Ann. Inst. Statist. Math.
Vol. 47, No. 2, 195–209 (1995)

LOCAL ASYMPTOTIC NORMALITY OF A SEQUENTIAL MODEL FOR MARKED POINT PROCESSES AND ITS APPLICATIONS

YOICHI NISHIYAMA

The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan

(Received April 11, 1994; revised June 27, 1994)

Abstract. This paper deals with statistical inference problems for a special type of marked point processes based on the realization in random time intervals $[0, \tau_u]$. Sufficient conditions to establish the local asymptotic normality (LAN) of the model are presented, and then, certain class of stopping times τ_u satisfying them is proposed. Using these stopping rules, one can treat the processes within the framework of LAN, which yields asymptotic optimalities of various inference procedures. Applications for compound Poisson processes and continuous time Markov branching processes (CMBP) are discussed. Especially, asymptotically uniformly most powerful tests for criticality of CMBP can be obtained. Such tests do not exist in the case of the non-sequential approach. Also, asymptotic normality of the sequential maximum likelihood estimators (MLE) of the Malthusian parameter of CMBP can be derived, although the non-sequential MLE is not asymptotically normal in the supercritical case.

Key words and phrases: Local asymptotic normality, stopping rule, marked point process, branching process, maximum likelihood estimation, test for criticality.

1. Introduction

It is well known that the asymptotic behavior of the continuous time Markov branching processes is drastically different for the subcritical, critical and supercritical cases. For this cause, there do not exist asymptotically uniformly most powerful tests for key parameters of the processes. Also, even if we restrict our attention to the supercritical case, the asymptotic distributions of the maximum likelihood estimators are non-normal. For such problems, several authors have developed the conditional inference procedures within the framework of locally asymptotically mixed normal (LAMN) families (see e.g. Basawa and Scott (1983)). On the other hand, the method discussed in this paper is based on a sequential approach. Our main goal is to propose certain stopping rules which allow us to treat such processes within the framework of locally asymptotically normal (LAN) families. The general theory for LAN families provides stronger optimality results than that for LAMN families.

Including the continuous time Markov branching processes, this paper deals with a class of marked point processes. After the preliminaries in Section 2, a statistical model for marked point processes is introduced in Section 3. Our formulation also contains the compound Poisson processes and the birth-and-death processes. A series of conditions to establish the LAN property of the model is given in Section 4. In Section 5, we present a class of stopping rules which leads us to the main goal. Section 6 consists of some applications of the results. Especially, based on our stopping rules, we can obtain the asymptotically uniformly most powerful tests for criticality of the continuous time Markov branching processes. The asymptotic normality of the sequential maximum likelihood estimator of the Malthusian parameter can also be derived.

Our work is motivated by Sørensen's (1986) one, who has considered exponential families of stochastic processes. However, as far as one considers point processes, our formulation and stopping rules are more general than his. Mel'nikov and Novikov (1989) have considered sequential inference problems for semimartingales, on the other hand, Svensson (1990) has investigated the sequential maximum likelihood estimation for multivariate point processes. Exact (not asymptotic) optimal estimation problems for the compound Poisson processes are discussed by Stefanov (1982), and for the birth-and-death processes by Franz (1982) and Manjunath (1984).

2. Preliminaries

In this section, we summarize the theory of marked point processes (see Brémaud (1981) or Jacod and Shiryaev (1987) for details) and give a central limit theorem which is used in the following sections.

Let E be a Polish space endowed with its Borel σ-field $\mathcal{E}$. Fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, where the filtration $(\mathcal{F}_t)$ is assumed to be increasing and right continuous. We do not ask the σ-fields to be complete. An *E-marked point process* is a double sequence $p = \{(T_i, Z_i); i = 1, 2, \ldots\}$ of random variables defined on $(\Omega, \mathcal{F}, P)$ where $0 < T_1 < T_2 < \cdots$, P-a.s. and $Z_i \in E$, $i = 1, 2, \ldots$. The measurable space $(E, \mathcal{E})$ is called the *mark space*.

For given E-marked point process p, we define the non-negative integer valued random measure $p(dt \times dz)$ on $[0, \infty) \times E$ by $p([0, t] \times A) = \#\{i; T_i \in [0, t], Z_i \in A\}$, $t \in [0, \infty)$, $A \in \mathcal{E}$. We also use the notations $N_t(A) := p([0, t] \times A)$ for $A \in \mathcal{E}$ and especially $N_t := p([0, t] \times E)$. Note that $N_0 = 0$, P-a.s. by definition. Here, we assume that p admits the $(P, (\mathcal{F}_t))$-intensity kernel $\lambda_t(dz)$; that is, $\lambda_t(\omega, dz)$ is a transition measure from $(\Omega \times [0, \infty), \mathcal{F} \otimes \mathcal{B}[0, \infty))$ into $(E, \mathcal{E})$, and for every $A \in \mathcal{E}$, $N_t(A)$ admits the $(P, (\mathcal{F}_t))$-predictable intensity $\lambda_t(A)$.

Next, let $H(t, z)$ be an E-indexed $(\mathcal{F}_t)$-predictable process such that the process $t \mapsto \int_0^t \int_E |H(s, z)|^2 \lambda_s(dz) ds$ is locally integrable. Then the stochastic integral of $H(t, z)$ with respect to $q(dt \times dz) = p(dt \times dz) - \lambda_t(dz)dt$ is well defined, and $t \mapsto \int_0^t \int_E H(s, z) q(ds \times dz)$ is a $(P, (\mathcal{F}_t))$-locally square integrable martingale with the quadratic variational process $\langle \int_0^t \int_E H(s, z) q(ds \times dz) \rangle = \int_0^t \int_E |H(s, z)|^2 \lambda_s(dz) ds$.

Finally, we give a central limit theorem for randomly stopped martingales derived from marked point processes. It is a slightly extended version of Kutoyants' work ((1984), Theorem 4.5.4), so we omit the proof. Although a similar result is given by Svensson ((1990), Theorem 5.2), certain a.s. uniformly boundedness condition imposed there is not desirable for us.

Let $\mathcal{U}$ be the totality of the positive integers $\{1, 2, \ldots\}$ or the positive real numbers $(0, \infty)$. Assume that, for each $u \in \mathcal{U}$, a mark space $(E^{(u)}, \mathcal{E}^{(u)})$, a filtered probability space $(\Omega^{(u)}, \mathcal{F}^{(u)}, (\mathcal{F}_t^{(u)}), P^{(u)})$ and an $E^{(u)}$-marked point process $p^{(u)}$ as above are given. Further, let τ_u be an $(\mathcal{F}_t^{(u)})$-stopping time such that $P^{(u)}(\tau_u < \infty) = 1$. Finally, let d be a fixed positive integer.

