Ann. Inst. Statist. Math.
Vol. 42, No. 3, 489-507 (1990)

ON THE ROBUST ESTIMATION IN POISSON PROCESSES WITH PERIODIC INTENSITIES

NAKAHIRO YOSHIDA[1*] AND TOSHIHARU HAYASHI[2]

[1]*Department of Applied Mathematics, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan*
[2]*Department of Mathematics and Related Fields, College of Integrated Arts and Sciences, University of Osaka Prefecture, Sakai, Osaka 591, Japan*

(Received January 17, 1989)

Abstract. Under some regularity conditions, it is well known that the maximum likelihood estimator (MLE) is asymptotically normal and efficient. However, if the observation is contaminated, the MLE is not always an appropriate estimator. In this paper, we treat M-estimators and study their asymptotic behavior. By choosing estimation equations, robust M-estimators are presented for phase parameters.

Key words and phrases: Efficiency, M-estimator, minimax robust, Poisson process, robustness.

1. Introduction

We consider a Poisson process $X(t)$ with a parametrized intensity $\lambda(t,\theta)$, where the parameter θ is to be estimated and belongs to a bounded open interval Θ of $\boldsymbol{R}$ (the real line). The log likelihood function based on the observation $(X(t); 0 \leq t \leq T)$ up to time T is given by

$$l(T,\theta) = \int_0^T \log \lambda(t,\theta) dX(t) - \int_0^T \lambda(t,\theta) dt .$$

The maximum likelihood estimator (MLE) maximizes the log likelihood $l(T,\theta)$ and is a solution of the likelihood equation

$$\int_0^T \frac{\dot{\lambda}(t,\theta)}{\lambda(t,\theta)} dX(t) - \int_0^T \dot{\lambda}(t,\theta) dt = 0$$

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

under some regularity conditions, where $\dot{\lambda}$ is the derivative of λ with respect to θ. Moreover, it is well known that the MLE is consistent, asymptotically normal and efficient (see, e.g., Kutoyants (1984)).

If the artificial model does not sufficiently reflect the generation mechanism of the data or the data are contaminated by noises, the true intensity $\mu(t)$ of the process X may not belong to the parametric model $\{\lambda(t,\theta); \theta \in \Theta\}$. In such circumstances, the MLE is not always an appropriate estimator of the parameter θ.

For example, we shall consider the following point process X. Let the true intensity $\mu(t)$ be $(1-\varepsilon)f(t)+\varepsilon c(t)$ and $\lambda(t,\theta)=f(t-\theta)$, where f and c are periodic even functions with the period 1, ε denotes the rate of contamination and the phase parameter θ ($\in \Theta = (-1/2, 1/2)$) is required to be estimated. If the data are not contaminated (i.e., $\varepsilon = 0$), the MLE is a very good estimator. However, its asymptotic efficiency diminishes for $\varepsilon > 0$. So, our purpose is to construct robust estimators in the sense that high efficiency is kept even if the data are contaminated.

The robust estimation problem has been studied by many statisticians. Huber (1981) and Hampel *et al.* (1986) sum up it in the independently and identically distributed case. In time series, it has been studied by Kleiner *et al.* (1979), Denby and Martin (1979), Künsch (1984), Martin and Yohai (1985, 1986), Bustos and Yohai (1986) and many other authors. Yoshida (1988) treats it in diffusion processes. They use the M-estimation and the GM-estimation to get robust estimators. Here, we treat an M-estimator which is a solution of a generalized likelihood equation. In Section 2, we examine its asymptotic behavior, that is, its consistency in a sense and asymptotic normality. In Section 3, we illustrate how to get a robust M-estimator for the above model with unknown phase parameter. Moreover, we show that our robust estimator has the minimax variance provided that the true intensity belongs to a suitable class.

2. Asymptotic behavior of the M-estimator

We treat the M-estimator defined by

DEFINITION 2.1. For functions $h(t,\theta)$ and $H(t,\theta)$, the solution of the equation

$$C(T,\theta)=\int_0^T h(t,\theta)dX(t)-\int_0^T H(t,\theta)dt=0$$

is called the M-estimator.

The MLE corresponds to the M-estimator for

$$h(t,\theta)=\frac{\dot{\lambda}(t,\theta)}{\lambda(t,\theta)}\,,\qquad H(t,\theta)=\dot{\lambda}(t,\theta)$$

when the parametric model is $\{\lambda(t,\theta);\ \theta\in\Theta\}$.

We assume the following conditions to show the consistency and asymptotic normality of the M-estimator $\hat{\theta}_T$.

(1) The true intensity $\mu(t)$ of the Poisson process X is a bounded measurable function with period τ (>0).

(2) The functions h and H are periodic in t with period τ for any $\theta\in\Theta$ and are absolutely continuous with respect to θ for any $t\ge 0$. Their Radon-Nikodym derivatives $\dot{h}$ and $\dot{H}$ are bounded in (t,θ).

(3) There exists a $\theta_1\in\Theta$ such that $\int_0^\tau\{H(t,\theta_1)-h(t,\theta_1)\mu(t)\}dt=0$.

(4)
$$\Gamma=\frac{1}{\tau}\int_0^\tau[\dot{H}(t,\theta_1)-\dot{h}(t,\theta_1)\mu(t)]dt>0\,,$$

$$\Phi=\frac{1}{\tau}\int_0^\tau h(t,\theta_1)^2\mu(t)dt>0\,.$$

(5) There exist constants C_1 and C_2 which are independent of θ and t, such that for any sufficiently small $\delta>0$,

$$\nu\left(\bigcup_{|\theta-\theta_1|\le\delta}\{t\in[0,\tau];\ |\dot{h}(t,\theta)-\dot{h}(t,\theta_1)|\ge C_1|\theta-\theta_1|\}\right)\le C_2\delta\,,$$

where $\nu(\cdot)$ denotes the Lebesgue measure.

