Lifetime Data Analysis, 10, 407–423, 2004
© 2005 Kluwer Academic Publishers. Printed in The Netherlands.

Survival Models Based on the Ornstein-Uhlenbeck Process

ODD O. AALEN o.o.aalen@basalmed.uio.no
Department of Statistics, Institute of Basic Medical Sciences, University of Oslo, P.O.Box 1122 Blindern, N-0317 Oslo, Norway

HÅKON K. GJESSING hakon.gjessing@fhi.no
Norwegian Institute of Public Health, P.O.Box 4404 Nydalen, N-0403 Oslo, Norway

Received December 12, 2003; Revised July 19, 2004; Accepted August 30, 2004

Abstract. When modelling survival data it may be of interest to imagine an underlying process leading up to the event in question. The Ornstein-Uhlenbeck process is a natural model to consider in a biological context because it stabilizes around some equilibrium point. This corresponds to the homeostasis often observed in biology, and also to some extent in the social sciences. First, we study the first-passage time distribution of an Ornstein-Uhlenbeck process, focussing especially on what is termed quasi-stationarity and the various shapes of the hazard rate. Next, we consider a model where the individual hazard rate is a squared function of an Ornstein-Uhlenbeck process. We extend known results on this model. The results on quasi-stationarity are relevant for recent discussions about mortality plateaus. In addition, we point out a connection to models for short-term interest rates in financial modeling.

Keywords: Cox-Ingersoll-Ross process, first-passage time, hazard rate, mortality plateau, Ornstein-Uhlenbeck process, quasi-stationary distribution, survival analysis

1. Introduction

Survival analysis, considered as a branch of statistics, is a thoroughly pragmatic field. The aim is to estimate survival curves and perform regression analyses, with no deeper concern for the stochastic processes that clearly must lie behind the occurrence of events. When time-dependent covariates occur, they are often included pragmatically into a Cox regression analysis, without much concern for the stochastic nature of these covariate processes. An exception to this, however, is found in the modelling of marker processes and in joint models, (see e.g., Lee et al., 2000).

Although it does not immediately solve any practical problems, it may be of conceptual interest to model an underlying process causing an event to occur. A number of such models, based on first-passage time in stochastic processes, were reviewed by Aalen and Gjessing (2001, 2003). Here it was pointed out that the phenomenon of quasi-stationarity may be of interest, among other things as a tool for understanding the various shapes that hazard rates can assume. In demography there is presently a discussion concerning so-called mortality plateaus, and quasi-stationarity has been presented as a means of understanding these phenomena, (see Steinsaltz and Evans, 2004).

Quasi-stationarity is a limiting state for the hazard distribution of survivors, resulting in a constant population hazard rate. A general result on asymptotic constancy of the hazard rate was given already by Keilson (1966). As he points out, the phenomenon is closely related to the process losing its memory when time passes. This is true for Markov processes and regenerative processes. If the process preserves the memory of early events, the hazard will generally not become constant. An example of this is the so-called frailty models, where a basic hazard rate is multiplied by an individual specific frailty. A high frailty indicates an increased vulnerability that is present from the start and does not change over time. This could for instance be a genetic defect. Frailty models often give hazards that eventually decrease. Hence, the asymptotic behavior of hazard rates depends on whether previous effects are retained in the system. In biology one would think that some effects, like genetic ones, are having long term influence, while others are more short-term. The distinction between short-term and long-term memory is also important in stochastic process theory. For instance, in a fractional Brownian motion the degree of long range dependence may be described by the so-called Hurst parameter, H , with $H = 1/2$ corresponding to ordinary Brownian motion. Models with long range dependence often result in distributions that are sub-exponential, that is, an eventually declining hazard rate, (see e.g., Samorodnitsky, 2002). The focus here, however, will be on short-term effects.

The Ornstein-Uhlenbeck process is the prototype of a process that combines diffusion with a point of equilibrium. One would assume that many natural processes has this property of diffusing back and forth, and at the same time tending to stabilize around a certain point. The phenomenon is termed homeostasis and is ubiquitous in biology and social contexts. The same concept appears in an equilibrium approach to term structure modeling in finance, as developed, for instance, in (Cox et al., 1985).

The aim of the present paper is to study survival distributions and quasi-stationarity when the underlying process is an Ornstein-Uhlenbeck process. Two situations will be considered. First, we assume that the event corresponds to the Ornstein-Uhlenbeck process crossing a certain limit. We apply known results about distributions of first-passage times for this process to derive new results about the quasi-stationary distributions. Second, we consider a model where the hazard rate of each individual is modelled directly by the square of an Ornstein-Uhlenbeck process. This model has already been studied by a number of authors (Woodbury and Manton, 1977; Myers, 1981; Yashin and Vaupel, 1986), but here we consider quasi-stationary distributions. The model is an example of the so-called "killing" models for diffusion processes. We furthermore extend the model to a situation where the individual hazard rate is not only dependent on the present value of the Ornstein-Uhlenbeck process, but also on the past.

We wish to present ideas and examples from a statistical point of view, and do not focus on formal mathematical derivation. Suitable references to the probabilistic literature are, however, given. The paper is theoretical in the sense of discussing ideas and principles. Analysis of data are therefore not included. Statistical analyses of

medical data, applying Ornstein-Uhlenbeck type processes may, for instance, be found in the book of Manton and Stallard (1988).

