AD-A089 117 INDIANA UNIV AT BLOOMINGTON DEPT OF MATHEMATICS F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION FOR STATIONARY POINT PROCESSES.(U)
AUG 80 P TUAN, M L PURI AFOSR-76-2927

UNCLASSIFIED AFOSR-TR-80-0688 NL

1 OF 1
AD
A089 117

END
DATE
FILMED
10-80
DTIC

AD A089117

(18) AFOSR-TR-80-0688 (19) LEVEL II (4)

(6)

MAXIMUM LIKELIHOOD ESTIMATION FOR STATIONARY POINT PROCESSES.

(11) Aug [redacted] 80 (9)

By (10) [redacted] Pham-Dinh Tuan Madan L. Puri

Université de Grenoble, France and Indiana University, Bloomington

(12) 21

(15) AFOSR-76-2927

DTIC ELECTE SEP 4 1980 S C

Summary. In this paper we derive the log likelihood function for point processes in terms of their stochastic intensities, using the martingale approach. For practical purposes we work with an approximate log likelihood function which is shown to possess the usual asymptotic properties of a log likelihood function. The resulting estimates are strongly consistent and asymptotically normal (under some regularity conditions). As a by-product, a strong law of large numbers and a central limit theorem for continuous martingale are derived.

[1]Work supported by the Air Force Office of Scientific Research, AFSC, USAF under Grant No. AFOSR 76-2927. Reproduction in whole or in part permitted for any purpose of the United States Government.

Key words and phrases. Maximum likelihood, compensator, stochastic intensity, martingale, natural increasing process, point process, predictable process, likelihood function, stochastic integral.

DDC FILE COPY

402522 80 9 2 221

Approved for public release; distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-0688	2. GOVT ACCESSION NO. AD-A089 117	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Maximum Likelihood Estimation for Stationary Point Processes		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Madan L. Puri and Pham-dinh Tuan		8. CONTRACT OR GRANT NUMBER(s) AFOSR 76-2927
9. PERFORMING ORGANIZATION NAME AND ADDRESS Indiana University Department of Mathematics Bloomington, Indiana 47405		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332		12. REPORT DATE August, 1980
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Maximum likelihood; compensator, stochastic intensity, martingale, natural increasing process, point process, predictable process, likelihood function, stochastic integral.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we derive the log likelihood function for point processes in terms of their stochastic intensities, using the martingale approach. For practical purposes we work with an approximate log likelihood function which is shown to possess the usual asymptotic properti of a log likelihood function. The resulting estimates are strongly consistent and asymptotically normal (under some regularity conditions). As a by-product, a

(over)

DD FORM 1 JAN 73 1473

UNCLASSIFIED

20. (cont'd.)
strong law of large numbers and a central limit theorem for continuous martingale are derived.

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
A

1. Introduction. The maximum likelihood estimates for point process model have been occasionally used especially for Poisson model (see e.g. Snyder (1975)). The log likelihood function has also been used in some cases in a somewhat heuristic manner (see e.g. Snyder (1975), Vere-Jone(1975)). Although it is expected that the maximum likelihood estimates possess the usual asymptotic properties, to the best of our knowledge, there has not been a rigorous proof for it. In this paper, we use some of the ideas of Jacod (1975) and Lipster and Shiryayev (1978) on martingale theory for point processes to derive the log likelihood function, and we prove under some regularity conditions the asymptotic properties of the log likelihood and the maximum likelihood estimators. The case of self exciting processes is of great interest since the log likelihood can be written down easily and hence is treated in greater detail.

2. The Log Likelihood Function for Point Process.

We shall be concerned with multitype point processes, that is, point events of r different types randomly occurring along the real line. This process can be described by a multivariate counting process $N(t) = \{N_1(t),\ldots,N_r(t)\}$, $t \in R$, defined on some probability space (Ω,A,P). Here $N_j(t) - N_j(s)$, $t > s$ denotes the number of point events of type j which occur in $(s,t]$. By convention $N_j(0) = 0$. We shall suppose that at each t , at most one event regardless of its type can occur. Let A_t , $t \geq 0$ be an increasing family of sub-σ fields such that $N(t)$ is A_t-measurable, $t > 0$. Then (Lipster and Shiryayev (1977),

p. 239) there is for each $j = 1,\ldots,r$ a natural increasing process $\Lambda_j(t)$, called the compensator of $N_j(t)$, relative to (A_t,P) such that $m_j(t) = N_j(t) - \Lambda_j(t)$ $t \geq 0$ is a A_t-local martingale. Here we shall be interested only in the case where the measure $d\Lambda_j(t)dP$ admits a density $\lambda_j(t)$ with respect to $dtdP$. The process $\lambda_j(t)$ can be chosen to be A_t-predictable, that is $(t,\omega) \to \lambda_j(t,\omega)$ is measurable with respect to the σ-field generated by all the A_t-adapted processes with left continuous sample paths, and is called the stochastic intensity of the $N(t)$ process. Intuitively $\lambda_j(t)$ can be interpreted as

$$\lim_{\Delta t \downarrow 0} P\{\Delta N_j(t) = 1 | A_t\} / \Delta t \tag{2.1}$$

where $\Delta N(t) = N(t + \Delta t) - N(t)$. Indeed if τ_n^j is the time of occurrence of the n-th event of type j after the origin, then $m_j(t \wedge \tau_n^j)$, $t \geq 0$ is a martingale and so for $s > t$, $A \in A_t$,

$$\begin{aligned} E[1_A\{N_j(s\wedge\tau_n^j) - N_j(t\wedge\tau_n^j)\}] &= E\{1_A\int_{t\wedge\tau_n^j}^{s\wedge\tau_n^j} d\Lambda_j(u)\} \\ &= \int_t^s E[1_A\, 1_{\{u \leq \tau_n^j\}}\lambda_j(u)]du . \end{aligned} \tag{2.2}$$

Heuristically if $s = t + \Delta t$ and Δt is small, then the left hand side of (2.2) is approximately $P(A \cap \{t \leq \tau_n^j\} \cap \{\Delta N(t) = 1\})$ and the right hand side of (2.2) is approximately $E[1_A\, 1_{\{t \leq \tau_n^j\}}\lambda(t)]\Delta t$. Since $\{\tau_n^j < t\} \in A_t$, we obtain the interpretation (2.1). Historically (2.1) was proposed as the definition of the stochastic intensity. This definition requires the existence of the limit in (2.1) and is equivalent to our

definition only under some regularity conditions.