(6)
$$\int_0^\tau|\dot{H}(t,\theta)-\dot{H}(t,\theta_1)|dt\to 0\quad\text{as}\quad\theta\to\theta_1\,.$$

It follows from the conditions (2) and (5) that

$$\int_0^\tau|\dot{h}(t,\theta)-\dot{h}(t,\theta_1)|dt\to 0\quad\text{as}\quad\theta\to\theta_1\,. \tag{2.1}$$

Under these conditions, the M-estimator $\hat{\theta}_T$ is near θ_1 with high probability for sufficiently large T. More precisely, the M-estimator $\hat{\theta}_T$ is consistent in the following sense.

THEOREM 2.1. *For any $T\ge 0$, there exist a positive number $\delta(T)$ ($\to 0$ as $T\to\infty$) and an event $A(T)$ such that $P(A(T))\to 1$ as $T\to\infty$ and an M-estimator $\hat{\theta}_T$ exists in $U(\delta(T))$ on the event $A(T)$, where $U(\delta)=\{\theta;\ |\theta-\theta_1|\le\delta\}$.*

PROOF. Let

$$m(T,\theta)=\frac{1}{T}\int_0^T h(t,\theta)(dX(t)-\mu(t)dt)\,,$$

$$G(T,\theta)=\frac{1}{T}\int_0^T \{H(t,\theta)-h(t,\theta)\mu(t)\}dt\,,$$

$$\dot{m}(T,\theta)=\frac{1}{T}\int_0^T \dot{h}(t,\theta)(dX(t)-\mu(t)dt)$$

and

$$\dot{G}(T,\theta)=\frac{1}{T}\int_0^T \{\dot{H}(t,\theta)-\dot{h}(t,\theta)\mu(t)\}dt\,.$$

Hereafter, ε denotes any fixed positive number. We have that for any $T\geq\tau$,

$$\begin{aligned}
&\sup_{\theta\in U(\delta)}|\dot{G}(T,\theta)-\dot{G}(T,\theta_1)|\\
&\quad\leq \sup_{\theta\in U(\delta)}\left\{\frac{1}{T}\int_0^T|\dot{H}(t,\theta)-\dot{H}(t,\theta_1)|dt\right\}\\
&\qquad+\|\mu\|_\infty \sup_{\theta\in U(\delta)}\left\{\frac{1}{T}\int_0^T|\dot{h}(t,\theta)-\dot{h}(t,\theta_1)|dt\right\}\\
&\quad\leq \sup_{\theta\in U(\delta)}\left\{\frac{2}{\tau}\int_0^\tau|\dot{H}(t,\theta)-\dot{H}(t,\theta_1)|dt\right\}\\
&\qquad+\|\mu\|_\infty \sup_{\theta\in U(\delta)}\left\{\frac{2}{\tau}\int_0^\tau|\dot{h}(t,\theta)-\dot{h}(t,\theta_1)|dt\right\},
\end{aligned}$$

where $\|\mu\|_\infty=\sup|\mu(t)|$. By the conditions (1), (6) and (2.1), the right-hand side converges to 0 as $\delta\to 0$. Hence, we obtain that

$$\sup_{\theta\in U(\delta)}|\dot{G}(T,\theta)-\dot{G}(T,\theta_1)|\to 0 \quad\text{as}\quad \delta\to 0 \tag{2.2}$$

uniformly in $T\geq\tau$. Since $\dot{G}(T,\theta_1)\to\Gamma$ as $T\to\infty$, we have that there exists a $T_1\geq\tau$ such that for any sufficiently small $\delta>0$ and any $T>T_1$,

$$\begin{aligned}
\inf_{\theta\in U(\delta)}\dot{G}(T,\theta)&\geq \inf_{\theta\in U(\delta)}\{\dot{G}(T,\theta_1)-|\dot{G}(T,\theta)-\dot{G}(T,\theta_1)|\}\\
&>\Gamma-2\varepsilon\,,
\end{aligned}\tag{2.3}$$

where Γ is a positive constant given in the condition (4). We easily see that for any $T \geq \tau$,

$$
\begin{aligned}
&|\dot{m}(T,\theta) - \dot{m}(T,\theta_1)| \\
&\quad \leq \frac{1}{T}\int_0^T |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dX(t) + \frac{1}{T}\int_0^T |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|\mu(t)dt \\
&\quad \leq \frac{1}{T}\int_{D_{\theta,T}} |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dX(t) + \frac{1}{T}\int_{[0,T]-D_{\theta,T}} |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dX(t) \\
&\qquad + \frac{2\|\mu\|_\infty}{\tau}\int_0^\tau |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dt\,,
\end{aligned}
$$

where $D_{\theta,T} = \{t \in [0,T];\ |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)| \geq C_1|\theta - \theta_1|\}$ and C_1 is the constant given in the condition (5). By (2.1), there exists a $\delta_1 > 0$ such that for any $\theta \in U(\delta_1)$, the last term of the right-hand side is less than ε. Since $X(T)$ conforms to the Poisson distribution with mean $\int_0^T \mu(t)dt$, we have that for any $\delta > 0$ and any $T > 0$,

$$
\begin{aligned}
&P\left\{\sup_{\theta \in U(\delta)} \frac{1}{T}\int_{[0,T]-D_{\theta,T}} |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dX(t) \geq \varepsilon\right\} \\
&\quad \leq P\left\{\frac{C_1\delta}{T} X(T) \geq \varepsilon\right\} \\
&\quad \leq \frac{C_1\delta}{T\varepsilon}\int_0^T \mu(t)dt \\
&\quad \leq \frac{C_1\|\mu\|_\infty}{\varepsilon}\delta\,.
\end{aligned}
$$

