APPLIED STOCHASTIC MODELS AND DATA ANALYSIS, VOL. 13, 1–13 (1997)

EVALUATING PARTIALLY OBSERVED SURVIVAL HISTORIES: RETROSPECTIVE PROJECTION OF COVARIATE TRAJECTORIES

ANATOLI I. YASHIN, KENNETH G. MANTON* AND GENE R. LOWRIMORE
Duke University, Center for Demographic Studies, 2117 Campus Drive, Durham, NC 27708, U.S.A.

SUMMARY

The use of maximum likelihood methods in analysing times to failure in the presence of unobserved randomly changing covariates requires constrained optimization procedures. An alternative approach using a generalized version of the EM-algorithm requires smoothed estimates of covariate values. Similar estimates are needed in evaluating past exposures to hazardous chemicals, radiation or other toxic materials when health effects only become evident long after their use. In this paper, two kinds of equation for smoothing estimates of unobserved covariates in survival problems are derived. The first shows how new information may be used to update past estimates of the covariates' values. The second can be used to project the covariates' trajectory from the present to the past. If the hazard function is quadratic in form, both types of smoothing equation can be derived in a closed analytical form. Examples of both types of equation are presented. Use of these equations in the extended EM-algorithm, and in estimating past exposures to hazardous materials, are discussed. © 1997 by John Wiley & Sons, Ltd.

Appl. Stochastic Models Data Anal., Vol. 13, 1–13 (1997)
(No. of Figures: 3 No. of Tables: 0 No. of Refs: 23)

KEY WORDS randomly changing covariates; survival analysis; smoothing equations; survival history; evaluation of exposure

1. INTRODUCTION

Survival analyses examine the rate at which morbid or lethal events occur in a population, and the relation of those rates to covariates. The goal is generally to predict future survival under specific covariate conditions. An alternative problem—to evaluate past covariate values given information about current morbid or lethal events—is less commonly done. Such analyses are important when events do not occur at the expected rate. In those circumstances one might wish retrospectively to evaluate unobserved covariate trajectories to determine if they operated to prevent, or stimulate, the occurrence of events. Such evaluations can be done if the form of the hazard function, its parameter values, and probabilistic properties of the covariates are known from theoretical, or other empirical, bases.

For example, the level of asbestos exposure, and the risk of mesothelioma, has been estimated in a number of studies of selected occupationally exposed populations with generally very high exposure levels (e.g. shipyard workers, heating and insulation workers[1]). Such estimates can be used, with appropriate models, to evaluate past exposure to asbestos among workers in other industries where the fact of asbestos exposure is known, but not the quantitative level of

*Correspondence to: K. G. Manton.

CCC 8755–0024/97/010001–13 $17.50

Received 18 April 1995
Revised 6 August 1996

exposure (e.g. construction workers). This is of practical, as well as scientific, consequence since such estimates may determine compensation for asbestos-exposed workers in occupations not intensively studied.[2,3]

Similar problems arise for other chemical, medication or pesticide exposures where significant adverse physiological or environmental effects can become evident long after their use. If the influence of exposure on survival is established in studies of select populations with measured exposure levels, those studies can be used to estimate exposure levels for persons whose exposure was not directly measured using parametrized hazard models where the functional dependency of the hazard on covariates is known—or can be rationalized by theoretical arguments (e.g. the Weibull hazard model for the multi-hit/stage model of carcinogenesis[4]).

A model appropriate for retrospective analyses of unobserved randomly changing covariates must both be able to probabilistically describe the processes of covariate change as well as be able to describe how covariates influence survival (i.e. times to failure). Proportional hazard or logistic regression models (which are often used to estimate regression parameters when covariates are measured at fixed times) do not contain the necessary structure and parameters to conduct retrospective analyses of unobserved stochastically evolving covariates. We present a model, comprising both a quadratic hazard function and a Gaussian–Markov process describing covariate evolution over time, which can be used to derive the necessary equations to analyse retrospectively the effects of latent covariate trajectories on mortality (or other discrete health changes). This model was initially suggested for use in biological applications by Woodbury and Manton,[5] and examined by Yashin[6] for analysis of longitudinal studies (e.g. the Framingham Heart Study.[7,8]

Below, we show how this two-component model can be used in retrospective covariate analyses by deriving the corresponding 'smoothing' (conditional expectations for different order moments) equations. Smoothing equations for diffusion type stochastic processes have been discussed by Lipster and Shiryayev.[9] They did not, however, describe stochastic processes with jump components (i.e. discrete changes—such as mortality) which are needed to analyse survival. Khametov and Yashin[10] derived smoothing equations for multivariate point processes with observed trajectories. Smoothing estimates have also been used to analyse observed diffusion processes.[11,12] Yashin and Manton[13] proved that smoothing estimates of covariate trajectories were needed to extend the EM-algorithm for survival problems with unobserved (or partially observed) randomly changing covariates. Below, two types of smoothing equations will be presented. One type, the backward smoothing equations, are appropriate for estimating the past values of covariates whose evolution is described by a certain general class of stochastic process. A second type, the forward smoothing equations, are a computationally efficient way of updating model parameter estimates when new survival information becomes available.