2. First-passage Time for the Ornstein-Uhlenbeck Process

The Wiener process, denoted W_t, is well known as the prototype of a random process with continuous sample paths. Its increment over a time interval is normally distributed with expectation 0 and variance proportional to the length of the interval. The Ornstein-Uhlenbeck (O.-U.) process, denoted X_t, is a modification of a Wiener process, with a drift towards an equilibrium state. It may be defined by the following stochastic differential equation:

$$dX_t = (a - bX_t)dt + \sigma\, dW_t. \tag{1}$$

We shall assume that the parameter $b > 0$. It is typically assumed that X_0 is Gaussian, possibly with variance zero. Eq. (1) means that the change in X_t in a small time interval has a drift towards a/b, but is disturbed by the noise embodied in dW_t (often called white noise). The attraction of the equilibrium point is referred to as the *mean-reverting* property of the O.-U. process. The unrestricted solution, in the absence of boundaries, to (1) can be written

$$X_t = \frac{a}{b} + \left(X_0 - \frac{a}{b}\right)e^{-bt} + \sigma \int_0^t e^{-b(t-s)} dW_s.$$

It is clear that X is Gaussian, and that $EX_t = a/b + (X_0 - a/b)\exp(-bt) \to a/b$ as $t \to \infty$. Furthermore, $\mathrm{Var}(X_t) = \sigma^2/(2b)(1 - \exp(-2bt)) \to \sigma^2/(2b)$. Apart from the run-in caused by $X_0 \neq a/b$, the Ornstein-Uhlenbeck is a stationary Gaussian process with an autocorrelation function that decays exponentially over time.

From now on we shall assume that the process is absorbed once it hits zero. We think of the O.-U process X_t as a process describing a latent, unobserved development influencing the hazard of a "unit" under study. For instance, X_t may represent the disease development of a particular patient, although this development may not be observable. At the moment the development reaches a certain level, an event occurs for that individual. Thus, the event (survival) time for a unit can be defined as the first time the unobserved process X_t strikes a barrier. For notational simplicity, we assume that the O.-U. process X_t starts in a positive non-random value x_0, i.e., $X_0 = x_0 > 0$, and that the absorbing boundary for the process is zero, and thus define

$$T = \inf_{t \geq 0}\{t : X_t = 0\}$$

to be the event time. In the present section we will discuss the distribution of T, and look at the properties of the corresponding hazard rate, defined as usual for continuous T by

$$\theta(t) = \frac{-d/dtP(T > t)}{P(T > t)}.$$

In particular, it may be shown that when $t \to \infty$, the hazard rate $\theta(t)$ (i.e., the rate of the first passage across the boundary) converges to a constant θ_0. This limiting rate will be derived below.

Let $\psi_t(x)$, $x > 0$, be the probability density for the process X_t at time t. Since X_t is absorbed in zero at time T and we consider only $x > 0$, we see that

$$\psi_t(x) = \frac{1}{dx} P(X_t \in (x, x + dx], T > t) = \frac{1}{dx} P(X_t \in (x, x + dx]),$$

i.e. ψ_t involves only the non-absorbed part of the population. The distribution ψ_t is thus "defect" in the sense that

$$\int_0^\infty \psi_t(x)dx = \int_0^\infty P(X_t \in (x, x + dx], T > t) = P(T > t) < 1.$$

It is clear that $\lim_{t\to\infty} \int_0^\infty \psi_t(x)dx = P(T = \infty) = 0$ since absorption is certain as long as the process is mean-reverting. However, if we instead of ψ_t consider the distribution *conditional* on survival, i.e., $\phi_t(x) = P(X_t \in (x, x + dx]|T > t)/dx$, then ϕ_t is a genuine probability density on $(0, \infty)$, with $\phi_t(x) = \psi_t(x)/P(T > t)$ and $\int_0^\infty \phi_t(x)dx = 1$ for all $t \geq 0$. There is then the following general connection between the hazard rate (rate of absorption) and the distribution of the survivors

$$\theta(t) = \frac{\sigma^2}{2} \phi_t'(0).$$

See, for instance, Aalen and Gjessing (2001).

The above relation between the distribution of survivors and the instantaneous rate of absorption indicates that convergence of $\theta(t)$ as $t \to \infty$ relates to the convergence of ϕ_t as a distribution. If $\lim_{t\to\infty} \phi_t(x) = \phi(x)$ exists for all $x > 0$ this is referred to as the quasi-stationary distribution. The existence of a quasi-stationary distribution is in effect a stabilization of the risk profile over time for the remaining part of the population. This phenomenon may manifest itself under a variety of conditions. However, it should be noted that at the outset the concept of a quasi-stationary distribution is a theoretical construct. In practice, for a population of limited size the importance of quasi-stationarity depends entirely on how fast the stationarity sets in. Sometimes it may occur almost immediately, whereas under other circumstances it may not set in until most of the observed population has died out. In the latter case quasi-stationarity is of little practical importance since it will not be observed in the population under study. For an illustration of this, (see Gjessing et al., 2003).

Somewhat surprisingly, the quasi-stationary distribution will in general not be unique. In fact, the form of the quasi-stationary distribution will often depend on the starting distribution of the underlying process, i.e., on the probability distribution of X_0. So by starting X in different distributions one may obtain a whole spectrum of different quasi-stationary distributions in the limit. However, it is typically the case that starting the process in a deterministic value $x_0 = X_0 > 0$ leads to one quasi-stationary distribution independent of x_0. This distribution is known as the *canonical*

distribution, and will sometimes be denoted $\phi_0(x)$ to emphasize the difference from a general quasi-stationary distribution $\phi(x)$.

