solved for ${}_na_0$.

III. CREDIBILITY FORMULAS OF THE UPDATING TYPE

In the preceding section it was shown that linear credibility formulas of form (7) require knowledge of only the first- and second-order moments of the distribution of total claims. We shall study the relationship between the expected-value vector and the covariance matrix on the one side and the type of credibility formula on the other side.

Definition. A linear credibility formula is said to be of the *updating type*, if there is a sequence $Z_1, Z_2, \ldots$ of real numbers such that

$$P_{n+1} = (1 - Z_n)P_n + Z_nS_n . \tag{11}$$

Remarks

1. A condition equivalent to equation (11) is that

$$P_{n+1} - P_n = Z_n(S_n - P_n) , \tag{12}$$

which shows that the premium adjustment from year n to year $n + 1$ is proportional to the excess of claims over premiums in year n.

2. The updating type of formula should be distinguished from the Markov type [13, 18], $P_{n+1} = (1 - Z_n)P_1 + Z_nS_n$.

The following two examples are of the updating type.

Example 1 (*geometric credibility weights*) [9]: If $Z_n = Z$ (independent of n), repeated application of formula (11) shows that

$${}_na_i = Z(1 - Z)^{n-i} \tag{13}$$

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

Example 2 (*uniform credibility weights*): If

$$Z_n = \frac{W}{nW + V} \tag{14}$$

for $n = 1, 2, \ldots$ and certain constants $V > 0$, $W > 0$, repeated application of formula (11) shows that

$$_na_1 = {_na_2} = \ldots = {_na_n} = \frac{W}{nW + V}. \tag{15}$$

The following theorem characterizes the expected-value vectors and the covariance matrices that lead to credibility formulas of the updating type.

THEOREM 1. *The credibility formula is of the updating type, if there exist a number μ and sequences $V_1, V_2, \ldots$ and $W_1, W_2, \ldots$ such that*

$$E[S_i] = \mu , \tag{16}$$

$$\begin{aligned} \operatorname{Cov}[S_i, S_j] &= V_i + W_i \quad &\text{if } i = j \\ &= W_i &\text{if } i < j , \end{aligned} \tag{17}$$

for $j = i, i + 1, \ldots$ and $i = 1, 2, \ldots$.

Proof:

1. For a credibility formula of the updating type, we want to show the validity of formulas (16) and (17). Comparing equations (7) and (8), we see that $E[S_{n+1}] = E[P_{n+1}]$ for all n. From these equalities and those obtained by taking expected values in equation (11), it follows that $E[S_{n+1}] = E[S_n]$ for all n, which is equivalent to formula (16). Formula (11) means that $_na_n = Z_n$ and that

$$_na_j = (1 - Z_n)\ {_{n-1}a_j} \tag{18}$$

for $j = 0, 1, \ldots, n - 1$. From these and formula (10), we obtain, for $j = 1, \ldots, n - 1$ and all n, that

$$\begin{aligned} &\operatorname{Cov}[S_j, S_{n+1}] \\ &= \sum_{i=1}^{n} {_na_i} \operatorname{Cov}[S_i, S_j] \\ &= Z_n \operatorname{Cov}[S_n, S_j] + (1 - Z_n) \sum_{i=1}^{n-1} {_{n-1}a_i} \operatorname{Cov}[S_i, S_j] \\ &= Z_n \operatorname{Cov}[S_n, S_j] + (1 - Z_n) \operatorname{Cov}[S_j, S_n] \\ &= \operatorname{Cov}[S_j, S_n] . \end{aligned}$$

But this means precisely that the covariance matrix can be written in the form (17).

2. Now assume that equations (16) and (17) hold. We want to show the validity of formula (11) or, equivalently, that

$$_na_j = (1 - {_na_n})\ {_{n-1}a_j} \tag{19}$$

for $j = 0, 1, \ldots, n - 1$. From equations (10) and (17) we find, for $j = 1, 2, \ldots, n - 1$, that

$$\begin{aligned}\sum_{i=1}^{n-1} {}_n a_i \operatorname{Cov}[S_i, S_j] \\ &= \operatorname{Cov}[S_j, S_{n+1}] - {}_n a_n \operatorname{Cov}[S_j, S_n] \\ &= (1 - {}_n a_n) \operatorname{Cov}[S_j, S_n] .\end{aligned} \tag{20}$$

Comparing this system of equations with

$$\sum_{i=1}^{n-1} {}_{n-1}a_i \operatorname{Cov}[S_i, S_j] = \operatorname{Cov}[S_j, S_n] , \tag{21}$$

we conclude that formula (19) holds for $j = 1, 2, \ldots, n - 1$. Its validity for $j = 0$ follows from equations (8) and (16). Q.E.D.

The following result shows how the Z_n's can be obtained recursively from the elements of the covariance matrix.

THEOREM 2a. *Under the conditions of Theorem 1,*

$$Z_1 = \frac{W_1}{W_1 + V_1} \tag{22}$$

and

$$Z_n = \frac{W_n - W_{n-1} + Z_{n-1}V_{n-1}}{W_n - W_{n-1} + Z_{n-1}V_{n-1} + V_n} \tag{23}$$

for $n = 2, 3, \ldots$.