2. SURVIVAL AND THE INFLUENCE OF UNOBSERVED COVARIATES; GENERAL PROPERTIES

The survival of members of a cohort may be influenced by an unobserved stochastic process. Assume this process operates independently for each individual. Denote the value of an unobserved time varying factor, possibly affecting survival at t, as Y_t. Assume also that the individual's hazard rate is influenced by a Y_t whose change is described by a stochastic process with initial condition Y_0, time dependence $a_0(t)$, state dependence $a(t)Y_t$ and diffusion

generated by a Wiener process with a time dependent scale parameter $b(t)$, i.e.

$$Y_t = Y_0 + \int_0^t (a_0(u) + a(u)Y_u)\,\mathrm{d}u + \int_0^t b(u)\,\mathrm{d}W_u \tag{1}$$

Y_0 is assumed normally distributed with initial mean m_0, and variance Γ_0. There is no explicit formulation of a stochastic risk factor process like (1) in standard survival models, e.g. Cox regression.[18] Consequently, those models implicitly assume that unobserved stochastically evolving factors with systematic drift (i.e. state dependence) do not influence survival so that the coefficient estimates can be assumed to be unbiased.

When we have multiple longitudinal studies with statistically significant, but discordant, results, the differences could be due to the influence of unobserved stochastic risk factor processes. To assess this, more general types of hazard model containing a covariate process like (1) must be used.

To produce unbiased estimates of stochastic risk factor process parameters for survivors to a given time, equation (1) must also be adjusted for systematic mortality selection. This requires (1) to be simultaneously estimated with the hazard function parameters. Consequently, the hazard function selected must be mathematically consistent with the form of the process as described by equation (1). If the time to death for individuals is indicated by random variable T, the survival function for an individual, conditional on the trajectory of Y_u over the interval $[0, t]$, can be written as a quadratic function with a symmetric coefficient matrix (u) (an asterisk represents the transposition operator), or

$$P(T > t \,|\, Y_0^t) = \exp\left(-\int_0^t Y_u^* Q(u) Y_u \,\mathrm{d}u\right) \tag{2}$$

where $Y_0^t = \{Y_u, 0 \leqslant u \leqslant t\}$.

The use of a quadratic hazard function can be empirically justified in many epidemiological studies, e.g. many longitudinal epidemiological studies show a U or J-shaped relation between cardiovascular risk factors (e.g. serum cholesterol, diastolic blood pressure, body mass index[14]) and the risk of total mortality.[15,16] The quadratic hazard function might also be numerically justified as the first two terms from a Taylor series expansion approximating a more general hazard function.[17] Practically, most data sets will have sufficient information only to estimate the first two terms of such an approximation. Though we restrict ourselves here to quadratic hazard functions, the form of the dependence between the hazard and higher-order terms (potentially affecting greater than second order moments of the risk factor process) can be derived using conditional semi-invariant procedures.[18]

The matrix of hazard coefficients $Q(u) = \| q_{i,j}(u) \|\ i, j = 1, 2, \ldots, n$ for each $u > 0$ is non-negative-definite. For any $t > 0$, elements of Q satisfy

$$\int_0^t \sum_{i,j=1}^{n} |\, q_{i,j}(u) \,|\, \mathrm{d}u < \infty \tag{3}$$

If (3) holds, the unconditional survival, $S(t) = P(T > t)$, is[6]

$$S(t) = \exp\left[-\int_0^t (\mathrm{tr}\, Q(u)\Gamma(u) + m^*(u)Q(u)m(u))\,\mathrm{d}u\right] \tag{4}$$

$m(u) = E(Y_u \,|\, T > u)$ is the mean of Y_u for survivors to u, and $\Gamma(u)$ is the variance—covariance matrix of the time dependent conditional Gaussian distribution of Y:

$$\frac{\partial}{\partial y} P(Y_u < y \,|\, T > u) = \frac{1}{(2\pi\,|\,\Gamma(u)\,|\,)^{1/2}} \exp[-(y - m(u))^* \Gamma^{-1}(u)(y - m(u))] \tag{5}$$

where $|\Gamma(u)|$ is the determinant of matrix $\Gamma(u)$ and $m(u)$ and $\Gamma(u)$ satisfy

$$\frac{dm(u)}{du} = a_0(u) + a(u)m(u) - 2\Gamma(u)Q(u)m(u), \qquad m(0) = m_0 \tag{6}$$

and

$$\frac{d\Gamma(u)}{du} = a(u)\Gamma(u) + \Gamma(u)a^*(u) + b(u)b^*(u) - 2\Gamma(u)Q(u)\Gamma(u), \qquad \Gamma(0) = \Gamma_0 \tag{7}$$

In (4), $S(t)$ is not only a quadratic function of $m(u)$ but also, in part, a function of the dispersion $\Gamma(u)$, of Y_u.

3. SMOOTHING EQUATIONS FOR RISK FACTOR PROCESS

Equations (6) and (7) are the filtration equations for the first-, m, and second-, Γ, order moments of the process Y_t. Once the filtration equations are estimated, they can be used to find the smoothing equations necessary to either estimate past values of Y_t—or to update parameter estimates with new survival information. Needed for either task are estimates of the means (m) and variance–covariance (Γ) of Y_t, conditional on survival or, for $s \leqslant t$,

$$m(s, t) = E(Y_s \mid T > t)$$
$$\Gamma_{12}(s, t) = E((Y_t - m(t))(Y_s - m(s, t))^* \mid T > t)$$
$$\Gamma_{22}((s, t) = E((Y_s - m(s, t))(Y_s - m(s, t))^* \mid T > t)$$
$$\Gamma_{21}(s, t) = E((Y_s - m(s, t))(Y_t - m(t))^* \mid T > t)$$

Yashin and Manton [13] proved that such estimates must be calculated for the E-step of an extended EM-algorithm to analyse survival in the presence of unobserved randomly changing covariates. Here we derive the forward and backward smoothing equations for estimating m and Γ and discuss their use in analysing covariate trajectories.

