

LINEAR ESTIMATION AND CREDIBILITY IN CONTINUOUS TIME¹

BY RAGNAR NORBERG

University of Copenhagen

ABSTRACT

The theory of linear filtering of stochastic processes provides continuous time analogues of finite-dimensional linear Bayes estimators known to actuaries as credibility methods. In the present paper a selfcontained theory is built for processes of bounded variation, which are of particular relevance to insurance. Two methods for constructing the optimal estimator and its mean squared error are devised. Explicit solutions are obtained in a continuous time variation of Hachemeister's regression model and in a homogeneous doubly stochastic generalized Poisson process. The traditional discrete time set-up is compared to the one with continuous time, and some merits of the latter are pointed out.

1. INTRODUCTION

A. Objective of the present study

The modern theory of linear filtering of stochastic processes provides powerful generalizations of standard results on linear estimation from a finite set of observations to the case where the observations are the values of a process recorded continuously through some time interval. This has recently been pointed out in the context of credibility theory by HISS (1991), who obtains the continuous time analogue of the credibility estimator in HACHEMEISTER'S (1975) random coefficient regression model by application of the so-called Kalman-Bucy filter, see e.g. KALLIANPUR and KARINDAKAR (1988) and RUYMGAART and SOONG (1988). Similar results appear in earlier works, notably the paper by NÄTHER (1984), who deals with more general random fields. In that paper the optimal estimator is constructed in a direct manner, minimizing the generalized mean squared loss by matrix algebraic methods.

The present paper undertakes to build a selfcontained theory, starting from first principles with Hilbert space methods. This can be done at a fairly simple mathematical level by restricting to processes of bounded variation, which are particularly apt to describe claims processes in insurance and other cash flows. (A function is of bounded variation if it is the difference between two non-decreasing finite-valued functions). It turns out that this together with a few more weak assumptions is sufficient to establish the basic results — no

¹ Presented to the XXII ASTIN Colloquium, Montreux, Sept. 90.

special distributional properties and no deep theory of stochastic integrals and differential equations are needed.

B. Outline of the paper

Section 2 presents some prerequisites from Hilbert space theory. In Section 3 these are applied to the finite-dimensional case (where Hilbert space methods can be dispensed with), and Hachemeister's random coefficient regression model is taken as an example. Section 4 deals with estimation on the basis of a stochastic process observed continuously over some time interval. A space of linear estimators is defined and, adopting mean squared estimation error as performance criterion, the normal equations determining the optimal estimator are derived. The continuous time version of the normal equations is an integral equation, which is a perfect analogue to the normal equations in the finite case. When the stochastic process is of bounded variation, a useful alternative form of the optimal estimator and the normal equations is obtained upon integrating by parts. The mean squared error of the optimal estimator is given by a compact expression. In Section 5 the continuous time random coefficient regression model is revisited, and in Section 6 a generalized Poisson process with stochastic intensity process is analysed. In Section 7 the two approaches, discrete versus continuous, are compared, and it is concluded that the latter is in a strong position from a practical as well from a theoretical point of view.

2. REVIEW OF SOME BASIC CONCEPTS AND RESULTS IN ESTIMATION THEORY

A. The Hilbert space of square integrable random variables

Let \mathcal{L}^2 be the family of all square integrable real random variables defined on some probability space (Ω, \mathcal{F}, P) . On identifying random variables that are equal a.s. (almost surely), \mathcal{L}^2 becomes a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\|\cdot\|$ defined by

$$\langle m, n \rangle = E(mn), \quad \|m\|^2 = Em^2,$$

$m, n \in \mathcal{L}^2$.

If $\langle m, n \rangle = 0$, then m and n are said to be *orthogonal*, written $m \perp n$. If $m \perp n$ for all $n \in \mathcal{M} \subset \mathcal{L}^2$, then m is said to be orthogonal to \mathcal{M} , written $m \perp \mathcal{M}$.

In general $\|m+n\|^2 = \|m\|^2 + 2\langle m, n \rangle + \|n\|^2$, and in particular

$$(2.1) \quad \|m+n\|^2 = \|m\|^2 + \|n\|^2 \quad \text{if } m \perp n.$$

B. Statement of the estimation problem

In an estimation context \mathcal{F} has a sub-sigmafield \mathcal{H} representing observable events. An *estimand* is a random variable m representing some unobservable quantity (not \mathcal{H} -measurable) that is to be estimated from the information

provided by \mathcal{H} . An estimator is an \mathcal{H} -measurable random variable \check{m} that is used as an approximation of m . The quality of the approximation is measured by the distance between m and \check{m} ,

$$(2.2) \quad \|m - \check{m}\|.$$

Thus, m and \check{m} are assumed to be in \mathcal{L}^2 , implying that \check{m} is restricted to the class of all \mathcal{H} -measurable random variables in \mathcal{L}^2 , henceforth denoted by $\check{\mathcal{M}}$.

