Ann. Inst. Statist. Math.
Vol. 48, No. 1, 61–74 (1996)

CURVED EXPONENTIAL FAMILIES OF STOCHASTIC PROCESSES AND THEIR ENVELOPE FAMILIES

UWE KÜCHLER[1] AND MICHAEL SØRENSEN[2]

[1] *Institut für Stochastik, Fachbereich Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin, Germany*
[2] *Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus, 8000 Aarhus C, Denmark*

(Received January 17, 1994; revised February 27, 1995)

Abstract. Exponential families of stochastic processes are usually curved. The full exponential families generated by the finite sample exponential families are called the envelope families to emphasize that their interpretation as stochastic process models is not straightforward. A general result on how to calculate the envelope families is given, and the interpretation of these families as stochastic process models is considered. For Markov processes rather explicit answers are given. Three examples are considered some in detail: Gaussian autoregressions, the pure birth process and the Ornstein-Uhlenbeck process. Finally, a goodness-of-fit test for censored data is discussed.

Key words and phrases: Censored data, diffusion processes, Gaussian autoregression, goodness-of-fit test, Markov processes, Ornstein-Uhlenbeck process, pure birth process.

1. Introduction

Many important statistical stochastic process models are exponential families in the sense that the likelihood function corresponding to observation of the process in the time interval $[0, t]$ has an exponential family representation of the same dimension for all $t > 0$. The exponential structure of the likelihood function implies several probabilistic properties of the processes in the family and statistical results for the model, see Küchler and Sørensen (1989, 1994*a*, 1994*b*) and Sørensen (1986). Thus the study of exponential families of stochastic processes casts light on basic problems of statistical inference for stochastic processes and reveals important structure of many particular types of statistical models for stochastic processes.

Most exponential families of stochastic processes are curved exponential families in the sense that the canonical parameter space is a curved submanifold of a Euclidean space. It is therefore important to develop statistical theory for general curved exponential families of processes. Some steps in this direction are taken in the present paper.

Several modern statistical techniques for curved exponential families use properties of the full exponential family generated by the curved model. Examples are methods based on differential geometric considerations or on approximately ancillary statistics. It is therefore, from a statistical point of view, of interest to study the full exponential families generated by an exponential family of stochastic processes and to investigate their interpretation as stochastic process models. This is the main purpose of the present paper. To emphasize the fact that a stochastic process interpretation of the full families is not straightforward, we propose to call these, in the stochastic process setting, envelope families. From a probabilistic point of view it is interesting that this statistical investigation provides a new way of deriving other stochastic processes from a given class of processes.

Some basic definitions are introduced in Section 2. In Section 3 we study the envelope families corresponding to a curved exponential family of stochastic processes. A general result on how to calculate the envelope families is given. Particular attention is given to the question in what sense the envelope families can be interpreted as stochastic process models. For Markov processes rather explicit answers can be given. In Section 4 three examples are considered in detail: Gaussian autoregressions, the pure birth processes and the Ornstein-Uhlenbeck processes. In Section 5 a goodness-of-fit test for censored data is studied using the techniques introduced in this paper.

2. Basic definitions

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$ be a filtered space where the filtration $\{\mathcal{F}_t : t \geq 0\}$ is supposed right-continuous and where $\mathcal{F} = \sigma(\mathcal{F}_t : t \geq 0)$. Consider a class $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^k$ of probability measures on $(\Omega, \mathcal{F})$. We will denote by μ^t the restriction of a measure μ to the σ-algebra $\mathcal{F}_t$.

The class $\mathcal{P}$ is called an *exponential family* on the filtered space if there exists a measure μ on $(\Omega, \mathcal{F})$ such that $P_\theta^t \ll \mu^t$, $t \geq 0$, $\theta \in \Theta$ and such that we have an exponential representation

$$\frac{dP_\theta^t}{d\mu^t} = \exp(\gamma_t(\theta)^T B_t - \phi_t(\theta)), \qquad \theta \in \Theta, \ t \geq 0, \tag{2.1}$$

where T denotes transposition. For fixed t this Radon-Nikodym derivative is the likelihood function corresponding to observation of events in $\mathcal{F}_t$. In (2.1) ϕ and $\gamma^{(i)}$, $i = 1, \ldots, m$, are non-random real functions of θ and t. The m-dimensional stochastic process B_t is adapted to $\{\mathcal{F}_t\}$ and is called a *canonical process*. Without loss of generality we can assume that $0 \in \Theta$ and that the dominating measure is P_0. If an exponential representation exists with γ independent of t, we call the exponential family *time-homogeneous*.

Time-homogeneous exponential families of stochastic processes can be parametrized by the set $\Gamma = \{\gamma(\theta) : \theta \in \Theta\}$. This parametrization is called a *canonical parametrization*. Typically the set Γ is a curved (i.e. non-affine) submanifold of $\mathbb{R}^m$, in which case we talk about a curved exponential family. In fact, for a minimal time-homogeneous representation, $\operatorname{int}\Gamma \neq \emptyset$ implies that the canonical process B has independent increments, see Küchler and Sørensen (1994*b*), so such models are essentially similar to repeated sampling from a classical exponential family of distributions. For many curved exponential families it is possible to find a representation of the form

$$\frac{dP_\theta^t}{dP_0^t} = \exp(\theta^T A_t - \alpha_t(\theta)^T S_t - \phi_t(\theta)), \qquad t \geq 0, \tag{2.2}$$

where $\alpha_t(\theta)$ is a $(m-k)$-dimensional vector with $\alpha_t(0) = 0$. Moreover, A_t and S_t are vectors of $\{\mathcal{F}_t\}$-adapted processes of dimension k and $(m-k)$, respectively. The natural exponential family generated by a semimartingale has a representation of this form; see Küchler and Sørensen (1994*a*, 1994*b*) and Sørensen (1993).

3. Envelope families

3.1 *Definition and interpretation*

Classical curved exponential families can be embedded in a corresponding full exponential family. In this section we study the problem of similarly extending a curved exponential family on a filtered space. Particular attention is given to stochastic process interpretations of the full family generated by a finite sample exponential family.