THEOREM 2.1. *For each $u \in \mathcal{U}$, let $H_i^{(u)}(t, z)$, $i = 1, \ldots, d$ be $E^{(u)}$-indexed $(\mathcal{F}_t^{(u)})$-predictable processes satisfying the following conditions* (i)–(iii):

(i) *For every $i = 1, \ldots, d$,*

$$\int_0^t \int_{E^{(u)}} |H_i^{(u)}(s, z)|^2 \lambda_s^{(u)}(dz) ds < \infty \quad \textit{for every } t \in [0, \infty),\ P^{(u)}\textit{-a.s.}$$

(ii) *For every $i, j = 1, \ldots, d$,*

$$\int_0^{\tau_u} \int_{E^{(u)}} H_i^{(u)}(s, z) H_j^{(u)}(s, z) \lambda_s^{(u)}(dz) ds \to \sigma_{ij} \quad \textit{in } P^{(u)}\textit{-probability as } u \to \infty,$$

and the $d \times d$ matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$ is positive definite.

(iii) *For any $\varepsilon > 0$,*

$$\lim_{u \to \infty} \boldsymbol{E}^{(u)} \left[\int_0^{\tau_u} \int_{E^{(u)}} \sum_{i=1}^d |H_i^{(u)}(s, z)|^2 \cdot I\left\{ \sum_{i=1}^d |H_i^{(u)}(s, z)|^2 > \varepsilon \right\} \lambda_s^{(u)}(dz) ds \right] = 0.$$

Then

$$\left(\int_0^{\tau_u} \int_{E^{(u)}} H_1^{(u)}(s, z) q^{(u)}(ds \times dz), \ldots, \int_0^{\tau_u} \int_{E^{(u)}} H_d^{(u)}(s, z) q^{(u)}(ds \times dz) \right)^{\top} \to N(\boldsymbol{0}, \boldsymbol{\Sigma}) \quad \textit{in } P^{(u)}\textit{-law as } u \to \infty.$$

3. Introduction of the model

Let us introduce a statistical model for marked point processes. Fix a mark space $(E, \mathcal{E})$ on which a σ-finite measure μ is defined. Fix a positive integer d and an open set $\boldsymbol{\Theta} \subset \boldsymbol{R}^d$. Let $\mathcal{U} = \{1, 2, \ldots\}$ or $(0, \infty)$. For each $u \in \mathcal{U}$, let $p^{(u)} = \{(T_i^{(u)}, Z_i^{(u)}); i = 1, 2, \ldots\}$ be an E-marked point process defined on a filtered probability space $(\Omega^{(u)}, \mathcal{F}^{(u)}, (\mathcal{F}_t^{(u)}), \{P_{\boldsymbol{\theta}}^{(u)}; \boldsymbol{\theta} \in \boldsymbol{\Theta}\})$, where the probability measures $\{P_{\boldsymbol{\theta}}^{(u)}; \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ are mutually absolutely continuous on $\mathcal{F}_t^{(u)}$ for every $t \in [0,\infty)$. We also assume that the filtration $(\mathcal{F}_t^{(u)})$ is of the form

$$
\begin{aligned}
&\mathcal{F}_t^{(u)} = \mathcal{F}_0^{(u)} \vee \mathcal{H}_t^{(u)}, \\
&\mathcal{H}_t^{(u)} = \sigma\{p^{(u)}((0,s] \times A); s \in (0,t], A \in \mathcal{E}\}, \qquad t \in (0,\infty)
\end{aligned}
\tag{3.1}
$$

and that all the measures $\{P_{\boldsymbol{\theta}}^{(u)}; \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ coincide on $\mathcal{F}_0^{(u)}$. Further, we suppose that $p^{(u)}$ admits the $(P_{\boldsymbol{\theta}}^{(u)}, (\mathcal{F}_t^{(u)}))$-intensity kernel

$$
\lambda_t^{(u)}(dz;\boldsymbol{\theta}) = g(\boldsymbol{\theta}) y_t^{(u)} \cdot f(z;\boldsymbol{\theta})\mu(dz),
\tag{3.2}
$$

where $g(\boldsymbol{\theta})$ is a positive real valued function on $\boldsymbol{\Theta}$, $t \mapsto y_t^{(u)}$ is a non-negative $(\mathcal{F}_t^{(u)})$-predictable process and $\{f(z;\boldsymbol{\theta})\mu(dz); \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ are probability measures on $(E,\mathcal{E})$. Then, a version of the likelihood ratio process $t \mapsto L_t^{(u)}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) := dP_{\boldsymbol{\theta}_1}^{(u)}|\mathcal{F}_t^{(u)}/dP_{\boldsymbol{\theta}_0}^{(u)}|\mathcal{F}_t^{(u)}$ is given by

$$
L_t^{(u)}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) = \exp\left(\int_0^t \int_E \log\frac{\alpha(z;\boldsymbol{\theta}_1)}{\alpha(z;\boldsymbol{\theta}_0)} p^{(u)}(ds \times dz) - (g(\boldsymbol{\theta}_1) - g(\boldsymbol{\theta}_0))Y_t^{(u)}\right),
$$

where

$$
\alpha(z;\boldsymbol{\theta}) := g(\boldsymbol{\theta}) f(z;\boldsymbol{\theta}) \quad \text{and} \quad Y_t^{(u)} := \int_0^t y_s^{(u)} ds.
$$

When an E-marked point process $p^{(u)} = \{(T_i^{(u)}, Z_i^{(u)}); i = 1,2,\ldots\}$ as above is given, the sequence $\{T_i^{(u)}; i = 1,2,\ldots\}$ forms the (univariate) point process which admits the $(P_{\boldsymbol{\theta}}^{(u)}, (\mathcal{F}_t^{(u)}))$-predictable intensity $g(\boldsymbol{\theta})y_t^{(u)}$, and the random variables $Z_1^{(u)}, Z_2^{(u)}, \ldots$ are I.I.D. with the common distribution $f(z;\boldsymbol{\theta})\mu(dz)$ on E. We call $f(z;\boldsymbol{\theta})\mu(dz)$ the *mark distribution.*