2.1. Laplace Transform of the First-passage Time

We will now discuss how the Laplace transform of the event time T can be obtained, and from it the limiting hazard θ_0. Assume the process starts in a positive value $x_0 = X_0 > 0$, and T is defined as above. Let

$$L(s; x_0, a, b, \sigma) = \mathrm{E}(e^{-sT})$$

be the Laplace transform of T. There exists a differential equation for this Laplace transform for quite general stochastic processes, including diffusion processes (see e.g., Paulsen and Gjessing, 1997). In our case, it can be shown that the Laplace transform is given by the formula

$$L(s; x_0, a, b, \sigma) = \frac{H_{-s/b}(\frac{-a+bx_0}{\sigma\sqrt{b}})}{H_{-s/b}(-\frac{a}{\sigma\sqrt{b}})}, \tag{2}$$

where $H_\nu(z)$ is the Hermite function (Abramowitz and Stegun, 1964). This result may be found in Ricciardo and Sato (1988, formula (1a) and (b)). In section 6 of their paper it is explained how the Laplace transform may be given as a spectral decomposition, where the countable number of eigenvalues correspond to the zeros of the denominator in Eq. (2). Each term in this spectral decomposition may be inverted to yield an exponential function, and so the probability density of T can be written as a weighted sum of exponential functions of type $\exp(\rho t)$ with ρ denoting the (negative) eigenvalues. In such a decomposition there is a dominant eigenvalue ρ_0, i.e., the one which is closest to zero, and when t increases then clearly the corresponding $\exp(\rho_0 t)$ will dominate. Hence, asymptotically, the distribution of the first-passage time T will be exponential with limiting hazard rate $\theta_0 = -\rho_0$, According to the theory of Ricciardo and Sato (1988), θ_0 is then the absolute value of the largest negative zero of $H_{-s/b}(-a/(\sigma\sqrt{b}))$, regarded as a function of s.

There exists an infinite family of quasi-stationary distributions for the Ornstein-Uhlenbeck process. One of these is termed the canonical one because it is the limiting distribution conditional on non-absorption, when starting out from a fixed state. The density of the canonical quasi-stationary distribution is the (suitably normalized) eigenfunction of the spectral decomposition corresponding to the dominant eigenvalue ρ_0. Below we shall derive it from a differential equation which also yields a more general result.

2.2. The Shape of the Hazard Rate

Exact formulas exist in the symmetric case ($a = 0$). In Ricciardo and Sato (1988, p. 46) we find the following probability density of time to absorption when starting in x (parameter values $a = 0$, $b = 1$, $\sigma^2 = 2$)

$$g(t) = \sqrt{\frac{2}{\pi}} x_0 \frac{e^{2t}}{(e^{2t}-1)^{3/2}} \exp\left(-\frac{x_0^2}{2(e^{2t}-1)}\right). \tag{3}$$

The corresponding survival function is

$$S(t) = 2\Phi\left(\frac{x_0}{\sqrt{e^{2t}-1}}\right) - 1, \tag{4}$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution function. The hazard rate is computed as $\theta(t) = g(t)/S(t)$.

The hazard rate starts from 0 and then converges towards a fixed level, which in this case equals 1. The convergence signals the approach towards quasi-stationarity of the underlying distribution on the state space, (see also Aalen and Gjessing, 2001). The shape of the hazard rate depends on whether the starting point x_0 is close to or further away from the point of absorption 0, where the distance is compared to the quasi-stationary distribution. Some typical hazard rates are shown in Figure 1. The upper hazard rate also starts out in 0, but apart from a strong initial increase it is generally decreasing. The picture that emerges is the following: When the initial state, x_0, is far from 0, the hazard rate is generally increasing. As the initial state gets closer to zero a unimodal hazard appears, while eventually, for small x_0, a mainly decreasing hazard occurs. (Actually, the hazards in this model will always be unimodal in principle, but here we discuss the dominating parts of their shapes.) Hence, essentially, the major hazard shapes are determined by the distance from the absorbing point (Aalen and Gjessing, 2001).

In general, it is not easy to invert the Laplace transform to deduce the density and survival functions. A closed-form symbolic inversion is hardly possible in general. In principle, the spectral decomposition referred to above can be used. However, one must compute a large number of eigenvalues to get a good

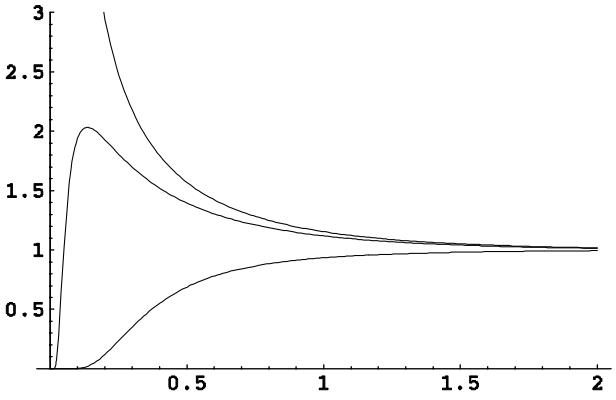

Figure 1. Hazard rates of time to absorption when initial value x_0 equals 0.2 -upper curve, 0.8 - middle curve, and 2.0 - lower curve (parameter values: $a = 0, b = 1, \sigma^2 = 2$).

approximation. On the other hand, one can use numerical inversion techniques for Laplace transforms.

2.3. *The Quasi-stationary Distribution*

Let us for a moment consider a general diffusion process X_t solving the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{5}$$

Typically for diffusion processes one has the following situation: For any θ in $(0, \theta_0]$ there exists a quasi-stationary distribution yielding θ as the hazard rate of absorption. For the O.-U. process this has been proved in the symmetric case, that is $a = 0$ in our setting (Lladser and San Martin, 2000). We will here assume it to be true also in the non-symmetric case. General results on quasi-stationarity for diffusion processes may be found in Mandl (1961) and Steinsaltz and Evans (2003).