C. Restricted optimal estimators

Often there are reasons for restricting \check{m} further to some subset $\check{\mathcal{M}}$ of $\check{\mathcal{M}}$. The estimation problem then amounts to minimizing (2.2) as \check{m} ranges in $\check{\mathcal{M}}$. If $\check{\mathcal{M}}$ is closed, then there exists a unique minimizing point \check{m} in $\check{\mathcal{M}}$, called the projection of m onto $\check{\mathcal{M}}$. The proof of this result, which is essentially a copy of the proof of the completeness of \mathcal{L}^2 (see e.g. BAUER, 1978, p. 78), gives no lead to the construction of \check{m} , which therefore will have to be devised ad hoc for each given specification of $\check{\mathcal{M}}$.

D. Optimal estimators in linear spaces

Things are simpler if restriction is made to some linear space. Let $\overline{\mathcal{M}}$ be some linear subspace of $\check{\mathcal{M}}$, not necessarily closed, and suppose there exists a point $\bar{m} \in \overline{\mathcal{M}}$ such that

$$(2.3) \quad m - \bar{m} \perp \overline{\mathcal{M}}.$$

For any $\check{m} \in \overline{\mathcal{M}}$ the linearity of $\overline{\mathcal{M}}$ entails $\bar{m} - \check{m} \in \overline{\mathcal{M}}$, and so (2.1) and (2.3) imply

$$\|m - \check{m}\|^2 = \|m - \bar{m}\|^2 + \|\bar{m} - \check{m}\|^2 \geq \|m - \bar{m}\|^2.$$

Thus, \bar{m} is a closest point to m in $\overline{\mathcal{M}}$, and also in this case it is called the projection since it is necessarily unique. The condition (2.3) can be spelled out as

$$(2.4) \quad E\{(m - \bar{m})\check{m}\} = 0, \quad \forall \check{m} \in \overline{\mathcal{M}},$$

the so-called normal equations (with respect to $\overline{\mathcal{M}}$).

The normal equations are indispensable constructive means in many situations with $\overline{\mathcal{M}}$ of infinite dimension. Some important cases are studied by DE VYLDER (1976). The following section treats in some detail estimation by finite-dimensional linear functions, the generalization of which is the topic of this paper.

3. FINITE-DIMENSIONAL LINEAR ESTIMATORS

A. The general form of the finite-dimensional linear estimator

The results obtained in this section are all well-known. The purpose is to motivate the approach and facilitate interpretation of the results in the remainder of the paper.

For a given set of random observables x_1, \dots, x_n in $\tilde{\mathcal{M}}$, consider the space $\overline{\mathcal{M}}$ of estimators of the form

$$(3.1) \quad \check{m} = a_0 + \sum_{j=1}^n a_j x_j,$$

with a_0, a_1, \dots, a_n non-random. Clearly $\overline{\mathcal{M}}$ is the linear space spanned by $\{1, x_1, \dots, x_n\}$, and it is an easy exercise to show that it is closed. Therefore, the $\overline{\mathcal{M}}$ -projection,

$$(3.2) \quad \bar{m} = \alpha_0 + \sum_{j=1}^n \alpha_j x_j,$$

say, exists and is uniquely determined by the normal equations (2.4), which now become

$$(3.3) \quad E \left(m - \alpha_0 - \sum_{i=1}^n \alpha_i x_i \right) = 0,$$

$$(3.4) \quad E \left\{ \left(m - \alpha_0 - \sum_{i=1}^n \alpha_i x_i \right) x_j \right\} = 0, \quad j = 1, \dots, n.$$

Upon multiplying by Ex_j in (3.3) and subtracting the result from (3.4), one obtains the equivalent

$$(3.5) \quad Em = \alpha_0 + \sum_{i=1}^n \alpha_i Ex_i,$$

$$(3.6) \quad \text{Cov}(m, x_j) = \sum_{i=1}^n \alpha_i \text{Cov}(x_i, x_j), \quad j = 1, \dots, n.$$

Solving (3.5) and (3.6) and putting $\alpha = (\alpha_1, \dots, \alpha_n)'$ and $x = (x_1, \dots, x_n)'$, the optimal linear estimator in (3.2) is found to be

$$(3.7) \quad \bar{m} = Em + \sum_{j=1}^n \alpha_j (x_j - Ex_j) = Em + \alpha' (x - Ex),$$

with

$$(3.8) \quad \alpha' = \text{Cov}(m, x') (\text{Var } x)^{-1}.$$

(It is assumed that the x_i are linearly independent so that $\text{Var } x$ is non-singular).