Proof: Formula (22) follows immediately from equation (10) for $n = j = 1$. Now, using equation (10) for $j = n$ and formula (18), we obtain

$$\operatorname{Cov}[S_n, S_{n+1}] = Z_n \operatorname{Var}[S_n] + (1 - Z_n) \sum_{i=1}^{n-1} {}_{n-1}a_i \operatorname{Cov}[S_i, S_n] . \tag{24}$$

By equation (17), the last summation may be written as

$${}_{n-1}a_{n-1}\{\operatorname{Cov}[S_{n-1}, S_n] - \operatorname{Var}[S_{n-1}]\} + \sum_{i=1}^{n-1} {}_{n-1}a_i \operatorname{Cov}[S_i, S_{n-1}] \tag{25}$$

and then, by equation (10), as

$$Z_{n-1}\{\operatorname{Cov}[S_{n-1}, S_n] - \operatorname{Var}[S_{n-1}]\} + \operatorname{Cov}[S_{n-1}, S_n] .$$

Substituting this in formula (24), we obtain in the notation of Theorem 1

$$W_n = Z_n(V_n + V_n) + (1 - Z_n)(W_{n-1} - Z_{n-1}V_{n-1}) , \tag{26}$$

which may be solved for Z_n to obtain the recurrence relationship (23). Q.E.D.

In terms of an auxiliary sequence $U_1, U_2, \ldots$, defined recursively by the formula

$$U_n = W_n - W_{n-1} + \frac{V_{n-1}U_{n-1}}{U_{n-1} + V_{n-1}} \tag{27}$$

with $U_1 = W_1$, Theorem 2a can be restated as follows:

THEOREM 2b. *For $n = 1, 2, \ldots$*

$$Z_n = \frac{U_n}{U_n + V_n}. \tag{28}$$

Section 5 contains a natural interpretation for the U_n's.

IV. A SPECIAL FAMILY

Theorem 1 shows that the family of covariance matrices that leads to a credibility formula of the updating type is quite rich. While a rich family is sometimes advantageous (e.g., in marriage!), it can make statistical estimation more difficult.

In this section we will restrict ourselves to the three-parameter family of covariance matrices of the type (17), where

$$W_1 = W > 0\,, \quad W_{j+1} - W_j = \delta^2 \geq 0\,, \quad V_j = V > 0\,. \tag{29}$$

Theorem 2a shows that in this case

$$Z_1 = \frac{W}{W + V}, \tag{30}$$

$$Z_n = F(Z_{n-1})\,, \tag{31}$$

where the function $F(x)$ is defined as

$$F(x) = \frac{\delta^2 + Vx}{\delta^2 + Vx + V}. \tag{32}$$

The resulting family of credibility formulas includes the two examples of the preceding section. For $\delta = 0$, formulas (30) and (31) lead to formula (14), and we are in the case of uniform credibility weights. This is not surprising, since in this case the covariance matrix is constant on the main diagonal and constant off the main diagonal. If δ is chosen such that

$$F(Z_1) = Z_1\,, \tag{33}$$

formula (31) shows that we arrive at geometric credibility weights.

Let us now consider the case of an arbitrary $\delta \neq 0$. From

$$F'(x) = \frac{V^2}{(\delta^2 + Vx + V)^2} \tag{34}$$

we see that $F'(x) < F'(0) < 1$ for $x > 0$. Since $F(0)$ is positive, the equation

$$F(x) = x \tag{35}$$

has a unique positive solution Z. Furthermore, the inequality $F(1) < 1$ implies that $Z < 1$.

THEOREM 3. *Suppose $\delta \neq 0$. Then Z_n converges monotonically to Z for $n \to \infty$.*

Proof: From formula (31) and the above remarks it is evident that the standard argument for the convergence of the iterative algorithm of numerical analysis is applicable [10, sec. 4.2]. This is best summarized by the graph in Figure 2.

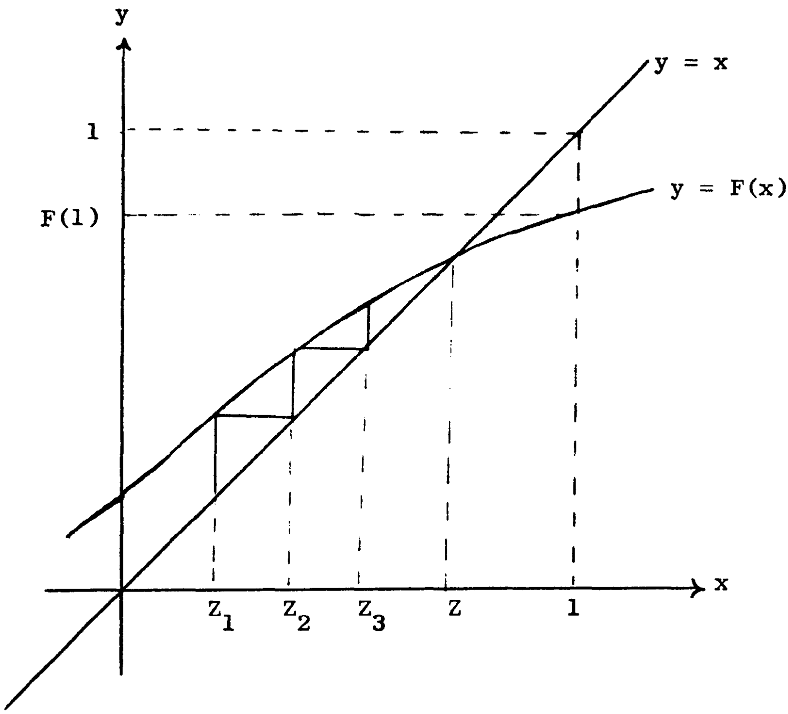

FIG. 2.—Graphical representation of the iterative algorithm

Remarks

1. The sequence $\{Z_n\}$ is increasing (decreasing), if and only if $Z_1 < Z$ $(Z_1 > Z)$.
2. If $\delta = 0$, the Z_n's obviously converge to zero.