If $A_t = \tilde{A}_t$, the sub σ-field generated by $N(s)$, $0 \le s \le t$, then the stochastic intensity, which we write now $\tilde{\lambda}_j(t)$ completely defines the probability distribution of the process (Jacod (1975), Lipster and Shiryayev (1977), p. 252). Now let π be the probability such that the $N_j(t)$, relative to π , are independent Poisson processes with unit rate. Thus, relative to π , the stochastic intensity of $N_j(t)$ is one. Observe that the random measure $\{j\} \times (s,t) \to \int_s^t \tilde{\lambda}_j(u)du$ on $\{1,\ldots,r\} \times R^+$ is precisely the dual predictable projection in the sense of Jacod (1975) of the random measure $\{j\} \times (s,t) \to N_j(t) - N_j(s)$, by the result of this paper, P is absolutely continuous with respect to π on any $\tilde{A}_T$, $T > 0$, with density

$$\frac{dP}{d\pi} = \exp\left[\int_0^T \sum_{j=1}^{r} \{\mathrm{Log}\ \tilde{\lambda}_j(t)\ dN(t) - \tilde{\lambda}_j(t)\ dt + dt\}\right] .$$

Let now $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions on (Ω,A) and let $\tilde{\lambda}_{\theta,j}(t)$ be the corresponding stochastic intensity of the $N_j(t)$ process. Then the above result shows that the log likelihood function corresponding to an interval of observation $[0,T]$ is, up to an additive constant:

$$(2.3) \qquad \tilde{L}_T(\theta) = \int_0^T \sum_{j=1}^{r} \{\mathrm{Log}\ \tilde{\lambda}_{\theta,j}(t)\ dN_j(t) - \tilde{\lambda}_{\theta,j}(t)\ dt\} .$$

<u>Remark</u>

The multitype process is a special case of the marked point process when the space of the marks is just $\{1,\ldots,r\}$. There is no difficulty to write down the log likelihood function for marked point process. Also, it is not necessary to suppose the existence of the stochastic intensity to write down the log likelihood func-

tion; it suffices that the measure $d\tilde{\Lambda}_{\theta,j}(t)\, dP_\theta$ be absolutely continuous with respect to some measure $d\nu(t)\, dP_\theta$, $\tilde{\Lambda}_{\theta,j}(t)$ being the natural increasing process of $N_j(t)$ relative to $\tilde{A}_t$ and to P_θ. The special case we considered is convenient for further developments concerning the asymptotic properties of the log likelihood function.

3. The Approximate Log Likelihood Function

The considered process is defined on the whole line although only an observation on $[0,T]$ is available. Denote by $\lambda_{\theta,j}(t)$ the stochastic intensity of the $N_j(t)$ process relative to P_θ and the sub σ-fields A_t, $t \in R$ generated by $N(s)$, $s \le t$. Models of point processes are usually described in terms of $\lambda_\theta(t) = (\lambda_{\theta,1}(t),\ldots,\lambda_{\theta,r}(t))$. For example, the self exciting process introduced by Hawkes (1972) can be defined by

$$\lambda_\theta(t) = \alpha_\theta + \int_{-\infty}^{t} g_\theta(t-s)\, dN(s) \tag{3.1}$$

where α_θ is a constant vector and $g_\theta(.)$ is some appropriate matrix function. Thus, it is desirable to obtain the log likelihood function in terms of $\lambda_\theta(t)$. We are led to the problem of computing $\tilde{\lambda}_\theta(t)$ in terms of $\lambda_{\theta,j}(t)$. Now, from the interpretation (2.1) of $\lambda_{\theta,j}(t)$ one can expect that $\tilde{\lambda}_{\theta,j}(t) = E_\theta\{\lambda_{\theta,j}(t)\,|\,A_t\}$ in case when $\lambda_{\theta,j}(t)$ is integrable. The rigorous result is

<u>Theorem 1.</u> <u>Let</u> τ_n^j <u>be the time of occurrence of the</u> n-<u>th event of type</u> j <u>after the origin.</u> <u>Then for almost all</u> t, <u>the random</u>

<u>variable</u> $1_{\{t \le \tau_n^j\}} \lambda_{\theta,j}(t)$ <u>is</u> P_θ-<u>integrable and</u>

$$1_{\{t \le \tau_n^j\}} \tilde{\lambda}_{\theta,j}(t) = E_\theta[1_{\{t \le \tau_n^j\}} \lambda_{\theta,j}(t) | \tilde{A}_t]$$

<u>almost surely</u>.

<u>Proof</u>.

From (2.2), we have for all $s > 0$

$$n \ge E_\theta\{N_j(s \wedge \tau_n^j)\} = \int_0^s E_\theta[1_{\{t \le \tau_n^j\}} \lambda_{\theta,j}(t)]dt$$

and hence the function

$$(t,\omega) \to X_n(t,\omega) = 1_{\{t \le \tau_n^j\}}(\omega)\ \lambda_{\theta,j}(\omega)$$

belongs to $L^1(R^+ \times \Omega, B_{R^+} \otimes A, dt\, dP_\theta)$. We shall show that there exists a A_t-predictable process $\tilde{X}_n(t)$ with $\tilde{X}_n(t) = E\{X_n(t)|\tilde{A}_t\}$ for almost all t , almost surely. Indeed, there exists a sequence of simple functions of the form

$$X_{n,k}(t,\omega) = \sum_{m=1}^{m_k} Z_m^{(k)}(\omega)\ 1_{(t_{m-1}^{(k)}, t_m^{(k)}]}(t) \qquad k = 1, 2, \ldots$$

which converges in L^1 to X_n . For $t_{m-1}^{(k)} < t \le t_m^{(k)}$, set $\tilde{X}_{n,k}(t)$ to be a version of $E_\theta\{Z_m^{(k)}|\tilde{A}_t\}$, such that the process $\tilde{X}_{n,k}(t)$ has left continuous sample path and is A_t-adapted, which is possible because of the martingale property of the $E_\theta\{Z_m^{(k)}|\ \tilde{A}_t\}$, $t_{m-1}^{(k)} < t \le t_m^{(k)}$. Hence $\tilde{X}_{n,k}(t)$ is $\tilde{A}_t$-predictable.

From

$$|\tilde{X}_{n,k}(t) - E_\theta\{X_n(t)|A_t\}| = |E_\theta[\{X_{n,k}(t) - X_n(t)\}|\tilde{A}_t]|$$

$$\le E_\theta\{|X_{n,k}(t) - X_n(t)||\tilde{A}_t\}$$

and the fact that $X_{n,k} \to X_n$ in L^1 as $k \to \infty$, we get that the sequence $\tilde{X}_{n,k}$, $k \geq 1$ is a Cauchy sequence in L^1 and hence converges to some $\tilde{A}_t$-predictable process $\tilde{X}_n$ which equals $E_\theta\{X_n(t)|\tilde{A}_t\}$ almost surely for almost all t.

Now, from the definition of $X_n(t)$ and the fact that $\{t \leq \tau_n^j\} \in \tilde{A}_t$, we have, almost surely

$$1_{\{t \leq \tau_n^j\}} E_\theta\{X_m(t) \mid \tilde{A}_t\} = E_\theta\{X_n(t)|\tilde{A}_t\} \quad \text{for} \quad m \geq n$$

and hence there is a $\tilde{A}_t$-predictable process $\tilde{X}(t)$ such that

$$1_{\{\tau_n^j > t\}} \tilde{X}(t) = \tilde{X}_n(t) = E_\theta[1_{\{\tau_n^j > t\}}\lambda_{\theta,j}(t) \mid \tilde{A}_t]$$

for almost all t, almost surely.