Any 1-1 linear transformation of $(1, x')$ forms a basis of the space $\overline{\mathcal{M}}$ and yields the same solution. Thus, from a theoretical point of view the choice of basis is immaterial. From a practical point of view it is not, however, since a judicious choice of basis may uncover possible structures of the moments

involved in (3.8) that can be utilized in the computations. To concretize, suppose x_j represents the total amount of claims paid up to time $j, j = 1, \dots, n$, for some risk insured at time 0 ($x_0 = 0$). The amounts $\Delta x_j = x_j - x_{j-1}$ paid in the periods $(j-1, j], j = 1, \dots, n$, are typically easier to work with, e.g. if they are assumed to be independent, possibly conditional on some latent risk factors. In terms of these observations (3.1) may be equivalently expressed as

$$(3.9) \quad \check{m} = g_0 + \sum_{j=1}^n g_j \Delta x_j = g_0 + g' \Delta x,$$

with $\Delta x = (\Delta x_1, \dots, \Delta x_n)'$, $g = (g_1, \dots, g_n)'$, and the optimal solution is

$$(3.10) \quad \bar{m} = Em + \sum_{j=1}^n \gamma_j (\Delta x_j - E\Delta x_j) = Em + \gamma' (\Delta x - E\Delta x)$$

obtained upon replacing α_0, α, x by $\gamma_0, \gamma, \Delta x$ in (3.3)-(3.8).

B. The credibility formula in the regression case

A suitable reference for the present paragraph is NORBERG (1986), Section 3. The random coefficient regression model specifies that

$$(3.11) \quad \Delta x_j = y_j' b + e_j, \quad j = 1, \dots, n,$$

where the y_j are q -vectors of non-random regressors, b is a q -vector of random regression coefficients, and the e_j are random disturbances. It is assumed that b, e_1, \dots, e_n are uncorrelated. Put

$$(3.12) \quad \beta = Eb, \quad A = \text{Var } b, \quad \text{Var } e_j = \varphi_j, \quad j = 1, \dots, n.$$

The \mathcal{H} -projection of $c' b$ (c a constant q -vector) is $c' \bar{b}$, where

$$(3.13) \quad \bar{b} = (I - Z) \left(A \sum_{j=1}^n \varphi_j^{-1} y_j \Delta x_j + \beta \right).$$

with

$$(3.14) \quad Z = AM(AM + I)^{-1} = (AM + I)^{-1} AM = I - (AM + I)^{-1},$$

$$(3.15) \quad M = \sum_{j=1}^n \varphi_j^{-1} y_j y_j'.$$

If M is non-singular, one easily verifies by use of (3.14) and (3.15) that (3.13) may be cast as

$$(3.16) \quad \bar{b} = Z\hat{b} + (I - Z)\beta,$$

with

$$(3.17) \quad \hat{b} = M^{-1} \sum_{j=1}^n \varphi_j^{-1} y_j \Delta x_j.$$

Formula (3.16) is a credibility weighted mean of the sample estimator of b in (3.17) and the expected value of b , which is the estimate of b based solely on "prior" information.

The quantities in (3.15) and (3.17) are usually expressed in matrix form, but the expressions given here are preferred since they relate more transparently to the continuous analogues to be given in Section 5. The regression model was introduced in credibility theory by HACHEMEISTER (1975), who worked with $w_j y_j$ in the place of y_j in (3.11) and with φ_j in (3.12) of the form φw_j . Thus, the w_j , which represent the amounts of risk exposed in the different periods, are absorbed in the regressors y_j in the present set-up.

4. CONTINUOUS LINEAR ESTIMATORS

A. General model assumptions

Working now in continuous time, let $x(t)$ be the total amount paid in respect of claims generated by a certain risk in the time interval $[0, t]$, $t \geq 0$. Alternatively, $x(t)$ may represent the number of claims in $[0, t]$. The function $\{x(t)\}_{t \geq 0}$, is modelled as a stochastic process satisfying

[i] $x(t) \in \mathcal{L}^2$ for each t ,

[ii] $x(0) = 0$ a.s.,

[iii] x is a.s. right-continuous and of bounded variation in every finite interval.

The bounded variation assumption ensures that certain operations like integration with respect to the process can be performed path by path. It covers all point processes and, thereby, all models commonly used to describe the development of insurance portfolios in continuous time, in life insurance, non-life insurance, and traditional risk theory. Processes of unbounded variation, e.g. nonstandard Brownian motion, might serve as approximate descriptions of large portfolios, but a treatment of those would require elaborate stochastic calculus and shall not be undertaken here.

B. Linear estimators

Suppose some unknown quantity $m \in \mathcal{L}^2$ is to be estimated by the information accrued in the time interval $[0, \tau]$, τ representing the present moment. Now, what is to be meant by a linear estimator based on the values $x(t)$, $0 < t \leq \tau$? A straightforward answer is

$$(4.1) \quad \hat{m} = a_0 + \int_{(0, \tau]} x(t) da(t),$$

where the scalar a_0 and the function a are non-random and, to make the expression in (4.1) a well defined random variable in \mathcal{L}^2 ,

[iv] a is right-continuous and has bounded variation on $[0, \tau]$.

By the general rule of integration by parts,

$$\begin{aligned} \int_{(0, \tau]} x(t) da(t) &= x(\tau) a(\tau) - x(0) a(0) - \int_{(0, \tau]} a(t-) dx(t) \\ &= \int_{(0, \tau]} \{a(\tau) - a(t-)\} dx(t), \end{aligned}$$

(recall [ii]), and so (4.1) can be cast as

$$(4.2) \quad \bar{m} = g_0 + \int_{(0, \tau]} g(t) dx(t),$$

where the scalar g_0 and the function g are non-random, and [iv'] g is left-continuous and has bounded variation on $(0, t]$.