For a measurable set B, let $X(B)$ denote $\int_B dX(t)$ which is the number of events occurring in B. Since $X(B)$ conforms to the Poisson distribution with mean $\int_B \mu(t)dt$, we see that for any $T \geq \tau$,

$$
\begin{aligned}
&P\left\{\sup_{\theta \in U(\delta)} \frac{1}{T}\int_{D_{\theta,T}} |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)|dX(t) \geq \varepsilon\right\} \\
&\quad \leq P\left\{\frac{2\|\dot{h}\|_\infty}{T} \sup_{\theta \in U(\delta)} X(D_{\theta,T}) \geq \varepsilon\right\} \\
&\quad \leq P\left\{\frac{2\|\dot{h}\|_\infty}{T} X\left(\bigcup_{\theta \in U(\delta)} D_{\theta,T}\right) \geq \varepsilon\right\}
\end{aligned}
$$

$$\le \frac{2\|\dot{h}\|_\infty}{T\varepsilon} E\left[X\left(\bigcup_{\theta \in U(\delta)} D_{\theta,T} \right)\right]$$

$$\le \frac{2\|\dot{h}\|_\infty \|\mu\|_\infty}{T\varepsilon} \frac{T+\tau}{\tau}$$

$$\cdot \nu\left(\bigcup_{|\theta-\theta_1|\le\delta} \{t \in [0,\tau]; |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)| \ge C_1|\theta-\theta_1|\} \right),$$

where $\nu(\cdot)$ denotes the Lebesgue measure. From the condition (5), we get that for any sufficiently small $\delta > 0$ and any $T \ge \tau$,

$$P\left\{ \sup_{\theta \in U(\delta)} \frac{1}{T}\int_{D_{\theta,T}} |\dot{h}(t,\theta) - \dot{h}(t,\theta_1)| dX(t) \ge \varepsilon \right\} \le \frac{4\|\dot{h}\|_\infty \|\mu\|_\infty C_2}{\varepsilon\tau} \delta .$$

Hence, we have that for any sufficiently small δ and any $T \ge \tau$,

$$P\left\{ \sup_{\theta \in U(\delta)} |\dot{m}(T,\theta) - \dot{m}(T,\theta_1)| \ge 3\varepsilon \right\} < \frac{C_3}{\varepsilon}\delta ,$$

where $C_3 = (C_1 + 4\|\dot{h}\|_\infty C_2/\tau)\|\mu\|_\infty$. Since $\dot{m}(T,\theta_1)$ converges in probability to 0 as T tends to infinity, we obtain that for any sufficiently small δ, there exists a $T_2(\delta) \ge \tau$ such that for any $T > T_2(\delta)$,

$$P\left\{ \sup_{\theta \in U(\delta)} |\dot{m}(T,\theta)| \ge 4\varepsilon \right\} < \frac{2C_3}{\varepsilon}\delta . \tag{2.4}$$

From the condition (2), we have, for any $\theta \in \Theta$,

$$C(T,\theta) = C(T,\theta_1) + \int_{\theta_1}^{\theta} \dot{C}(T,u)du , \tag{2.5}$$

where

$$\dot{C}(T,u) = \int_0^T \dot{h}(t,u)dX(t) - \int_0^T \dot{H}(t,u)dt .$$

Let $A_1(T,\delta)$ be an event $\left\{ \omega;\ \sup_{\theta \in U(\delta)} |\dot{m}(T,\theta)| < \Gamma/3 \right\}$, where Γ is a positive constant given in the condition (4). Then, we have that for any sufficiently small δ and any $T > T_2(\delta)$,

$$P(A_1(T,\delta)) \ge 1 - \frac{24C_3}{\Gamma}\delta \quad (= 1 - C_4\delta,\ \text{say})$$

by using (2.4) for $\varepsilon = \Gamma/12$. We easily see that for any sufficiently small δ and $T > T_1$,

$$\inf_{u \in U(\delta)} \left\{ -\frac{1}{T} \dot{C}(T,u) \right\} = \inf_{u \in U(\delta)} \{ -\dot{m}(T,u) + \dot{G}(T,u) \} \tag{2.6}$$

$$> \frac{\Gamma}{2}$$

on the event $A_1(T,\delta)$ by using (2.3) for $\varepsilon = \Gamma/12$. Since $(1/T)C(T,\theta_1) = m(T,\theta_1) - G(T,\theta_1) \to 0$ in probability as $T \to \infty$, for any sufficiently small δ, there exists a $T_3(\delta)$ ($\geq T_2(\delta)$) such that for any $T > T_3(\delta)$,

$$P(A_2(T,\delta)) \geq 1 - 2C_4\delta\ ,$$

where $A_2(T,\delta)$ denotes the event $\{\omega; |C(T,\theta_1)|/T \leq \Gamma\delta/4\} \cap A_1(T,\delta)$. From (2.5) and (2.6), we easily see that for any sufficiently small δ and $T > T_1$, $C(T,\theta_1 + \delta) < 0$ and $C(T,\theta_1 - \delta) > 0$ on the event $A_2(T,\delta)$, which implies that there exists a $\hat{\theta}_T \in U(\delta)$ such that $C(T,\hat{\theta}_T) = 0$. Consequently, we obtain that for any sufficiently small $\delta > 0$, there exist an event $A_2(T,\delta)$ and a $T_0(\delta)$ ($= \max\{T_1, T_3(\delta)\}$) such that $P(A_2(T,\delta)) \geq 1 - 2C_4\delta$ for any $T > T_0(\delta)$ and an M-estimator $\hat{\theta}_T$ exists in $U(\delta)$ on the event $A_2(T,\delta)$. We can take a monotone increasing sequence $\{T_n\}$ ($T_n \to \infty$) such that for any $T \geq T_n$, $P(A_2(T,1/n)) \geq 1 - 2C_4/n$. Hence, we obtain the conclusion of this theorem by setting $\delta(T) = 1/n$ for $T_n \leq T < T_{n+1}$ and $A(T) = A_2(T,\delta(T))$.