We now show that $\tilde{X}(t)$ is $\tilde{\lambda}_{\theta,j}(t)$. For this let $A \in \tilde{A}_t$, then the right hand side of (2.2) is equal to

$$\begin{aligned} E_\theta[\int_t^s 1_A X_n(u)\, du] &= \int_t^s E_\theta[1_A\, E_\theta\{X_n(u) \mid \tilde{A}_u\}]du \\ &= \int_t^s E_\theta\{1_A \tilde{X}_n(u)\}\, du \\ &= E_\theta\{1_A \int_t^{s\wedge\tau_n^j} \tilde{X}(u)\, du\}\,. \end{aligned}$$

Hence, by (2.2) the process $N(t) - \int_0^t \tilde{X}(u)\, du$, $t \geq 0$ is a local martingale, which gives the desired result.

<u>Corollary</u>. <u>Suppose that</u> $\lambda_{\theta,j}(t)$ <u>is integrable for almost all</u> t. <u>Then</u> $\tilde{\lambda}_{\theta,j}(t) = E_\theta\{\lambda_{\theta,j}(t)|\tilde{A}_t\}$ <u>almost surely, for almost all</u> t.

Although the above result provides a means of computing $\tilde{\lambda}_\theta(t)$

in terms of $\lambda_\theta(t)$, the actual computation is not easy. So we are led to approximate $\lambda_\theta(t)$ by some $\hat{\lambda}_\theta(t)$ which depend only on $N(s)$, $0 \le s \le t$. The approximate log likelihood function is then

$$(3.2) \quad L_T(\theta) = \int_0^T \{ \sum_{j=1}^{r} \text{Log } \hat{\lambda}_{\theta,j}(t)\, dN_j(t) - \hat{\lambda}_{\theta,j}(t)\, dt\}$$

Since $T^{-1} \tilde{L}_T(\theta)$ depends essentially on the values of $\tilde{\lambda}_\theta(t)$ for large t if T is large, we would expect that L_T is a good approximation to $\tilde{L}_T$ for large T if $\hat{\lambda}_\theta(t)$ is a good approximation to $\tilde{\lambda}_\theta(t)$ for large t . But by the corollary of Theorem 1 and the stationarity of $N(t)$, $E\| \tilde{\lambda}_\theta(t) - \lambda_\theta(t)\| \to 0$ as $t \to \infty$. Therefore one could expect that L_T is good approximation to $\tilde{L}_T$ for large T if $\hat{\lambda}_\theta(t)$ is a good approximation to $\lambda_\theta(t)$ for large t . We will make our assumptions on $\hat{\lambda}_\theta(t)$ precise later on.

As example, consider the self exciting process when $\lambda_\theta(t)$ is given by (3.1). In order that $\lambda_{\theta,j}(t) \ge 0$ for all t , we shall assume that $\alpha_{\theta,j\ell} \ge 0$, $g_{\theta,j\ell}(t) \ge 0$ for all t . If the $g_{\theta,j\ell}$ are integrable and $N(t)$ is of stationary increments with $EN_j(1) = \mu_{\theta,j} < \infty$, then the integral in (3.1) exists almost surely. Indeed

$$E_\theta \{\int_{-\infty}^{t} g_{\theta,j\ell}(t - s)\, dN_\ell(s)\} = \int_{-\infty}^{t} g_{\theta,j\ell}(t - s)\mu_{\theta,\ell} ds$$
$$\le \mu_{\theta,\ell} \int_{-\infty}^{\infty} g_{\theta,j\ell}(t)\, dt < +\infty$$

and hence the integral in the above left hand side is finite almost surely. As an approximation to $\lambda_\theta(t)$, one might consider

$$(3.3) \qquad \hat{\lambda}_\theta(t) = \alpha_\theta + \int_0^t g_\theta(t - s)\, dN(s)$$

which is evidently $\tilde{A}_t$-measurable and would be a good approximation to $\lambda_\theta(t)$ for large t if $g_\theta(t) \to 0$ sufficiently fast.

4. Asymptotic Properties of the Approximate Log Likelihood

From now on, we shall suppose that the $N(t)$ process is of stationary increments and metrically transitive in the sense of Doob (1953, p. 510). Let $\tilde{\Omega}$ be the space of $\tilde{\omega} = (\tilde{\omega}_1, \ldots, \tilde{\omega}_r)$ where the $\tilde{\omega}_i$ are non-decreasing integral valued functions on $(-\infty, \infty)$, $\tilde{A}$ the σ-field generated by the projections $\pi_{st} : \tilde{\omega} \to \tilde{\omega}(t) - \tilde{\omega}(s)$, $s < t$, $\tilde{P}_\theta$ the restriction to $\tilde{A}$ of the image of P_θ by the application $N : \omega \to N(.)$, and T_h the shift operator $(T_h\tilde{\omega})(t) = \tilde{\omega}(t + h)$. Then stationarity means that T_h conserves the probability $\tilde{P}_\theta$, that is $\tilde{P}_\theta(T_h^{-1}A) = \tilde{P}(A)$, $\forall A \in \tilde{A}$, and metric transitivity means that the invariant sets, that is, those sets $A \in \tilde{A}$ for which $T_h^{-1}A = A$, a.s., have probabilities zero or one. Now, from the fact that $N(t) - N(s) = \pi_{st} \circ N$ and that T_h conserves the probability $\tilde{P}_\theta$, one can show that $\lambda_\theta(t) = \bar{\lambda}_\theta(t) \circ N$ with $\bar{\lambda}_\theta(t + h) = \bar{\lambda}_\theta(t) \circ T_h$. Hence if $\lambda_\theta(t)$ is integrable

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \lambda_\theta(t)\, dt = E\{\lambda_\theta(0)\} ,$$

almost surely. (see Doob, 1953, p. 515). Here the expectation is computed with respect to the true probability

We suppose in the sequel that $\Theta \subset R^k$ and $\lambda_\theta(t)$, $\hat{\lambda}_\theta(t)$

are twice continuously differentiable with respect to θ, almost surely. We shall use the notation $x_\theta^{(1)}$, $x_\theta^{(2)}$ to denote the vector and the matrix of first and second derivatives of x_θ with respect to θ. Let $\psi(t)$ be a function of the $\lambda_{\theta,j}(t)$, $\lambda_{\theta,j}^{(i)}(t)$, $i = 1,2$ such that $\psi(t)$ is integrable. Then by the same argument as above

$$\frac{1}{T}\int_0^T \psi(t)dt \to E\psi(0) \text{ , a.s. } (T \to \infty) \tag{4.1}$$