Formula (11) shows that

$$ {}_{n}a_{n-m} = (1 - Z_n)(1 - Z_{n-1}) \ldots (1 - Z_{n-m})Z_{n-m-1} \,. \tag{36} $$

Therefore, a by-product of Theorem 3 is that the credibility weights are asymptotically geometric:

COROLLARY. *Suppose* $\delta \neq 0$. *Then, for any* m,

$$ \lim_{n\to\infty} {}_{n}a_{n-m} = Z(1 - Z)^m \,. \tag{37} $$

V. THE CONCEPT OF RISK PARAMETERS

In the case where $V_j > 0$ and $0 < W_j \leq W_{j+1}$ for $j = 1, 2, \ldots,$ formulas (16) and (17) can be explained by the concept of risk parameters [6, chap. 3].

Suppose that the intrinsic (but not directly observable) quality of our policyholder is given by the random variables $\theta_1, \theta_2, \theta_3, \ldots$ (the risk parameters). Intuitively, θ_n is the parameter that is effective for year n.

We require that the sequence $\{\theta_n\}$ have independent increments, with

$$ E[\theta_{j+1} - \theta_j] = 0 \,, \tag{38} $$

$$ \operatorname{Var}[\theta_{j+1} - \theta_j] = W_{j+1} - W_j \,, \tag{39} $$

and set

$$ E[\theta_1] = \mu \,, \qquad \operatorname{Var}[\theta_1] = W_1 \,. \tag{40} $$

Furthermore, for given θ_j, S_j and all future claims S_{j+h} $(h = 1, 2, \ldots)$ should be essentially independent. More precisely, we require that

$$ E[S_j \mid \theta_j] = \theta_j \,, \tag{41} $$

$$ \operatorname{Var}[S_j \mid \theta_j] = V_j \,, \tag{42} $$

$$ E[S_j S_{j+h} \mid \theta_j] = \theta_j^2 \,. \tag{43} $$

From the assumptions (38), (40), and (41) we obtain

$$ E[S_j] = E[E[S_j \mid \theta_j]] = E[\theta_j] = \mu \,, \tag{44} $$

which is equation (16). Next, using the well-known decomposition formula for the variance, assumptions (39), (41), and (42) lead to

$$ \begin{aligned} \operatorname{Var}[S_j] &= E[\operatorname{Var}[S_j \mid \theta_j]] + \operatorname{Var}[E[S_j \mid \theta_j]] \\ &= V_j + \operatorname{Var}[\theta_j] = V_j + W_j \,. \end{aligned} \tag{45} $$

Finally, for $h = 1, 2, \ldots,$

$$ \begin{aligned} E[S_j S_{j+h}] &= E[E[S_j S_{j+h} \mid \theta_j]] \\ &= E[\theta_j^2] = \operatorname{Var}[\theta_j] + \mu^2 = W_j + \mu^2 \,, \end{aligned} \tag{46} $$

and, therefore, $\operatorname{Cov}[S_j, S_{j+h}] = W_j$.

Remarks about Parametrization

1. Under the additional assumption that the risk parameters $\theta_1, \theta_2, \theta_3, \ldots$ are normally distributed and also that the conditional distributions of the S_j's are normal, the Bayesian analysis can be carried out explicitly and leads to an identical result:
$$P_{n+1} = E[S_{n+1} | S_1, \ldots, S_n] . \tag{47}$$
Furthermore, the auxiliary parameter U_n of Section III can now be interpreted as the conditional variance of θ_{n+1}, given the observations $S_1, S_2, \ldots, S_n$ [17, p. 250].
2. If the credibility formula should take into account only the frequency of the claims, one might assume that the conditional distribution of S_j is Poisson with parameter θ_j. Consequently, condition (42) would have to be replaced by
$$\text{Var}\,[S_j | \theta_j] = \theta_j . \tag{48}$$
In the following, only formula (45) has to be modified. It has to be replaced by
$$\text{Var}\,[S_j] = \mu + W_j . \tag{49}$$
It is remarkable that the resulting covariance matrix is still of the form (17), namely, now with $V_j = \mu$.

VI. REINTERPRETATION OF GEOMETRIC CREDIBILITY WEIGHTS

In this section we shall present deterministic and probabilistic properties of
$$P_{n+1} = (1 - Z)^n \mu + \sum_{i=1}^{n} Z(1 - Z)^{n-i} S_i \tag{50}$$
with $0 < Z < 1$. These properties motivate the use of such a formula even in cases where the underlying model does not justify it on statistical grounds.

First let $P_{n+k}(S_n)$ denote the part of P_{n+k} that is due to the occurrence of $S_n (k = 1, 2, \ldots)$. From formula (50),
$$P_{n+k}(S_n) = Z(1 - Z)^{k-1} S_n ; \tag{51}$$
therefore,
$$\sum_{k=1}^{\infty} P_{n+k}(S_n) = S_n . \tag{52}$$
Thus, regardless of what the claims of year n are, they will be fully repaid by future premiums. This, and the following arguments, are particularly meaningful if the S_n's are measured in indexed monetary units.