Two theorems describe the derivation of smoothing estimators for the stochastic process specified in (1). The first theorem describes the forward smoothing equation which can be used with the initial risk factor conditions to calculate their future changes given new information on survival (i.e. updates on the number of events occurring over time).

Theorem 1

Let the random variable T and stochastic process, $Y = (Y_t)$, $t \geqslant 0$ satisfy (1) and (2). Assume $s \leqslant t$. The forward smoothing equations for the mean $m(s, t)$ and covariances $\Gamma_{12}(s, t)$. $\Gamma_{21}(s, t)$, $\Gamma_{22}(s, t)$ are the integral forms

$$m(s,t) = m(s) - 2\int_s^t \Gamma_{21}(s,u)Q(u)m(u)\,du \tag{8}$$

$$\Gamma_{12}(s,t) = \Gamma(s) + \int_s^t a(u)\Gamma_{12}(u,s)\,du - 2\int_s^t \Gamma(u)Q(u)\Gamma_{12}(u,s)\,du \tag{9}$$

$$\Gamma_{21}(s,t) = \Gamma(s) + \int_s^t \Gamma_{21}(s,u)a^*(u)\,du - 2\int_s^t \Gamma_{21}(s,u)Q(u)\Gamma(u)\,du \tag{10}$$

$$\Gamma_{22}(s,t) = \Gamma(s) - 2\int_s^t \Gamma_{12}(u,s)Q(u)\Gamma_{21}(s,u)\,du \tag{11}$$

where $m(u)$ and $\Gamma(u)$ are estimated from the filtration equations (6) and (7).

The proof that these are the appropriate forward smoothing estimators is in the Appendix. These equations apply when s (the time the study was started, i.e. initial conditions are known) is fixed and t is *increasing*.

To estimate covariate values when t is *fixed* and s is *decreasing* requires '*backward* smoothing equations', i.e. one has fixed observations on a time interval $[0, t]$ and wishes to estimate prior values of Y. This is the type of equation needed to make backward projections of risk factor trajectories, i.e. to project such trajectories prior to the initial conditions.

Theorem 2

Let the random variable T and stochastic process $Y = (Y_t)$, $t \geq 0$, be as specified above, with $s < t$. The *backward* smoothing estimates of the mean $m(s, t)$ and elements of the covariance matrix $\Gamma_{12}(s, t)$, $\Gamma_{21}(s, t)$, $\Gamma_{22}(s, t)$ for s years in the past are the derivatives with respect to s of the moments, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}s} m(s,t) = a_0(s) + a(s)m(s,t) + b(s)b^*(s)\Gamma^{-1}(s)(m(s,t) - m(s)); \qquad m(t,t) = m(t) \tag{12}$$

$$\frac{\mathrm{d}}{\mathrm{d}s} \Gamma_{21}(s,t) = a(s)\Gamma_{21}(s,t) + b(s)b^*(s)\Gamma^{-1}(s)\Gamma_{21}(s,t); \qquad \Gamma_{21}(t,t) = \Gamma(t) \tag{13}$$

$$\frac{\mathrm{d}}{\mathrm{d}s} \Gamma_{12}(t,s) = \Gamma_{12}(t,s)a^*(s) + \Gamma_{12}(s,t)\Gamma^{-1}(s)b(s)b^*(s)]; \qquad \Gamma_{12}(t,t) = \Gamma(t) \tag{14}$$

$$\frac{\mathrm{d}}{\mathrm{d}s} \Gamma_{22}(s,t) = a(s)\Gamma_{22}(s,t) + \Gamma_{22}(s,t)a^*(s) - b(s)b^*(s)$$
$$+b(s)b^*(s)\Gamma^{-1}(s)\Gamma_{22}(s,t) + \Gamma_{22}(s,t)\Gamma^{-1}(s)b(s)b^*(s); \qquad \Gamma_{22}(t,t) = \Gamma(t) \tag{15}$$

The proof of the second theorem is also in the Appendix. These equations produce smoothing estimates for a fixed observation interval.

4. EXAMPLE

To evaluate the properties of the smoothing equations we provide an example where full information about the process is available, i.e. the filtration equations can be estimated directly from the available data. Under these conditions we can show that the forward and backward equations provide identical results. Then we discuss differences in the application of the two types of equation when only specific partial information on the process is available.

Consider the one-dimensional equation for the process Y_t:

$$\mathrm{d}Y_t = (a_0 - a_1 Y_t)\,\mathrm{d}t + b\,\mathrm{d}W_t \tag{1'}$$

where the coefficients fulfil the conditions $a_0 > 0$; $a_1 > 0$; $b^2 > 0$. We also assume the quadratic hazard coefficient $q > 0$. We designate $m(0)$, and $\Gamma(0)$ as the initial mean and variance of the process. The filtration equations (6) and (7) for this case are:

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = a_0 - a_1 m(t) - 2q\Gamma(t)m(t); \qquad m(0) \tag{6'}$$

$$\frac{\mathrm{d}\Gamma(t)}{\mathrm{d}t} = -2a_1\Gamma(t) - 2\Gamma^2(t)q + b^2; \qquad \Gamma(0) > 0 \tag{7'}$$

The stationary solutions for m and Γ are

$$\Gamma = \frac{a_1}{2q}\left(\left[1 + \frac{2qb^2}{a_1^2}\right]^{1/2} - 1\right) > 0$$

$$m = \frac{a_0}{a_1 + 2\Gamma q} > 0$$

Graphs of the solutions of the filtration equations for the mean (m) and variance (Γ) as a function of time, from 0 to 2·5, are presented in Figure 1.