Consider a time-homogeneous exponential family $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ on a filtered space with a general representation of the form (2.1). For fixed $t \geq 0$ we can define the full exponential family generated by P_0^t and B_t in the classical way. Specifically, for every $t \geq 0$ we denote the domain of the Laplace transform of B_t under P_0 by $\tilde{\Gamma}_t$. The full exponential family $\mathcal{Q}_t = \{Q_\gamma^{(t)} : \gamma \in \tilde{\Gamma}_t\}$ of probability measures on $\mathcal{F}_t$ is given by

$$\frac{dQ_\gamma^{(t)}}{dP_0^t} = \exp(\gamma^T B_t - \Psi_t(\gamma)), \qquad \gamma \in \tilde{\Gamma}_t, \tag{3.1}$$

with $\Psi_t(\gamma) = \log E_0(\exp(\gamma^T B_t))$. Here $E_\theta(\cdot)$ denotes expectation under P_θ. For fixed γ the class of measures $\{Q_\gamma^{(t)} : t \geq 0\}$ need not be consistent (i.e. need not be in accordance with our observation scheme given by the filtration $\{\mathcal{F}_t\}$), as appears from the following discussion. To emphasize this fact we call the class $\mathcal{Q}_t$ (with t fixed) the *envelope exponential family of* $\mathcal{P}$ *on* $\mathcal{F}_t$.

Obviously, $Q_\gamma^{(t)}$ only exists for all $t > 0$ if γ belongs to the set

$$\tilde{\Gamma} = \bigcap_{t \geq 0} \tilde{\Gamma}_t, \tag{3.2}$$

which is non-empty because $\Gamma \subseteq \tilde{\Gamma}$. If Γ is a curved sub-manifold of $\mathbb{R}^m$, the set $\tilde{\Gamma}$ is necessarily strictly larger than Γ, because $\tilde{\Gamma}$ is a convex set.

Fix $\gamma \in \tilde{\Gamma}$. It is well-known that the class of probability measures $\{Q_\gamma^{(t)} : t \geq 0\}$ is consistent if and only if $\{dQ_\gamma^{(t)}/dP_0^t : t \geq 0\}$ is a P_0-martingale. If this is the case, and if $(\Omega, \mathcal{F})$ is standard measurable, then there exists a probability measure P_γ on $(\Omega, \mathcal{F})$ such that $Q_\gamma^{(t)}$ is the restriction of P_γ to $\mathcal{F}_t$ for all $t \geq 0$; see

Ikeda and Watanabe ((1981), p. 176). This can only be the case for all $\gamma \in \tilde{\Gamma}$ if the canonical process has independent increments, i.e. when we are essentially in an i.i.d. situation. Specifically, let Γ^* denote the set of γ-values in $\tilde{\Gamma}$ for which $dQ_\gamma^{(t)}/dP_0^t$ is a P_0-martingale. Then $\operatorname{int}\Gamma^* \neq \emptyset$ implies that the canonical process has independent increments under P_γ for all $\gamma \in \Gamma^*$. This follows from Theorem 3.1 in Küchler and Sørensen (1994*b*).

Because the measures $\{Q_\gamma^{(t)} : t \geq 0\}$ are typically not consistent, we need the following more complicated approach to obtain a stochastic process interpretation of the envelope family on $\mathcal{F}_t$. For fixed $t > 0$ and $\gamma \in \tilde{\Gamma}_t$ we consider the restriction $Q_\gamma^{(t,s)}$ of $Q_\gamma^{(t)}$ to $\mathcal{F}_s$, $s \leq t$, and note that

$$(3.3) \qquad \frac{dQ_\gamma^{(t,s)}}{dP_0^s} = E_0\left(\frac{dQ_\gamma^{(t)}}{dP_0^t} \;\middle|\; \mathcal{F}_s\right) = \exp(\gamma^T B_s + C_s^{(t)}(\gamma) - \Psi_t(\gamma)),$$

where $C_s^{(t)}(\gamma) = \log E_0[\exp(\gamma^T(B_t - B_s)) \mid \mathcal{F}_s]$. Suppose B is a semimartingale under P_0 and that $\{\mathcal{F}_s\}$ is generated by observing a semimartingale X. Then $\{X_s : s \leq t\}$ is also a semimartingale under $Q_\gamma^{(t)}$, and its local characteristics under $Q_\gamma^{(t)}$ can be determined from (3.3) by Theorem 3.3 in Jacod and Mémin (1976). This gives an interpretation of the envelope family on $\mathcal{F}_t$ as a stochastic process model.

Note incidentally that $Q_\gamma^{(t,s)} = Q_\gamma^{(s)}$ only when $C_s^{(t)}(\gamma) = \Psi_t(\gamma) - \Psi_s(\gamma)$, which happens only when B has independent increments under P_0. In general the class of probability measures $\{Q_\gamma^{(t,s)} : \gamma \in \tilde{\Gamma}_t\}$ is not an exponential family. If $\mathcal{P}$ is a time-homogeneous curved exponential family with representation (2.2), it follows easily that for $\gamma = (\gamma_1, \gamma_2)$ with γ_1 k-dimensional

$$(3.4) \quad C_u^{(t)}(\gamma) = \log(E_{\gamma_1}\{\exp[(\gamma_2 + \alpha(\gamma_1))^T(S_t - S_u)] \mid \mathcal{F}_u\}) + \phi_t(\gamma_1) - \phi_u(\gamma_1).$$