Now consider the loss (or gain) of the insurer, say L_n, for the n-year period:

$$L_n = \sum_{i=1}^{n} (S_i - P_i) , \qquad (53)$$

which may be rearranged as

$$L_n = \sum_{i=1}^{n} (1 - Z)^{n-i} S_i - \frac{\mu}{Z} [1 - (1 - Z)^n] . \qquad (54)$$

This shows that a great degree of financial stability can be accomplished by the use of geometric credibility weights. Suppose that $S_1, S_2, \ldots$ are independent and identically distributed with $E[S_n] = \mu$ and $\text{Var}\,[S_n] = \sigma^2$. From formula (54)

$$E[L_n] = 0 , \qquad (55)$$

$$\text{Var}\,[L_n] = \frac{1 - (1 - Z)^{2n}}{1 - (1 - Z)^2} \sigma^2 . \qquad (56)$$

As $n \rightarrow \infty$,

$$\text{Var}\,[L_n] \rightarrow (2Z - Z^2)^{-1} \sigma^2 . \qquad (57)$$

Furthermore, the distribution of L_n converges to the distribution of the random variable L, where

$$L = \sum_{i=0}^{\infty} (1 - Z)^i S_{i+1} - \frac{\mu}{Z} , \qquad (58)$$

which can be interpreted as a discounted sum. On the other hand, a purely statistical argument would have led to constant premiums, $P'_n = \mu$ for all n, with resulting aggregate losses

$$L'_n = \sum_{i=1}^{n} S_i - \mu n . \qquad (59)$$

While the expected value of the loss at any time is still zero, we see that

$$\text{Var}\,[L'_n] = \sigma^2 n , \qquad (60)$$

which implies divergence of its variance and consequently of its distribution.

As an illustration, we have simulated the outcomes of 100 periods under the assumption that the S_n's were independent and identically distributed, each assuming only the values 0 and 2, with equal probabilities. Figure 3

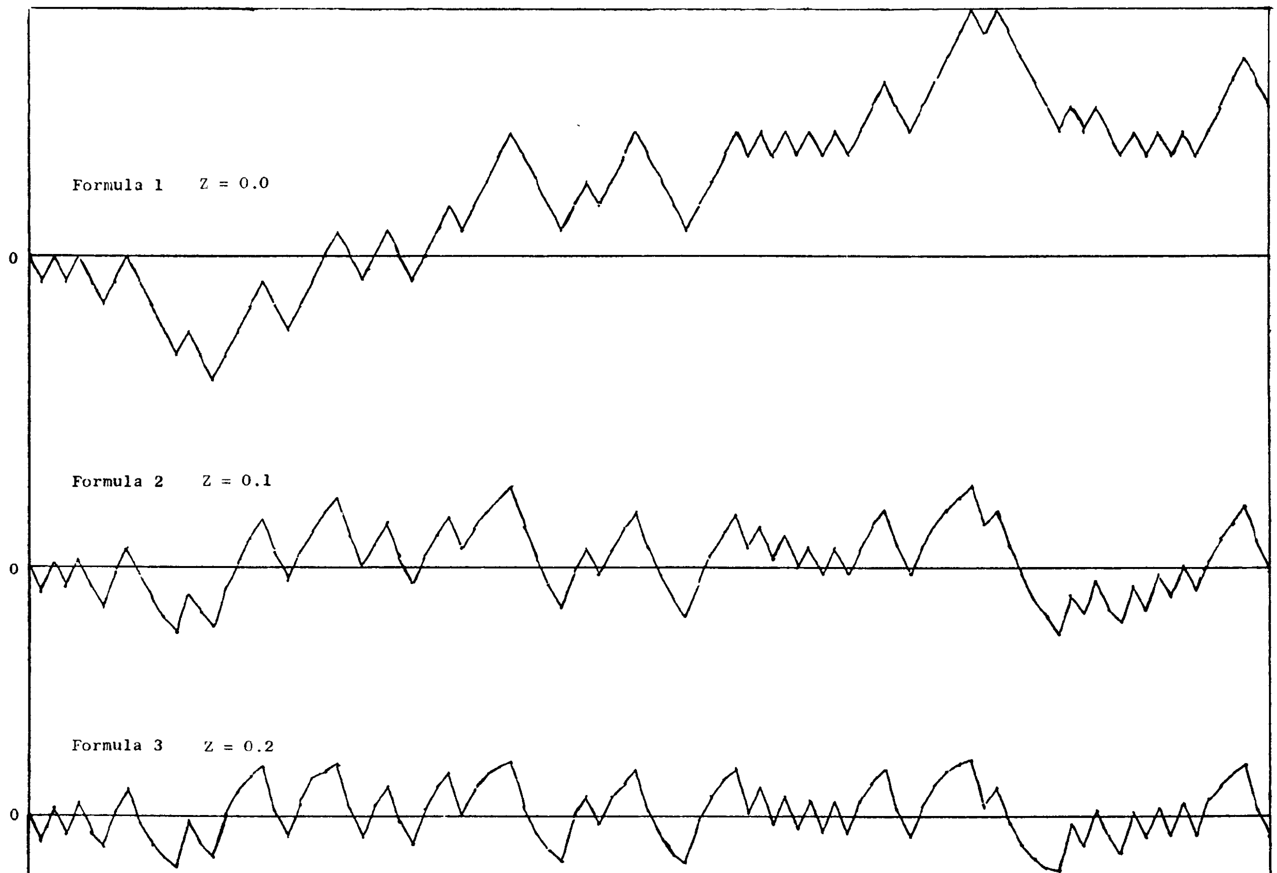

FIG. 3.—Simulated sample paths of the aggregate losses

compares the aggregate losses as a function of time under three different credibility formulas. The formulas are

Formula 1: $P_n = 1$ for all n (purely statistical approach),

Formula 2: P_n according to formula (50) with $Z = 0.1$,

Formula 3: P_n according to formula (50) with $Z = 0.2$,

As expected, formula 3 produced the most stable result for the insurance company.