In this section, we are interested in the limiting behavior of $L_T(\theta)$, $L_T^{(i)}(\theta^*)$, $i = 1,2$ as $T \to \infty$, θ^* being the true value of θ. We have, omitting the subscript θ when $\theta = \theta^*$ and putting $\hat{\phi}_{\theta,j}(t) = \text{Log } \hat{\lambda}_{\theta,j}(t)$, $dm_j = dN_j(t) - \lambda_j(t)dt$:

$$\begin{aligned} L_T(\theta) &= \sum_{j=1}^{r} \{\int_0^T \hat{\phi}_{\theta,j}(t)dN_j(t) - \int_0^T \hat{\lambda}_{\theta,j}(t)dt\} \\ &= \sum_{j=1}^{r} [\int_0^T \hat{\phi}_{\theta,j}(t)dm_j(t) \\ &\quad + \int_0^T \{\hat{\phi}_{\theta,j}(t)\lambda_j(t) - \hat{\lambda}_j(t)\}dt] \end{aligned} \tag{4.2}$$

$$\begin{aligned} L_T^{(1)}(\theta) &= \sum_{j=1}^{r} \{\int_0^T \hat{\phi}_{\theta,j}^{(1)}(t)dN_j(t) - \int_0^T \hat{\lambda}_{\theta,j}^{(1)}(t)dt\} \\ &= \sum_{j=1}^{r} \int_0^T \hat{\phi}_{\theta,j}^{(1)}(t)\{\lambda_j(t) - \hat{\lambda}_{\theta,j}(t)\}dt \\ &\quad + \int_0^T \hat{\phi}_{\theta,j}^{(1)}(t)dm_j(t)] \end{aligned} \tag{4.3}$$

$$(4.4) \qquad L_T^{(2)}(\theta) = \sum_{j=1}^{r} \{ \int_0^T \hat{\phi}_{\theta,j}^{(2)}(t)dN_j(t) - \int_0^T \hat{\lambda}_{\theta,j}^{(2)}(t)dt\}$$

$$= \sum_{j=1}^{r} [\int_0^T \hat{\phi}_{\theta,j}^{(2)}(t)\{\lambda_j(t) - \hat{\lambda}_{\theta,j}(t)\}dt$$

$$+ \int_0^T \hat{\phi}_{\theta,j}^{(2)}(t)dm_j(t)$$

$$- \int_0^T \hat{\phi}_{\theta,j}^{(1)}(t)\hat{\phi}_{\theta,j}^{(1)}(t)'\hat{\phi}_{\theta,j}(t)dt]$$

In order to use the result (4.1) we are led to replace $\hat{\lambda}_{\theta,j}(t)$, $\hat{\phi}_{\theta,j}(t),\ldots$ by $\lambda_{\theta,j}(t)$, $\phi_{\theta,j}(t) = \log \lambda_{\theta,j}(t),\ldots$. The following result, which can be easily proved, is quite useful.

<u>Lemma 1</u>. <u>Let</u> ε_t, x_t <u>be such that</u> $\varepsilon_t/x_t \to 0$ <u>as</u> $t \to \infty$. <u>Also let</u> $\int_0^T |\varepsilon_t| d\mu(t) < \infty$ <u>and</u> $\mu(T)^{-1} \int_0^T |x_t| d\mu(t)$ <u>be bounded where</u> μ <u>is a non-decreasing function with</u> $\mu(\infty) = \infty$. <u>Then</u> $\mu(T)^{-1} \int_0^T \varepsilon_t \, d\mu(t) \to 0$ <u>as</u> $T \to \infty$.

To obtain convergence results for the stochastic integrals with respect to $dm_j(t)$ in (4.2) - (4.4), we will need the following result which is of independent interest.

<u>Lemma 2</u>. <u>Let</u> M_t , $t \geq 0$ <u>be a locally square integrable martingale with continuous natural increasing process</u> $\langle M\rangle_t$. <u>Let</u> g_t , $t > 0$ <u>be a non-decreasing left continuous function with</u> $g_\infty = \infty$, <u>such that</u> $\langle M\rangle_t = O(g_t^{2-\varepsilon})$, $\varepsilon > 0$, <u>almost surely as</u> $t \to \infty$. <u>Then</u> $M_T/g_T \to 0$ <u>almost surely as</u> $T \to \infty$.

Proof

Let $a > 0$ be such that $g_a > 0$. Let c be an arbitrary positive constant. Define the stopping time

$$\tau = \operatorname{Inf}\{t : \int_a^t g_s^{-2}\, d\langle M\rangle_s \geq c\} .$$

Set $\tilde{M}_t = M_{t\wedge\tau}$, and

$$Z_t = \int_a^t g_s^{-1}\, d\tilde{M}_s = \int_a^{t\wedge\tau} g_s^{-1}\, dM_s .$$

The process Z_t, $t \geq a$ is a martingale with natural increasing process

$$\langle Z\rangle_t = \int_a^{t\wedge\tau} g_s^{-2}\, d\langle M\rangle_s \leq c , \quad t \geq a .$$

Since $EZ_t^2 \leq c$, $t \geq a$, we know (Doob, 1953, p. 354, 361) that almost surely Z_{a+}, Z_{t-}, Z_{t+}, $t \geq a$, Z_∞ exist and Z_t is bounded on any finite interval. Set $t_i^{(k)} = a + (T - a)\, i/k$ and write

$$g_T Z_T = \sum_{i=1}^{k} [\{g_{t_i^{(k)}} - g_{t_{i-1}^{(k)}}\} Z_{t_i^{(k)}} + g_{t_{i-1}^{(k)}} \{Z_{t_i^{(k)}} - Z_{t_{i-1}^{(k)}}\}]$$

$$= \int_0^T Z_t^{(k)}\, dg_t + \int_a^T g_t^{(k)}\, dZ_t$$

where $Z_t^{(k)} = Z_{t_i^{(k)}}$, $g_t^{(k)} = g_{t_{i-1}^{(k)}}$ for $t_{i-1}^{(k)} \leq t < t_i^{(k)}$.

By the Lebesque dominated convergence Theorem and the property of stochastic integral, almost surely as $k \to \infty$

$$\int_a^{T-} Z_t^{(k)}\, dg_t \to \int_a^{T-} Z_{t+}\, dg_t$$

and

$$\int_a^T g_t^{(k)}\, dZ_t \to \int_a^T g_t\, dZ_t = \int_a^T d\tilde{M}_t = \tilde{M}_T - \tilde{M}_a$$

Therefore

$$g_T^{-1}(\tilde{M}_T - \tilde{M}_a) = Z_T - g_T^{-1} \int_a^{T-} Z_{t+}\, dg_t$$

$$= Z_T - (1 - g_a g_T^{-1})\, Z_\infty - g_T^{-1} \int_a^{T-} (Z_{t+} - Z_\infty)\, dg_t$$

Since $g_T \uparrow \infty$ as $T \to \infty$, from Lemma 1, we get $g_T^{-1}(\tilde{M}_T - M_a) \to 0$ and hence $g_T^{-1} M_T \to 0$, almost surely as $T \to \infty$.

Now, we have $\tilde{M}_t = M_t$ for all t on the set

$$\{\tau = \infty\} = \{\int_a^\infty g_{s-}^2\, d\langle M\rangle_s \le c\}.$$

Since c is arbitrary, we obtain $g_T^{-1} M_T \to 0$ almost surely as $T \to \infty$ on the set

$$\{\int_a^\infty g_s^{-2}\, d\langle M\rangle_s < \infty\}$$

By assumption $\langle M\rangle_s \le \text{const.}\ g_s^{2-\varepsilon}$, for all $s \ge b$, for some $b \ge a$, almost surely. So

$$\int_b^\infty g_s^{-2}\, d\langle M\rangle_s \le \text{const.} \int_b^\infty \langle M\rangle_s^{-2/(2-\varepsilon)}\, d\langle M\rangle_s$$

$$= \text{const.} \int_{\langle M\rangle_b}^\infty x^{-2/(2-\varepsilon)}\, dx < \infty$$

since the image of the measure $d\langle M\rangle_s$ by the application $s \to \langle M\rangle_s$ is just the Lebesque measure. Hence, almost surely

$$\int_a^\infty g_s^{-2}\, d\langle M\rangle_s < \infty,$$

since g_s^{-2} is bounded on $[a,b]$. The proof is completed.