3.2 *Markov processes*

Let us consider the case where $\{\mathcal{F}_t\}$ is generated by observation of a Markov process X with state space E. We will look at the conditional exponential family, where we condition on $X_0 = x$, see Küchler and Sørensen (1991). It is useful to make the initial condition x explicit in the notation, so we replace P_θ by $P_{\theta,x}$, $Q_\gamma^{(t)}$ by $Q_{\gamma,x}^{(t)}$, et cetera. In (3.1) we replace $\Psi_t(\gamma)$ by $\Psi_t(\gamma, x)$. We assume that X is a Markov process under $\{P_{0,x} : x \in E\}$ and that B is a right-continuous additive functional with respect to X and $\{P_{0,x} : x \in E\}$. Then X is a Markov process under $\{P_{\theta,x} : x \in E\}$ for every $\theta \in \Theta$, see Küchler and Sørensen (1991). Under these assumptions

$$(3.5) \qquad C_s^{(t)}(\gamma) = \log E_{\theta,X_s}(\exp\{\gamma^T B_{t-s}\}) = \Psi_{t-s}(\gamma, X_s),$$

so

$$(3.6) \qquad \frac{dQ_{\gamma,x}^{(t,s)}}{dP_{0,x}^s} = \exp(\gamma^T B_s + \Psi_{t-s}(\gamma, X_s) - \Psi_t(\gamma, X_0)).$$

This is only an exponential family when $\Psi_u(\gamma, y) = \sum_i f_u^{(i)}(\gamma) g_u^{(i)}(y)$, $u \leq t$. For exponential families of Markov processes the key to studying the envelope families is the function $\psi_u(\gamma, y)$. Before giving results on how to determine $\psi_t(\gamma)$ for general curved exponential families, we shall first consider two important classes of Markov processes.

Example 3.1. Suppose we observe a diffusion process X which under P_θ solves the stochastic differential equation

$$dX_t = \theta\mu(X_t)dt + \sigma(X_t)dW_t, \qquad X_0 = x, \;\; \theta \in \Theta, \tag{3.7}$$

where W is a Wiener process, $\Theta \subseteq \mathbb{R}$ and $\sigma > 0$. It is well known that, provided

$$P_\theta\left(\int_0^t \mu^2(X_s)\sigma^{-2}(X_s)ds < \infty\right) = 1,$$

this model is an exponential family of stochastic processes with likelihood function

$$\frac{dP_\theta^t}{dP_0^t} = \exp\left(\theta\int_0^t \frac{\mu(X_u)}{\sigma^2(X_u)}dX_u - \frac{1}{2}\theta^2\int_0^t \frac{\mu^2(X_u)}{\sigma^2(X_u)}du\right), \tag{3.8}$$

which is of the form (2.2). The envelope family on $\mathcal{F}_t$ is given by

$$\begin{aligned}\log\frac{dQ_\gamma^{(t,s)}}{dP_0^s} = &\int_0^s \frac{\gamma_1\mu(X_u) + \frac{\partial}{\partial y}\Psi_{t-u}(\gamma, X_u)\sigma^2(X_u)}{\sigma^2(X_u)}dX_u \\ &+ \int_0^s \left\{\frac{\gamma_2\mu^2(X_u)}{\sigma^2(X_u)} + \frac{\partial}{\partial u}\Psi_{t-u}(\gamma, X_u)\right. \\ &\qquad \left. + \frac{1}{2}\frac{\partial^2}{\partial y^2}\Psi_{t-u}(\gamma, X_u)\sigma^2(X_u)\right\}du,\end{aligned} \tag{3.9}$$

for $\gamma = (\gamma_1, \gamma_2) \in \tilde{\Gamma}_t$, and under $Q_\gamma^{(t)}$, the process X solves the equation

$$dX_s = d_t(\gamma, X_s, s)ds + \sigma(X_s)dW_s, \qquad X_0 = x, \;\; s \leq t, \tag{3.10}$$

where

$$d_t(\gamma, y, s) = \gamma_1\mu(y) + \frac{\partial}{\partial y}\psi_{t-s}(\gamma, y)\sigma^2(y), \qquad s \leq t. \tag{3.11}$$

The result (3.9) follows by applying Ito's formula to the function $(y, s) \to \Psi_{t-s}(\gamma, y)$. The second result follows from Theorem 3.3 in Jacod and Mémin (1976).

Example 3.2. Next let the observed Markov process X be a counting process with intensity $\lambda_t(\theta) = (1-\theta)F(X_{t-})$, where $\theta \in (-\infty, 1)$ and F is a mapping

$\mathbb{N} \to (0, \infty)$ satisfying that $F(x) \leq a + bx$ for some $a \geq 0$ and $b \geq 0$. Then X is non-explosive for all θ (see Jacobsen (1982), p. 115), and

$$\frac{dP_\theta^t}{dP_0^t} = \exp\left\{\theta \int_0^t F(X_s)ds + \log(1-\theta)(X_t - X_0)\right\}. \tag{3.12}$$

Under $Q_\gamma^{(t)}$, $\gamma \in \tilde{\Gamma}_t$, given by (3.1) the process $\{X_s : s \leq t\}$ has almost surely sample paths like a counting process because $Q_\gamma^{(t)}$ is dominated by P_0^t. In order to find the intensity of $\{X_s : s \leq t\}$ under $Q_\gamma^{(t)}$ note that it follows from (3.6) that for $u < s \leq t$

$$\begin{aligned} &Q_\gamma^{(t)}(X_s = i \mid \mathcal{F}_u) \\ &\quad = E_0\left[1_{\{i\}}(X_s)\exp\left\{\gamma_1 \int_u^s F(X_v)dv + \gamma_2(X_s - X_u)\right.\right. \\ &\qquad\qquad\qquad \left.\left. + \Psi_{t-s}(\gamma, X_s) - \psi_{t-u}(\gamma, X_u)\right\} \,\middle|\, X_u\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} &(s-u)^{-1}Q_\gamma^{(t)}(X_s = i+1 \mid X_u = i) \\ &\quad = \exp(\gamma_2 + \psi_{t-s}(\gamma, i+1) - \psi_{t-u}(\gamma, i)) \\ &\quad\quad \times (s-u)^{-1}E_0\left(1_{\{i+1\}}(X_s)\exp\left[\gamma_1 \int_u^s F(X_v)dv\right] \,\middle|\, X_u = i\right) \\ &\quad \underset{s\downarrow u}{\to} \exp(\gamma_2 + \psi_{t-u}(\gamma, i+1) - \psi_{t-u}(\gamma, i))F(i), \end{aligned}$$

where we have used that the intensity of X under P_0 is $F(X_{t-})$. We have, for simplicity, assumed that $\psi_t(\gamma, x)$ is a left-continuous function of time. The intensity under $Q_\gamma^{(t)}$ is thus given by

$$\lambda_s^{(t)}(\gamma) = \exp(\gamma_2 + \psi_{t-s}(\gamma, X_{s-} + 1) - \psi_{t-s}(\gamma, X_{s-}))F(X_{s-}), \qquad s \leq t. \tag{3.13}$$

3.3 *Explicit calculations*

We conclude this section by giving some results about how to calculate the function $\Psi_t(\gamma)$ explicitly for general curved exponential families. We assume that the family has a representation (2.2) with $\operatorname{int}\Theta \neq \emptyset$.