For each $u \in \mathcal{U}$, let τ_u be an $(\mathcal{F}_t^{(u)})$-stopping time such that $P_{\boldsymbol{\theta}}^{(u)}(\tau_u < \infty) = 1$ for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Then, the probability measures $\{P_{\boldsymbol{\theta}}^{(u)}; \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ are mutually absolutely continuous on $\mathcal{F}_{\tau_u}^{(u)}$, and $L_{\tau_u}^{(u)}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_0) = dP_{\boldsymbol{\theta}_1}^{(u)}|\mathcal{F}_{\tau_u}^{(u)}/dP_{\boldsymbol{\theta}_0}^{(u)}|\mathcal{F}_{\tau_u}^{(u)}$ holds (see Jacod and Shiryaev (1987), p. 153). We are interested in the statistical inference problems for this model based on the realization of $p^{(u)}$ in the random time interval $[0,\tau_u]$. Also, we will consider certain class of stopping times τ_u in Section 5. Before proceeding with them, we mention a typical example which is included in the present model.

Example 3.1. *Birth-and-death process* (*large initial population case*). Let $\boldsymbol{\Theta} = (0,\infty)^2$, $E = \{+1,-1\}$ and $\mu(dz)$ be the counting measure on E. We assume that the mark distribution in (3.2) is given by

$$
f(z;\boldsymbol{\theta}) = \begin{cases} \dfrac{\theta_1}{\theta_1+\theta_2}, & z = +1, \\[2mm] \dfrac{\theta_2}{\theta_1+\theta_2}, & z = -1, \end{cases}
$$

and that $g(\boldsymbol{\theta}) = \theta_1 + \theta_2$, for $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top \in \boldsymbol{\Theta}$. Further, we set $\mathcal{U} = \{1, 2, \ldots\}$, and for each $u \in \mathcal{U}$

$$y_t^{(u)} = u + N_{t-}^{(u)}(\{+1\}) - N_{t-}^{(u)}(\{-1\}).$$

Then, the process $p^{(u)}$ forms the linear birth-and-death process with the birth rate θ_1, the death rate θ_2 and the initial population size u. Here $N_t^{(u)}(\{+1\})$ and $N_t^{(u)}(\{-1\})$ denote the numbers of births and deaths in the time interval $(0, t]$, respectively. Keiding (1975) has considered this process in his Section 3. He has supposed that the process is observed in the fixed time interval $[0, t_0]$, and investigated the asymptotic properties of maximum likelihood estimators as the initial population size $u \to \infty$. We consider the more general situation that the observation of $p^{(u)}$ is given in the random time interval $[0, \tau_u]$, and let $u \to \infty$.

Next, we mention that our formulation includes the *one realization case*, in which we drop the super-index $^{(u)}$ in all notations; that is, we denote the $(P_{\boldsymbol{\theta}}, (\mathcal{F}_t))$-intensity kernel of $p = \{(T_i, Z_i); i = 1, 2, \ldots\}$ by

$$\lambda_t(dz; \boldsymbol{\theta}) = g(\boldsymbol{\theta}) y_t \cdot f(z; \boldsymbol{\theta}) \mu(dz) \tag{3.3}$$

instead of (3.2). Note that stopping times τ_u and constants c_u (which will appear in the following sections) are indexed by $u \in \mathcal{U}$. A typical example is as follows.

Example 3.2. *Birth-and-death process (fixed initial population case)*. We assume that the mark distribution and $g(\boldsymbol{\theta})$ in (3.3) are given as in Example 3.1. Further, we set

$$y_t = x_0 + N_{t-}(\{+1\}) - N_{t-}(\{-1\}),$$

where x_0 is a fixed positive integer. Then, the process p forms the linear birth-and-death process with the birth rate θ_1, the death rate θ_2 and the initial population size x_0. Keiding (1975) has supposed in his Section 4 that the process is observed in the non-random time interval $[0, u]$, and let $u \to \infty$. We consider the more general situation that the observation is given in the random time interval $[0, \tau_u]$, and let $u \to \infty$.

In Section 6, we refer to our formulation (3.2) as the *general case*, contrasted with the term *one realization case*.

4. LAN property of the model

In this section, we give a series of conditions to establish the LAN property of the model introduced in the previous section. Fix a point $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$. Hereafter, we use the notation $\frac{\partial h(\boldsymbol{\theta}_0)}{\partial \theta_i}$ for $\frac{\partial h(\boldsymbol{\theta})}{\partial \theta_i}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$.

CONDITION 4.1. (A1) *There exists a neighborhood $\boldsymbol{\Theta}_0$ of $\boldsymbol{\theta}_0$ such that $\alpha(z; \boldsymbol{\theta}) = g(\boldsymbol{\theta}) f(z; \boldsymbol{\theta})$ is three times continuously differentiable on $\boldsymbol{\Theta}_0$ for μ-almost all z.*

(A2) *For every $i, j = 1, \dots, d$,*

$$\int_E \left| \frac{\partial \log \alpha(z; \boldsymbol{\theta}_0)}{\partial \theta_i} \right|^2 \alpha(z; \boldsymbol{\theta}_0) \mu(dz) < \infty \qquad \textit{and}$$

$$\int_E \left| \frac{\partial^2 \log \alpha(z; \boldsymbol{\theta}_0)}{\partial \theta_i \partial \theta_j} \right|^2 \alpha(z; \boldsymbol{\theta}_0) \mu(dz) < \infty.$$

Further, the $d \times d$ matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$ where

$$\sigma_{ij} = \int_E \frac{\partial \log \alpha(z; \boldsymbol{\theta}_0)}{\partial \theta_i} \frac{\partial \log \alpha(z; \boldsymbol{\theta}_0)}{\partial \theta_j} \alpha(z; \boldsymbol{\theta}_0) \mu(dz), \qquad i, j = 1, \dots, d$$

is positive definite.