Theorem 2. Suppose that

(i) $P\{\lambda_{\theta,j}(t) \geq c\} = 0$ for some $c > 0$ (c is independent of t by stationarity) and $\hat{\lambda}_{\theta,j}(t) - \lambda_{\theta,j}(t)$, $\hat{\lambda}^{(i)}_{\theta,j}(t) - \lambda^{(i)}_{\theta,j}(t)$, $i = 1,2$ tend to zero almost surely as $t \to \infty$.

(ii) $\lambda_{\theta,j}(t)$ is integrable, and $\phi_{\theta,j}(t)$ and the elements of $\phi^{(1)}_{\theta,j}(t)$, $\phi^{(2)}_{\theta,j}(t)$ are square integrable with respect to the measure $\lambda_j(t)dP$.

Then almost surely as $T \to \infty$

$$T^{-1} L_T(\theta) \to E[\sum_{j=1}^{r} \phi_{\theta,j}(0)\lambda_j(0) - \lambda_{\theta,j}(0)] = \Lambda(\theta)$$

$$T^{-1} L^{(1)}_T(\theta^*) \to 0$$

$$T^{-1} L^{(2)}_T(\theta^*) \to -E\{\sum_{j=1}^{r} \phi^{(1)}_j(0)\phi^{(1)}_j(0)'\phi_j(0) = -J$$

Proof

Consider the right hand side of (4.2). By (i)

$$\hat{\phi}_{\theta,j}(t)\lambda_j(t) - \hat{\lambda}_{\theta,j}(t) = \phi_{\theta,j}(t)\lambda_{\theta,j}(t) - \lambda_{\theta,j}(t) + o(t)$$

where in this proof $o(t)$ denotes a quantity tending to zero almost surely as $T \to \infty$. By lemma 1 and the fact that $\phi_{\theta,j}(t)\lambda_j(t)$ is integrable since $\phi_{\theta,j}(t)(\lambda_j(t))^{\frac{1}{2}}$ is square integrable, we obtain from (4.1) that T^{-1} times the last term of the extreme right hand side of (4.2) tends to $\Lambda(\theta)$ almost surely as $T \to \infty$. On the other hand, $m_j(t)$ is a martingale with natural increasing process given by $d\langle m_j\rangle_t = \lambda_j(t)dt$. Therefore

$$M_{j,t} = \int_0^t \hat{\phi}_{\theta,j}(t)\ dm_j(t)$$

is also a martingale with natural increasing process

$$\langle M_j \rangle_t = \int_0^t \hat{\phi}^2_{\theta,j}(t)\ \lambda_j(t)dt$$

If $\langle M_j \rangle_t = O(t)$ almost surely as $t \to \infty$, then by lemma 2, $M_{j,T}/T \to 0$ and hence $T^{-1} L_T(\theta) \to \Lambda(\theta)$ almost surely as $T \to \infty$. Since from (i)

$$\hat{\phi}^2_{\theta,j}(t)\lambda_j(t) = \phi^2_{\theta,j}(t)\lambda_j(t) + \phi_{\theta,j}(t)\lambda_j(t)o(t) + \lambda_j(t)o(t)$$

and therefore by (4.1) and Lemma 1, $M_{j,T}/T \to E\{\phi^2_{\theta,j}(0)\lambda_j(0)\}$ almost surely as $T \to \infty$, we obtain the result

The proof for the convergence of $T^{-1}L_T^{(1)}(\theta^*)$ and $T^{-1}L_T^{(2)}(\theta^*)$ uses the same idea. We have from (i)

$$\begin{aligned}
\hat{\phi}_j^{(1)}(t)\{\lambda_j(t) - \hat{\lambda}_j(t) &= \{\phi_j^{(1)}(t) + \phi_j^{(1)}(t)o(t) + o(t)\}o(t) \\
&= \phi_j^{(1)}(t)o(t) + o(t) \\
\hat{\phi}_j^{(2)}(t)\{\lambda_j(t) - \hat{\lambda}_j(t)\} &= \{\hat{\lambda}_j^{(2)}(t)/\hat{\lambda}_j(t) - \hat{\phi}_j^{(1)}(t)\hat{\phi}_j^{(1)}(t)'\}o(t) \\
&= \{\phi_j^{(2)}(t) + \lambda_j^{(2)}(t)o(t) + \phi_j^{(1)}(t)o(t) \\
&\quad + \phi_j^{(1)}(t)\phi_j^{(1)}(t)'o(t) + o(t)\}o(t) \\
&= \phi_j^{(2)}(t)o(t) + \lambda_j^{(2)}(t)o(t) + \phi_j^{(1)}(t)o(t) + o(t)
\end{aligned}$$

Since $\lambda_j(t)$ is bounded below, from (ii), $\phi_j^{(i)}(t)$, $\lambda_j^{(i)}(t)$, $i = 1,2$ are integrable. By lemma 1 and (4.1), the first term of the extreme right hand side of (4.3), (4.4), divided by T

tend to zero as $T \to \infty$.

Finally by a similar argument as above, the second terms in the extreme right hand side of (4.3) and (4.4) are $T\,o(T)$ and the last term in the right hand side of (4.4) divided by T converge to $-J$ almost surely as $T \to \infty$. The proof is completed.

Remark

Condition (i) of the theorem is introduced for convenience. The result might hold under weaker conditions. In fact, it suffices that

$$\frac{1}{T}\int_0^T \{\hat{\lambda}_j(t) - \lambda_j(t)\}dt \ , \ \frac{1}{T}\int_0^T \{\hat{\phi}_{\theta,j}(t) - \phi_{\theta,j}(t)\}\lambda_j(t)dt \ ,$$

$$\frac{1}{T}\int_0^T \hat{\phi}_j^{(i)}(t)\{\hat{\lambda}_j(t) - \lambda_j(t)\}dt \ , \ i = 1,2,$$

$$\frac{1}{T}\int_0^T \{\hat{\phi}_j^{(1)}(t)\ \hat{\phi}_j^{(1)}(t)'\hat{\lambda}_j(t) - \phi_j^{(1)}(t)\phi_j^{(1)}(t)'\lambda_j(t)\}dt$$

tend to zero almost surely as $T \to \infty$ and

$$\frac{1}{T}\int_0^T \hat{\phi}^2_{\theta,j}(t)\lambda_j(t)dt \ , \ \frac{1}{T}\int_0^T \|\hat{\phi}_j^{(i)}(t)\|^2\lambda_j(t)dt \ , \ i = 1,2$$

are bounded almost surely, and condition (ii) to obtain the result.