We examine $C(T,\theta)$ to obtain the asymptotic normality of the M-estimator $\hat{\theta}_T$. From the condition (3), we get

$$\frac{1}{\sqrt{T}} C(T,\theta_1) = \frac{1}{\sqrt{T}} \int_0^T h(t,\theta_1)(dX(t) - \mu(t)dt) + o_p(1)\ .$$

The first term of the right-hand side converges in distribution to the normal distribution $N(0,\Phi)$ by the central limit theorem for martingales, where Φ is a positive constant given in the condition (4).

On the other hand, we get

$$\frac{1}{\sqrt{T}} C(T,\theta_1) = \sqrt{T} \int_{\hat{\theta}}^{\theta_1} \frac{1}{T} \dot{C}(T,u)du$$

$$= \sqrt{T} \left(\int_{\hat{\theta}}^{\theta_1} \dot{m}(T,u)du - \int_{\hat{\theta}}^{\theta_1} \dot{G}(T,u)du \right)$$

$$= \sqrt{T}\left(\int_{\hat\theta}^{\theta_1} \dot m(T,u)du - \int_{\hat\theta}^{\theta_1} \{\dot G(T,u) - \dot G(T,\theta_1)\}du + (\hat\theta - \theta_1)\dot G(T,\theta_1)\right),$$

where $\hat\theta = \hat\theta_T$ is a zero point of $C(T,\theta)$, that is, the M-estimator. We easily see that

$$\left|\int_{\hat\theta}^{\theta_1} \dot m(T,u)du\right| \le |\hat\theta - \theta_1| \sup\{|\dot m(T,\theta)|; \theta \in U(|\hat\theta - \theta_1|)\}$$

and

$$\left|\int_{\hat\theta}^{\theta_1} \{\dot G(T,u) - \dot G(T,\theta_1)\}du\right| \le |\hat\theta - \theta_1| \sup\{|\dot G(T,\theta) - \dot G(T,\theta_1)|; \theta \in U(|\hat\theta - \theta_1|)\}.$$

Hence, we have that

$$\frac{1}{\sqrt{T}} C(T,\theta_1) = \sqrt{T}(\hat\theta - \theta_1)(\Gamma + o_P(1)),$$

from (2.2), (2.4) and the previous theorem. Consequently, we obtain the following theorem.

THEOREM 2.2. *The M-estimator $\hat\theta_T$ is asymptotically normal*:

$$\sqrt{T}(\hat\theta_T - \theta_1) \Rightarrow N(0, \Phi\Gamma^{-2})$$

in distribution as $T \to \infty$, where Φ and Γ are positive constants given in the condition (4).

3. Minimax robust M-estimator

We shall estimate the phase parameter θ in the periodic intensity $\lambda(t,\theta) = f(t-\theta)$, where f is a C^2-class, strictly positive and even function with period 1 and $\theta \in \Theta = (-1/2, 1/2)$. For simplicity, we assume that $\int_0^1 f(t)dt = 1$. We suppose that the score function $S(t,\theta)$ $(= S(t-\theta)) = (\partial/\partial\theta)\log f(t-\theta)$ is concave in $t \in [\theta, \theta + 1/2]$ and that $\int_0^1 S(t-\theta)^2 f(t-\theta)dt > 0$. Note that $(\partial/\partial\theta)S(t-\theta) = -(\partial/\partial t)S(t-\theta) = -S'(t-\theta)$.

The true intensity is given by

$$\mu(t) = (1-\varepsilon)f(t-\theta_0) + \varepsilon c(t-\theta_0)\,, \tag{3.1}$$

where $\theta_0 \in \Theta$, $\varepsilon \in [0, M_\varepsilon]$, $0 < M_\varepsilon < 1$ and c is a periodic, even, bounded and measurable function. Without loss of generality, we can assume $\theta_0 = 0$. Let $h(t,\theta) = \psi(t-\theta)$ and $H(t,\theta) = \varphi(t-\theta)$, where φ is odd and both ψ and φ are periodic functions for which the conditions (1)–(6) in the previous section hold with $\theta_1 = 0$ ($= \theta_0$). Then, we easily see that Γ in the condition (4) is equal to $\int_0^1 \psi'(t)\mu(t)dt$. Hence, the asymptotic variance is a measure of goodness of estimation and is related to ψ only. The authors will discuss the robust estimation for asymmetric contaminations elsewhere.

We shall construct an M-estimator which has the minimax variance provided that the true intensity μ belongs to a suitable class. First, we shall look for an intensity μ_0 minimizing the information $I(\mu)$ defined by

$$I(\mu) = \sup_{\psi \in Q_1} \frac{\left(\int_0^1 \psi'(t)\mu(t)dt\right)^2}{\int_0^1 \psi(t)^2\mu(t)dt}\,, \tag{3.2}$$

where Q_1 is the class of all periodic and continuously differentiable functions ψ with $\int_0^1 \psi(t)^2\mu(t)dt > 0$.