Let $P_{\theta,1}^t$ and $P_{\theta,2}^t$ denote the marginal distributions of A_t and S_t, respectively, under P_θ. Further, define for all $t \geq 0$ and $\theta \in \Theta$ the Laplace transforms

$$c_1(w; \theta, t) = E_\theta(e^{w^T A_t}) \quad \text{and} \quad c_2(w; \theta, t) = E_\theta(e^{w^T S_t}), \tag{3.14}$$

and denote by $D_i(\theta, t)$ the domain of $c_i(\cdot; \theta, t)$, $i = 1, 2$.

PROPOSITION 3.1. *The envelope exponential family on $\mathcal{F}_t$ contains the measures given by*

$$\frac{dQ_{\theta,\varphi}^{(t)}}{dP_0^t} = \exp[\theta^T A_t + \varphi^T S_t - \Psi_t(\theta, \varphi)] \tag{3.15}$$

where

$$\Psi_t(\theta,\varphi) = \log c_2(\varphi + \alpha_t(\theta); \theta, t) + \phi_t(\theta) \tag{3.16}$$

and

$$(\theta,\varphi) \in \mathcal{M}_t = \{(\theta,\varphi) : \theta \in \Theta, \varphi + \alpha_t(\theta) \in D_2(\theta,t)\}. \tag{3.17}$$

Suppose $\operatorname{int}\mathcal{M}_t \neq \emptyset$, *and let* $\bar{\mathcal{M}}_t$ *denote the largest subset of* $\mathbb{R}^m$ *to which* $\Psi_t(\theta,\varphi)$ *can be extended by analytic continuation. Then* $\tilde{\Gamma}_t = \bar{\mathcal{M}}_t$ *and the measures in the envelope family are given by* (3.15) *with* $\Psi_t(\theta,\varphi)$ *defined by analytic continuation.*

Remark. It is well-known that $\bar{\mathcal{M}}_t$ is a convex set and that the convex hull of $\{(\theta, -\alpha_t(\theta)) : \theta \in \Theta\}$ is contained in $\bar{\mathcal{M}}_t$. For a discussion of how to determine $\bar{\mathcal{M}}_t$, see Hoffmann-Jørgensen (1994).

PROOF. Let $\bar{P}^t_\theta$ denote the conditional distribution under P_θ of A_t given S_t, and set

$$f_t(x;\theta) = \frac{dP^t_{\theta,2}}{dP^t_{0,2}}(x).$$

Then

$$\frac{d\bar{P}^t_\theta}{d\bar{P}^t_0} = \exp[\theta^T A_t - \alpha_t(\theta)^T S_t - \phi_t(\theta) - \log f_t(S_t;\theta)],$$

from which we see that

$$E_0(e^{\theta^T A_t} \mid S_t) = \exp[\alpha_t(\theta)^T S_t + \phi_t(\theta) + \log f_t(S_t;\theta)].$$

Therefore

$$\begin{aligned} E_0(e^{\theta^T A_t + \varphi^T S_t}) &= E_0(e^{\varphi^T S_t} E_0(e^{\theta^T A_t} \mid S_t)) \\ &= E_0(e^{(\varphi+\alpha_t(\theta))^T S_t} f_t(S_t;\theta))e^{\phi_t(\theta)} \\ &= E_\theta(e^{(\varphi+\alpha_t(\theta))^T S_t})e^{\phi_t(\theta)} \\ &= c_2(\varphi + \alpha_t(\theta); \theta, t)e^{\phi_t(\theta)} \end{aligned}$$

provided $\theta \in \Theta$ and $\varphi + \alpha_t(\theta) \in D_2(\theta,t)$.

The extension to $\bar{\mathcal{M}}_t$ follows from well-known properties of the Laplace transform, see e.g. Hoffmann-Jørgensen (1994). The idea of exploiting the above expression for the conditional Laplace transform of A_t given S_t was first used by Jensen (1987) to obtain conditional expansions. □

Using arguments similar to those for Proposition 3.1 we can prove the following result.

PROPOSITION 3.2. *Suppose the function $\theta \to \alpha_t(\theta)$ is invertible on $\Theta_t^* \subseteq \Theta$ and set $\Lambda_t = -\alpha_t(\Theta_t^*)$. Then*

$$\mathcal{M}_t^* = \{(\theta,\varphi) : \varphi \in \Lambda_t, \theta - \alpha_t^{-1}(-\varphi) \in D_1(\alpha_t^{-1}(-\varphi),t)\} \subseteq \tilde{\Gamma}_t, \tag{3.18}$$

and for $(\theta,\varphi) \in \mathcal{M}_t^*$ the function $\Psi_t(\theta,\varphi)$ in (3.15) *can be expressed as*

$$\Psi_t(\theta,\varphi) = \log[c_1(\theta - \alpha_t^{-1}(-\varphi); \alpha_t^{-1}(-\varphi),t)] + \phi_t(\alpha_t^{-1}(-\varphi)). \tag{3.19}$$

If $\operatorname{int}\mathcal{M}_t^* \neq \emptyset$, *the whole envelope family can be obtained by analytic continuation as described in Proposition* 3.1.