The coefficients in (4.1) and (4.2) are related by

$$(4.3) \quad g_0 = a_0,$$

$$(4.4) \quad g(t) = a(\tau) - a(t-), \quad t \in (0, \tau].$$

The equivalent forms (4.1) and (4.2) are the continuous time analogues of (3.1) and (3.9), respectively.

C. The normal equations

Consider now the class $\bar{\mathcal{M}}$ of estimators of the form (4.1), with a satisfying [iv]. Clearly, $\bar{\mathcal{M}}$ is a linear subspace of \mathcal{L}^2 . Let

$$(4.5) \quad \bar{m} = \alpha_0 + \int_{(0, \tau]} x(t) d\alpha(t)$$

be the candidate for solving the normal equations (2.4), which now become

$$E \left[\left\{ m - \alpha_0 - \int_{(0, \tau]} x(s) d\alpha(s) \right\} \left\{ \alpha_0 + \int_{(0, \tau]} x(t) da(t) \right\} \right] = 0,$$

for all real α_0 and all functions a satisfying [iv]. Now proceed essentially as in Paragraph 3A. First put $\alpha_0 = 1$ and $a = 0$, then put $\alpha_0 = 0$ and a a unit mass at t , and use Fubini to switch the order of the expectation and the integral, to obtain

$$(4.6) \quad Em = \alpha_0 + \int_{(0, \tau]} Ex(s) d\alpha(s),$$

$$(4.7) \quad \text{Cov}(m, x(t)) = \int_{(0, \tau]} \text{Cov}(x(s), x(t)) d\alpha(s), \quad t \in (0, \tau],$$

which are the continuous time analogues of (3.5) and (3.6).

On subtracting (4.6) from (4.5), the optimal estimator can be written as

$$(4.8) \quad \bar{m} = Em + \int_{(0, \tau]} \{x(t) - Ex(t)\} d\alpha(t),$$

with α determined by (4.7), which thus is the important part of the normal equations.

D. An alternative form of the normal equations

Alternatively, one can work with the form (4.2) and put

$$(4.9) \quad \bar{m} = \gamma_0 + \int_{(0, \tau]} \gamma(t) dx(t),$$

where γ_0 and γ are related to α_0 and α by (4.3) and (4.4). Glossing over some technicalities for the time being, argue informally that the optimal estimator in (4.9) must be orthogonal to 1 and to all increments $dx(t)$, $t \in (0, \tau]$, and proceed as in the previous paragraph to obtain

$$(4.10) \quad Em = \gamma_0 + \int_{(0, \tau]} \gamma(s) E dx(s),$$

$$(4.11) \quad \text{Cov}(m, dx(t)) = \int_{s \in (0, \tau]} \gamma(s) \text{Cov}(dx(s), dx(t)), \quad t \in (0, \tau),$$

which are the continuous time analogue of the normal equations underlying (3.10). The optimal estimator is

$$(4.12) \quad \bar{m} = Em + \int_{(0, \tau]} \gamma(t) d\{x(t) - Ex(t)\},$$

with γ determined by (4.11).

As will be demonstrated by examples in Sections 5 and 6, (4.12) is usually the more convenient form of the optimal estimator, and the corresponding form (4.11) of the normal equations is the more direct route to the solution. It is therefore, appropriate to add some rigour to the informal reasoning above.

Under the conditions [i]-[iii] the function $\text{Cov}(x(s), x(t))$ is right-continuous and of bounded variation. (This is a consequence of the dominated convergence theorem. It suffices to look at a non-negative, non-decreasing process x satisfying [i]-[iii], and to observe that for $r, s, t \in (0, \tau]$, $\lim_{r \downarrow s} x(r) = x(s)$ and $\lim_{r \downarrow s} x(r)x(t) = x(s)x(t)$, and that $x(r)$ and $x(r)x(t)$ are bounded by $x(\tau)$ and $x^2(\tau)$, respectively, which both have finite expected value.) Then integration by parts is allowed on the right of (4.7),

which turns the expression into (note that $\text{Cov}(x(0), x(t)) = 0$ due to [ii])

$$\begin{aligned} \text{Cov}(x(\tau), x(t)) \alpha(\tau) &= \int_{s \in (0, \tau]} \alpha(s-) d_s \text{Cov}(x(s), x(t)) \\ &= \int_{s \in (0, \tau]} \{\alpha(\tau) - \alpha(s-)\} d_s \text{Cov}(x(s), x(t)). \end{aligned}$$

This together with (4.4) shows that (4.7) is equivalent to

$$(4.13) \quad \text{Cov}(m, x(t)) = \int_{s \in (0, \tau]} \gamma(s) d_s \text{Cov}(x(s), x(t)).$$

By the same argument as before, the function $\text{Cov}(m, x(t))$ is right-continuous and of bounded variation. Thus, (4.13) can be cast as

$$(4.14) \quad d_t \text{Cov}(m, x(t)) = d_t \int_{s \in (0, \tau]} \gamma(s) d_s \text{Cov}(x(s), x(t)).$$

Provided that differentiation with respect to t on the right of (4.14) can be performed under the integral sign, this gives precise contents to (4.11).