Putting, for $0 \le \beta \le \max_t S(t)$,

$$a = \inf\left\{ t \in \left[0, \frac{1}{2}\right]; S(t) \ge \beta \right\} \tag{3.3}$$

and

$$b = \sup\left\{ t \in \left[0, \frac{1}{2}\right]; S(t) \ge \beta \right\}, \tag{3.4}$$

We easily see that a and b are solutions of the equation $S(t) = \beta$ and differentiable with respect to β $\left(0 \le \beta < \max_t S(t)\right)$ and that $\{t \in [0, 1/2]; S(t) \ge \beta\} = [a, b]$ by the concavity of $S(t)$ on $[0, 1/2]$. Let

$$\mu_\beta(t) = \begin{cases} (1-M_\varepsilon)f(t) & |t| \le a\,, \\ (1-M_\varepsilon)f(a)\exp\{-\beta(t-a)\} & a \le |t| \le b\,, \\ (1-M_\varepsilon)\dfrac{f(a)}{f(b)}\exp\{-\beta(b-a)\}f(t) & b \le |t| \le \dfrac{1}{2}\,, \\ \text{periodic} & \text{otherwise}\,. \end{cases} \tag{3.5}$$

Then its score function is given by

$$S_\beta(t) = -\frac{\mu_\beta'(t)}{\mu_\beta(t)} = \max\{-\beta, \min\{S(t), \beta\}\}$$

$$= \begin{cases} S(t) & |t| < a,\ b < |t| \le \frac{1}{2}, \\ \beta & a < t < b, \\ -\beta & -b < t < -a, \\ \text{periodic} & \text{otherwise}. \end{cases}$$

For a constant ξ (> 1, near by 1), we determine β by the equation

$$\int_0^1 \mu_\beta(t)\,dt = \xi,$$

equivalently,

$$\int_0^a f(t)\,dt + f(a)\int_a^b \exp\{-\beta(t-a)\}\,dt + \frac{f(a)}{f(b)}\exp\{-\beta(b-a)\}\int_b^{1/2} f(t)\,dt = \frac{\xi}{2(1-M_\varepsilon)}. \tag{3.6}$$

Since the derivative of the left-hand side of (3.6) with respect to β is negative, it is decreasing in β and its maximal value and minimal value are $f(0)/2$ and $1/2$, respectively. Hence, for $M_\varepsilon \in [0, 1-\xi/f(0))$, equation (3.6) has a unique solution β_0. Hereafter, we abbreviate the intensity μ_{β_0} and its score function S_{β_0} as μ_0 and S_0, respectively. Let M be a class of all periodic, even and measurable functions $\mu(t)$ satisfying that

$$\int_0^1 \mu(t)\,dt \le \int_0^1 \mu_0(t)\,dt \quad (=\xi),$$

and for any t,

$$(1-M_\varepsilon)f(t) \le \mu(t) \le (1-M_\varepsilon)\frac{f(a_0)}{f(b_0)}\exp\{-\beta_0(b_0-a_0)\}f(t)$$

where a_0 and b_0, respectively, denote a and b given by (3.3) and (3.4) for $\beta = \beta_0$ and ξ is explained as the upper bound of the average number of the events occurring during one period where the observation is contaminated. From $\xi > 1$, we easily see that the intensity f of the model belongs to the class M of the contaminated intensities.

Under the condition

$$\beta_0^2 + 2S'(t) - S(t)^2 \le 0 \tag{3.7}$$

for any $t \in [b_0, 1/2]$, we shall show that

$$I(\mu_0) = \min_{\mu \in M} I(\mu) . \tag{3.8}$$

As in Chapter 4 of Huber (1981), it is sufficient to check that for any $\mu_1 \in M$ with $I(\mu_1) < \infty$,

$$\frac{d}{ds} I(\mu_s)\bigg|_{s=0} = \int_0^1 (2S_0'(t) - S_0(t)^2)(\mu_1(t) - \mu_0(t))\,dt \ge 0 , \tag{3.9}$$

where $\mu_s = (1-s)\mu_0 + s\mu_1$, $s \in [0, 1]$. We easily see that

$$\begin{aligned}
\frac{1}{2}\frac{d}{ds} I(\mu_s)\bigg|_{s=0} &= \int_0^{1/2} (\beta_0^2 + 2S_0'(t) - S_0(t)^2)(\mu_1(t) - \mu_0(t))\,dt \\
&\quad - \beta_0^2 \int_0^{1/2} (\mu_1(t) - \mu_0(t))\,dt \\
&\ge \int_0^{a_0} (\beta_0^2 + 2S'(t) - S(t)^2)(\mu_1(t) - \mu_0(t))\,dt \\
&\quad + \int_{b_0}^{1/2} (\beta_0^2 + 2S'(t) - S(t)^2)(\mu_1(t) - \mu_0(t))\,dt \\
&= \mathrm{I} + \mathrm{II} \quad (\text{say}) .
\end{aligned}$$

We see that $\mathrm{I} \ge 0$ because for $t \in [0, a_0]$, $\beta_0^2 - S(t)^2 \ge 0$, $\mu_1(t) - \mu_0(t) \ge 0$ and $S'(t) \ge 0$. Since $\mu_1(t) - \mu_0(t) \le 0$ for $t \in [b_0, 1/2]$, we get $\mathrm{II} \ge 0$ by (3.7). Consequently, we obtain (3.9).