Let ψ_t, ϕ_t and ϕ be defined as above, and let $A^\star$ be the (formal) adjoint infinitesimal operator of X, defined as the second-order differential operator

$$(A^\star f)(x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x)f(x)\right) - \frac{\partial}{\partial x}(\mu(x)f(x)).$$

The function ψ_t will solve the Kolmogorov forward equation

$$\frac{\partial}{\partial t}\psi_t(x) = (A^\star\psi_t)(x).$$

(Øksendal, 1998). Assuming a constant limiting rate θ, we can intuitively write $\psi_t(x) \sim \exp(-\theta t)\phi(x)$ when $t \to \infty$, where ϕ is a quasi-stationary distribution, with $\int_0^\infty \phi(x)dx = 1$. Entering this in the forward equation, it is seen that ϕ solves the following eigenvalue problem (see e.g., Aalen and Gjessing, 2001)

$$-\theta\phi(x) = (A^\star\phi)(x).$$

In particular, specifying $\sigma^2(x) = \sigma^2$ and $\mu(x) = a - bx$ for the O.-U process, gives

$$\frac{1}{2}\sigma^2\phi''(x) - (a - bx)\phi'(x) + b\phi(x) = -\theta\phi(x). \tag{6}$$

Convenient boundary conditions are

$$\phi(0) = 0, \qquad \phi'(0) = \frac{2\theta}{\sigma^2}.$$

The second condition follows from Aalen and Gjessing (2001, formula (4)).

Using Mathematica (Wolfram, 1999) the following solution is found:

$$\phi(x) = \frac{\sqrt{b}}{\sigma}e^{\frac{x(2a-bx)}{\sigma^2}} \times \frac{H_{\theta/b}\left(\frac{-a+bx}{\sigma\sqrt{b}}\right){}_1F_1\left(\frac{-\theta}{2b},\frac{1}{2},\frac{a^2}{b\sigma^2}\right) - H_{\theta/b}\left(-\frac{a}{\sigma\sqrt{b}}\right){}_1F_1\left(\frac{-\theta}{2b},\frac{1}{2},\frac{(a-bx)^2}{b\sigma^2}\right)}{H_{(\theta/b)-1}\left(-\frac{a}{\sigma\sqrt{b}}\right){}_1F_1\left(\frac{-\theta}{2b},\frac{1}{2},\frac{a^2}{b\sigma^2}\right) - \frac{a}{\sigma\sqrt{b}}H_{\theta/b}\left(-\frac{a}{\sigma\sqrt{b}}\right){}_1F_1\left(1-\frac{\theta}{2b},\frac{3}{2},\frac{a^2}{b\sigma^2}\right)}, \tag{7}$$

where ${}_1F_1$ denotes the confluent hypergeometric function (Abramowitz and Stegun, 1964; Lebedev, 1972).

Note that this is a valid distribution on $(0, \theta_0]$. In the canonical case, with $\theta = \theta_0$, it follows from the above discussion that $H_{\theta_0/b}(-a/(\sigma\sqrt{b})) = 0$ and hence the canonical stationary distribution is given by

$$\phi_0(x) = \frac{\sqrt{b}}{\sigma} e^{\frac{x(2a-bx)}{\sigma^2}} \frac{H_{\theta_0/b}\left(\frac{-a+bx}{\sigma\sqrt{b}}\right)}{H_{(\theta_0/b)-1}\left(-\frac{a}{\sigma\sqrt{b}}\right)}. \tag{8}$$

Note that the formula for θ_0 can easily be solved numerically, e.g. in Mathematica (Wolfram, 1999).

The canonical distribution in (8) is a simple modification of a normal distribution to handle the limitation to the positive real axis. From the formula found on the web page in reference Wolfram (2004a) it follows that the Hermite function part increases asymptotically as $x^{\theta_0/b}$, and therefore the tail of the distribution is not much more heavy than that of a normal distribution. For the general distribution in (3) the tail is dominated by the hypergeometric function. From the formula found in reference Wolfram (2004b) it follows that the tail of $\phi(x)$ is proportional to $x^{-1-\theta/b}$. Hence, the distribution is extremely heavy-tailed, which is reasonable since probability mass has to hide away, so to speak, in the upper reaches of the state space to yield a low rate of absorption.

Example 1 A special case. Consider the special case $a = b = \sigma^2 = 1$. Then θ_0 is the smallest positive zero of $H_s(-1)$. Using the program Mathematica (Wolfram, 1999) one finds $\theta_0 \approx 0.23423$. The canonical quasi-stationary distribution has the form

$$\phi_0(x) = e^{x(2-x)} \frac{H_{\theta_0}(x-1)}{H_{\theta_0-1}(-1)}.$$

This fits well with a normal distribution with expectation 1 (which is what you would have if there was no absorption), but which is skewed due to the Hermite function. The distribution is shown in Figure 2. The general quasi-stationary distribution for θ in $(0, \theta_0]$ is given by

$$\phi(x) = e^{x(2-x)} \frac{H_\theta(x-1)\,{}_1F_1(-\theta/2, 1/2, 1) - H_\theta(-1)\,{}_1F_1(-\theta/2, 1/2, (1-x)^2)}{H_{\theta-1}(-1)\,{}_1F_1(-\theta/2, 1/2, 1) - H_\theta(-1)\,{}_1F_1(1-\theta/2, 3/2, 1)}. \tag{9}$$

As mentioned, the quasi-stationary distributions are, in general, strongly skewed compared with the canonical one. An example, showing the heavy tail, is given in Figure 3.

Example 2 The symmetric case. Consider the situation $a = 0$. Then θ_0 is the smallest positive zero of $H_{s/b}(0)$. This zero occurs for $s/b = 1$, hence $\theta_0 = b$. A notable feature is that the limiting hazard rate is independent of the diffusion coefficient σ^2. From Eq. (8) one gets the canonical quasi-stationary distribution

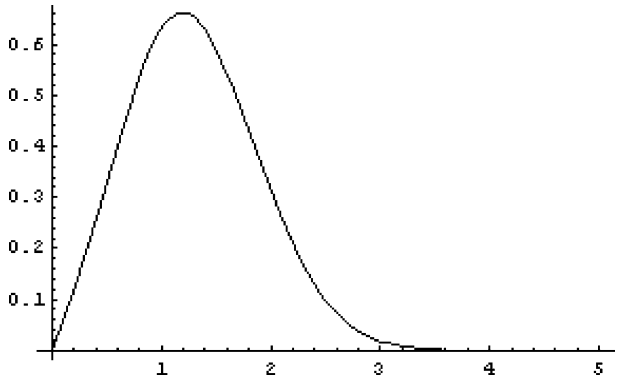

Figure 2. The canonical quasi-stationary distribution $\phi_0(x)$ for the Ornstein-Uhlenbeck process absorbed in zero, for the special case $a = b = \sigma^2 = 1$.