Note that it follows from Proposition 3.1 or Proposition 3.2 that in order to show that an element of our exponential family of processes belongs to the interior of the envelope exponential family, we need only know something about the tail behaviour of the distribution under P_θ of A_t or of S_t.

4. Examples

In this section we will study some examples of curved exponential families of stochastic processes and their envelope families.

4.1 *The Gaussian autoregression*

The Gaussian autoregression of order one is defined by

$$X_i = \theta X_{i-1} + Z_i, \qquad i = 1,2,\dots, \tag{4.1}$$

where $\theta \in \mathbb{R}$, $X_0 = x_0$ and where the Z_i's are independent standard normal distributed random variables. This is a curved exponential family of processes with the representation

$$\frac{dP_\theta^t}{dP_0^t} = \exp\left[\theta \sum_{i=1}^t X_i X_{i-1} - \frac{1}{2}\theta^2 \sum_{i=1}^t X_{i-1}^2\right]. \tag{4.2}$$

The envelope family on $\mathcal{F}_t$ has a representation of the form

$$\frac{dQ_{\theta,\varphi}^{(t)}}{dP_0^t} = \exp\left[\theta \sum_{i=1}^t X_i X_{i-1} + \varphi \sum_{i=1}^t X_{i-1}^2 - \Psi_t(\theta,\varphi;x_0)\right], \tag{4.3}$$

where $(\theta,\varphi) \in \tilde{\Gamma}_t$. The function $\Psi_t(\theta,\varphi;x_0)$ is easily found by direct calculation. Indeed,

$$\begin{aligned}
&\exp(\Psi_t(\theta,\varphi;x_0)) \\
&\quad= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[\theta \sum_{i=1}^t x_i x_{i-1} + \varphi \sum_{i=1}^t x_{i-1}^2\right] (2\pi)^{-t/2} \\
&\quad\quad \times \exp\left[-\frac{1}{2}\sum_{i=1}^t x_i^2\right] dx_t \cdots dx_1 \\
&\quad= \exp[x_0^2(\varphi - \theta^2/(4A_t))](2\pi)^{-t/2} \\
&\quad\quad \times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[\sum_{i=1}^t A_{t-i+1}(x_i + \theta x_{i-1}/(2A_{t-i+1}))^2\right] dx_t \cdots dx_1,
\end{aligned}$$

where the quantities $A_1, \ldots, A_t$ are functions of θ and φ defined iteratively by

$$A_1 = -\frac{1}{2} \quad \text{and} \quad A_i = \varphi - \frac{1}{2} - \theta^2/(4A_{i-1}). \tag{4.4}$$

Clearly, $\Psi_t(\theta, \varphi; x_0)$ is finite if and only if $A_i(\theta, \varphi) < 0$ for $i = 1, \ldots, t$, and if this is the case,

$$\Psi_t(\theta, \varphi; x_0) = x_0^2 \left[A_{t+1}(\theta, \varphi) + \frac{1}{2} \right] - \frac{1}{2} \sum_{i=1}^{t} \log(-2A_i(\theta, \varphi)). \tag{4.5}$$

Explicit, but complicated, expressions for the A_i's can be derived from results in White (1958).

The set $\tilde{\Gamma}_t = \{(\theta, \varphi) : A_i(\theta, \varphi) < 0, i = 1, \ldots, t\}$ is not easy to characterize in an explicit way. However, because $\tilde{\Gamma}_t$ is a convex set containing $\{(\theta, -\frac{1}{2}\theta^2) : \theta \in \mathbb{R}\}$, it follows that $\{(\theta, \varphi) : \varphi \leq -\frac{1}{2}\theta^2\} \subseteq \tilde{\Gamma}_t$ for all $t \geq 1$. Moreover, from the inequality $A_2 = \varphi - \frac{1}{2} + \frac{1}{2}\theta^2 < 0$ we see that $\tilde{\Gamma}_t \subseteq \{(\theta, \varphi) : \varphi < -\frac{1}{2}\theta^2 + \frac{1}{2}\}$ for $t \geq 2$. An elementary, but somewhat involved, analysis of the iteration formula (4.4) reveals that

$$\bigcap_{t>0} \tilde{\Gamma}_t = \left\{ (\theta, \varphi) : \varphi \leq \frac{1}{2} - |\theta|, \varphi < \frac{1}{2}, |\theta| \leq 1 \right\} \cup \left\{ (\theta, \varphi) : \varphi \leq -\frac{1}{2}\theta^2 \right\} \tag{4.6}$$

and that $(\theta, -\frac{1}{2}\theta^2) \in \operatorname{int} \tilde{\Gamma}_t$ for $|\theta| \neq 1$ for all $t \geq 1$. For $|\theta| = 1$ the points $(\theta, -\frac{1}{2}\theta^2) \in bd\tilde{\Gamma}_t$ for t large enough. By (3.6) and (4.5) the restriction of $Q_{\theta,\varphi}^{(t)}$ to $\mathcal{F}_s$ $(s \leq t)$ is given by

$$\begin{aligned} \frac{dQ_{\theta,\varphi}^{(t,s)}}{dP_0^s} = \exp \Bigg[& \theta \sum_{i=1}^{s} X_i X_{i-1} + \varphi \sum_{i=1}^{s} X_{i-1}^2 + \left(A_{t-s+1}(\theta, \varphi) + \frac{1}{2} \right) X_s^2 \\ & - \left(A_{t+1}(\theta, \varphi) + \frac{1}{2} \right) x_0^2 + \frac{1}{2} \sum_{i=t-s+1}^{t} \log(-2A_i(\theta, \varphi)) \Bigg]. \end{aligned} \tag{4.7}$$

Note that $\{Q_{\theta,\varphi}^{(t,s)} : (\theta, \varphi) \in \tilde{\Gamma}_t\}$ is an exponential family for all $s \leq t$ so that $\{Q_{\theta,\varphi}^{(t)} : (\theta, \varphi) \in \tilde{\Gamma}_t\}$ defines an exponential family of stochastic processes which is not time-homogeneous.