VII. ANALOGY WITH THE THEORY OF BROWNIAN MOTION [2, 4, 16]

The Ornstein-Uhlenbeck theory of Brownian motion is concerned with the velocity $V(t)$ and the position $X(t)$ of a particle as a function of time t. Then, from physics, the Langevin equation is

$$dV(t) = -bV(t)dt + dW(t), \quad b > 0, \tag{61}$$

where $W(t)$ denotes a Wiener process. (Without the last term we would have merely a deterministic slow-down of the particle; the purpose of the last term is to make the motion oscillate.) Formal integration of the last equation leads to

$$V(t) = e^{-bt}V(0) + \int_0^t e^{b(t-s)}dW(s), \tag{62}$$

and from this we obtain

$$X(t) = X(0) + \int_0^t V(s)ds. \tag{63}$$

Interestingly enough, the same construction is followed when a formula of type (50) is applied. Assuming a continuous time model, the analogue of the difference equation (12) with Z_n = constant is

$$dP(t) = c[dS(t) - P(t)dt], \tag{64}$$

where $P(t)$ is the premium density at time t and $S(t)$ is the aggregate claims at time t and $c > 0$. Obviously equation (64) is of the form (61). The only difference is that now the Wiener process has been replaced by the aggregate claims process $S(t)$ (for example, a compound Poisson process). Formal integration of the stochastic equation (64) leads to

$$P(t) = e^{-ct}P(0) + c\int_0^t e^{-c(t-u)}dS(u). \tag{65}$$

Finally, the aggregate loss at time t is

$$L(t) = \int_0^t e^{-c(t-u)}dS(u) - \frac{P(0)}{c}(1 - e^{-ct}). \tag{66}$$

The last two formulas should be viewed as the continuous analogues of formulas (50) and (54).

APPENDIX

CREDIBILITY FORMULAS OF THE UPDATING TYPE IN THE LIGHT OF FUNCTIONAL ANALYSIS

The purpose of this Appendix is to present an alternative proof for Theorem 1 by use of Hilbert space techniques. Intuitively, a Hilbert space is an infinite-dimensional generalization of Euclidean space of finite dimension [3]. The concepts of distance and orthogonality carry over into Hilbert space theory, and, of course, then so do projections and the theory of least squares. The idea that Hilbert space methods might have applications in credibility theory is not new, but their actual use has not appeared; this is the raison d'être for this Appendix.

Let H be a Hilbert space with elements ("vectors") $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, ... For any $\boldsymbol{a}, \boldsymbol{b} \in H$ an inner product $(\boldsymbol{a}, \boldsymbol{b})$ is defined. Let U be a given subspace of H, and let $\boldsymbol{x}$ and $\boldsymbol{y}$ be any two vectors in H. We denote by V the subspace that is spanned by U and $\boldsymbol{x}$ (i.e., the smallest subspace of H that contains U and $\boldsymbol{x}$), and by T_1 and T_2 the projection operators onto U and V, respectively.

THEOREM 4. *There is a real number Z such that*

$$T_2\boldsymbol{y} = Z\boldsymbol{x} + (1 - Z)T_1\boldsymbol{x} \tag{67}$$

if and only if

$$(\boldsymbol{a}, \boldsymbol{y}) = (\boldsymbol{a}, \boldsymbol{x}) \tag{68}$$

for all $\boldsymbol{a} \in U$.

Proof:

1. Assume that there is a constant Z such that equation (67) holds. Since $\boldsymbol{y} - T_2\boldsymbol{y}$ is orthogonal to V, we obtain, for any $\boldsymbol{a} \in V$,

$$\begin{aligned} 0 &= (\boldsymbol{a}, \boldsymbol{y} - T_2\boldsymbol{y}) \\ &= (\boldsymbol{a}, \boldsymbol{y}) - Z(\boldsymbol{a}, \boldsymbol{x}) - (1 - Z)(\boldsymbol{a}, T_1\boldsymbol{x}) . \end{aligned} \tag{69}$$

In particular this is true for all $\boldsymbol{a} \in U \subseteq V$. But for $\boldsymbol{a} \in U$, $(\boldsymbol{a}, T_1\boldsymbol{x}) = (\boldsymbol{a}, \boldsymbol{x})$, and therefore

$$0 = (\boldsymbol{a}, \boldsymbol{y}) - (\boldsymbol{a}, \boldsymbol{x}) , \tag{70}$$

which shows the validity of equation (68).

2. Now assume that equation (68) holds. Since $T_2\boldsymbol{y} \in V$, it is of the form

$$T_2\boldsymbol{y} = Z\boldsymbol{x} + \boldsymbol{u} \tag{71}$$

for some constant Z and some vector $u \in U$. The difference $y - T_2 y$ is orthogonal to V. Thus, for any $a \in U$, we obtain

$$\begin{aligned} 0 &= (a, y - T_2 y) \\ &= (a, y) - Z(a, x) - (a, u) \\ &= (1 - Z)(a, x) - (a, u) . \end{aligned} \tag{72}$$

It follows that $u = (1 - Z)T_1 x$. Q.E.D.