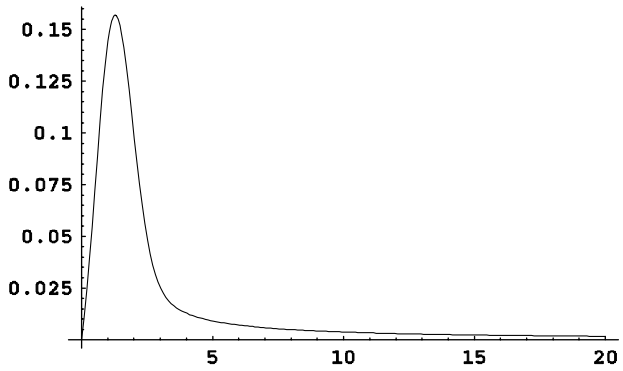

Figure 3. One of the non-canonical quasi-stationary distributions $\phi(x)$ for the O. -U process absorbed in zero, for the special case $a = b = \sigma^2 = 1$. The limiting hazard rate is $\theta = 0.05$.

$$\phi_0(x) = \frac{\sqrt{b}}{\sigma} e^{\frac{-bx^2}{\sigma^2}} H_1\left(\frac{bx}{\sigma\sqrt{b}}\right) = 2\frac{b}{\sigma^2} x e^{\frac{-bx^2}{\sigma^2}}, \quad x \geq 0. \tag{10}$$

3. Modelling the Hazard Rate as the Square of an Ornstein-Uhlenbeck Process

This far we have studied the hazard rate of the time to barrier crossing for an O.-U process. We shall now use the same process in a different setting, namely when modelling the hazard rate directly. This has been proposed many years ago (Woodbury and Manton, 1977; Myers, 1981; Yashin, 1985), but has not received much attention in the biostatistical literature. We will also show that there is a close relationship between the two approaches.

3.1. *General Diffusions*

Let again X_t be a general diffusion process defined by (5). Let us assume that the individual hazard at time t is $k(X_t)$, where k is a non-negative function, and let T be the corresponding event time. Thus, conditional on X the survival function of T is

$$P(T > t|X) = \exp\left(-\int_0^t k(X_s)\ ds\right),$$

and consequently the population survival is

$$S(t) = P(T > t) = E\exp\left(-\int_0^t k(X_s)\ ds\right).$$

It can be shown in general that the population hazard can be computed as

$$\theta(t) = E[k(X_t)|T > t]$$

i.e., the expected hazard of the survivors at time t (Yashin and Manton, 1997).

This is completely equivalent to letting T be a *killing time* for the diffusion (Karlin and Taylor 1981; Øksendal 1998), i.e., that the process X enters a state Δ at time T, where Δ is an arbitrary "coffin state" not in the original state space. The adjoint infinitesimal operator for the killed process is $A^\star - k$. Thus, the Kolmogorov forward equation becomes

$$\frac{\partial}{\partial t}\psi_t(x) = (A^\star\psi_t)(x) - k(x)\psi_t(x).$$

Assuming the existence of a quasi-stationary distribution ϕ we again write $\psi_t(x) \sim \exp(-\theta t)\phi(x)$ and insert it in the forward equation, obtaining the second order eigenvalue problem for ϕ (suppressing the dependence on x)

$$\frac{1}{2}\sigma^2\phi'' + (2\sigma\sigma' - \mu)\phi' + \left(\frac{1}{2}(\sigma^2)'' - \mu' - k\right)\phi = -\theta\phi. \tag{11}$$

The boundary conditions are $\phi \geq 0$ and $\int_{-\infty}^{\infty} \phi = 1$.

It is interesting to note that killing and barrier hitting models are closely related, and that barrier hitting provides a concrete representation of the event time T. Let R be an exponentially distributed random variable with $ER = 1$, and assume R is independent of X. Define

$$T = \inf\left\{t : \int_0^t k(X_s)ds = R\right\}, \tag{12}$$

i.e., the time when the cumulative hazard hits the random barrier R. Conditional on X,

$$P(T > t|X) = P(R > \int_0^t k(X_s)ds|X) = e^{-\int_0^t k(X_s)ds}.$$

Thus, X is killed when cumulative hazard strikes the (randomized) barrier R, and (12) provides a representation of the event time. For more information about killing

of diffusion processes, see Karlin and Taylor (1981), and Singpurwalla (1995) in the survival setting.

3.2. *The Hazard Rate as the Square of the Ornstein-Uhlenbeck Process*

Let then X_t be the Ornstein-Uhlenbeck process defined by (1). It is mathematically convenient to assume a model whereby the individual hazard rate (killing rate) is given as $Z_t = k(X_t) = X_t^2$. An interesting feature of this model is that the process X, which is Gaussian, is still Gaussian conditional on survival (Yashin, 1985). Consequently, it is enough to find its conditional mean and variance to describe the conditional distribution of X. Using this, we shall show that also this model yields quasi-stationarity for X with a Gaussian distribution as its limit. Usually, the O.-U process is restricted to $b \geq 0$, but in the present case we will use any value of b, positive or negative.