This is about as far as general arguments can bring us to the solution of the normal equations. From here on the course must be adapted to the special features of the model at hand.

E. The mean squared error of the optimal estimator

The performance criterion defined by (2.2) is equivalent to the mean squared estimation error, $\|m - \check{m}\|^2 = E(m - \check{m})^2$. For the optimal estimator (4.8) it is

$$\begin{aligned} \|m - \bar{m}\|^2 &= \left\| m - Em - \int_{(0, \tau]} \{x(t) - Ex(t)\} d\alpha(t) \right\|^2 \\ &= \|m - Em\|^2 - 2 \left\langle\left\langle m - Em, \int_{(0, \tau]} \{x(t) - Ex(t)\} d\alpha(t) \right\rangle\right\rangle + \\ &\quad + \left\| \int_{(0, \tau]} \{x(t) - Ex(t)\} d\alpha(t) \right\|^2 \\ &= \text{Var } m - 2 \int_{(0, \tau]} \text{Cov}(m, x(t)) d\alpha(t) + \\ &\quad + \int_{t \in (0, \tau]} \int_{s \in (0, \tau]} \text{Cov}(x(s), x(t)) d\alpha(s) d\alpha(t), \end{aligned}$$

hence, by use of (4.7),

$$(4.15) \quad \|m - \bar{m}\|^2 = \text{Var } m - \int_{(0, \tau]} \text{Cov}(m, x(t)) \, d\alpha(t).$$

In a similar way one gets from (4.10) and (4.11) the equivalent form

$$(4.16) \quad \|m - \bar{m}\|^2 = \text{Var } m - \int_{(0, \tau]} \gamma(t) \text{Cov}(m, dx(t)),$$

with qualifications as at the end of the previous section.

5. A CONTINUOUS TIME REGRESSION MODEL

A. A Poisson regression model

Assume for the time being that x is a generalized Poisson process with claim intensity $y'(t) a(\theta)$ at time t and claim size distribution $G(\cdot|\theta)$. Here θ represents some unknown risk characteristics, and $y(t)$ is a q -vector of observable risk characteristics at time t . Thus, the intensity of claims at time t is regressed on the observable risk characteristics, and the regression coefficient vector $a(\theta)$ depends on hidden risk factors. For fixed value of θ , the first and second order moments of the process are

$$(5.1) \quad E(x(t)|\theta) = \int_0^t y'(\vartheta) \, d\vartheta \, b(\theta),$$

$$(5.2) \quad \text{Cov}(x(s), x(t) | \theta) = \int_0^{s \wedge t} v(\vartheta, \theta) \, d\vartheta,$$

with

$$(5.3) \quad b(\theta) = a(\theta) \int y \, dG(y|\theta),$$

$$(5.4) \quad v(t, \theta) = y'(t) a(\theta) \int y^2 \, dG(y|\theta).$$

B. A general random coefficient regression model

The special structure of $v(t, \theta)$, given by (5.4), is disguised in (5.2), and so (5.1)-(5.2) serve as the moment structure for a wider class of models than the Poisson case.

Assume now that also θ is stochastic, whereby (5.1) and (5.2) are moments in the conditional model for given θ . Put

$$(5.5) \quad \beta = E b(\theta), \quad A = \text{Var } b(\theta), \quad \varphi(t) = E v(t, \theta).$$

The unconditional moments of the risk process are easily obtained as

$$(5.6) \quad Ex(t) = \int_0^t y'(\vartheta) d\vartheta \beta,$$

$$(5.7) \quad \text{Cov}(x(s), x(t)) = \int_0^s y'(\vartheta) d\vartheta A \int_0^t y(\vartheta) d\vartheta + \int_0^s \varphi(\vartheta) d\vartheta.$$

C. Credibility estimation of the regression coefficients

Suppose

$$(5.8) \quad m = c' b$$

(c constant) is to be estimated. Then the expression on the left of (4.7) is

$$\text{Cov}(m, x(t)) = c' A \int_0^t y(\vartheta) d\vartheta.$$

Inserting this and (5.7) into (4.14), gives

$$(5.9) \quad c' Ay(t) = \int_0^\tau \gamma(s) y'(s) ds Ay(t) + \varphi(t) \gamma(t).$$