For the alternative proof of Theorem 1, let the Hilbert space H be the set of all square-integrable random variables, with the inner product defined as

$$(X, Y) = E[XY] . \tag{73}$$

Let U be the subspace that is spanned by

$$S_0 = 1, S_1, \ldots, S_n . \tag{74}$$

(Thus, U is the set of all random variables of the form (5).) The roles of x and y are played by S_{n+1} and S_{n+2}. Since P_{n+1} is the projection of S_{n+1} on U, and P_{n+2} the projection of S_{n+2} on V, statement (67) reads that there is a constant Z (call it Z_{n+1}) such that

$$P_{n+2} = Z_{n+1} S_{n+1} + (1 - Z_{n+1}) P_{n+1} . \tag{75}$$

The equivalent condition (68) now reads

$$E[S_i S_{n+1}] = E[S_i S_{n+2}] \tag{76}$$

for $i = 0, 1, \ldots, n$. For $i = 0$ this gives us

$$E[S_{n+1}] = E[S_{n+2}] . \tag{77}$$

Because of that, the equations for $i = 1, \ldots, n$ become

$$\text{Cov}\,[S_i, S_{n+1}] = \text{Cov}\,[S_i, S_{n+2}] . \tag{78}$$

Theorem 1 is now an immediate application of Theorem 4.

REFERENCES

1. BAILEY, A. L. "A Generalized Theory of Credibility," *Proceedings of the Casualty Actuarial Society*, XXXII (1945), 13–20.
2. BEEKMAN, J. A. "A New Collective Risk Model," *TSA*, XXV (1973), 573.
3. BERBERIAN, S. K. *Introduction to Hilbert Space*. New York: Oxford University Press, 1961.
4. BREIMAN, L. *Probability*. Reading, Mass.: Addison-Wesley, 1968.
5. BÜHLMANN, H. "Experience Rating and Credibility," *ASTIN Bulletin*, IV, Part 3 (1967), 199–207.
6. BÜHLMANN, H. *Mathematical Methods in Risk Theory*. New York: Springer-Verlag, 1970.

7. BÜHLMANN, H., AND STRAUB, E. "Glaubwürdigkeit für Schadensätze," *Mitteilungen der Vereinigung Schweizer Versicherungsmathematiker*, LXX, 111–33. Distributed by *ARCH*, 1972.2.
8. DEGROOT, M. H. *Optimal Statistical Decisions*. New York: McGraw-Hill Book Co., 1970.
9. GERBER, H. U., AND JONES, D.A. "Credibility Formulae with Geometric Weights," in *Proceedings of the Business and Economics Section, American Statistical Association*, 1973.
10. HENRICI, P. *Elements of Numerical Analysis*. New York: John Wiley & Sons, 1964.
11. HICKMAN, J. C. Discussion of Margolin's "On the Credibility of Group Insurance Claim Experience," *TSA*, XXIII (1971), 254–57.
12. JEWELL, S. W. "Multi-dimensional Credibility." ORC Report 73-7. Berkeley: Operations Research Center, University of California, 1973.
13. JONES, D. A. "On the Pure Premium." Paper presented to the Credibility Seminar sponsored by the Casualty Actuarial Society and the Society of Actuaries, 1965.
14. MARGOLIN, M. H. "On the Credibility of Group Insurance Claim Experience," *TSA*, XXIII (1971), 229–38.
15. MAYERSON, A. L. "A Bayesian View of Credibility," *Proceedings of the Casualty Actuarial Society*, LI (1964), 85–104.
16. NELSON, E. *Dynamical Theories of Brownian Motion*. Mathematical Notes. Princeton, N.J.: Princeton University Press, 1967.
17. RAIFFA, H., AND SCHLAIFER, R. *Applied Statistical Decision Theory*. Cambridge, Mass.: Harvard University Press, 1961.
18. SHUR, W. Discussion of Margolin's "On the Credibility of Group Insurance Claim Experience," *TSA*, XXIII (1971), 240–51.

DISCUSSION OF PRECEDING PAPER

JAMES C. HICKMAN AND ROBERT B. MILLER:*

Cramér [2] has given the classic definition of risk theory: "The object of the theory of risk is to give a mathematical analysis of the random fluctuations in an insurance business and to discuss the various means of protection against their inconvenient effects." The subject of this paper might be stated as the adaptive estimation of the parameters of risk-theory models. Estimation is of obvious importance to any discussion of "means of protection against the inconvenient effects" of deviations from expected results.

Jones and Gerber have contributed to the solution of this estimation problem. They have adopted a general multivariate model and have developed in detail the linear least-squares method of estimating conditional means. Rather than embellish their impressive results, we would like to raise some questions about their basic model.

Other things being equal, simple models are to be preferred to complex models. Yet it seems to us that the appealing simplicity of the multivariate model for the random vector of claims used by Jones and Gerber may be somewhat illusory. Their model requires the estimation of a great many covariances. Because of the obvious impossibility of iterating a sequence of claim results for a particular case, it seems clear that satisfactory estimates will require the use of prior and ancillary information. Since credibility theory is devoted to the adaptive estimation of risk parameters, it does not seem inappropriate to consider somewhat more elaborate models, in which the steps in the estimation procedure are indicated. Consequently, our discussion will consider the formulation of a multivariate model that is suggested by plausible economic assumptions and explicitly requires the introduction of ancillary information as well as past claims for the particular case under consideration into the parameter estimation process. We will be developing the first Remark at the end of Section V of the paper (p. 40).