We shall show that the M-estimator corresponding to $h(t, \theta) = S_0(t - \theta)$ has the minimax asymptotic variance, that is,

$$\inf_{\psi \in Q_2} \sup_{\mu \in M} V(\mu, \psi) = V(\mu_0, S_0) , \tag{3.10}$$

where Q_2 is the class of all periodic functions ψ with the Radon-Nikodym derivative ψ', for which the conditions (2)–(5) in the previous section hold with $\theta_1 = 0$ ($= \theta_0$) and

$$V(\mu, \psi) = \frac{\int_0^1 \psi(t)^2 \mu(t)\,dt}{\left(\int_0^1 \psi'(t)\mu(t)\,dt\right)^2}$$

is the asymptotic variance of the M-estimator corresponding to $h(t,\theta) = \psi(t-\theta)$. From Lemma 4.4 in Huber (1981), $1/V(\mu_s, S_0)$ is convex in s, where $\mu_s = (1-s)\mu_0 + s\mu_1$, $s \in [0,1]$ and $\mu_1 \in M$. We see that for any $\mu_1 \in M$,

$$\frac{d}{ds}\left(\frac{1}{V(\mu_s, S_0)}\right)\Bigg|_{s=0} = \int_0^1 (2S_0'(t) - S_0(t)^2)(\mu_1(t) - \mu_0(t))\,dt \geq 0$$

because the last inequality is valid by (3.9). Hence, we have that

$$\begin{aligned} V(\mu_0, S_0) &= \sup_{\mu \in M} V(\mu, S_0) \\ &\geq \inf_{\psi \in Q_2} \sup_{\mu \in M} V(\mu, \psi)\,. \end{aligned} \tag{3.11}$$

We get

$$\begin{aligned} V(\mu_0, S_0) &= \frac{1}{I(\mu_0)} \\ &= \inf_{\psi_1 \in Q_1} V(\mu_0, \psi_1) \end{aligned}$$

(see, e.g., Huber (1981)). For any $\psi_2 \in Q_2$ and any $\varepsilon > 0$, we can find a $\psi_1 \in Q_1$ such that $|\psi_2(t) - \psi_1(t)| < \varepsilon$ for any $t \in [0,1]$ from Weierstrass' theorem. For any $\psi_i \in Q_i$ $(i = 1,2)$, $\int_0^1 \psi_i'(t)\mu_0(t)\,dt = -\int_0^1 \psi_i(t)\mu_0'(t)\,dt$ because both ψ_i and μ_0 are periodic. Hence, for any $\psi_2 \in Q_2$, we can approximate $V(\mu_0, \psi_2)$ by $V(\mu_0, \psi_1)$ for some $\psi_1 \in Q_1$ from the boundedness of ψ_1, ψ_2, μ_0 and μ_0'. Furthermore, we have that

$$\begin{aligned} V(\mu_0, S_0) &= \inf_{\psi_1 \in Q_1} V(\mu_0, \psi_1) \\ &\leq \inf_{\psi_2 \in Q_2} V(\mu_0, \psi_2) \\ &\leq \inf_{\psi_2 \in Q_2} \sup_{\mu \in M} V(\mu, \psi_2)\,. \end{aligned}$$

From (3.11) and the above inequality, we obtain the following theorem.

THEOREM 3.1. *Under the condition* (3.7), *the M-estimator corresponding to* $h(t,\theta) = S_0(t-\theta)$ *has the minimax asymptotic variance, that is,*

$$\inf_{\psi \in Q_2} \sup_{\mu \in M} V(\mu, \psi) = V(\mu_0, S_0) ,$$

where $\mu_0 = \mu_{\beta_0}$ *is given by* (3.5), β_0 *is a unique solution of the equation* (3.6) *and* $S_0(t) = -\mu_0'(t)/\mu_0(t)$.

In the first half of this section, we have constructed the M-estimator which has the minimax variance provided that the true intensity μ belongs to the class M. In the latter half, we shall consider the minimax problem when the true intensity $\mu(t)$ is given by (3.1). The class of all functions given by (3.1) is wider than the class M but we impose a restriction on the function $h(t, \theta)$. More precisely, for a fixed function ψ satisfying conditions below, the function h is given by

$$h(t, \theta) = \beta\psi\left(\frac{S(t-\theta)}{\beta}\right),$$

where β is a positive constant, $S(t-\theta) = (\partial/\partial\theta) \log f(t-\theta)$ is the score function and $\psi(x)$ is a piece-wise continuously differentiable, continuous, monotone increasing, odd function and is concave on the $[0, \infty)$. Then the asymptotic variance of the M-estimator is given by

$$V(c, \varepsilon, \beta) = \frac{\Phi(c, \varepsilon, \beta)}{\Gamma(c, \varepsilon, \beta)^2},$$

where

$$\Phi(c, \varepsilon, \beta) = \int_0^1 \beta^2 \psi\left(\frac{S(t)}{\beta}\right)^2 \{(1-\varepsilon)f(t) + \varepsilon c(t)\} dt$$

and

$$\Gamma(c, \varepsilon, \beta) = \int_0^1 S'(t)\psi'\left(\frac{S(t)}{\beta}\right)\{(1-\varepsilon)f(t) + \varepsilon c(t)\} dt .$$

Let C be a class of all periodic, even and measurable functions c with $0 \le c(t) \le M_c$ for any t and B be a class of all β satisfying that for any $c \in C$ and $0 \le \varepsilon \le M_\varepsilon$, $\Gamma(c, \varepsilon, \beta) > 0$, where M_ε is the bound of ε. We suppose that the class B is non-empty.

Our purpose is to determine the $\beta \in B$ which minimizes $\max_{\substack{c \in C \\ 0 \le \varepsilon \le M_\varepsilon}} V(c, \varepsilon, \beta)$.

First, we shall show the following lemma.

LEMMA 3.1. *For* $0 \le \alpha \le 1/2$, *let*

$$c_\alpha(t) = \begin{cases} M_c & \alpha < |t| \le \dfrac{1}{2}, \\ 0 & |t| \le \alpha, \\ \textit{periodic} & \textit{otherwise}. \end{cases}$$

For $c = c_\alpha$, we abbreviate the asymptotic variance $V(c, \varepsilon, \beta)$ as $V(\alpha, \varepsilon, \beta)$. Then for any $\beta \in B$, there exists an $\alpha^* \in [0, 1/2]$ such that

$$\max_{c \in C} V(c, \varepsilon, \beta) = V(\alpha^*, \varepsilon, \beta). \tag{3.12}$$

Moreover, for any $\alpha \in [0, 1/2]$ and $\beta \in B$, $V(\alpha, \varepsilon, \beta)$ is monotone increasing in ε.