At this step continuity of $y(t)$ was required for $t \in (0, \tau)$, at least piecewise. Multiply by $\varphi^{-1}(t) y'(t)$ from the right on both sides in (5.9), then integrate over t from 0 to τ , and solve

$$(5.10) \quad \int_0^\tau y'(t) \gamma(t) dt = c' Z,$$

where Z is defined by (3.14), now with

$$(5.11) \quad M = \int_0^\tau \varphi^{-1}(t) y(t) y'(t) dt.$$

Substituting (5.10) into (5.9) and then dividing by $\varphi(t)$, gives

$$(5.12) \quad \gamma(t) = c' (I - Z) \Lambda \varphi^{-1}(t) y(t).$$

Now insert $\gamma(t)$ from (5.12), $dEx(t) = y'(t) dt\beta$ from (5.6), and $Em = c' \beta$ into (4.12) and rearrange by use of (3.14) to obtain the following continuous time analogue of the discrete solution given by (3.12)-(3.17): the optimal linear estimator of $c' b$ is $c' \bar{b}$, where

$$(5.13) \quad \bar{b} = (I - Z) \left\{ \Lambda \int_{(0, \tau]} \varphi^{-1}(t) y(t) dx(t) + \beta \right\},$$

with Z defined by (3.14) and (5.11). If M is nonsingular, then

$$(5.14) \quad \bar{b} = Z\hat{b} + (I - Z)\beta,$$

with

$$(5.15) \quad \hat{b} = M^{-1} \int_{(0, \tau]} \varphi^{-1}(t) y(t) dx(t).$$

If $y(t) = w(t)$, a positive function representing the amount of risk exposed per time unit at time t , $b(\theta)$ is scalar-valued, and $v(t, \theta) = w(t)v(\theta)$, then (5.14), (5.15), and (3.14) specialize to

$$\bar{b} = \zeta \hat{b} + (1 - \zeta)\beta,$$

$$\hat{b} = x(\tau) \left| \int_0^\tau w(t) dt \right|,$$

$$\zeta = \lambda \int_0^\tau w(t) dt \left| \left(\lambda \int_0^\tau w(t) dt + \varphi \right) \right|,$$

where $\beta = Eb(\theta)$, $\lambda = \text{Var } b(\theta)$, $\varphi = Ev(\theta)$. This is the continuous version of the BÜHLMANN and STRAUB (1970) credibility formula.

The Poisson process in Paragraph A increases by jumps, and there is a.s. a finite number of them in the time interval $(0, \tau]$. The same is the case for any bounded variation process with (conditional on θ) independent increments and $x(\tau) < \infty$ a.s. Let t_1, t_2, \dots denote the jump epochs of the process recorded in chronological order. Then \hat{b} in (5.15) can be written as

$$\hat{b} = M^{-1} \sum_{j: t_j \leq \tau} \varphi^{-1}(t_j) y(t_j) \{x(t_j) - x(t_j^-)\}.$$

D. The mean squared estimation error

Entering (5.12) and

$$\text{Var } m = c' \Delta c,$$

$$\text{Cov}(m, dx(t)) = y'(t) \Lambda c dt = c' \Lambda y(t) dt$$

into (4.16) and making use of (3.14), yields

$$(5.16) \quad \begin{aligned} \|m - \bar{m}\|^2 &= c' \Delta c - \int_0^\tau c' (I - Z) \Lambda \varphi^{-1}(t) y(t) y'(t) \Lambda c dt \\ &= c' (\Lambda - (I - Z) \Lambda M \Lambda) c \\ &= c' \Delta c, \end{aligned}$$

where

$$(5.17) \quad \Delta = (I - Z) \Lambda.$$

The formulas (5.16) and (5.17) are the same as in the discrete case (NORBERG, 1986) and, accordingly, Δ can be termed the *risk matrix*.

6. A GENERALIZED POISSON PROCESS WITH STOCHASTIC INTENSITY

A. Basic model assumptions

Let $\{x(t)\}_{t \geq 0}$ be a risk process of generalized Poisson type with claim intensity at time t of the form $w(t)\theta(t)$, where $w(t)$ is the size of the risk at time t and $\theta(t)$ is a basic claim intensity per unit amount of risk exposure at time t . Taking this as the conditional model for fixed intensity process, assume now that $\{\theta(t)\}_{t \geq 0}$ is a stationary stochastic process, representing fluctuating basic risk conditions. Put

$$(6.1) \quad \beta = E\theta(t), \quad \rho(|t-s|) = \text{Cov}(\theta(s), \theta(t)).$$

Assume furthermore that the single claim amounts are i.i.d. replicates of a square integrable random variable y , and that they are independent of the number of claims and of the process θ .

B. Estimation

Consider the problem of predicting $m = x(\infty) - x(\tau)$, the total future claim amount in respect of claims for which the insurer has assumed liability by time τ . It is assumed that $\int_{\tau}^{\infty} w(\vartheta) d\vartheta < \infty$. The problem is to estimate m by a linear function of the continuously recorded observations of the process x up to the present moment τ .