Condition (i) is also not very restrictive. In case of the self-exciting process (3.1) with $g_{\theta,jk} \geq 0$, then $\lambda_{\theta,j}(t)$ is bounded below by a $\alpha_{\theta,j}$, which we assume to be strictly positive. If we also assume that

$$\bar{g}_\theta(t) = \sup_{h\geq 0}\|g_\theta(t+h)\|$$

is integrable, then

$$\|\hat{\lambda}_\theta(t) - \lambda_\theta(t)\| \le \text{const.}\int_{-\infty}^{0} \|g_\theta(t-s)\| \sum_{j=1}^{r} dN(s)$$
$$\le \text{const.}\int_{-\infty}^{0} \bar{g}_\theta(t-s) \sum_{j=1}^{r} dN(s)$$

where the last integral is almost surely finite since it has a finite expectation and converges to zero as $t \to \infty$ by the monotonous convergence Theorem. In the same way, if $g_\theta(t)$ is twice differentiable with respect to θ with derivatives satisfying the same condition as $g_\theta(t)$ as above, then $\|\hat{\lambda}^{(i)}_{\theta,j}(t) - \lambda^{(i)}_{\theta,j}(t)\|$, $i = 1,2$, tend to zero almost surely as $t \to \infty$.

Consider now the asymptotic distribution of $T^{-\frac{1}{2}} L_T(\theta^*)$. From (4.3), we expect that the above is asymptotically distributed like

$$Z_T = T^{-\frac{1}{2}}\int_0^T \sum_{j=1}^{r} \phi_j^{(1)}(t)\,dm_j(t) \tag{4.5}$$

The asymptotic distribution of Z_T can be obtained from the following result, which is of independent interest.

<u>Lemma 3</u>. <u>Let</u> $M_t, A_t, t \ge 0$ <u>be a square integrable martingale with natural increasing process</u> $\langle M_t \rangle$ <u>satisfying</u> $d\langle M\rangle_t = X_t dt$. <u>Suppose that there is a semi group of shift operator</u> $T_s, s \ge 0$, <u>conserving the probability and metrically transitive, such that</u> $T_h^{-1} A_t = A_{t+h}$, $(M_t - M_s) \circ T_h = M_{t+h} - M_{s+h}$. <u>Then as</u> $T \to \infty$, $T^{-\frac{1}{2}} M_T$ <u>is asymptotically normal with zero mean and variance</u> EX_0 .

<u>Proof</u>.

Let n_T be integers such that $\Delta_T = T/n_T \to 0$ as $T \to \infty$.

Set

$$Y_{T,j} = T^{-\frac{1}{2}} (M_{j\Delta T} - M_{(j-1)\Delta T})$$

Then from the result of Durett and Resnick (1978), as $T \to \infty$,

$$T^{-\frac{1}{2}} M_T = T^{-\frac{1}{2}} M_0 + \sum_{j=1}^{n_T} Y_{T,j}$$

is asymptotically normal with zero mean and variance EX_0, provided

$$\text{(i)} \quad \sum_{j=1}^{n_T} E\{Y_{T,j}^2 \,|A_{(j-1)\Delta T}\} \to EX_0 \quad \text{in probability}$$

$$\text{(ii)} \quad \sum_{j=1}^{n_T} E[Y_{T,j}^2 \, 1_{\{Y_{T,j}^2 > \varepsilon\}}] \to 0 , \quad \varepsilon > 0 .$$

To verify (i), observe that the sum in (i) is equal to

$$\frac{1}{T} \int_0^T X_t^{(T)} \, dt = \langle M \rangle_T / T + \frac{1}{T} \int_0^T (X_t^{(T)} - X_t) \, dt$$

where $X_t^{(T)} = E(X_t | A_{j\Delta T})$ for $j\Delta_T \le t < (j+1)\Delta_T$. Clearly $X_{t+h} = X_t \circ T_h$ and $E(X_{t+h} | A_{s+h}) = E(X_t | A_s) \circ T_h$. Therefore, as $\Delta_T \to 0$,

$$\sup_{0 \le t \le T} E|X_t^{(T)} - X_t| \le \sup_{0 \le u \le \Delta T} E|E(X_1 | A_{1-u}) - X_1| \to 0$$

Note that X_t is A_t-measurable because this process is predictable. Thus, the sum in (i) differs from $\langle M \rangle_T / T$ by a term tending to 0 in the mean and since $\langle M \rangle_T / T \to EX_0$ by the metric transitivity of T_h, we obtain the result.

To verify (ii), observe that the $Y_{T,j}$, $j = 1,\ldots,n_T$ have the same distribution. Put $\xi(t) = M_t^2 - M_0$, we are led to verify that:

$$n_T[E(Y_{T,1}^2 \, 1_{\{Y_{T,1}^2 > \varepsilon\}}] = \Delta_T^{-1} E[\xi(\Delta_T) \, 1_{\{\xi(\Delta_T) > \varepsilon T\}}]$$

tend to zero as $T \to \infty$. This is true if $\Delta_T \to 0$ sufficiently slow. Indeed, $\xi(\Delta)$, $0 \le \Delta \le 1$ being a positive sub-martingale is uniformly integrable (Doob, 1953, p. 359), that is:

$$\alpha(T) = \sup_{0 \le \Delta \le 1} E[\xi(\Delta) \, 1_{\{\xi(\Delta) > \varepsilon T\}}] \to 0 .$$

as $T \to \infty$. So all we have to do is to choose n_T such that $\alpha(T)/\Delta_T \to 0$ as $T \to \infty$.

We now show that the difference between $T^{\frac{1}{2}} L_T(\theta^*)$ and Z_T, defined by (4.3), (4.5) tend to zero in probability as $T \to \infty$.

<u>Lemma 4</u>. <u>Let</u> M_t, $t \ge 0$ <u>be a locally integrable martingale with continuous natural increasing process</u> $\langle M \rangle_t$. <u>Let</u> $h_t \ge 0$, <u>if</u> $\langle M \rangle_T / h_T^2 \to 0$ <u>in probability as</u> $T \to \infty$, <u>then so is</u> M_T/h_T.

<u>Proof</u>

Let $\varepsilon > 0$. Define the stopping time σ_t to be the value of s such that $\langle M \rangle_s = \varepsilon h_t^2$. Then as $T \to \infty$,

$$P\{M_{T \wedge \sigma_T} \ne M_T\} \to 0$$

since $P\{\sigma_T \le T\} = P\{\langle M \rangle_T \ge \varepsilon h_T^2\} \to 0$. On the otherhand

$$E(M^2_{T \wedge \sigma_T}) = E\langle M \rangle_{T \wedge \sigma_T} \le E\langle M \rangle_{\sigma_T} = \varepsilon h_T^2$$

and hence by Tchebycheff inequality

$$P\{|M_T/h_T| > \delta\} \le P\{M_{T \wedge \sigma_T} \ne M_T\} + \varepsilon/\delta^2$$

from which the result follows.

Theorem 3. Under the condition of Theorem 2 and suppose that

(i) $$T^{-\frac{1}{2}}\int_0^T \hat{\phi}_j^{(1)}(t)\{\hat{\lambda}_j(t) - \lambda_j(t)\}dt \to 0$$

in probability as $T \to \infty$. Then $T^{-\frac{1}{2}} L_T(\theta^*)$ is asymptotically normal with zero-mean and covariance matrix J given in Theorem 2.

Proof.