(A3) *There exists an $\mathcal{E}$-measurable function $H(z)$ on E (possibly depending on $\boldsymbol{\theta}_0$) such that the relation*

$$\left| \frac{\partial^3 \log \alpha(z; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| \leq H(z)$$

holds for every $i, j, k = 1, \dots, d$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ and μ-almost all z, and that

$$\int_E |H(z)|^2 \alpha(z; \boldsymbol{\theta}_0) \mu(dz) < \infty.$$

(B) *There exist a family of constants $\{c_u; u \in \mathcal{U}\}$ such that $c_u \uparrow \infty$ as $u \to \infty$ and a positive constant η satisfying*

$$\lim_{u \to \infty} \boldsymbol{E}^{(u)}_{\boldsymbol{\theta}_0} \left[\left| \frac{1}{c_u} Y^{(u)}_{\tau_u} - \eta \right| \right] = 0.$$

Although the conditions (A1)–(A3) may seem to be classical and strong, they will serve a rich class of models covering most applications. On the other hand, the condition (B) is essential in our context. The process $t \mapsto N^{(u)}_t(\cdot)$ is non-ergodic as far as $\frac{1}{t} Y^{(u)}_t$ does not converge to a positive constant as $t \to \infty$. The condition (B) requires the existence of appropriate random time change to make the model be LAN with respect to $u \in \mathcal{U}$ even if the process is non-ergodic.

Set $L^{(u)}_t(\boldsymbol{\theta}) := L^{(u)}_t(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Under the conditions (A1)–(A3), we have

$$\begin{aligned} U^{(u)}_i(t; \boldsymbol{\theta}) &:= \frac{\partial \log L^{(u)}_t(\boldsymbol{\theta})}{\partial \theta_i} \\ &= \int_0^t \int_E \frac{\partial \log \alpha(z; \boldsymbol{\theta})}{\partial \theta_i} p^{(u)}(ds \times dz) - \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_i} Y^{(u)}_t, \qquad i = 1, \dots, d, \\ I^{(u)}_{ij}(t; \boldsymbol{\theta}) &:= \frac{\partial^2 \log L^{(u)}_t(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \end{aligned}$$

$$= \int_0^t \int_E \frac{\partial^2 \log \alpha(z;\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j} p^{(u)}(ds \times dz) - \frac{\partial^2 g(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j} Y_t^{(u)}, \qquad i,j = 1,\dots,d,$$

$$R_{ijk}^{(u)}(t;\boldsymbol{\theta}) := \frac{\partial^3 \log L_t^{(u)}(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j \partial\theta_k}$$

$$= \int_0^t \int_E \frac{\partial^3 \log \alpha(z;\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j \partial\theta_k} p^{(u)}(ds \times dz) - \frac{\partial^3 g(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j \partial\theta_k} Y_t^{(u)},$$

$$i,j,k = 1,\dots,d$$

for every $u \in \mathcal{U}$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ and $t \in [0,\infty)$. The maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}_u$ based on the realization of $p^{(u)}$ in the random time interval $[0,\tau_u]$ is defined as the solution of the likelihood equations

$$U_i^{(u)}(\tau_u;\boldsymbol{\theta}) = 0, \qquad i = 1,\dots,d. \tag{4.1}$$

THEOREM 4.1. *Suppose that Condition* 4.1 (A1)–(A3) *and* (B) *are satisfied. Then we have*:

(i) *Consistency of MLE*: *with $P_{\boldsymbol{\theta}_0}^{(u)}$-probability tending to* 1, *the likelihood equations* (4.1) *have exactly one consistent solution, and this solution provides a local maximum of the likelihood ratio $L_{\tau_u}^{(u)}(\boldsymbol{\theta})$.*

(ii) *Asymptotic normality of MLE*: *let $\hat{\boldsymbol{\theta}}_u$ be the consistent solution of the likelihood equations* (4.1), *then*

$$\sqrt{c_u}(\hat{\boldsymbol{\theta}}_u - \boldsymbol{\theta}_0) \to N(\mathbf{0}, (\eta\boldsymbol{\Sigma})^{-1}) \quad \textit{in } P_{\boldsymbol{\theta}_0}^{(u)}\textit{-law as } u \to \infty.$$

(iii) *LAN property of the model*: *with* $\boldsymbol{U}^{(u)}(\tau_u;\boldsymbol{\theta}_0) := (U_1^{(u)}(\tau_u;\boldsymbol{\theta}_0),\dots, U_d^{(u)}(\tau_u;\boldsymbol{\theta}_0))^\top$, *we have for every* $\boldsymbol{h} \in \boldsymbol{R}^d$

$$\log L_{\tau_u}^{(u)}(\boldsymbol{\theta}_u) - \frac{1}{\sqrt{c_u}}\boldsymbol{h}^\top \boldsymbol{U}^{(u)}(\tau_u;\boldsymbol{\theta}_0) + \frac{1}{2}\boldsymbol{h}^\top(\eta\boldsymbol{\Sigma})\boldsymbol{h} \to 0$$

$$\textit{in } P_{\boldsymbol{\theta}_0}^{(u)}\textit{-probability as } u \to \infty,$$

where $\boldsymbol{\theta}_u = \boldsymbol{\theta}_0 + \frac{1}{\sqrt{c_u}}\boldsymbol{h}$ *and*

$$\frac{1}{\sqrt{c_u}}\boldsymbol{U}^{(u)}(\tau_u;\boldsymbol{\theta}_0) \to N(\mathbf{0}, \eta\boldsymbol{\Sigma}) \quad \textit{in } P_{\boldsymbol{\theta}_0}^{(u)}\textit{-law as } u \to \infty.$$

PROOF. Although our model is not a special case of the one in Borgan (1984) or van Pul (1992), the conclusions follows from the same arguments as there. As pointed out by them, it suffices to show that

$$\frac{1}{\sqrt{c_u}}\boldsymbol{U}^{(u)}(\tau_u;\boldsymbol{\theta}_0) \to N(\mathbf{0}, \eta\boldsymbol{\Sigma}) \quad \text{in } P_{\boldsymbol{\theta}_0}^{(u)}\text{-law as } u \to \infty, \tag{4.2}$$

$$\frac{1}{c_u} I_{ij}^{(u)}(\tau_u;\boldsymbol{\theta}_0) \to -\eta\sigma_{ij} \quad \text{in } P_{\boldsymbol{\theta}_0}^{(u)}\text{-probability as } u \to \infty, \ i,j = 1,\dots,d \tag{4.3}$$

and

$$\lim_{u\to\infty} P^{(u)}_{\boldsymbol{\theta}_0}\left(\left|\frac{1}{c_u}R^{(u)}_{ijk}(\tau_u;\boldsymbol{\theta})\right| < M\right) = 1, \quad \boldsymbol{\theta}\in\boldsymbol{\Theta}_0,\ i,j,k=1,\ldots,d, \tag{4.4}$$

where M is a positive constant not depending on $\boldsymbol{\theta}\in\boldsymbol{\Theta}_0$.