The parameter values assumed in the filtration calculations are

$$a_0 = 0{\cdot}5; \qquad a_1 = 2; \qquad b = 1; \qquad q = 1; \qquad m(0) = 1; \qquad \Gamma(0) = 0{\cdot}1$$

The stationary solutions for this example are:

$$\Gamma = 0{\cdot}225, \qquad m = 0{\cdot}204$$

For this example, the stationary values were reached after $t = 1{\cdot}5$.

The continuous time solutions of the filtration equations and forward smoothing equations (8), (10), (11) over time 2·5 to 3·5 are presented in Figure 2 for the same parameter values used to generate the filtration equation solutions in Figure 1.

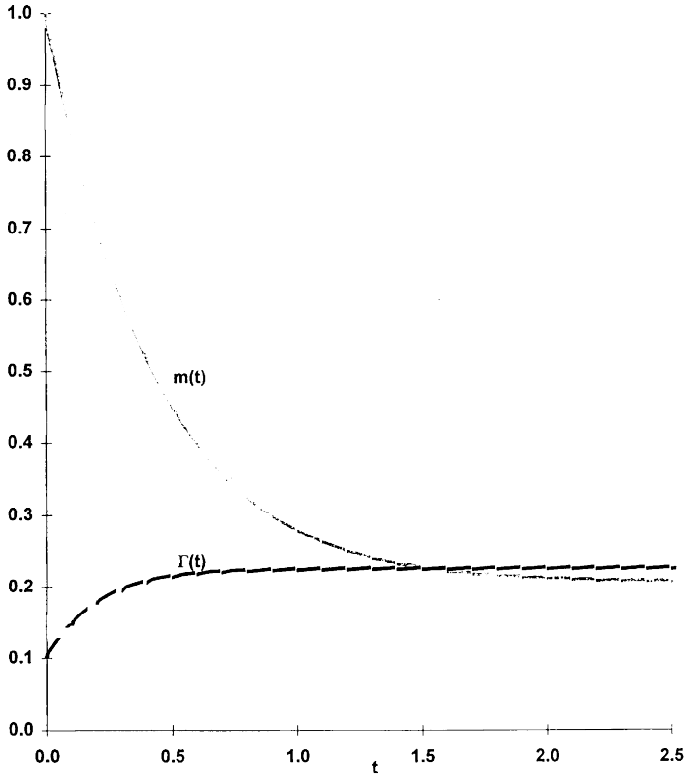

Figure 1. Graphs of $m(t)$ and $\Gamma(t)$ given by filtration equations (6′) and (7′) with initial conditions $m(0) = 1$, $\Gamma(0) = 0{\cdot}1$ and parameter values $a_0 = 0{\cdot}5$; $a_1 = 2$, $b = 1$, $q = 1$

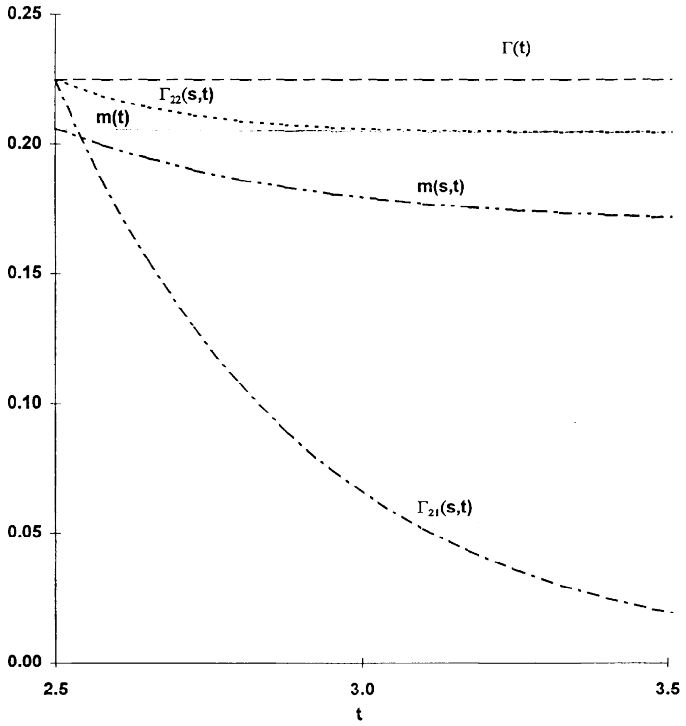

Figure 2. Graphs of $m(s, t)$, $\Gamma_{21}(s, t)$ and $\Gamma_{22}(s, t)$ calculated from forward smoothing equations (8), (10) and (11) with $s = 2{\cdot}5$ together with the solutions of filtration equations (6′) and (7′). Parameter values are the same as in Figure 1.

The smoothed estimate $m(s, t)$ of the mean tends to a stationary value as t (i.e. the period of follow-up or observation of the process) increases past $s(= 2{\cdot}5)$. The smoothed estimate of the mean m of Y_s *declines when* t increases from $t = 2{\cdot}5$ to $3{\cdot}5$. This was expected, since observation of a lethal event $\{T > t)$ at a later time makes it likely to be associated with a lower value (assuming the coefficients for the factor in Q are positive) of the risk factor in the past, i.e. before time t. As expected, the covariance $\Gamma_{21}(s, t)$ of Y_s and Y_t also declines as $(t - s)$ increases, i.e. the relation of risk factor values separated by time decreases as the interval increases.

Figure 3 shows solutions of the backward smoothing equations (12), (13) and (15) and solutions of the filtration equations (6′) and (7′).