Let α be a vector and Z_T be given by (4.5). Then $M_T = T^{\frac{1}{2}} \alpha' Z_T$, $T \geq 0$ is a martingale with natural increasing process

$$\langle M\rangle_t = \int_0^t \sum_{j=1}^r \{\alpha'\phi_j^{(1)}(s)\}^2 \lambda_j(s)ds$$

since (Lipster and Shiryayev, 1977, p. 269)

$$\begin{aligned} M_t^2 &= \int_0^t 2M_{s-}dM_s + \sum_{s\leq t}(\Delta M_s)^2 \\ &= \int_0^t 2M_{s-}dM_s + \sum_{j=1}^r \{\alpha\phi_j^{(1)}(s)\}^2 dN_j(s) \\ &= \int_0^t [2M_{s-}dM_s + \sum_{j=1}^r \{\alpha'\phi_j^{(1)}(s)\}^2 dm_j(s)] + \langle M\rangle_t \end{aligned}$$

where the first term of the last expression is a martingale and the second is a natural process. By Lemma 3, $\alpha' Z_T$ is asymptotically normal with zero mean and variance $\alpha' J\alpha$. On the other hand, by (i) and Lemma 4 with $dM_t = \sum_j \alpha'\{\hat{\phi}_j^{(1)}(t) - \phi_j^{(1)}(t)\}dm_j(t)$, $h_t = \sqrt{t}$, we see that $T^{-\frac{1}{2}}L_T^{(1)}(\theta^*) - Z_T \to 0$ in probability as $T \to \infty$. The result follows.

Remark

Condition (i) of Theorem 3 is satisfied if, for example, $E\|\hat{\phi}_j^{(1)}(t)\|^2$ is bounded and

$$\int_0^\infty \{E|\hat{\lambda}_j(t) - \lambda_j(t)|^2\}^{\frac{1}{2}}\, dt < +\infty \tag{4.6}$$

since by Schwartz inequality, the first absolute moment of the expression in (i) is bounded by

$$T^{-\frac{1}{2}} \int_0^T \{E\|\hat{\phi}_j^{(1)}(t)\|^2\, E|\hat{\lambda}_j(t) - \lambda_j(t)|^2\}^{\frac{1}{2}}\, dt$$

In case of the self exciting process with $\lambda_j(t)$, $\hat{\lambda}_j(t)$ given by (3.1), (3.3) we have

$$\begin{aligned}\hat{\lambda}_j(t) - \lambda_j(t) &= \sum_k \int_{-\infty}^0 g_{ik}(t - s)\, dN_k(s) \\ &= \sum_k \{\int_{-\infty}^0 g_{jk}(t - s)\lambda_k(s)\, ds + \int_{-\infty}^0 g_{jk}(t - s)\, dm_k(s)\}\end{aligned}$$

Denote by $\|\cdot\|_2$ the L^2 norm, by the triangular inequality

$$\begin{aligned}\|\hat{\lambda}_j(t) - \lambda_j(t)\|_2 \le \sum_k \int_{-\infty}^0 |g_{jk}(t - s)|\,\|\lambda_k(s)\|_2\, ds \\ + [\sum_k \int_{-\infty}^0 g_{jk}^2(t - s)\; E\{\lambda_k(s)\}ds]^{\frac{1}{2}}\end{aligned}$$

Note that as in the proof of Theorem 3, the martingale M_s defined by $dM_s = \sum_k g_{jk}(t - s)\, dm_k(s)$ has the natural increasing process given by $d\langle M\rangle_s = \sum_k g_{jk}^2(t - s)\lambda_k(s)\, ds$. Therefore, the process $\lambda_k(t)$ being stationary

$$\begin{aligned}&\int_0^\infty \|\hat{\lambda}_j(t) - \lambda_j(t)\|_2 dt \\ &\le \text{const.} \sum_k \int_0^\infty [\int_t^\infty |g_{jk}(s)|\, ds + \{\int_t^\infty g^2{}_{jk}(s)\, ds\}^{\frac{1}{2}}]dt\end{aligned}$$

Observe that

$$(4.7) \qquad \int_0^\infty \int_t^\infty |g_{jk}(s)|\,ds \le \int_0^\infty t|g_{jk}(t)|\,dt$$

and, by Schwartz inequality, for any positive function h on $(0,\infty)$ which integrates to one

$$\int_0^\infty \{\int_t^\infty g_{jk}^2(s)\,ds\}^{\frac{1}{2}}dt \le \{\int_0^\infty \frac{1}{h(t)} \int_t^\infty g_{jk}^2(s)\,ds\,dt\}^{\frac{1}{2}}$$
$$= \int_0^\infty \{t g_{jk}^2(t)/h(t)\}\,dt$$

Take $h(t) \sim t^{-1-\alpha}$, $\alpha > 0$. Then (4.6) holds if the right hand size of (4.7) and $\int_0^\infty t^{2+\alpha}\, g_{jk}^2(t)\,dt$ are finite. In the same way one can show that $\|\hat{\lambda}_j^{(1)}(t) - \lambda_j^{(1)}(t)\|_2 \to 0$ as $t \to \infty$ if the $g_{jk}^{(1)}(t)$ are integrable and square integrable. Hence $E\|\phi_j^{(1)}(t)\|^2$ are bounded if the above conditions hold and $\lambda_j(t)$ are bounded below. Thus condition (i) of the Theorem is not restrictive.

5. Asymptotic Properties of the Maximum Likelihood Estimator

We are interested in the asymptotic properties of the estimate θ_T, which maximize L_T in Θ. We shall use this general result for which the proof is quite standard.

Theorem 4. 1. Let Λ_T be a random function on $\Theta \subset R^k$, satisfying

(i) $\Lambda_T(\theta) \to \Lambda(\theta)$ almost surely as $T \to \infty$, with Λ being continuous, admitting a unique maximum θ^*

(ii) For any $\theta \ne \theta^*$

$$\limsup_{T\to\infty} \sup_{\theta'\in U(\theta)} \{\Lambda_T(\theta') - \Lambda_T(\theta)\} \to 0 \qquad \text{a.s.}$$

as the neighborhood $U(\theta)$ of θ shrinks to θ.

Then any θ_T realising the maximum of Λ_T on a compact C of Θ containing θ^*, converges almost surely to θ^* as $T \to \infty$.

2. Suppose that θ^* is an interior point of Θ and Λ admits continuous first and second derivatives with respect to θ, denoted by the vector $\Lambda_T^{(1)}$ and the matrix $\Lambda_T^{(2)}$, satisfying

(iii) As $T \to \infty$, $\Lambda_T^{(2)}(\theta^*) \to -J$ in probability and $\sqrt{T}\,\Lambda_T^{(1)}(\theta)$ is asymptotically normal with zero mean and covariance matrix J.

(iv) For every $\varepsilon > 0$

$$\liminf_{T\to\infty} P\{|\Lambda_T^{(2)}(\theta) - \Lambda_T^{(2)}(\theta^*)|| < \varepsilon, \ \forall\theta\in U(\theta^*)$$

increases to 1 as the neighborhood $U(\theta^*)$ of θ^* shrinks to θ^*.

Then $\hat{\theta}_T$ of 1) is asymptotically normal with mean θ^* and covariance matrix $T^{-1}J^{-1}$. Moreover, if θ_T is $T^{\frac{1}{2}}$-consistent, that is the distributions of $\sqrt{T}(\theta_T - \theta^*)$ are tight, then $\hat{\theta}_T - \{\theta_T - \Lambda_T^{(2)}(\theta_T)^{-1}\Lambda_T^{(1)}(\theta_T)\} \to 0$ in probability as $T \to \infty$.