The distribution of X_t at time 0 is supposed to be a normal distribution with mean $m(0)$ and variance $\gamma(0)$. The distributions of survivors is normally distributed at any time t, with mean $m(t) = E[X_t|T > t]$ and variance $\gamma(t) = Var[X_t|T > t]$ that solve the Ricatti equations

$$m'(t) = a - bm(t) - 2m(t)\gamma(t), \tag{13}$$

$$\gamma'(t) = \sigma^2 - 2b\gamma(t) - 2\gamma^2(t) \tag{14}$$

(see Yashin, 1985). The population hazard rate for this model is then given by

$$\theta(t) = E[Z_t|T > t] = m^2(t) + \gamma(t). \tag{15}$$

To find m and γ, we solve for γ first, noticing that the equation is separable. Let g_1 and g_2 be the solutions to the equation $0 = \sigma^2 - 2bg - 2g^2$, i.e. $g_1 = (-b - B)/2$ and $g_2 = (-b + B)/2$ with $B = \sqrt{b^2 + 2\sigma^2}$, so that $g_1 < 0 < g_2$. The solution is

$$\gamma(t) = -\frac{b}{2} + \frac{B}{2}\tanh(Bt + c),$$

where $c = 1/2 \ \log|(\gamma(0) - g_1)/(\gamma(0) - g_2)|$. Then solving for $m(t)$ we obtain

$$m(t) = \frac{\cosh(c)}{\cosh(Bt + c)}m(0) + \frac{\sinh(Bt + c) - \sinh(c)}{\cosh(Bt + c)}\frac{a}{B}.$$

Related results may be found in Wenocur (1990).

3.3. *Quasi-stationarity*

From the results in Section 3.2 it is clear that when t increases the mean and variance of the normal distribution of survivors converge to a limit. This is seen from the solutions for m and γ, which yield $\lim_{t\to\infty} m(t) = a/B$, and $\lim_{t\to\infty} \gamma(t) = g_2$. Thus, the limiting distribution is normal, $N(a/B, g_2)$. Note that the limit values for m and γ could have been obtained without solving the equations explicitly, by assuming

$\lim_{t\to\infty} m'(t) = 0$ and $\lim_{t\to\infty} \gamma'(t) = 0$ in Eqs. (13) and (14), and solving for the limits. We also note that the limiting value a/B for $m(t)$ is always smaller than the mean-revertion value a/b, and the larger the diffusion σ, the more the quasi-stationary distribution of X is forced towards zero, due to large positive and negative values being "killed off".

Note that from (11) the quasi-stationary distribution should solve

$$\frac{1}{2}\sigma^2\phi''(x) - (a - bx)\phi'(x) + (b - x^2)\phi(x) = -\theta\phi(x).$$

It is not hard to verify that the normal distribution $N(a/B, g_2)$ is a solution when $\theta = (a/B)^2 + g_2$.

Example 3. As an illustration, Figure 4 shows the curves for m and γ when $a = 1$, $b = 1$ and $\sigma = 1$, with starting values $m(0) = a/b = 1$ and $\gamma(0) = 0$, i.e., the process starts in the point of mean revertion, a/b. The hazard $\theta(t) = m^2(t) + \gamma(t)$ is also included.

3.4. *The Cox-Ingersoll-Ross Model and Zero-coupon Bond Prices*

In the symmetric case, i.e., when $a = 0$, the process $Z_t = X_t^2$ in the above discussion has been extensively studied in its own right. By a simple application of Itô's formula we have

$$dZ_t = (\sigma^2 - 2bZ_t)dt + 2\sigma\sqrt{Z_t}\, dW_t.$$

The solution to this equation is known as the *Cox–Ingersoll–Ross* (CIR) process. From the above discussion it is clear that the CIR process has a non-central chi-

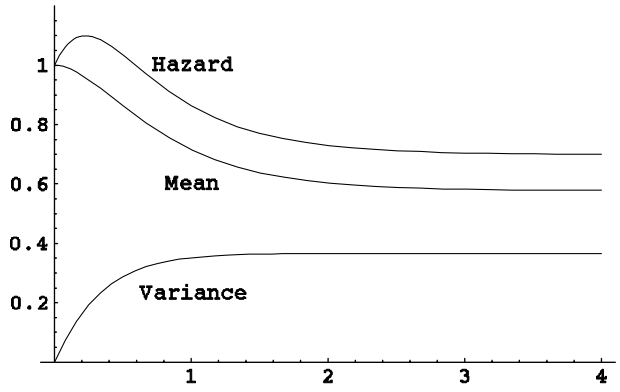

Figure 4. The mean function $m(t) = E[X_t|T > t]$ and variance function $\gamma(t) = Var[X_t|T > t]$ for the O.- U process X_t conditional on survival, and the population hazard function $\theta(t) = m^2(t) + \gamma(t)$. The individual hazard is X_t^2, the process starts in $X_0 = 1$, and $a = b = \sigma = 1$.

squared distribution at time t. Killing the O.-U. process with a quadratic rate is equivalent to killing the CIR process with a linear rate.

Because of its non-negativity and long-run stability the CIR process is frequently used as a model for short-term interest rates in financial modelling. It should be noted that computing the price of a zero-coupon bond for this short-rate model is equivalent to finding the survival function in our setting. For more details, (see Baxter and Rennie, 1997; Musiela and Rutkowski, 1997, Section 12.3).