First we prove (4.2). Noting $\int_E(\partial/\partial\theta_i)f(z;\boldsymbol{\theta}_0)\mu(dz)=0$, we can see that for every $i=1,\ldots,d$

$$\begin{aligned}U^{(u)}_i(t;\boldsymbol{\theta}_0) &= \int_0^t\int_E \frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}p^{(u)}(ds\times dz)\\ &\quad -\int_0^t\int_E\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}\alpha(z;\boldsymbol{\theta}_0)\mu(dz)y^{(u)}_s ds\\ &= \int_0^t\int_E\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}q^{(u)}(ds\times dz).\end{aligned}$$

Thus the condition (ii) in Theorem 2.1 is satisfied by (A2) and (B). Next, it is seen that

$$\begin{aligned}&\boldsymbol{E}^{(u)}_{\boldsymbol{\theta}_0}\Bigg[\int_0^{\tau_u}\int_E\sum_{i=1}^d\left|\frac{1}{\sqrt{c_u}}\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}\right|^2\\ &\qquad\cdot I\left\{\sum_{i=1}^d\left|\frac{1}{\sqrt{c_u}}\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}\right|^2>\varepsilon\right\}\lambda^{(u)}_s(dz;\boldsymbol{\theta}_0)ds\Bigg]\\ &=\int_E\sum_{i=1}^d\left|\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}\right|^2 I\left\{\sum_{i=1}^d\left|\frac{1}{\sqrt{c_u}}\frac{\partial\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i}\right|^2>\varepsilon\right\}\alpha(z;\boldsymbol{\theta}_0)\mu(dz)\\ &\quad\cdot\frac{1}{c_u}\boldsymbol{E}^{(u)}_{\boldsymbol{\theta}_0}[Y^{(u)}_{\tau_u}]\end{aligned}$$

which converges to 0 as $u\to\infty$ by the dominated convergence theorem and (B), so the condition (iii) is satisfied. Hence Theorem 2.1 is applicable and we obtain (4.2).

As to (4.3) and (4.4), we only give the key relations to prove them:

$$\begin{aligned}I^{(u)}_{ij}(t;\boldsymbol{\theta}_0) &= \int_0^t\int_E\frac{\partial^2\log\alpha(z;\boldsymbol{\theta}_0)}{\partial\theta_i\partial\theta_j}q^{(u)}(ds\times dz)-\sigma_{ij}Y^{(u)}_t, \qquad i,j=1,\ldots,d,\\ |R^{(u)}_{ijk}(t;\boldsymbol{\theta})| &\le \left|\int_0^t\int_E H(z)q^{(u)}(ds\times dz)\right|\\ &\quad+\left(\int_E H(z)\alpha(z;\boldsymbol{\theta}_0)\mu(dz)+\left|\frac{\partial^3 g(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|\right)Y^{(u)}_t,\\ &\qquad\qquad i,j,k=1,\ldots,d,\quad\boldsymbol{\theta}\in\boldsymbol{\Theta}_0.\ \square\end{aligned}$$

As consequences of the above results, we can see that a lot of statistical procedures based on the likelihood ratio have good operating characteristics. See e.g. Andersen *et al.* ((1993), Chapter VIII) for details.

5. Stopping rules

For the practical use of the results given in the previous section, it will be the most important problem to construct the stopping times $\{\tau_u; u \in \mathcal{U}\}$ satisfying Condition 4.1 (B). For this aim, we introduce the $(\mathcal{F}_t^{(u)})$-stopping time

$$\tag{5.1} \begin{aligned} \tau_u &= \inf\left\{t : \int_0^t \int_E \xi(z) p^{(u)}(ds \times dz) + vY_t^{(u)} > c_u\right\} \\ &= \inf\left\{t : \sum_{i=1}^{N_t^{(u)}} \xi(Z_i^{(u)}) + vY_t^{(u)} > c_u\right\}, \end{aligned}$$

where $\xi(z)$ is a real valued $\mathcal{E}$-measurable function and v is a prescribed constant. Various choice of $\xi(z)$ and v brings us a wide class of stopping rules. Here we give a sufficient condition for $P_{\boldsymbol{\theta}}^{(u)}(\tau_u < \infty) = 1$.

LEMMA 5.1. *For a fixed $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, assume that $\xi(Z)$ has the finite first moment $\phi_1(\boldsymbol{\theta}) = \int_E \xi(z) f(z; \boldsymbol{\theta}) \mu(dz)$ and that*

$$\tag{5.2} \psi(\boldsymbol{\theta}) := g(\boldsymbol{\theta})\phi_1(\boldsymbol{\theta}) + v > 0.$$

For each $u \in \mathcal{U}$, if $Y_t^{(u)} \uparrow \infty$ as $t \to \infty$, $P_{\boldsymbol{\theta}}^{(u)}$-a.s., then $P_{\boldsymbol{\theta}}^{(u)}(\tau_u < \infty) = 1$.

PROOF. Fix $u \in \mathcal{U}$. By the random time change theorem (see Kallenberg (1990), Theorem 2.1), the random double sequence $\{(Y_{T_i^{(u)}}^{(u)}, Z_i^{(u)}); i = 1, 2, \ldots\}$ forms the compound Poisson process which admits the intensity kernel $g(\boldsymbol{\theta}) f(z; \boldsymbol{\theta}) \mu(dz)$. Hence the random variables

$$\tag{5.3} W_i^{(u)} := \xi(Z_i^{(u)}) + v(Y_{T_i^{(u)}}^{(u)} - Y_{T_{i-1}^{(u)}}^{(u)}), \qquad i = 1, 2, \ldots,$$

where $T_0^{(u)} \equiv 0$, are I.I.D. with the first moment $\phi_1(\boldsymbol{\theta}) + \frac{v}{g(\boldsymbol{\theta})} > 0$. Therefore, it follows from the strong law of large numbers that

$$\sum_{i=1}^n \xi(Z_i^{(u)}) + vY_{T_n^{(u)}}^{(u)} = \sum_{i=1}^n W_i^{(u)} \to \infty \qquad \text{as } n \to \infty, \quad P_{\boldsymbol{\theta}}^{(u)}\text{-a.s.}$$

which completes the proof. □

The next result, concerning the properties of the random variables $\{Y_{\tau_u}^{(u)}; u \in \mathcal{U}\}$, plays the key role in our context. Note that the limiting constant and distribution in the latter conclusions (5.5) and (5.6) respectively, do not depend on the behavior of the processes $t \mapsto Y_t^{(u)}$.