The solutions for the backward smoothing equations are found when s decreases from $s = 3{\cdot}5$ to $2{\cdot}5$ and $t = 3{\cdot}5$. The smoothing estimate of the covariate's mean $m(s, t)$, variance $\Gamma_{22}(s, t)$ and covariance $\Gamma_{21}(s, t)$ decline when s decreases (i.e. the difference between s and t increases). One can see that the trajectories in Figures 2 and 3 are different: Figure 2 shows the evolution of $m(s, t)$, $\Gamma_{22}(s, t)$ and $\Gamma_{21}(s, t)$ as t increases. Figure 3 shows the evolution of $m(s, t)$, $\Gamma_{21}(s, t)$ and $\Gamma_{21}(s, t)$ as s decreases. However, the values of $m(s, t)$, $\Gamma_{21}(s, t)$ and $\Gamma_{21}(s, t)$ in Figures 2 and 3 coincide at $(s, t) = (2{\cdot}5, 3{\cdot}5)$ as expected.

Note that despite the fact that forward and backward equations describe the same smoothing estimates as functions of the two variables s and t, in general their use in analysing problems will be different. Forward smoothing equations update the estimate of the covariate's value at a fixed time s in the past when new information becomes available with t increasing. Backward

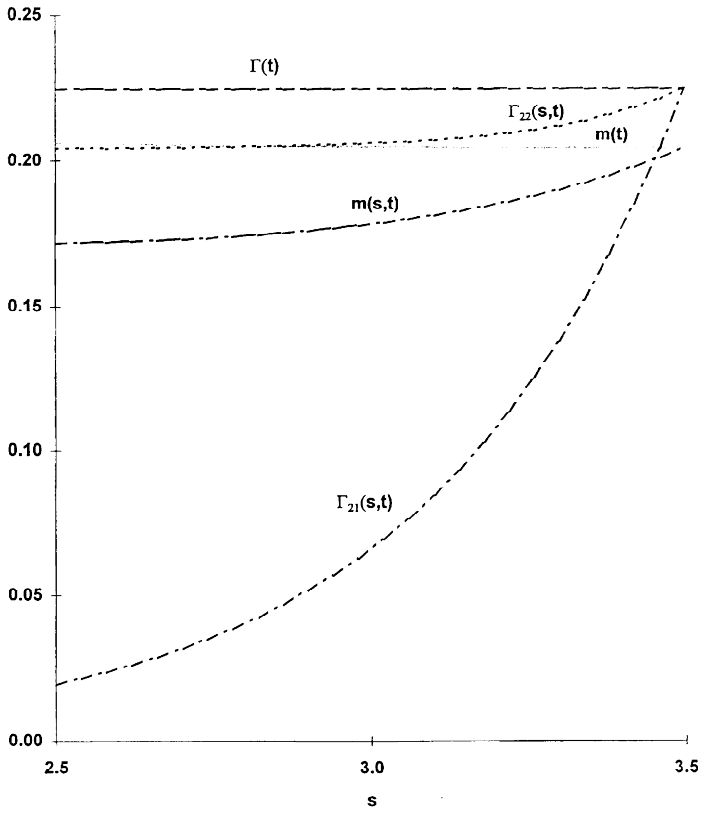

Figure 3. Graphs of $m(s, t)$, $\Gamma_{21}(s, t)$, and $\Gamma_{22}(s, t)$ calculated from backward smoothing equations (12), (13) and (15) with $t = 3{\cdot}5$ together with the solutions of filtration equations (6′) and (7′). Parameter values are the same as in Figure 1

smoothing equations re-evaluate the entire past unobserved covariate trajectory when information up to time t is fixed. These equations can also be used for updating the covariate's values at a fixed point s in the past when t is changing. However, for this purpose one needs to solve the backward equations iteratively starting from the boundary conditions $t_1, t_2, \ldots, t_n$ up to time s instead of one forward equation. Thus, forward equations are preferable for updating information about covariate's values at a fixed point in the past.

5. DISCUSSION

Results for the analysis of survival data with a stochastic process model show that the EM-algorithm can be extended to estimate the parameters of a process of unobserved randomly changing covariates.[13] The generalization of the EM-algorithm required estimating past covariate values. The smoothing estimators presented above can be used for this purpose assuming the hazard function is quadratic. Although both forward and backward smoothing equations are nonlinear differential equations without closed analytic solutions, their numerical evaluation is not difficult. As our example showed, the calculation of both types of equation, assuming full information, required approximately the same effort. Hence, both could be used in extended versions of the EM-algorithm—which of the two types of equation is more

appropriate for use in the algorithm depends on the details of the specific longitudinal observational plan.

Such estimators may be useful for other types of analytic problems. In engineering, reliability and clinical trial applications, one often needs to know what caused, or prevented, system failure. Since many 'pathologic' or 'failure' processes have a long latency, the observed health changes, or the observed evolution of system damage, may be the only measurable signs of unobserved influential processes starting at a given time in the past. For example, the health consequences of exposure to toxic chemicals may be revealed many years after exposure ended. Many health effects of asbestos may only begin to be manifest after 15 years, and may continue to emerge for 50 or more years after exposure.[19]

In randomized clinical trials, to analyse the effectiveness of medical or surgical therapies one often needs to evaluate why an expected health effect did *not* occur. For example, drug trials for lowering cholesterol often show highly significant reductions in coronary heart disease but little effect on total mortality.[20] One explanation of this result is that the observation period was not long enough for total mortality effects to be manifest. Alternately, low cholesterol has been found, in some studies, to be associated with higher cancer risks.[15,20] One hypothesis is that the metabolic effects of cancer lowers cholesterol and that the relation would disappear as one tracks cholesterol values for individuals over time. An alternative hypothesis is that very low cholesterol values damage cell membrane integrity and make cells more susceptible to carcinogenic exposures.