The simultaneous Laplace transform under $Q_{\theta,\varphi}^{(t)}$ of the random variables $W_i = X_i + \frac{1}{2}\theta A_{t-i+1}^{-1}(\theta, \varphi) X_{i-1}$, $i = 1, \ldots, t$, can be found by direct calculation. This shows that the random variables W_i, $i = 1, \ldots, t$ are independent, and that $W_i \sim N(0, -\frac{1}{2}A_{t-i+1}^{-1})$, $i = 1, \ldots, t$. Under $Q_{\theta,\varphi}^{(t)}$ the process $\{X_i : i = 1, \ldots, t\}$ is thus the autoregression

$$X_i = -\frac{1}{2}\theta A_{t-i+1}^{-1}(\theta, \varphi) X_{i-1} + W_i, \tag{4.8}$$

where the regression parameter as well as the variance of W_i depend on i.

If we restrict the parameter set to $\Gamma_t^* = \tilde{\Gamma}_t \backslash \{(\theta, \varphi) : |\theta| \leq 1, \varphi > -\frac{1}{2}\theta^2\}$, the process (4.8) can be extended beyond t in a natural way. This is done by defining A_{-i}, $i = 0, 1, 2, \ldots$, iteratively such that they are related by (4.4). Considerations like those for $i \geq 1$ show that for $(\theta, \varphi) \in \Gamma_t^*$ we have $A_{-i}(\theta, \varphi) < 0$, $i = 0, 1, 2, \ldots$, while for $(\theta, \varphi) \in \tilde{\Gamma}_t \backslash \Gamma_t^*$ there exists an $i \geq 0$ such that $A_{-i}(\theta, \varphi) > 0$. The likelihood function for the extended process is given by (4.7) for all $s \in \mathbb{N}$.

4.2 *The pure birth process*

The pure birth processes are counting processes with intensity λX_{t-} where $\lambda > 0$. We assume that $X_0 = x_0$ is given. The likelihood function is

$$\frac{dP_\theta^t}{dP_0^t} = \exp\left[\theta \int_0^t X_s ds + \log(1-\theta)(X_t - x_0)\right], \tag{4.9}$$

where $\theta = 1 - \lambda < 1$. To determine the envelope families by means of Proposition 3.1, we use that the Laplace transform of $S_t = X_t - x_0$ is

$$E_\theta(e^{zS_t}) = [e^{(1-\theta)t} - e^z(e^{(1-\theta)t} - 1)]^{-x_0}$$

with domain $z < -\log[1 - \exp((\theta - 1)t)]$. Hence

$$\frac{dQ_{\theta,\varphi}^{(t)}}{dP_0^t} = \exp\left[\theta \int_0^t X_s ds + \varphi(X_t - x_0) - x_0\beta_t(\theta, \varphi)\right], \tag{4.10}$$

where

$$\beta_t(\theta, \varphi) = -\log[e^{(1-\theta)t} - e^\varphi(1-\theta)^{-1}(e^{(1-\theta)t} - 1)] \tag{4.11}$$

and $\varphi < \log[(1-\theta)/(1 - \exp((\theta - 1)t))]$. Here we have used that $\beta_t(\theta, \varphi)$ is also defined for $\theta \geq 1$ provided φ is as specified. Note that the canonical parameter set of the class of linear birth processes $\Gamma = \{(\theta, \log(1-\theta)) : \theta < 1\}$ is contained in $\tilde{\Gamma}_t$, which is open for all $t > 0$. Note also that $\tilde{\Gamma} = \bigcap_{t>0} \tilde{\Gamma}_t = \operatorname{conv} \Gamma$, where $\operatorname{conv} \Gamma$ denotes the convex hull of Γ. The simultaneous cumulant transform of $X_t - x_0$ and $\int_0^t X_s ds$ appearing in (4.10) was first calculated by Puri (1966), see also Keiding (1974).

The family $\{Q_{\theta,\varphi}^{(t,s)} : (\theta, \varphi) \in \tilde{\Gamma}_t\}$ obtained by restriction to $\mathcal{F}_s$ $(s < t)$ is an exponential family, which is not time-homogeneous. By (3.6) and (4.11) we see that

$$\frac{dQ_{\theta,\varphi}^{(t,s)}}{dP_0^s} = \exp\left\{\theta \int_0^s X_u du + h_{t-s}(\theta, \varphi)X_s - x_0 h_t(\theta, \varphi)\right\} \tag{4.12}$$

with $h_u(\theta, \varphi) = \varphi + \beta_u(\theta, \varphi)$. In Example 3.2 we saw that under $Q_{\theta,\varphi}^{(t)}$ the process $\{X_s : s \leq t\}$ is a counting process with intensity $\lambda_s^{(t)}(\theta, \varphi) = \exp[h_{t-s}(\theta, \varphi)]X_{s-}$.

For every $(\theta, \varphi) \in \tilde{\Gamma}$ the function $\beta_u(\theta, \varphi)$ is not only defined for $u \in [0, t]$, but also for $u < 0$. The function $\exp[h_{t-s}(\theta, \varphi)]$ is thus defined for all $s > 0$ and remains bounded for $s \to \infty$ for all $(\theta, \varphi) \in \tilde{\Gamma}_t$. Hence $\lambda_s^{(t)}(\theta, \varphi)$ defines a non-exploding counting process for all $s > 0$, and for each $(\theta, \varphi) \in \tilde{\Gamma}_t$ there exists a measure $P_{\theta,\varphi}^{(t)}$ on $\mathcal{F}$ the restriction of which to $\mathcal{F}_s$ is given by (4.12) for all $s > 0$. For all $s > 0$ the measure $Q_{\theta,\varphi}^{(t,s)}$ belongs to the exponential family $\{Q_{\theta,\varphi}^{(s)} : (\theta, \varphi) \in \tilde{\Gamma}_t\}$. The curve $(\theta, h_{t-s}(\theta, \varphi))$ tends monotonically to $(\theta, \log(1-\theta))$ for $\theta < 1$, i.e. to the curve Γ corresponding to the original counting process model. For $\theta \geq 1$ the function $h_{t-s}(\theta, \varphi)$ decreases to $-\infty$ for $s \to \infty$. For $s = t$ the curve passes through the point (θ, φ) for all $\theta \in \mathbb{R}$.