THEOREM 5.1. *For a fixed $\boldsymbol{\theta} \in \Theta$, assume that $P_{\boldsymbol{\theta}}^{(u)}(\tau_u < \infty) = 1$ for every $u \in \mathcal{U}$. If $\xi(Z)$ has the finite first moment $\phi_1(\boldsymbol{\theta}) = \int_E \xi(z) f(z;\boldsymbol{\theta})\mu(dz)$ and if (5.2) is satisfied, then for every $u \in \mathcal{U}$,*

$$\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[Y_{\tau_u}^{(u)}] < \infty. \tag{5.4}$$

Further, if $\xi(Z)$ has the finite second moment $\phi_2(\boldsymbol{\theta}) = \int_E |\xi(z)|^2 f(z;\boldsymbol{\theta})\mu(dz)$, then

$$\lim_{u\to\infty} \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\frac{1}{c_u}Y_{\tau_u}^{(u)} - \frac{1}{\psi(\boldsymbol{\theta})}\right|\right] = 0 \tag{5.5}$$

and

$$\sqrt{c_u}\left(\frac{1}{c_u}Y_{\tau_u}^{(u)} - \frac{1}{\psi(\boldsymbol{\theta})}\right) \to N\left(0, \frac{g(\boldsymbol{\theta})\phi_2(\boldsymbol{\theta})}{\psi(\boldsymbol{\theta})^3}\right) \tag{5.6}$$

in $P_{\boldsymbol{\theta}}^{(u)}$-law as $u \to \infty$.

Remark. We do not require here that $Y_t^{(u)} \uparrow \infty$ as $t \to \infty$, $P_{\boldsymbol{\theta}}^{(u)}$-a.s. (see also Subsection 6.2), however, we may assume it without loss of generality. If not, it is enough for our purpose to introduce new E-marked point processes $\tilde{p}^{(u)}$ having the intensity kernels $g(\boldsymbol{\theta})\tilde{y}_t^{(u)} f(z;\boldsymbol{\theta})\mu(dz)$ where

$$\tilde{y}_t^{(u)} = \begin{cases} Y_t^{(u)}, & t \in [0,\tau_u], \\ 1, & t \in (\tau_u, \infty). \end{cases}$$

The aim of this modification is to reduce $\tilde{p}^{(u)}$ to the compound Poisson process as in the proof of Lemma 5.1. Hereafter, we use the original notations $p^{(u)}$ and $y_t^{(u)}$ instead of the modified $\tilde{p}^{(u)}$ and $\tilde{y}_t^{(u)}$ respectively.

PROOF. As seen in the proof of Lemma 5.1, the random variables $W_i^{(u)}$, $i = 1, 2, \ldots$ in (5.3) are I.I.D. with the finite positive first moment. For each $u \in \mathcal{U}$, we define

$$\mathcal{N}_u := \inf\left\{n : \sum_{i=1}^n \xi(Z_i^{(u)}) + vY_{T_n^{(u)}}^{(u)} > c_u\right\} = \inf\left\{n : \sum_{i=1}^n W_i^{(u)} > c_u\right\},$$

$$R_u := \sum_{i=1}^{\mathcal{N}_u} \xi(Z_i^{(u)}) + vY_{T_{\mathcal{N}_u}}^{(u)} - c_u = \sum_{i=1}^{\mathcal{N}_u} W_i^{(u)} - c_u$$

and

$$D_u := \int_0^{\tau_u}\int_E \xi(z)p^{(u)}(ds \times dz) + vY_{\tau_u}^{(u)} - c_u. \tag{5.7}$$

Then, we have $\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[\mathcal{N}_u] < \infty$ by Lemma 2.1 in Woodroofe (1982). So it follows from the Wald equation that

$$\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[R_u] = \left(\phi_1(\boldsymbol{\theta}) + \frac{v}{g(\boldsymbol{\theta})}\right) \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[\mathcal{N}_u] - c_u < \infty.$$

Since $0 \le D_u \le R_u$, $P_{\boldsymbol{\theta}}^{(u)}$-a.s., we obtain $\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[D_u] < \infty$. Hence the inequality

$$\begin{aligned}
&\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\int_0^{\tau_u \wedge S_n^{(u)}} \int_E \xi(z) p^{(u)}(ds \times dz) + vY^{(u)}_{\tau_u \wedge S_n^{(u)}}\right] \\
&\quad \le \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\int_0^{\tau_u} \int_E \xi(z) p^{(u)}(ds \times dz) + vY^{(u)}_{\tau_u}\right] = \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[c_u + D_u] < \infty
\end{aligned}$$

holds, where $\{S_n^{(u)}; n = 1, 2, \ldots\}$ is an appropriate sequence of localizing $(\mathcal{F}_t^{(u)})$-stopping times. Thus it follows from the optional sampling theorem that

$$\psi(\boldsymbol{\theta}) \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[Y^{(u)}_{\tau_u \wedge S_n^{(u)}}] \le c_u + \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[D_u] < \infty. \tag{5.8}$$

Noting that $0 \le Y^{(u)}_{\tau_u \wedge S_n^{(u)}} \uparrow Y^{(u)}_{\tau_u}$ as $n \to \infty$, $P_{\boldsymbol{\theta}}^{(u)}$-a.s., we obtain the first conclusion (5.4).

Next we prove (5.5). Since

$$\int_0^t \int_E \xi(z) p^{(u)}(ds \times dz) = \int_0^t \int_E \xi(z) q^{(u)}(ds \times dz) + g(\boldsymbol{\theta}) \phi_1(\boldsymbol{\theta}) Y_t^{(u)}$$

holds for every $t \in [0, \infty)$, we have from (5.7) that

$$Y^{(u)}_{\tau_u} = \frac{1}{\psi(\boldsymbol{\theta})} \left\{ c_u + D_u - \int_0^{\tau_u} \int_E \xi(z) q^{(u)}(ds \times dz) \right\}. \tag{5.9}$$

When the finite second moment $\phi_2(\boldsymbol{\theta})$ exists, it follows from Theorem 2.4 in Woodroofe (1982) that the family $\{R_u; u \in \mathcal{U}\}$ of random variables is uniformly integrable. Thus the family $\{D_u; u \in \mathcal{U}\}$ is so too. In particular, we have

$$\lim_{u \to \infty} \frac{1}{c_u} \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[D_u] = 0. \tag{5.10}$$