The following conditional moments need no explanation:

$$\begin{aligned} E(m|\theta) &= \int_{\tau}^{\infty} w(\vartheta) \theta(\vartheta) d\vartheta Ey, \\ E(x(t)|\theta) &= \int_0^t w(\vartheta) \theta(\vartheta) d\vartheta Ey, \\ \text{Cov}(x(s), x(t)|\theta) &= \text{Var}(x(s \wedge t)|\theta) \\ &= \int_0^{s \wedge t} w(\vartheta) \theta(\vartheta) d\vartheta Ey^2, \\ \text{Cov}(m, x(t)|\theta) &= 0. \end{aligned}$$

The unconditional moments needed in the linear estimator are

$$\begin{aligned}
 Em &= \int_{\tau}^{\infty} w(\vartheta) d\vartheta \beta Ey, \\
 Ex(t) &= \int_0^t w(\vartheta) d\vartheta \beta Ey, \\
 \text{Cov}(x(s), x(t)) &= \int_0^s \int_0^t w(\vartheta) w(\psi) \rho(|\vartheta - \psi|) d\psi d\vartheta E^2 y + \\
 &\quad + \int_0^{s \wedge t} w(\vartheta) d\vartheta \beta Ey^2, \\
 \text{Cov}(m, x(t)) &= \int_{\tau}^{\infty} \int_0^t w(\vartheta) w(\psi) \rho(|\vartheta - \psi|) d\psi d\vartheta E^2 y.
 \end{aligned}$$

It is easy to check that the normal equation (4.14) now becomes

$$(6.2) \quad \int_{\tau}^{\infty} w(s) \rho(s-t) ds = \int_0^{\tau} w(s) \rho(|s-t|) \gamma(s) ds + \gamma(t) \beta Ey^2 / E^2 y,$$

$t \in (0, \tau)$. This is an integral equation of the Fredholm type. It can be solved numerically, and this is actually the only way unless ρ and w are particularly nice functions. Modelling ρ by a simple parametric function would usually not evoke any misgivings, but the size of the portfolio is not likely to be governed by some simple law that could justify a parametric specification of w .

C. The case with exponential covariance function for the intensity process

Assume now that

$$(6.3) \quad \rho(\vartheta) = \lambda e^{-\kappa\vartheta}, \quad \vartheta \geq 0,$$

where $\lambda, \kappa \geq 0$. This covariance function arises from many models. One possibility is to let $\theta(t)$ be a stationary, continuous time Markov chain with two states and constant transition intensities.

Upon inserting (6.3) in (6.2) and then multiplying by $e^{-\kappa t}$, one obtains

$$\begin{aligned}
 (6.4) \quad \int_{\tau}^{\infty} w(s) e^{-\kappa s} ds &= \int_0^t w(s) e^{\kappa s} \gamma(s) ds e^{-2\kappa t} + \\
 &\quad + \int_t^{\tau} w(s) e^{-\kappa s} \gamma(s) ds + \eta e^{-\kappa t} \gamma(t),
 \end{aligned}$$

with

$$(6.5) \quad \eta = (\beta Ey^2) / (\lambda E^2 y).$$

Differentiation with respect to t in (6.4) and rearranging a bit, gives

$$\int_0^t w(s) e^{\kappa s} \gamma(s) ds e^{-2\kappa t} = (\eta/2\kappa) \frac{d}{dt} (e^{-\kappa t} \gamma(t)).$$

Substituting this in the first term on the right of (6.4) and introducing

$$(6.6) \quad v(t) = e^{-\kappa t} \gamma(t),$$

one arrives at

$$(6.7) \quad \int_{\tau}^{\infty} w(s) e^{-\kappa s} ds = (\eta/2\kappa) v'(t) + \int_t^{\tau} w(s) v(s) ds + \eta v(t).$$

At this point the appearance of the function w becomes decisive. Just as an example, suppose $w(t) = w$ (constant) for $t \in [0, \tau]$. Then differentiation of (6.7) yields

$$0 = (\eta/2\kappa) v''(t) + \eta v'(t) - wv(t),$$

a homogeneous second order differential equation with constant coefficients. Its general solution is of the form

$$(6.8) \quad v(t) = k_1 e^{r_1 t} + k_2 e^{r_2 t},$$

where k_1 and k_2 are some constants, and

$$(6.9) \quad r_1 = -\kappa + \sqrt{\kappa^2 + 2w\kappa/\eta}, \quad r_2 = -\kappa - \sqrt{\kappa^2 + 2w\kappa/\eta}$$

are the (necessarily distinct if $\kappa > 0$) roots of the quadratic equation

$$0 = (\eta/2\kappa) r^2 + \eta r - w.$$

Combining (6.6) and (6.8), one obtains

$$(6.10) \quad \gamma(t) = k_1 e^{t\sqrt{\kappa^2 + 2w\kappa/\eta}} + k_2 e^{-t\sqrt{\kappa^2 + 2w\kappa/\eta}}.$$

Finally, the constants k_1 and k_2 are determined upon substituting (6.10) into (6.4) with $w(s) = w$ and inserting two distinct values of t , e.g. the boundary points 0 and τ , yielding two simple linear equations.

In the form (4.12) the optimal estimator becomes

$$(6.11) \quad \begin{aligned} \bar{m} &= \int_{\tau}^{\infty} w(t) dt \beta Ey + \int_0^{\tau} \gamma(t) \{dx(t) - w\beta Ey dt\} \\ &= \left(\int_{\tau}^{\infty} w(t) dt - w \int_0^{\tau} \gamma(t) dt \right) \beta Ey + \sum_{j: t_j \leq \tau} \gamma(t_j) y_j, \end{aligned}$$

where the t_j are the jump epochs of the process x and $y_j = x(t_j) - x(t_j^-)$ is the amount of the j -th claim.