4.3 *The Ornstein-Uhlenbeck process*

Consider the class of solutions to the stochastic differential equations

$$dX_t = \theta X_t dt + dW_t, \qquad X_0 = x_0, \tag{4.13}$$

for $\theta \in \mathbb{R}$. The likelihood function corresponding to observation of X in $[0, t]$ is

$$L_t(\theta) = \exp\left\{\theta[X_t^2 - x_0^2]/2 - \frac{1}{2}\theta^2 \int_0^t X_s^2 ds - \frac{1}{2}\theta t\right\}. \tag{4.14}$$

The envelope families can be determined by Proposition 3.2 and are given by

$$\frac{dQ_{\theta,\varphi}^{(t)}}{dP_0^t} = \exp\left\{\theta[X_t^2 - x_0^2]/2 + \varphi \int_0^t X_s^2 ds - \Psi_t(\theta, \varphi; x_0)\right\} \tag{4.15}$$

with

$$\begin{aligned}\Psi_t(\theta, \varphi; x_0) = & -\frac{1}{2}\log\{\cosh(t\sqrt{-2\varphi}) - \theta\sinh(t\sqrt{-2\varphi})/\sqrt{-2\varphi}\} \\ & + \frac{x_0^2\left(\frac{1}{2}\theta^2 + \varphi\right)}{\sqrt{-2\varphi}\coth(t\sqrt{-2\varphi}) - \theta}\end{aligned} \tag{4.16}$$

and with parameter space $\tilde{\Gamma}_t$ given by $\varphi < \frac{1}{2}\pi^2 t^{-2}$, $\theta < \sqrt{-2\varphi}\coth(t\sqrt{-2\varphi})$, for details see Sørensen (1995).

The family $\{Q_{\theta,\varphi}^{(t,s)} : (\theta, \varphi) \in \tilde{\Gamma}_t\}$ is an exponential family for all $s \leq t$ which is not time-homogeneous. By (3.6) and (4.16) we see that

$$\begin{aligned}\frac{dQ_{\theta,\varphi}^{(t,s)}}{dP_0^s} = \exp\Big\{ & h(\theta, \varphi; t-s)X_s^2 + \varphi \int_0^s X_u^2 du \\ & + m(\theta, \varphi; t, s) - h(\theta, \varphi; t)x_0^2\Big\},\end{aligned} \tag{4.17}$$

where

$$h(\theta, \varphi; u) = \frac{1}{2}\theta + \frac{\frac{1}{2}\theta^2 + \varphi}{\sqrt{-2\varphi}\coth(u\sqrt{-2\varphi}) - \theta} \tag{4.18}$$

and

(4.19) $m(\theta, \varphi; t, s)$

$$= \frac{1}{2} \log \left\{ \frac{\cosh(t\sqrt{-2\varphi}) - \theta \sinh(t\sqrt{-2\varphi})/\sqrt{-2\varphi}}{\cosh((t-s)\sqrt{-2\varphi}) - \theta \sinh((t-s)\sqrt{-2\varphi})/\sqrt{-2\varphi}} \right\}.$$

By results in Example 3.1 it follows that under $Q^{(t)}_{\theta,\varphi}$ the process X solves the stochastic differential equation

$$dX_s = 2h(\theta, \varphi; t-s)X_s ds + dW_s, \qquad X_0 = x_0, \;\; s \leq t. \tag{4.20}$$

For $\varphi \leq 0$ and $\theta \geq -\sqrt{-2\varphi}$ the function h is well-defined and bounded for $u < 0$, so (4.20) has a solution for all $s \geq 0$. This is not the case if $\varphi \leq 0$ and $\theta < -\sqrt{-2\varphi}$ or if $\varphi > 0$. In these cases the drift tends to infinity (or minus infinity) at a finite time larger than t.

5. Goodness-of-fit tests

A possible test of the appropriateness of a stochastic process model, which for observation in $[0, t]$ is a curved exponential family, is the likelihood ratio test of the curved model against the full envelope family on $\mathcal{F}_t$. For an interpretation of this test and an evaluation of its relevance, the results in Section 3 are useful. For the Ornstein-Uhlenbeck process, for instance, the drift under the alternative model is not strictly proportional to the state of the process, but a certain temporal variation of the constant of proportionality is allowed. Similar remarks hold for the Gaussian autoregression and the pure birth process. The following simple, but interesting, example illustrates the main ideas.

Example 5.1. (A model for censored data) Consider the following well-known model for censored observation of a random variable with hazard function $(1-\theta)h$ $(\theta < 1)$ defined on $(0, \infty)$. We suppose that $h > 0$ and that it is integrable on $(0, t)$ for all $t > 0$. Let U and V be two independent random variables concentrated on $(0, \infty)$ such that the hazard function of U is $(1-\theta)h$, and denote the cumulative distribution function of V by G. Define two counting processes N and M by

$$N_t = 1_{\{U \leq t \wedge V\}}, \qquad M_t = 1_{\{V \leq t \wedge U\}}.$$

The intensity of N with respect to the filtration given by $\mathcal{F}_t = \sigma(N_s, M_s : s \leq t)$ is

$$\lambda_t(\theta) = (1-\theta)h(t)1_{\{t \leq V\}}1_{\{N_{t-}=0\}}, \qquad \theta < 1. \tag{5.1}$$

Observation of N and M in the time interval $[0, t]$ is equivalent to observation of a random variable U censored at time $V \wedge t$.