PROOF. Since $V(\alpha, \varepsilon, \beta)$ is continuous in α, there exists an $\alpha_0 \in [0, 1/2]$ such that

$$\max_\alpha V(\alpha, \varepsilon, \beta) = V(\alpha_0, \varepsilon, \beta).$$

It is sufficient to show that

$$\max_c V(c, \varepsilon, \beta) = V(\alpha_0, \varepsilon, \beta). \tag{3.13}$$

Let

$$\Psi(t) = \beta \psi\left(\frac{S(t)}{\beta}\right).$$

Since ψ is monotone increasing and concave on $[0, \infty)$ and $S(t)$ is concave on $[0, 1/2]$, $\Psi(t)$ is concave on $[0, 1/2]$. Hence, $\Psi'(t)$ is decreasing on $[0, 1/2]$. Accordingly, putting

$$t_0 = \sup\left\{ t \in \left[0, \frac{1}{2}\right]; \Psi'(t) \ge 0 \right\},$$

we have that $\Psi'(t) \ge 0$ for $t \in [0, t_0]$ and $\Psi'(t) \le 0$ for $t \in [t_0, 1/2]$. We can easily check that

$$\frac{1}{2} \Phi(c, \varepsilon, \beta) = \int_0^{1/2} \Psi(t)^2 \{(1 - \varepsilon) f(t) + \varepsilon c(t)\} dt > 0, \tag{3.14}$$

because $\beta \in B$. Since $S(t)$ is periodic and odd, $S(0) = S(1/2) = 0$, which implies $\Psi(0)=\Psi(1/2)=0$. From (3.14), there exists $t_1\in(0,1/2)$ such that $\Psi(t_1)\neq 0$. It follows from the concavity of Ψ and $\Psi(t_1)>0$ and for any $t\in(0,1/2)$, $\Psi(t)>0$. Moreover, we see that $0<t_0<1/2$.

Let

$$p(c)=\frac{1}{2}\int_0^1 \beta^2\psi\left(\frac{S(t)}{\beta}\right)^2 c(t)dt$$

and

$$g(c)=\frac{1}{2}\int_0^1 S'(t)\psi'\left(\frac{S(t)}{\beta}\right)c(t)dt\,.$$

Then we easily check that

$$p(c)=\int_0^{1/2}\Psi(t)^2c(t)dt$$

and

$$g(c)=\int_0^{1/2}\Psi'(t)c(t)dt\,.$$

For $c=c_\alpha$, we abbreviate $p(c)$ and $g(c)$ as $p(\alpha)$ and $g(\alpha)$, respectively. Since $p(\alpha)$ is continuous in α, for any function $c\in C$, there exists an α_1 such that $p(c)=p(\alpha_1)$; equivalently,

$$\Phi(c,\varepsilon,\beta)=\Phi(c_{\alpha_1},\varepsilon,\beta)\quad(\geq 0)\,. \tag{3.15}$$

Furthermore, we have that $g(c)\geq g(\alpha_1)$; equivalently,

$$\Gamma(c,\varepsilon,\beta)\geq\Gamma(c_{\alpha_1},\varepsilon,\beta)\quad(>0)\,. \tag{3.16}$$

Indeed, if $\alpha_1\in(0,t_0]$, we get

$$\begin{aligned}g(c)-g(\alpha_1)&=\int_0^{\alpha_1}\Psi'(t)c(t)dt+\int_{\alpha_1}^{1/2}\Psi'(t)(c(t)-M_c)dt\\&\geq\int_0^{\alpha_1}\Psi'(t)c(t)dt+\int_{\alpha_1}^{t_0}\Psi'(t)(c(t)-M_c)dt\\&\geq\Psi'(\alpha_1)\left[\int_0^{\alpha_1}c(t)dt+\int_{\alpha_1}^{t_0}(c(t)-M_c)dt\right],\end{aligned}$$

because Ψ' is monotone decreasing on $[0,1/2]$ and $\Psi'(t)\leq 0$ for $t\in[t_0,1/2]$.

On the other hand, we have

$$\begin{aligned} 0 &= p(c) - p(\alpha_1) \\ &\leq \int_0^{\alpha_1} \Psi(t)^2 c(t)\,dt + \int_{\alpha_1}^{t_0} \Psi(t)^2 (c(t) - M_c)\,dt \\ &\leq \Psi(\alpha_1)^2 \left[\int_0^{\alpha_1} c(t)\,dt + \int_{\alpha_1}^{t_0} (c(t) - M_c)\,dt \right], \end{aligned}$$

because Ψ is non-negative and monotone increasing on $[0, t_0]$. Since $\Psi(\alpha_1) > 0$ and $\Psi'(\alpha_1) \geq 0$ by definition of t_0, we obtain $g(c) - g(\alpha_1) \geq 0$. If $\alpha_1 \in [t_0, 1/2)$, $g(c) \geq g(\alpha_1)$ is similarly shown. We easily see that $g(c) = g(\alpha_1) = 0$ if $\alpha_1 = 0$ or $1/2$. Hence, for any $c \in C$, we can find an α_1 satisfying (3.15) and (3.16). Consequently we obtain that

$$V(c, \varepsilon, \beta) \leq V(\alpha_1, \varepsilon, \beta) \leq V(\alpha_0, \varepsilon, \beta),$$

which implies (3.13).