We have been impressed with the efficient market theory developed to study price changes in speculative markets. One form of this theory postulates that, in a market where the participants have access to a common body of information and where transaction and information costs

* Mr. Miller, not a member of the Society, is associate professor of statistics and business, University of Wisconsin, Madison.

are insignificant, current prices fully reflect all information and one may expect price changes to be independently distributed [4]. There is an impressive amount of statistical evidence that the commodity and common stock markets are approximately efficient.

Efficient market theory has led to new insights into many markets; may it be applied to help us understand the insurance market? We are not certain of the answer to this question. Nevertheless, we would like to suggest that in an efficient insurance market a common pool of economic and risk classification information, plus competitive pressures, will tend to produce pure premiums (expected claims) such that deviations from expected claims will be mutually independent. If the deviations from expected claims are not independent, there will be dependencies in the distribution that could be exploited by the insurance company or the insured to take advantage of the other party to the contract. Observed dependencies should be a signal to management to update its estimates of the risk model parameters or to improve its classification system, or to take both actions.

Let us illustrate our idea with an example that employs the tractable multinormal distribution [1]. An element S_i of the random vector $\boldsymbol{S} = (S_1, S_2, \ldots, S_{n+1})'$ may be interpreted as the aggregate claims in year i. Given the risk parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{n+1})'$, we will assume that $\boldsymbol{S}$ has a multinormal distribution with mean vector $h\boldsymbol{\theta}$ and variance-covariance matrix $r^2\boldsymbol{I}_{(n+1,n+1)}$, where $\boldsymbol{I}_{(n+1,n+1)}$ is the $(n+1,\ n+1)$ identity matrix. The constants h and r^2 are identified with the particular case under study. They adjust the risk parameters for the level and variability of claim experience for the case under consideration. Note that, given the risk parameters $h\boldsymbol{\theta}$, the elements of $\boldsymbol{S}$ are mutually independent. Thus our model assumes that the insurance market is efficient.

In addition, we will assume that all prior and ancillary information about the risk parameters $\boldsymbol{\theta}$ has been summarized in a conjugate multinormal distribution with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{A}$ [3]. For example, broad economic trends, such as inflation-induced shifts in claim amounts, might be reflected in the vector $\boldsymbol{\mu}$, and the fact that these shifts are not independent over time would be reflected in $\boldsymbol{A}$. The determination of the distribution of $\boldsymbol{\theta}$ might well involve a time series analysis of a broadly based claim index.

From the distribution assumptions made about $\boldsymbol{S}$ and $\boldsymbol{\theta}$, it may be shown that $(\boldsymbol{\theta}, \boldsymbol{S})'$ has a multinormal distribution with mean vector $(\boldsymbol{\mu}, h\boldsymbol{\mu})'$ and variance-covariance matrix

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{A} + r^2\boldsymbol{I} \end{pmatrix}.$$

It may also be shown that the distribution of S_{n+1}, given $S_1, S_2, \ldots, S_n$, $\boldsymbol{\mu}$, h, r^2, and A, is normal with mean

$$h\mu_{n+1} + \Delta'(A + r^2I)^{-1}_{(n,n)}(S - h\boldsymbol{\mu})_{(n,1)}$$

and variance

$$(a_{(n+1,n+1)} + r^2) - \Delta'(A + r^2I)^{-1}_{(n,n)}\Delta\ ,$$

where

$$\Delta' = (a_{n+1,1}, a_{n+1,2}, \ldots, a_{n+1,n})\ .$$

In each of these expressions $a_{i,j}$ is an element of the matrix A.

This model incorporates ideas from efficient market theory and permits broad economic information, as well as claim data for the particular case, to enter into the parameter estimation process. It is contemplated that tests of independence on the sequence of claim amounts would be periodically performed as a general test of the adequacy of the model and the quality of the parameter estimates and the classification system. We are not so naïve as to believe that all claim distributions are normal. We cheerfully acknowledge that other multivariate distributions should be explored. Nevertheless, we believe that even the crude probability statements available from the normal distribution would be useful in many risk management problems.

REFERENCES

1. Anderson, T. W. *An Introduction to Multivariate Statistical Analysis.* New York: John Wiley & Sons, Inc., 1953. Chapter 2 of this classic text contains all the distribution theory used in this discussion.
2. Cramér, H. *On the Mathematical Theory of Risk.* Skandia Jubilee Volume. Stockholm, 1930.
3. DeGroot, M. H. *Optimal Statistical Decisions.* New York: McGraw-Hill Book Co., Inc., 1970. Section 9.9 discusses the Bayesian analysis of multinormal distributions using conjugate prior distributions.
4. Fama, E. E., "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, Vol. XXV (1970).

MYRON H. MARGOLIN:

Many actuaries need to apply credibility procedures to practical problems such as group insurance experience rating. We welcome the paper of Professors Jones and Gerber as an effort to furnish us with theoretical models which may find practical application. Equations (16) and (17) define a very "rich" family of models, which lead to credibility formulas of the updating type. The restrictions at the beginning of Section V define a subset of this family in which the models can be interpreted in terms of "risk parameters." Then equation (29) defines a yet narrower subset, for which the credibility factors converge to a geometric series. If for convenience we denote these families respectively by F_1, F_2, F_3, we observe that $F_3 \subset F_2 \subset F_1$.