The likelihood function for the model based on observation of N and M in $[0, t]$ is given by

$$\frac{dP^t_\theta}{dP^t_0} = \exp\left\{\theta \int_0^{t\wedge V \wedge U} h(s)ds + \log(1-\theta)N_t\right\}. \tag{5.2}$$

We see that the model is a curved exponential family of stochastic processes. A possible test that the hazard function of the observed random variable belongs to the class $(1-\theta)h$, $\theta < 1$, is the likelihood ratio test for the curved family (5.2) against the envelope family on $\mathcal{F}_t$. This test was proposed in the particular case $h \equiv 1$ (censored exponential distribution) by Væth (1980). He gave an interpretation of the alternative hypothesis by means of a biased sampling scheme (Væth (1982)). Here we obtain a different interpretation by considering the envelope family as a stochastic process model.

It is easy to see that $E_\theta(e^{wN_t}) = \beta_t(\theta) + e^w(1-\beta_t(\theta))$, where $\beta_t(\theta) = P_\theta(N_t = 0) = \int_0^\infty (1-F_\theta(v \wedge t))dG(v)$ is the probability of obtaining a censored observation, and $F_\theta(x) = 1 - \exp\{-(1-\theta)\int_0^x h(v)dv\}$ is the cumulative distribution function corresponding to the hazard function $(1-\theta)h$. By Proposition 3.1 the envelope family is given by

$$\frac{dQ_{\theta,\varphi}^{(t)}}{dP_0^t} = \exp\left\{\theta \int_0^{t\wedge V\wedge U} h(s)ds + \varphi N_t - \Psi_t(\theta,\varphi)\right\} \tag{5.3}$$

with $\varphi \in \mathbb{R}$ and $\Psi_t(\theta,\varphi) = \log[\beta_t(\theta) + (1-\theta)^{-1}e^\varphi(1-\beta_t(\theta))]$.

The process N is not a Markov process (except for $h \equiv 1$), so the conclusions in Example 3.2 do not apply directly, but we can proceed in a very similar way. Thus we find that under $Q_{\theta,\varphi}^{(t)}$ the process $(N_s : s \leq t)$ is a counting process that makes at most one jump. Its intensity with respect to $\{\mathcal{F}_u\}$ is

$$\{\gamma_u^{(t)}(\theta)(e^{-\varphi} - (1-\theta)^{-1}) + (1-\theta)^{-1}\}^{-1} h(u) 1_{\{u \leq V\}} 1_{\{N_{u-}=0\}},$$

for $u \leq t$, so under $Q_{\theta,\varphi}^{(t)}$ observation of N and M in $[0,t]$ is equivalent to censored observation of a random variable with hazard function

$$\{\gamma_s^{(t)}(\theta)(e^{-\varphi} - (1-\theta)^{-1}) + (1-\theta)^{-1}\}^{-1} h(s), \qquad s \leq t. \tag{5.4}$$

Consider the situation where the censoring distribution G is concentrated on $[t,\infty)$ (type 1 censoring). Then $\gamma_s^{(t)}(\theta)$ is given by

$$\gamma_s^{(t)}(\theta) = \frac{1-F_\theta(t)}{1-F_\theta(s)}$$

which is an increasing function of s. The factor modifying h in (5.4) is in this situation increasing or decreasing depending on whether $\varphi > \log(1-\theta)$ or $\varphi < \log(1-\theta)$.

As another example suppose $h \equiv 1$ and $G(x) = 1 - e^{-\mu x}$. Then

$$\gamma_s^{(t)}(\theta) = \frac{\mu}{1-\theta+\mu} + \frac{1-\theta}{1-\theta+\mu} e^{-(1-\theta+\mu)s}.$$

Here $\gamma_s^{(t)}(\theta)$ is a decreasing function of s, and the hazard function given by (5.4) is monotonically increasing or decreasing depending on whether $\varphi < \log(1-\theta)$ or $\varphi > \log(1-\theta)$.

REFERENCES

Hoffmann-Jørgensen, J. (1994). *Probability with a View to Statistics*, Chapman and Hall, New York.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam.

Jacobsen, M. (1982). Statistical analysis of counting processes, *Lecture Notes in Statist.*, **12**, Springer, New York.

Jacod, J. and Mémin, J. (1976). Caractéristiques locales et conditions de continuité absolue pour les semi-martingales, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **35**, 1–37.

Jensen, J. L. (1987). On asymptotic expansions in non-ergodic models, *Scand. J. Statist.*, **14**, 305–318.

Keiding, N. (1974). Estimation in the birth process, *Biometrika*, **61**, 71–80.

Küchler, U. and Sørensen, M. (1989). Exponential families of stochastic processes: A unifying semimartingale approach, *Internat. Statist. Rev.*, **57**, 123–144.

Küchler, U. and Sørensen, M. (1991). On exponential families of Markov processes, Research Report, No. 233, Department of Theoretical Statistics, Aarhus University, Denmark.

Küchler, U. and Sørensen, M. (1994*a*). Exponential families of stochastic processes with time-continuous likelihood functions, *Scand. J. Statist.*, **21**, 421–431.

Küchler, U. and Sørensen, M. (1994*b*). Exponential families of stochastic processes and Lévy processes, *J. Statist. Plann. Inference*, **39**, 211–237.

Puri, P. S. (1966). On the homogeneous birth-and-death process and its integral, *Biometrika*, **53**, 61–71.

Sørensen, M. (1986). On sequential maximum likelihood estimation for exponential families of stochastic processes, *Internat. Statist. Rev.*, **54**, 191–210.

Sørensen, M. (1993). The natural exponential family generated by a semimartingale, Research Report, No. 269, Department of Theoretical Statistics, Aarhus University, *Proceedings of the Fourth Russian-Finnish Symposium on Probability Theory and Mathematical Statistics*, TVP, Moscow (to appear).

Sørensen, M. (1995). On conditional inference for the Ornstein-Uhlenbeck process (in preparation).

Væth, M. (1980). A test for the exponential distribution with censored data, Research Report, No. 60, Department of Theoretical Statistics, Aarhus University, Denmark.

Væth, M. (1982). A sampling experiment leading to the full exponential family generated by the censored exponential distribution, Research Report, No. 78, Department of Theoretical Statistics, Aarhus University, Denmark.

White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case, *Ann. Math. Statist.*, **29**, 1188–1197.