We easily see that for any $\alpha \in [0, 1/2]$ and $\beta \in B$,

$$\begin{aligned} \Gamma(c_\alpha, \varepsilon, \beta) &= 2\int_0^{1/2} \Psi'(t)\{(1-\varepsilon)f(t) + \varepsilon c_\alpha(t)\}\,dt \\ &= 2(1-\varepsilon)g(f) + 2\varepsilon g(\alpha) \quad (>0) \end{aligned}$$

and

$$\frac{\partial}{\partial \varepsilon} \Gamma = -2g(f) + 2g(\alpha) \quad (\leq 0),$$

because $g(f) \geq 0$ and $g(\alpha) \leq 0$. Similarly, we see that

$$\Phi(c_\alpha, \varepsilon, \beta) = 2(1-\varepsilon)p(f) + 2\varepsilon p(\alpha)$$

and

$$\frac{\partial}{\partial \varepsilon} \Phi = -2p(f) + 2p(\alpha).$$

Therefore

$$\frac{\partial}{\partial \varepsilon} V(\alpha, \varepsilon, \beta) = \left[\left(\frac{\partial}{\partial \varepsilon} \Phi \right) \Gamma - 2\Phi \left(\frac{\partial}{\partial \varepsilon} \Gamma \right) \right] \Big/ \Gamma^3$$

$$\geq \left[\left(\frac{\partial}{\partial \varepsilon}\Phi\right)\Gamma - \Phi\left(\frac{\partial}{\partial \varepsilon}\Gamma\right)\right]\Big/\Gamma^3$$

$$= 4[p(\alpha)g(f) - p(f)g(\alpha)]/\Gamma^3$$

$$\geq 0 .$$

Consequently, we obtain the conclusion of this lemma.

Let $\alpha^*(\beta)$ denote an α maximizing $V(\alpha, M_\varepsilon, \beta)$ and β_* denote a β minimizing $V(\alpha^*(\beta), M_\varepsilon, \beta)$, where M_ε is the bound of ε. From the previous lemma, we have that

$$V(\alpha^*(\beta_*), M_\varepsilon, \beta_*) = \min_{\beta \in B} \max_{\substack{c \in C \\ 0 \leq \varepsilon \leq M_\varepsilon}} V(c, \varepsilon, \beta) .$$

As an example, let

$$f(x) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-k)^2}{2\sigma^2}\right\} \qquad (\sigma = 0.1)$$

and

$$\psi(x) = \begin{cases} 1 & x > 1 \\ x & |x| \leq 1 \\ -1 & x < -1 . \end{cases}$$

We give the tables of asymptotic variance V and asymptotic relative efficiency (ARE) which is the reciprocal of the variance divided by that of the MLE in $M_\varepsilon = 0$.

Table 1. Model ($M_\varepsilon = 0$).

β	12.6	14.4	15.1	15.3	16.7
$V(\times 10^2)$	1.07	1.05	1.05	1.04	1.03
ARE	.941	.960	.966	.967	.976

β	17.6	17.8	19.6	21.0	ML
$V(\times 10^2)$	1.03	1.03	1.02	1.02	1.01
ARE	.981	.982	.989	.992	

Table 2. $M_\varepsilon = 0.01$.

M_c	1		2		3	
	$\beta_* = 21.0$	ML	$\beta_* = 19.6$	ML	$\beta_* = 17.6$	ML
α^*	.142	.147	.143	.152	.143	.157
$V(\times 10^2)$	1.06	1.09	1.09	1.17	1.13	1.25
ARE	.956	.927	.924	.863	.896	.807

M_c	4		5	
	$\beta_* = 16.7$	ML	$\beta_* = 15.1$	ML
α^*	.143	.161	.142	.166
$V(\times 10^2)$	1.16	1.33	1.19	1.42
ARE	.870	.757	.846	.713

Table 3. $M_\varepsilon = 0.05$.

M_c	0.5		1	
	$\beta_* = 17.8$	ML	$\beta_* = 15.3$	ML
α^*	.144	.155	.143	.167
$V(\times 10^2)$	1.16	1.27	1.25	1.50
ARE	.869	.795	.807	.675

M_c	1.5		2	
	$\beta_* = 14.4$	ML	$\beta_* = 12.6$	ML
α^*	.144	.179	.126	.190
$V(\times 10^2)$	1.34	1.73	1.42	1.97
ARE	.757	.584	.712	.512

REFERENCES

Bustos, O. and Yohai, V. J. (1986). Robust estimators for ARMA models, *J. Amer. Statist. Assoc.*, **81**, 155–168.

Denby, I. and Martin, R. D. (1979). Robust estimation of the first order autoregressive parameter, *J. Amer. Statist. Assoc.*, **74**, 140–146.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics*, Wiley, New York.

Huber, P. J. (1981). *Robust Statistics*, Wiley, New York.

Kleiner, B., Martin, R. D. and Thomson, D. J. (1979). Robust estimation of power spectra, *J. Roy. Statist. Soc. Ser. B*, **41**, 313–351.

Künsch, H. (1984). Infinitesimal robustness for autoregressive processes, *Ann. Statist.*, **12**, 843–863.

Kutoyants, Yu. A. (1984). *Parameter Estimation for Stochastic Processes*, (translated by B. S. L. Prakasa Rao), Heldermann Verlag, Berlin.

Martin, R. D. and Yohai, V. J. (1985). Robustness in time series and estimating ARMA models, *Handbook of Statistics*, Vol. 5, 119–155, North-Holland, Amsterdam.
Martin, R. D. and Yohai, V. J. (1986). Influence functionals for time series, *Ann. Statist.*, **14**, 781–818.
Yoshida, N. (1988). Robust *M*-estimator in diffusion processes, *Ann. Inst. Statist. Math.*, **40**, 799–820.