It might be expected that, since only the process counting the number of claims contains information about the θ and this process can be constructed from the continuously observed process x , the claim amounts should not

appear in an optimal estimator. They never the less do, and this is due to the very choice of the form of the linear estimator. In the present model it would be better to put the claim number process in the role of $\{x(t)\}_{0 < t \leq \tau}$, whereby $y_j = Ey = Ey^2 = 1$ is to be inserted in (6.5) and (6.11).

D. Comparison with a discrete time model

The covariance structure (6.3) is the continuous time analogue of the one introduced by SUNDT (1981) in an attempt to modify the classical Bühlmann-Straub model so as to account for fluctuating basic risk conditions. Sundt did not obtain an explicit formula for the discrete time estimator, but he devised a handy procedure for recursive calculation.

7. CONCLUDING REMARKS

A. Computational feasibility of continuous and discrete estimators

In Sections 5 and 6 some traditional discrete time models for the risk process were reformulated to fit into a continuous time framework, and explicit formulas were obtained for the optimal linear estimators. It turned out that, in the situations investigated here, the problem of solving the normal equations in the continuous case presents no greater difficulties than solving those of the discrete case. In fact, the continuous variant may be the simpler, as is illustrated in Paragraph 6D. In the discrete case one has to solve the normal equation (3.6), which amounts to inverting the matrix $\text{Var } x$. In the regression model studied in Paragraph 3B this is no problem: the special covariance structure implies that only matrices of order $q \times q$ (q is the dimension of the regression coefficient vector) have to be inverted. In the discrete time case one is in general compelled to look for model assumptions that effectively bound the dimensionality of the normal equations or that allow for recursive calculations. Of course, in the continuous time case the "dimension" of the normal equations is infinite unless the moments involved have a finite parametrization as in the examples studied above. However, the integral equation (4.11) can in general be solved numerically, no matter how ugly the moments are, if they only are smooth functions. This points to an advantage of the continuous approach over the traditional discrete one.

B. Continuous versus discrete models

Another important aspect of the continuous approach concerns modelling. Actually, the risk process is really developing continuously, and so it is the discrete model that represents an idealization. It is interesting to note that the discrete regression model in Paragraph 3B complies with the continuous one in Section 5, but already in the slightly more complex situation considered in Section 6 the discrete time model does not arise from a discretization of the observations in the continuous time model.

Finally, observe that discretized estimators of the form (3.1) are obtained in the continuous set-up by taking a in (4.1) to have a jump of size a_j at each $t_j, j = 1, \dots, n$ and to be constant elsewhere, or, equivalently, let g in (4.2) be constant and equal to g_j in each interval $(t_{j-1}, t_j]$. For a fixed partitioning $0 = t_0 < t_1 < \dots < t_n = \tau$ the class of such estimators is a linear subspace of the space of continuous linear estimators defined in Section 4. Therefore, in a continuous time model the optimal solution found in Section 4 is at least as good as the discretized one found in Section 3. In this trivial sense continuous estimators are superior to discrete, but the argument is void if a discrete time model is chosen, of course.

REFERENCES

- BAUER, H. (1978) *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie*. De Gruyter, Berlin.
- BÜHLMANN, H. and STRAUB, E. (1970) Glaubwürdigkeit für Schadensätze. *Mitteil. Schweiz. Verein. Vers. math.* **70**, 111–133.
- DE VYLDER, F. (1976) Geometrical credibility. *Scand. Actuarial J.* **1976**, 121–149.
- HACHEMEISTER, C. (1975) Credibility for regression models with application to trend. In *Credibility: Theory and Applications* (ed. P. M. KAHN), 129–163, Academic Press, New York.
- HISS, K. (1991) Lineare Filtration und Kreditabilitätstheorie. *Mitteil. Schweiz. Verein. Vers. math.* **1991**, 85–103.
- KALLIANPUR, G. and KARINDAKAR, R. L. (1988) *White Noise Theory of Prediction, Filtering and Smoothing*. Gordon and Breach Science Publishers, New York.
- NÄTHER, W. (1984) Bayes estimation of the trend parameter in random fields. *Statistics* **15**, 553–558.
- NORBERG R. (1986). Hierarchical credibility: analysis of a random effect linear model with nested classification. *Scand. Actuarial J.* **1986**, 204–222.
- RUYMGAART, R. A. and SOONG T. T. (1988) *The Mathematics of Kalman-Bucy Filtering*. Springer-Verlag, Berlin, Heidelberg, New York.
- SUNDT, B. (1981) Recursive credibility estimation. *Scand. Actuarial J.* **1981**, 3–21.

RAGNAR NORBERG

*Laboratory of Actuarial Mathematics, Universitetsparken 5,
DK-2100 Copenhagen Ø, Denmark.*