We apply the above results with $\Lambda_T = T^{-1}L_T$. By Theorem 2, condition (i) is satisfied except the continuity of $\Lambda(\theta)$, which we shall assume. To see that θ^* realises the maximum of $\Lambda(\theta)$, write

$$\Lambda(\theta) - \Lambda(\theta^*) = \sum_{j=1}^{r}[E\ \mathrm{Log}\{\lambda_{\theta,j}(0)\lambda_j^{-1}(0)\}\lambda_j(0) + \lambda_j(0) - \lambda_{\theta,j}(0)]$$

and note that $\log x \le x - 1$ with equality if and only if $x = 1$, we get $\Lambda(\theta) \le \Lambda(\theta^*)$ with equality if and only if $\lambda_{\theta,j}(0) = \lambda_j(0)$ almost surely, implying, by stationarity $\lambda_{\theta,j}(t) = \lambda_j(t)$ almost surely. If the parametrization is such that for $\theta' \neq \theta$, $\lambda_{\theta,j}(t)$ is not equal to $\lambda_{\theta',j}(t)$ for all t, almost surely, then θ^* is the unique maximum of Λ.

Theorems 2 and 3 show that condition (iii) is satisfied. So all we need is to verify conditions (ii) and (iv). This would require rather strong assumptions on $\lambda_\theta(t)$ and $\lambda_\theta(t)$. A sufficient set of assumptions is

A0: <u>For any compact</u> C <u>of</u> Θ, <u>there is a</u> $c > 0$ <u>such that</u> $\lambda_{\theta,j}(t) \ge c$ <u>almost surely for all</u> t, <u>all</u> $\theta \in C$.

A1: <u>For any compact</u> C of Θ

$$\lim_{t\to\infty} \operatorname*{Sup}_{\theta\in C} |\hat{\lambda}_{\theta,j}(t) - \lambda_{\theta,j}(t)| = 0 \quad \text{a.s.}$$

$$E\ \{\operatorname*{Sup}_{\theta\in C} \phi^2_{\theta,j}(t)\}\lambda_j(t) < \infty$$

A2: <u>For some compact neighborhood</u> U of θ^*,

$$\lim_{t\to\infty} \operatorname*{Sup}_{\theta\in U} \|\hat{\lambda}^{(i)}_{\theta,j}(t) - \lambda^{(i)}_{\theta,j}(t)\| = 0 \quad \text{a.s.,}\ i = 1,2,$$

$$E[\{\operatorname*{Sup}_{\theta\in U} \|\phi^{(2)}_{\theta,j}(t)\|^2\}\ \lambda_j(t)] < \infty$$

$$E\{\operatorname*{Sup}_{\theta\in U} \|\lambda^{(2)}_{\theta,j}(t)\|\} < \infty$$

<u>Theorem 5.</u> <u>Under the assumptions</u> A0, A1 <u>condition</u> (ii) <u>of</u>

Theorem 4 is satisfied, and under the assumptions A0, A2, condition (iv) of Theorem 4 is satisfied.

Proof.

Let $U(\theta)$ be a compact neighborhood of θ. Then from (4.2),

$$\sup_{\theta' \in U(\theta)} T^{-1} L_T(\theta') \le T^{-1} \sum_{j=1}^{r} [\int_0^T \{ \sup_{\theta' \in U(\theta)} \hat{\phi}_{\theta',j}(t) \} dN_j(t) - \int_0^T \{ \inf_{\phi' \in U(\theta)} \hat{\lambda}_{\theta,j}(t) \} dt] .$$

By the same argument as in the proof of Theorem 2, the above right hand side is seen to converge almost surely as $T \to \infty$ to

$$\sum_{j=1}^{r} [E[\{ \sup_{\theta' \in U(\theta)} \phi_{\theta',j}(t) \} \lambda_j(t)] - E\{ \inf_{\theta' \in U(\theta)} \lambda_{\theta,j}(t) \}]$$

By the monotonous convergence theorem, as $U(\theta)$ shrinks to θ the above expression converges to $\Lambda(\theta)$ and hence

$$\limsup_{T \to \infty} \{ \sup_{\theta' \in U(\theta)} T^{-1} L_T(\theta') \} \to \Lambda(\theta) \quad \text{a.s.} ,$$

which gives the result.

To verify condition (iv) of Theorem 4, from (4.4) observe that $T^{-1} | L_T^{(2)}(\theta) - L_T^{(2)}(\theta^*) |$ is bounded for all $\theta \in U$ by

$$T^{-1} \sum_{j=1}^{r} [\int_0^T \{ \sup_{\theta \in U} \| \hat{\phi}_{\theta,j}^{(2)}(t) - \phi_j^{(2)}(t) \| \} dN_j(t) + \int_0^T \{ \sup_{\theta \in U} \| \hat{\lambda}_{\theta,j}^{(2)}(t) - \hat{\lambda}_j^{(2)}(t) \| \} dt]$$

Again, by a similar argument, the above expression converges almost surely as $T \to \infty$ to

$$\sum_{j=1}^{r} E \sup_{\theta \in U} \|\phi_{\theta,j}^{(2)}(t) - \phi_j^{(2)}(t)\|\}\lambda_j(t)$$

$$+ E\{\sup_{\theta \in U} \| \lambda_{\theta,j}^{(2)}(t) - \lambda_j^{(2)}(t) \|\}$$

Note that we have used the fact that

$$\{\sup_{\theta \in U} \|\phi_{\theta,j}^{(2)}(t) - \phi_j(t)\|\}\sqrt{\lambda_j(t)} \ , \ \sup_{\theta \in U} \|\lambda_{\theta,j}(t) - \lambda_j(t)\|$$

are square integrable and integrable which follows easily from A2. Again, by the monotonous convergence Theorem, the above expectations converge to zero as U shrinks to θ^*. The proof is completed.

REFERENCES

DOOB, J.L., (1953) Stochastic processes. Wiley, New York.

JACOD, J., (1975) Multivariate point process predictable projection, Radon Nikodym derivatives, Representation of Martingales. Z. Wahrscheinlichkeitstheorie und Verw. Geb. 31, 235-53.

HAWKES, A.G., (1972) Spectra of some mutually exciting point processes with associated variables. In Stochastic point processes, 261-271, ed. P.W.A. LEWIS, Wiley, New York.

DURETT, R., RESNICK, S.I. (1978) Functional limit Theorems for dependent variables. Ann. Prob. 6, 829-46.

LIPSTER, R.S., SHRYAYEV, A.N., (1978) Statistic of Random processes II. Springer-Verlag, Berlin.

SNYDER, D.L., (1975) Random point processes. Wiley, New York.

VERE-JONES, D.,(1975) On updating algorithms and Inference for stochastic processes. In Perspective in Probability and Statistics, 239-59. Ed. G. Gani, Academic Press, New York.

Université de Grenoble
Mathématiques Appliquées
38041 Grenoble cédex, France

Department of Mathematics
Indiana University
Bloomington, Indiana 47401