What are the criteria by which the actuary should decide whether to adopt one type of credibility formula—such as an updating formula based on F_1 or a set of geometric weights based on F_3—in preference to another? Virtually all students of the credibility problem, the authors and myself included, have agreed that a necessary condition is the satisfaction of the least-squares principle, formula (1) or formula (6). Beyond that, some actuaries may be motivated by the subjective appeal of the formula. For example, there is a certain appeal to the notion that, with a formula based on F_3, the claims of year n will be fully repaid by the premiums of years $n + 1$ through ∞. However, since insurance companies and risks do not operate in infinite time, I would suggest that this notion is decidedly secondary to the least-squares principle.

The formulas based on F_1 or F_3 will satisfy the least-squares criterion only if the actual behavior of the S_i's satisfies the defining equations of the model. The authors present no empirical data to support either the fairly general conditions of F_1 or the tighter ones of F_3. Accordingly, all we can do is to consider whether they are plausible.

One point of view questions whether it is proper to ascribe a mathematical distribution, or even the first two moments of a distribution, to the actual S_i's. The symbols μ and σ and the concepts of mathematical expectation and variance have abstract mathematical significance, but whether they have objective counterparts in the actual behavior of insurance risks is dubious, according to this view. The same doubts would apply to the risk parameters of F_2. Perhaps these questions are "philosophical"; but the absence of a demonstration that these mathematical quantities exist should be considered in evaluating the plausibility of F_1 and hence also of F_2 and F_3.

A more telling argument against the plausibility of F_1 is that equation (16) demands the existence of a "privileged" manual premium rating system [1]. Thus, if actuary A expresses the S_i's as loss ratios of claim dollars to one set of manual premiums, and if actuary B calculates loss ratios with respect to a different set of manual premiums, then equation (16) generally cannot hold for both A and B. The privileged system is the one in which equation (16) holds. We have no reason to believe that such a system exists, and, if it does, we do not know how to find it.

Perhaps least plausible is the special case F_3, involving geometric credibility weights. The variance of S_i is a monotonically increasing function of i, implying that there once was a time when the variance equaled zero and that in the future both the variance and S_i itself will diverge without bound. Moreover, we observe that the correlation coefficient between S_i and S_{i+1} can be expressed as $1/\{(1 + V/W_i)[1 + (V + \delta^2)/W_i]\}^{1/2}$, a monotonically increasing function of i. It follows that, in the case where only the previous year's claim experience is known, the credibility of that experience is an increasing function of i; thus, 1975 experience is a better indicator of 1976 experience than 1965 experience was of 1966. Intuitively, these several anisotropisms with respect to time are highly implausible.

Indeed, considerations like these would seem to rule out almost any type of model which characterizes the behavior of individual risks (by postulating a formula for the parameters of the individual risk claim distribution). Any such formula calls for a privileged manual premium rating system. Moreover, if the formula is to be kept reasonably simple while the parameters are not to diverge, the latter must be either constant or periodic. The case of constant parameters, or homogeneity in time, is contrary to both common sense and a fair amount of empirical evidence. This leaves only the rather implausible possibility that one or more of the parameters is periodic.

Conversely, a plausible model should characterize only the *aggregate* behavior of a set of risks; it should not require the existence of a privileged manual premium rating system; and, where appropriate, it should be isotropic with respect to time. The simplest (minimal) set of assumptions consistent with these plausibility considerations is the following:

1. With the S_i's expressed as loss ratios to manual premium, the manual premiums should be incremented each year to compensate for aggregate trend (such as health care inflation or overall mortality reduction). Then the average loss ratio (for the set of risks) is the same in all years.
2. The aggregate variance of all S_i's in each year is independent of i.
3. The correlation between S_i and S_j is a function of $(j - i)$, not of i.

Precisely this minimal set of assumptions consists of three of the four underlying the empirical approach to credibility proposed in reference [2]. It is easy to test all three empirically. The fourth assumption, that credibility factors should progress smoothly by size, is not really essential to the basic approach.

REFERENCES

1. MARGOLIN, M. H. "Credibility Theory and Privileged Frames of Reference." To appear in *ARCH*, 1975.2.
2. ———. "On the Credibility of Group Insurance Claims Experience," *TSA*, XXIII (1971), 229.

(AUTHORS' REVIEW OF DISCUSSION)

HANS U. GERBER AND DONALD A. JONES:

We appreciate the interest in our paper and its subject matter that has been shown by Messrs. Hickman, Miller, and Margolin. We are particularly indebted to Hickman and Miller for expanding on our remark concerning Bayesian analysis of the multinormal model. A part of their motivation was to simplify the problem of estimating the number of covariances indicated in our paper. Another method to accomplish this would be to adopt a special family of covariance matrices from which the estimation method would select one.

Some of the points of Margolin's discussion are not so clear to us, due in part to his ambiguous use of the technical terms of probability theory. For example, what is "the aggregate variance of all S_i's" or "the actual behavior of the [stochastic process] S_i"?

Moreover, Margolin's interpretation of the time index for the process S_i as absolute led him to the conclusions in his sixth paragraph, which could be explained by what probabilists call a stationary sequence of S_i's. However, we have difficulties in accepting the assumptions that led to these conclusions. For example, the isotropism axiom seems to us as artificial in the context of economics and statistical decision-making as it is plausible in the context of physics.

Finally, we would like to update our bibliography. In a forthcoming paper ("Two Classes of Covariance Matrices Giving Simple Linear Forecasts") William Jewell will discuss a general family of credibility formulas that contains formulas of the updating as well as of the Markov type as special cases.