Analysis of such hypotheses about the temporal effects of risk factors requires evaluating the effects of influential, possibly unobserved, processes at multiple times in the past—and doing so in a way to produce parameter estimates not biased by systematic mortality. Such analyses can be done with the procedures presented above. Their most general formulation (i.e. without a specific parametrization of the hazard function or of the cross temporal density functions) involves stochastic differential equations for conditional multivariate probability densities[9] which generalize the well-known Kolmogorov–Fokker–Plank equations for unconditional densities. Those equations, however, do not represent the effects on the distribution of possible loss of density (i.e. mortality). These effects were taken into account by Woodbury and Manton[5] who considered the evolution of the multivariate distribution function generated by individual random walks in a multidimensional spice where probabilistic 'killing' terms are functions of an individual's state space location. Yashin *et al.*[22] included such 'killing' terms in the general nonlinear filtration schemes discussed by Lipster and Shiryayev.[9].

For the specific case of Gaussian processes, the smoothing estimators may be written for the first two moments of the *conditional* Gaussian distribution. Thus, instead of dealing with partial differential equations, one can deal with ordinary differential equations with more manageable, closed form, analytic solutions. This is because a quadratic hazard preserves the multivariate Gaussian distribution of state variables among survivors; i.e., the distribution of unobserved covariates among survivors is Gaussian (with changed mean and variance) if the marginal distribution of the covariates is Gaussian.[23] These equations can be used, for example, to extend the EM-algorithm to estimate the parameters of processes for unobserved randomly changing covariates.[13] If conditions require the quadratic hazard to be generalized (e.g. adding cubic and quartic terms) closed-form solutions are not possible.

6. CONCLUSION

We presented two strategies to evaluate covariate trajectories in a stochastic process model of mortality and aging when partial information on the process is available. The 'backward' and

'forward' smoothing equations presented use properties of the Gaussian distribution of covariate values among survivors assuming a quadratic hazard function to derive a computationally tractable solution to the problem. Forward smoothing equations allow for continuous evaluation of unobserved influential processes at fixed times in the past when survival time is increasing, i.e. as new information about events (or their absence) is forthcoming (e.g. 'monitoring' or 'updating' equations). Backward equations allow the past trajectory, or history, of influential factors to be evaluated when information is 'fixed,' i.e., survival times do not change. Those smoothing equations can be used either for 'reverse' health projections or for evaluation of the past exposure to hazardous materials. They are also crucial in implementing a generalized version of the EM-algorithm for survival influenced by an unobserved stochastic process.

ACKNOWLEDGEMENT

Professor Yashin's effort in this research was supported by NIA Grant No. P01 AG08791. Professors Manton's and Lowrimore's efforts were supported by NIA Grant No. AG01159.

APPENDIX

Proof of Theorem 1

To estimate the forward smoothing equations, one first uses (6) and (7) to solve the filtration problem for a stochastic process. Consider a two-component process satisfying

$$\mathrm{d}Y_1(u) = (a_0(u) + a(u)Y_1(u))\,\mathrm{d}u + b(u)\,\mathrm{d}W_u \tag{A1}$$

and

$$\mathrm{d}Y_2(u) = (a_0(u)I(u \leqslant s) + a(u)I(u \leqslant s)Y_1(u))\,\mathrm{d}u + b(u)I(u \leqslant s)\,\mathrm{d}W_u \tag{A2}$$

Note that for $s < t$, $Y_1(t) = Y_t$ and $Y_2(t) = Y_s$. Let the hazard be a quadratic function of Y_1 and time, or

$$\mu(Y_1(u), Y_2(u), u) = Y_1^*(u)Q(u)Y_1(u) \tag{A3}$$

The application of (6) and (7) and using estimates of Q from (A3)[6] yields the conditional means $m_1(t)$ and $m_2(t)$, $t > s$ as,

$$m_1(t) = m_0 + \int_0^t (a_0(u) + a(u)m_1(u) - 2\Gamma_{11}(u)Q(u)m_1(u)\,\mathrm{d}u \tag{A4}$$

and

$$m_2(t) = m_1(s) - 2\int_s^t \Gamma_{21}(u)Q(u)m_1(u)\,\mathrm{d}u \tag{A5}$$

and the conditional covariance elements, $\Gamma_{11}(t)$, $\Gamma_{12}(t)$, $\Gamma_{21}(t)$ and $\Gamma_{22}(t)$ as

$$\Gamma_{11}(t) = \Gamma_{11}(0) + \int_0^t [a(u)\Gamma_{11}(u) + \Gamma_{11}(u)a^*(u) + b(u)b^*(u) - 2\Gamma_{11}(u)Q(u)\Gamma_{11}(u)]\,\mathrm{d}u \tag{A6}$$

$$\Gamma_{12}(t) = \Gamma_{11}(s) + \int_s^t a(u)\Gamma_{12}(u)\,\mathrm{d}u - 2\int_s^t \Gamma_{11}(u)Q(u)\Gamma_{12}(u)\,\mathrm{d}u \tag{A7}$$

$$\Gamma_{21}(t) = \Gamma_{11}(s) + \int_s^t \Gamma_{21}(u)\,a^*(u)\,\mathrm{d}u - 2\int_s^t \Gamma_{21}(u)Q(u)\Gamma_{11}(u)\,\mathrm{d}u \tag{A8}$$

and

$$\Gamma_{22}(t) = \Gamma_{11}(s) - 2\int_s^t \Gamma_{21}(u)Q(u)\Gamma_{12}(u)\,\mathrm{d}u \tag{A9}$$

according to the definitions $m_1(t) = m(t)$; $m_2(t) = m(s, t)$; $\Gamma_{11}(t) = \Gamma(t)$, $\Gamma_{12}(t) = \Gamma_{12}(s, t)$, $\Gamma_{21}(t) = \Gamma_{21}(s, t)$, $\Gamma_{22}(t) = \Gamma_{22}(s, t)$. This proves Theorem 1. □

To prove theorem 2, for the 'backward' smoothing equations, we first prove the following lemma.