Therefore, combining (5.10) with (5.9), it suffices for (5.5) to show that

$$\lim_{u \to \infty} \frac{1}{c_u} \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\int_0^{\tau_u} \int_E \xi(z) q^{(u)}(ds \times dz)\right|\right] = 0. \tag{5.11}$$

Since $\langle \int_0^t \int_E \xi(z) q^{(u)}(ds \times dz) \rangle = g(\boldsymbol{\theta}) \phi_2(\boldsymbol{\theta}) Y_t^{(u)}$, we have

$$\begin{aligned}
\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\int_0^{\tau_u \wedge S_n^{(u)}} \int_E \xi(z) q^{(u)}(ds \times dz)\right|^2\right] &= g(\boldsymbol{\theta}) \phi_2(\boldsymbol{\theta}) \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[Y^{(u)}_{\tau_u \wedge S_n^{(u)}}] \\
&\le g(\boldsymbol{\theta}) \phi_2(\boldsymbol{\theta}) \cdot \frac{1}{\psi(\boldsymbol{\theta})} \{c_u + \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[D_u]\}
\end{aligned}$$

by the optional sampling theorem and (5.8). Thus it follows from the Fatou's lemma that

$$
\begin{aligned}
(5.12) \qquad & \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\int_0^{\tau_u}\int_E \xi(z)q^{(u)}(ds\times z)\right|^2\right] \\
&\leq \liminf_{n\to\infty} \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\int_0^{\tau_u\wedge S_n^{(u)}}\int_E \xi(z)q^{(u)}(ds\times dz)\right|^2\right] \\
&\leq g(\boldsymbol{\theta})\phi_2(\boldsymbol{\theta})\cdot\frac{1}{\psi(\boldsymbol{\theta})}\{c_u+\boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}[D_u]\}.
\end{aligned}
$$

Hence we obtain from (5.10) and (5.12) that

$$
\sup_{u\in\mathcal{U}'} \boldsymbol{E}_{\boldsymbol{\theta}}^{(u)}\left[\left|\frac{1}{\sqrt{c_u}}\int_0^{\tau_u}\int_E \xi(z)q^{(u)}(ds\times dz)\right|^2\right] < \infty,
$$

where $\mathcal{U}' := \{u \in \mathcal{U} : c_u \geq 1\}$. Therefore, the family $\{\frac{1}{\sqrt{c_u}}\int_0^{\tau_u}\int_E \xi(z)q^{(u)}(ds \times dz); u \in \mathcal{U}'\}$ of random variables is uniformly integrable (see e.g. Brémaud (1981), p. 286). This yields (5.11), hence the conclusion (5.5) is established.

Finally, we prove (5.6). From (5.9), we have

$$
\sqrt{c_u}\left(\frac{1}{c_u}Y_{\tau_u}^{(u)}-\frac{1}{\psi(\boldsymbol{\theta})}\right)=\frac{1}{\psi(\boldsymbol{\theta})}\left(\frac{1}{\sqrt{c_u}}D_u-\frac{1}{\sqrt{c_u}}\int_0^{\tau_u}\int_E \xi(z)q^{(u)}(ds\times dz)\right).
$$

Here, it easily follows from the uniform integrability of $\{D_u; u \in \mathcal{U}\}$ that

$$
\frac{1}{\sqrt{c_u}}D_u \to 0 \quad \text{in } P_{\boldsymbol{\theta}}^{(u)}\text{-probability as } u\to\infty,
$$

on the other hand, we can see

$$
\frac{1}{\sqrt{c_u}}\cdot\frac{1}{\psi(\boldsymbol{\theta})}\int_0^{\tau_u}\int_E \xi(z)q^{(u)}(ds\times dz) \to N\left(0,\frac{g(\boldsymbol{\theta})\phi_2(\boldsymbol{\theta})}{\psi(\boldsymbol{\theta})^3}\right) \quad \text{in } P_{\boldsymbol{\theta}}^{(u)}\text{-law as } u\to\infty
$$

from Theorem 2.1 with (5.5). These facts yield (5.6). □

6. Applications

6.1 *Compound Poisson process*

Here, we consider the *one realization case* with $\mathcal{U} = (0,\infty)$. When we set $y_t \equiv 1$ in (3.3), the process p forms the compound Poisson process with the intensity $g(\boldsymbol{\theta})$ and the compounding distribution $f(z;\boldsymbol{\theta})\mu(dz)$. Since $Y_{\tau_u} = \tau_u$ in this case, Condition 4.1 (B) is satisfied if we take $\tau_u \equiv u$ and $c_u = u$ for every $u \in \mathcal{U}$. However, in many practical situations, one may want to use other stopping rules for some reasons. For this purpose, the stopping time (5.1) which equals in the present case to

$$\tau_u = \inf\left\{ t : \sum_{i=1}^{N_t} \xi(Z_i) + vt > c_u \right\} \tag{6.1}$$

may be useful. Typically, it has the following interpretation:

(a) $\xi(Z)$ represents the damage by the event with its mark Z;

(b) v represents the running cost per unit time;

(c) we stop the experiment when the total sum of the damages and running cost exceeds the prescribed value c_u.

Stefanov (1982) has given a characterization of stopping rules which enable the *efficient* estimation of unknown parameters. Related with this problem, in view of Theorems 4.1 and 5.1, we assert that the *asymptotically efficient* inference is possible for the more general stopping rules (6.1).

6.2 *Pure death (birth) process*

Here, we consider the *general case* with $\mathcal{U} = \{1, 2, \ldots\}$. Let $E = \{-1\}$, hence the mark distribution in (3.2) is degenerated. We suppose that the (univariate) point process $p^{(u)}$ admits the $(P_\theta, (\mathcal{F}_t^{(u)}))$-predictable intensity

$$\lambda_t^{(u)}(\theta) = \theta y_t^{(u)}, \qquad \theta > 0, \tag{6.2}$$

where $y_t^{(u)} = u - N_{t-}^{(u)}$. Then the process $p^{(u)}$ forms the linear pure death process with the initial population size u. In this case, it does not hold that $Y_t^{(u)} \uparrow \infty$ as $t \to \infty$, $P_\theta^{(u)}$-a.s., however, all the assumptions in Theorem 5.1 are satisfied if we take e.g. $\xi \equiv 1$, $v \geq 0$ and $c_u = u$ in (5.1). This fact illustrates that the essential requirement is not "$Y_t^{(u)} \uparrow \infty$ as $t \to \infty$" but "$Y_{\tau_u}^{(u)} \uparrow \infty$ as $u \to \infty$".