4. Time-varying Coefficients

A rather more flexible model will be obtained when the coefficients in the equation for X may vary as functions of time

$$dX_t = (a_t - b_t X_t)dt + dW_t. \tag{16}$$

Furthermore, we can introduce a basic hazard rate λ_t, and define the individual hazard rate to be $Z_t = \lambda_t X_t^2$. Notice that we have set $\sigma = 1$ since the diffusion coefficient becomes redundant when including λ. As before, the distribution of X_0 is supposed to be a normal distribution with mean $m(0)$ and variance $\gamma(0)$. Following Yashin (1985), the conditional distribution of X given survival is still Gaussian, with mean $m(t) = E[X_t|T > t]$ and variance $\gamma(t) = Var[X_t|T > t]$ that solve

$$m'(t) = a_t - b_t m(t) - 2\lambda_t m(t)\gamma(t),$$

$$\gamma'(t) = 1 - 2b_t\gamma(t) - 2\lambda_t\gamma^2(t)$$

with the initial conditions defined by $m(0)$ and $\gamma(0)$. Typically, a numerical solution must be found. As before, the population hazard rate is

$$\theta(t) = \lambda_t(m^2(t) + \gamma(t)),$$

and the survival function is

$$S(t) = \exp\left(-\int_0^t \lambda_s(m^2(s) + \gamma(s))\ ds\right). \tag{17}$$

5. A Cumulated Hazard Rate Model

We shall now extend the model by assuming that Z_t does not depend only on the value of X at time t, but instead depends on the accumulated effect of X in some sense. To be specific, we assume that X is defined by (16), and

$$Z_t = \int_0^t w(s,t)X_s^2\ ds \tag{18}$$

for some weight function w. We see that cumulated hazard rate for this extended model is

$$\int_0^t Z_s\,ds = \int_0^t \int_0^s w(u,s)X_u^2\,du\,ds$$
$$= \int_0^t X_u^2 \int_u^t w(u,s)\,ds\,du = \int_0^t f_t(u)X_u^2\,du,$$

where $f_t(u) = \int_u^t w(u,s)\,ds$. Let T be the event time for the cumulated model, i.e., T a the random time whose distribution is determined, conditional on X, by

$$P[T > t|X] = \exp\left(-\int_0^t f_t(s)X_s^2\,ds\right)$$

with the population survival function

$$S(t) = E\exp\left(-\int_0^t f_t(s)X_s^2\,ds\right).$$

In the simpler model described in Section 4 the cumulated hazard rate is $\int_0^t \lambda_s X_s^2\,ds$, so the extension consists of replacing λ_s by $f_t(s)$. The occurrence of the variable t here is a complication, but otherwise the situation is very similar for the two models.

To help analyzing the cumulated model we define, for fixed τ, the random times T_τ by

$$P[T_\tau > t|X] = \exp\left(-\int_0^t f_\tau(s)X_s^2\,ds\right).$$

Thus, for fixed τ, T_τ is just the event time studied in Section 4, with survival function S_τ as computed in (16). Since

$$S(\tau) = P(T > \tau) = P(T_\tau > \tau) = S_\tau(\tau), \tag{19}$$

the results in Section 4 can be used directly to find the survival function in the cumulated model, and we have proved

PROPOSITION 4 *The survival function $S(\tau)$ of the cumulated model (18) is given by*

$$S(\tau) = \exp\left(-\int_0^\tau f_\tau(s)(m_\tau^2(s) + \gamma_\tau(s))\,ds\right), \tag{20}$$

where $m_\tau(s) = E[X_s|T_\tau > s]$ and $\gamma_\tau = Var[X_s|T_\tau > s]$ are solutions to the differential equations

$$m_\tau'(t) = a_t - b_t m_\tau(t) - 2f_\tau(t)m_\tau(t)\gamma_\tau(t), \tag{21}$$

$$\gamma_\tau'(t) = 1 - 2b_t\gamma_\tau(t) - 2f_\tau(t)\gamma_\tau^2(t). \tag{22}$$

Initial values are as before $m_\tau(0) = m(0)$ and $\gamma_\tau(0) = \gamma(0)$.

Note that the formula (19) requires solving the Eqs. (21) and (22) for all relevant τ's, yielding $m_\tau(s)$ and $\gamma_\tau(s)$ for all combinations of s and τ, $s \le \tau$.

In principle, the population hazard can then be obtained from S by numeric differentiation.

In addition to the survival function, it would be interesting to look at the behavior (distribution) of the survivors. Interestingly, the process X once again has a Gaussian-distribution conditional on survival.

PROPOSITION 5 *Let* $M \in \sigma(X_s, s \geq 0)$. *Then* $P(M|T > \tau) = P(M|T_\tau > \tau)$. *In particular, X is Gaussian conditional on* $T > t$. *Also, the conditional mean and variance are* $E[X_\tau|T > \tau] = E[X_\tau|T_\tau > \tau] = m_\tau(\tau)$ *and* $Var[X_\tau|T > \tau] = Var[X_\tau|T_\tau > \tau] = \gamma_\tau(\tau)$, *where* m_τ *and* γ_τ *solve* (21) *and* (22).

Proof.

$$\begin{aligned} P(M|T > \tau) &= \frac{P(T > \tau|M)P(M)}{P(T > \tau)} \\ &= \frac{E[P(T > \tau|X)|M]P(M)}{P(T > \tau)} \\ &= \frac{E[P(T_\tau > \tau|X)|M]P(M)}{P(T_\tau > \tau)} \\ &= P(M|T_\tau > \tau). \end{aligned}$$

As mentioned in Section 4, Yashin (1985) has proved that X is Gaussian conditional on $T_\tau > \tau$, and the result follows. □

Finally, we will show how to determine the population hazard using the conditional mean and variances rather than differentiating the survival function.