Lemma A.1

If

$$g(s, t) = \Gamma_{21}(s, t)\Gamma^{-1}(t) \tag{A10}$$

then $g(s, t)$ satisfies,

$$\frac{\mathrm{d}g(s,t)}{\mathrm{d}t} = -g(s,t)(a(t) + b(t)b^*(t)\Gamma^{-1}(t)), \qquad g(s,s) = I \tag{A11}$$

where I is the identity matrix.

Proof
Differentiating $\Gamma^{-1}(t)\,\Gamma(t) = 1$ with respect to t, and after simple transformations, one obtains,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Gamma^{-1}(t)) = -\Gamma^{-1}(t)a(t) - a^*(t)\Gamma^{-1}(t) - \Gamma^{-1}(t)b(t)b^*(t)\Gamma^{-1}(t) + 2Q(t) \tag{A12}$$

Using (A12), we differentiate $g(s, t) = \Gamma_{21}(s, t)\,\Gamma^{-1}(t)$ over t to get (A11). □

Remark

Note a solution of a system of linear equations $g(s, t)$ satisfies

$$g(s, t) = g(0, s)^{-1}g(0, t), \qquad s \leqslant t$$

The equation for $g(0, s)^{-1}$ can be found by differentiating

$$g(0,s)^{-1}g(0, s) = I$$

This is,

$$\frac{\mathrm{d}}{\mathrm{d}s}(g(0,s)^{-1}) = (a(s) + b(s)b^*(s)\Gamma^{-1}(s))g(0,s)^{-1} \tag{A13}$$

Similarly, one can show

$$g^*(0,t) = I - \int_0^t (a^*(u) + b(u)b^*(u)\Gamma^{-1}(u))g^*(0,u)\,\mathrm{d}u \tag{A14}$$

and

$$g^*(0,t)^{-1} = I + \int_0^t g^*(0,u)^{-1}(a^*(u) + b(u)b^*(u)\Gamma^{-1}(u))\,\mathrm{d}u \tag{A15}$$

where I is the identity matrix.

Proof of Theorem 2

Consider the equations for covariance elements $\Gamma_{21}(s,t)$ and $\Gamma_{22}(s,t)$. Putting $\Gamma_{21}(s,t) = g(s,t)\Gamma(t)$ on the r.h.s. of (8), for $m(s,t)$, produces

$$m(s,t) = m(s) - 2g(0,s)^{-1}\int_s^t g(0,u)\Gamma(u)Q(u)m(u)\,\mathrm{d}u \tag{A16}$$

Differentiating (A16) with respect to s we get

$$\frac{\mathrm{d}m(s,t)}{\mathrm{d}s} = a_0(s) + a(s)m(s,t) - 2[a(s) + b(s)b^*(s)\Gamma^{-1}(s)]\int_s^t \Gamma_{21}(s,u)Q(u)m(u)\,\mathrm{d}u \tag{A17}$$

Since, in accordance with (8),

$$-2\int_s^t \Gamma_{21}(s,u)Q(u)m(u)\,\mathrm{d}u = m(s,t) - m(s)$$

(A17) may be transformed to

$$\frac{\mathrm{d}m(s,t)}{\mathrm{d}s} = a_0(s) + a(s)m(s,t) + b(s)b^*(s)\Gamma^{-1}(s)[m(s,t) - m(s)]$$

which is (12). The boundary condition is $m(t,t) = m(t)$.

Using (A8), we have for $\Gamma_{21}(s,t) = g(s,t)\Gamma(t)$

$$\Gamma_{21}(s,t) = \Gamma(s) + g(0,s)^{-1}\int_s^t g(0,u)\Gamma(u)a^*(u)\,\mathrm{d}u - 2g(0,s)^{-1}\int_s^t g(0,u)\Gamma(u)Q(u)\Gamma(u)\,\mathrm{d}u \tag{A18}$$

After differentiating both parts of (A18) with respect to s, taking into account (A6) and (A13) for $\Gamma(s)$ and $g(0,s)^{-1}$, and using the equality

$$\int_s^t (\Gamma_{21}(s,u)a^*(u) - 2\Gamma_{21}(s,u)Q(u)\Gamma(u))\,\mathrm{d}u = \Gamma_{21}(s,t) - \Gamma(s)$$

which follows from (10), we get the differential equation,

$$\frac{\mathrm{d}}{\mathrm{d}s}\Gamma_{21}(s,t) = a(s)\Gamma_{21}(s,t) + b(s)b^*(s)\Gamma^{-1}(s)\Gamma_{21}(s,t)$$

for $\Gamma_{21}(s,t)$, with boundary condition $\Gamma_{21}(t,t) = \Gamma(t)$ which is (13).