When we set $E = \{+1\}$ and $y_t^{(u)} = u + N_{t-}^{(u)}$ in (6.2), the process $p^{(u)}$ forms the linear pure birth process with the initial population size u. More general processes, including the birth-and-death processes, are discussed in the next subsection.

6.3 *Branching process with immigration*

Here, we consider the *one realization case* with $\mathcal{U} = (0, \infty)$. Let $E = \{0, 1, 2, \ldots\}$ and $\mu(dz)$ be the counting measure on E. We suppose that the process p admits the $(P_{\boldsymbol{\theta}}, (\mathcal{F}_t))$-intensity kernel (3.3) with

$$y_t = X_{t-} + \sum_{i=1}^{N_{t-}} (Z_i - 1),$$

where $t \mapsto X_t$ is a right-continuous, positive integer valued non-decreasing process. The process $t \mapsto X_t$ may be either random or non-random, however, note that it is assumed to be $(\mathcal{F}_t)$-adapted where the filtration $(\mathcal{F}_t)$ is of the form (3.1). Then, the process p forms the continuous time Markov branching process with the split intensity $g(\boldsymbol{\theta})$, the offspring distribution $\{f(z; \boldsymbol{\theta}), z = 0, 1, 2, \ldots\}$ and the immigration process $t \mapsto X_t$. It is a common knowledge that we can interpret X_0 as the initial population size of p, and that $X_t - X_0$ represents the number of immigrations in the time interval $(0, t]$. We must require $Y_t \uparrow \infty$ as $t \to \infty$, $P_{\boldsymbol{\theta}}$-a.s. for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. It is sufficient for this aim that we assume $X_t \uparrow \infty$ as $t \to \infty$ $P_{\boldsymbol{\theta}}$-a.s. for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, but it is not necessary; for example, $X_t = X_0 + \sum_{i=1}^{N_t} I\{y_{T_i} = 1, Z_i = 0\}$ also will do.

When we adopt inference procedures based on the realization of p in the non-random time intervals $[0, u]$, the results are drastically different for the subcritical, critical and supercritical cases. Concerning this difficulty, several authors have assumed in advance either subcriticality or supercriticality. However, this assumption is not desirable in some situations. Indeed, testing of criticality itself is an important problem. Basawa and Becker (1983) have discussed it by a sequential approach; they have used the stopping time $\tau_u = \inf\{t : N_t = u\}$ where u is a positive integer, and obtained the *uniformly most powerful* test for criticality. However, when we use more general stopping times (5.1), the model has the LAN property and the usual procedure for constructing the *asymptotically uniformly most powerful* tests works well. See e.g. Theorem VIII.1.4 of Andersen *et al.* (1993).

When we restrict our attention to the supercritical case, the model forms a LAMN family if the process p is observed in the non-random time intervals $[0, u]$. Thus, the maximum likelihood estimators of $\boldsymbol{\theta}$ are asymptotically *mixed* normal, and the normalizing constants depend on the true value of $\boldsymbol{\theta}$ (see e.g. Athreya and Keiding (1977)). This phenomenon may be unpleasant in some practical situations. However, it does not appear in the sequential approach based on the stopping times (5.1). Further, we are able to consider the model for the subcritical, critical and supercritical cases in the unified context. This merit is important for the estimation problem of the Malthusian parameter $\gamma(\boldsymbol{\theta}) = g(\boldsymbol{\theta})(\sum_{z=0}^{\infty} z f(z; \boldsymbol{\theta}) - 1)$.

Finally, it should be noted that other specifications of y_t are possible; for example, $y_t = \bar{y}(X_{t-} + \sum_{i=1}^{N_{t-}} (Z_i - 1))$ where

$$\bar{y}(x) = x^{\alpha}, \quad xe^{-\beta x} \quad \text{or} \quad x \wedge k \quad \text{etc.} \tag{6.3}$$

The last one in (6.3) with $E = \{0\}$ makes p the $G/M/k$ queueing process. Also, we can treat these processes in the *general case* as in Example 3.1 and Subsection 6.2.

Acknowledgements

The author would like to thank Prof. N. Inagaki and Dr. T. Hayashi for their valuable suggestions and comments.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer, New York.

Athreya, K. B. and Keiding, N. (1977). Estimation theory for continuous-time branching processes, *Sankhyā Ser. A*, **39**, 101–123.

Basawa, I. V. and Becker, N. (1983). Remarks on optimal inference for Markov branching processes: a sequential approach, *Austral. J. Statist.*, **25**, 35–46.

Basawa, I. V. and Scott, D. J. (1983). Asymptotic optimal inference for non-ergodic models, *Lecture Notes in Statist.*, **17**, Springer, New York.

Borgan, Ø. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data, *Scand. J. Statist.*, **11**, 1–16 (Correction: *ibid.* (1984), **11**, p. 275).

Brémaud, P. (1981). *Point Processes and Queues*, Springer, New York.

Franz, J. (1982). Sequential estimation and asymptotic properties in birth-and-death processes, *Mathematische Operationsforschung und Statistik*, *Statistics*, **13**, 231–244.

Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*, Springer, Berlin.

Kallenberg, O. (1990). Random time change and an integral representation for marked stopping times, *Probab. Theory Related Fields*, **86**, 167–202.

Keiding, N. (1975). Maximum likelihood estimation in the birth-and-death process, *Ann. Statist.*, **3**, 363–372.

Kutoyants, Yu. A. (1984). *Parameter Estimation for Stochastic Processes*, Heldermann, Berlin.

Manjunath, S. M. (1984). Optimal sequential estimation for ergodic birth-death processes, *J. Roy. Statist. Soc. Ser. B*, **46**, 412–418.

Mel'nikov A. V. and Novikov, A. A. (1989). Sequential inferences with prescribed accuracy for semimartingales, *Theory Probab. Appl.*, **33**, 446–459.

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

Stefanov, V. T. (1982). Sequential estimation for compound Poisson process I, II, *Serdica*, **8**, 183–189, 255–261.

Svensson, Å. (1990). Asymptotic estimation in counting processes with parametric intensities based on one realization, *Scand. J. Statist.*, **17**, 23–33.

van Pul, M. C. (1992). Asymptotic properties of a class of statistical models in software reliability, *Scand. J. Statist.*, **19**, 235–253.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, SIAM, Philadelphia.