PROPOSITION 6 *The population hazard* $\theta(\tau)$ *in the cumulated model (18) can be computed from*

$$\theta(\tau) = \int_0^\tau w(s, \tau)(\tilde{m}_\tau^2(s) + \tilde{\gamma}_\tau(s))\, ds,$$

where $\tilde{m}_\tau(s) = E[X_s|T > \tau]$ *and* $\tilde{\gamma}_\tau(s) = Var[X_s|T > \tau]$ *solve, for fixed* τ, *the equations*

$$\tilde{m}'_\tau(s) = a_s - b_s\tilde{m}_\tau(s) + (\tilde{m}_\tau(s) - m_\tau(s))/\gamma_\tau(s), \tag{23}$$

$$\tilde{\gamma}'_\tau(s) = -1 + 2(1/\gamma_\tau(s) - b_s)\tilde{\gamma}_\tau(s). \tag{24}$$

Terminal conditions are $\tilde{m}_\tau(\tau) = m_\tau(\tau)$ and $\tilde{\gamma}_\tau(\tau) = \gamma_\tau(\tau)$. *Note that* $m_\tau(s)$ *and* $\gamma_\tau(s)$ *have been found as solutions to Eqs. (21) and (22).*

Proof. The population hazard is

$$\begin{aligned} \theta(\tau) &= E[Z_\tau|T > \tau] = E\left[\int_0^\tau w(s, \tau)X_s^2\, ds|T_\tau > \tau\right] \\ &= \int_0^\tau w(s, \tau)E[X_s^2|T_\tau > \tau]\, ds \\ &= \int_0^\tau w(s, \tau)(\tilde{m}_\tau^2(s) + \tilde{\gamma}_\tau(s))\, ds. \end{aligned}$$

The added complexity here is that one needs the *retrospective* moments of X_s at time τ, $s \leq \tau$. But according to Proposition 7, Yashin and Manton (1997), these moments satisfy Eqs. (23) and (24). □

References

O. O. Aalen and H. K. Gjessing, "Understanding the shape of the hazard rate," *Statistical Science,* vol. 16 pp. 1–22, 2001.

O. O. Aalen and H. K. Gjessing, "A look behind survival data: Underlying processes and quasi-stationarity," in *Mathematical and Statistical Methods in Reliability*, (B. H. Lindqvist and K. A. Doksum, eds.) World Scientific Publishing: Singapore, pp. 221–34, 2003.

M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications: New York, 1964.

M. Baxter and A. Rennie, *Financial Calculus. An Introduction to Derivative pricing*, Cambridge University Press: New York, 1997.

J. C. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica* vol. 53 pp. 385–407, 1985.

H. K. Gjessing, O. O. Aalen, and N. L. Hjort, "Frailty models based on Lévy processes," *Advances in Applied Probability,* vol. 35 pp. 532–550, 2003.

S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes,* Academic Press: New York, 1981.

J. Keilson, "A limit theorem for passage times in ergodic regenerative processes," *Annals of Mathematical Statistics* vol. 37 pp. 866–870, 1966.

M.-L. T. Lee, V. DeGruttola, and D. Schoenfeld, "A model for markers and latent health status," *Journal of the Royal Statistical Society B* vol. 62 pp. 747–762, 2000.

N. N. Lebedev, *Special Functions and Their Applications,* Dover Publications: New York, 1972.

M. Lladser and J. San Martin, "Domain of attraction of the quasi-stationary distributions for the Ornstein-Uhlenbeck process," *Journal of Applied Probability*, vol. 37 pp. 511–520, 2000.

K. G. Manton and E. Stallard, *Chronic Disease Modelling*, Charles Griffin & Company Limited: London, 1988.

M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling* . Springer-Verlag: New York, 1997.

L. E. Myers, "Survival functions induced by stochastic covariate processes," *Journal of Applied Probability* vol. 18 pp. 523–529, 1981.

B. Øksendal, *Stochastic differential equations,* Springer-Verlag: Berlin, 1998.

J. Paulsen and H. K. Gjessing, "Ruin theory with stochastic return on investments," *Advances in Applied Probability*, vol. 29 pp. 965–985, 1997.

L. M. Ricciardi and S. Sato, "First-passage-time density and moments for the Ornstein-Uhlenbeck process," *Journal of Applied Probability* vol. 25 pp. 43–57, 1988.

G. Samorodnitsky, *Long Range Dependence, Heavy Tails and Rare Events,* Lecture Notes Maphysto, Centre for Mathematical Physics and Stochastics: Aarhus, 2002.

N. D. Singpurwalla, "Survival in dynamic environments," *Statistical Science* vol. 10 pp. 86–103, 1995.

D. Steinsaltz and S. N. Evans, "Markov mortality models: Implications of quasistationarity and varying initial distributions," *Theoretical Population Biology,* vol. 65 pp. 319–337, 2004.

D. Steinsaltz and S. N. Evans, "Quasistationary distributions for one-dimensional diffusions with killing," Technical report, Department of Statistics, University of California: Berkeley, 2003.

M. L. Wenocur, "Ornstein-Uhlenbeck processes with quadratic killing," *Journal of Applied Probability*, vol. 27 pp. 707–712, 1990.

S. Wolfram, *The Mathematica book*, 4th edition, Wolfram Media/ Cambridge University Press, 1999.

S. Wolfram (2004a). http://functions.wolfram.com/07.01.06.0009.01

S. Wolfram (2004b). http://functions.wolfram.com/07.20.06.0009.01

M. A. Woodbury and K. G. Manton, "A random walk model of human mortality and aging," *Theoretical Population Biology* vol. 11 pp. 37–48, 1977.

A. I. Yashin, "Dynamics in survival analysis: conditional Gaussian property versus Cameron-Martin formula," in: *Statistics and control of stochastic processes* (N. V. Krylov, R. S. Liptser, and A. A. Novikov, eds.) Springer: New York, pp. 466–475, 1985.

A. I. Yashin and K. G. Manton, "Effects of unobserved and partially observed covariate processes on system failure: A review of models and estimation strategies," *Statistical Science* vol. 12 pp. 20–34, 1997.

A. I. Yashin and J. W. Vaupel, "Measurement and estimation in heterogeneous populations," in Lecture Notes in *Biomathematics 65: Immunology and Epidemiology* (Hoffman GW and Hraba T, eds.) pp. 198–206, 1986.