Now, consider the equation for $\Gamma_{22}(s,t)$:

$$\Gamma_{22}(s,t) = \Gamma(s) - 2g(0,s)^{-1}\left[\int_s^t g(s,u)\Gamma(u)Q(u)\Gamma(u)g^*(s,u)\,\mathrm{d}u\right]g^*(0,s)^{-1}$$

Differentiating both parts of this equality with respect to s and using the equality

$$-2\int_s^t \Gamma_{12}(s,u)Q(u)\Gamma_{21}(s,u)\,\mathrm{d}u = \Gamma_{22}(s,t) - \Gamma(s)$$

which follows from (11), we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\Gamma_{22}(s,t) = a(s)\Gamma_{22}(s,t) + \Gamma_{22}(s,t)a^*(s) - b(s)b^*(s) + b(s)b^*(s)\Gamma^{-1}(s)\Gamma_{22}(s,t) + \Gamma_{22}(s,t)\Gamma^{-1}(s)b(s)b^*(s)$$

with boundary condition $\Gamma_{22}(t,t) = \Gamma(t)$, which proves Theorem 2. □

REFERENCES

1. K. G. Manton, 'An evaluation of strategies for forecasting the implications of occupational exposure to asbestos', Report prepared for the Congressional Research Service, Government Division, 1985.
2. A. M. Walker, 'Projections of asbestos-related disease 1980–2009', Final Report, Epidemiology Resources, Inc., Chestnut Hill, Massachusetts, U.S.A, 1982.
3. I. Selikoff, Correspondence to the Honorable Jack B. Weinstein, U.S. District Court, Brooklyn, New York. Mount Sinai Medical Center, 7 February 1991.
4. P. Armitage and R. Doll, 'Stochastic models for carcinogenesis', in J. Neyman (ed), *Proceedings of the 11th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, California, U.S.A., 1961, pp. 19–38.
5. M. A. Woodbury and K. G. Manton, 'A random walk model of human mortality and aging', *Theoretical Population Biology*, **11**, 37–48 (1977).
6. A. I. Yashin, 'Dynamics in survival analysis. conditional Gaussian property versus Cameron–Martin formula', in N. V. Krylov, R. S. Lipster and A. A. Novikov (eds), *Statistic and Control of Stochastic Processes* (Proceedings of the Stector Seminar), Springer-Verlag, New York, 1985.
7. K. G. Manton and E. Stallard, *Chronic Disease Modeling*: *Measurement and Evaluation of the Risks of Chronic Disease Processes*, Charles Griffin, London, U.K., 1988.
8. L. Cupples, R. D'Agostino, K. Anderson and W. Kannel, 'Comparison of baseline and repeated measure covariate techniques in the Framingham Heart Study,' *Statistics in Medicine*, **7**, 205–218 (1988).
9. R. Lipster and A. Shiryayev, *Statistics of Random Processes*, Vol. I, *General Theory*, Springer-Verlag, New York, 1977.
10. V. M. Khametov and A. I. Yashin, 'The effective smoothing of random processes under jumping observations', *Problems of Information and Transmission*, **2**, 38–51 (1983).
11. A. Dembo and O. Zeitouni, 'Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm', *Stochastic Processes and Their Applications*, **23**, 91–113 (1986).
12. F. Campilo and F. Le Gland, 'MLE for partially observed diffusions: Direct minimization vs. the EM algorithm', *Stochastic Processes and Their Applications*, **33**, 245–274 (1989).
13. A. I. Yashin and K. G. Manton, 'Modification of the EM algorithm for survival influenced by unobserved stochastic processes', *Stochastic Processes and Their Applications*, **54**, 257–274 (1994).
14. K. G. Manton, E. Stallard, M. A. Woodbury and J. E. Dowd, 'Time varying covariates of human mortality and aging: multidimensional generalization of the Gompertz', *Journal of Gerontology*: *Biological Science*, **49**, B169–B190 (1994).
15. D. Jacobs, H. Blackburn, M. Higgin, *et al.*, 'Report of the conference on low blood cholesterol: mortality association', *Circulation*, **86**, 1046–1060 (1992).
16. S. Jacobsen, D. Freedman, R. Hoffmann, H. Gruchow, A. Anderson and J. Barboriak, 'Cholesterol and coronary artery disease: age as an effect modifier', *Journal of Clinical Epidemiology*, **45**, 1053–1059 (1992).
17. K. G. Manton, E. Stallard and B. H. Singer, 'Projecting the future size and health status of the U.S. elderly population', in D. Wise (ed.), *Studies of the Economics of Aging*, National Bureau of Economic Research, University of Chicago Press, Chicago, Illinois, U.S.A., Chapter 2, 1994, pp. 41–77.
18. A. I. Yashin, K. G. Manton, M. A. Woodbury and E. Stallard, 'The effects of health histories on stochastic process models of aging and mortality', *Journal of Mathematical Biology*, **33**, 1–16 (1995).
19. I. Selikoff and M. Greenbert, 'A landmark case in asbestosis', *Journal of the American Medical Association*, **265**, 898–901 (1991).
20. U. Ravnskov, 'Cholesterol lowering trials on coronary heart disease: frequency of citation and outcome', *British Medical Journal*, **305**, 15–19 (1992).
21. J. D. Neaton, H. Blackburn, D. Jacobs, L. Kuller, D. J. Lee, R. Sherwin, J. Shih, J. Stamler and D. Wentworth, 'Multiple risk factor intervention trial research group: serum cholesterol level and mortality findings for men screened in the multiple risk factor intervention trial', *Archives of International Medicine*, **152**, 1490–1500 (1992).
22. A. I. Yashin, K. G. Manton and J. W. Vaupel, 'Mortality and aging in a heterogeneous population: a stochastic process model with observed and unobserved variables', *Theoretical Population Biology*, **27**, 145–175 (1985).
23. A. I. Yashin, 'Unobserved covariates in survival models: smoothing estimates', Research Report 91-04-01, Center for Population Analysis and Policy, University of Minnesota, Minneapolis, Minnesota, U.S.A., 1991.

