The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 76–87

ESTIMATION OF THE LOCATION OF THE CUSP OF A CONTINUOUS DENSITY[1]

By B. L. S. Prakasa Rao

University of California, Berkeley

1. Introduction and summary. Chernoff and Rubin [1] and Rubin [5] investigated the problem of estimation of the location of a discontinuity in density. They have shown that the maximum likelihood estimator (MLE) is hyper-efficient under some regularity conditions on the density and that asymptotically the estimation problem is equivalent to that of a non-stationary process with unknown center of non-stationarity. We have obtained here similar results for a family of densities $f(x, \vartheta)$ which are continuous with a cusp at the point ϑ. In this connection, it is worth noting that Daniels [2] has obtained a modified MLE for the family of densities $f(x, \vartheta) = C(\lambda) \exp\{-|x - \vartheta|^\lambda\}$, for λ such that $\frac{1}{2} < \lambda < 1$, where $C(\lambda)$ is a constant depending on λ and he has shown that this estimator is asymptotically efficient. In this paper, we shall show that the MLE of ϑ is hyper-efficient for the family of densities $f(x, \vartheta)$ given by

$$(1.1) \qquad \begin{aligned} \log f(x, \vartheta) &= \epsilon(x, \vartheta)|x - \vartheta|^\lambda + g(x, \vartheta) && \text{for} \quad |x| \leqq A \\ &= g(x, \vartheta) && \text{for} \quad |x| > A \end{aligned}$$

where A is a constant greater than zero,

$$(1.2) \qquad \begin{aligned} \epsilon(x, \vartheta) &= \beta(\vartheta) && \text{if} \quad x < \vartheta \\ &= \gamma(\vartheta) && \text{if} \quad x > \vartheta, \end{aligned}$$

$$(1.3) \qquad 0 < \lambda < \tfrac{1}{2} \quad \text{and}$$

$$(1.4) \qquad \vartheta \,\varepsilon\, (a, b) \quad \text{where} \quad -A < a < b < A,$$

under some regularity conditions on $f(x, \vartheta)$ and we shall derive the asymptotic distribution of the MLE implicitly.

Section 2 contains the regularity conditions imposed on the family of densities $f(x, \vartheta)$. Section 3 contains some results related to the asymptotic properties of the MLE. The estimation problem is reduced to that of a stochastic process in Section 4. The asymptotic distribution of MLE is obtained in Section 5.

2. Regularity conditions. We shall assume that the following regularity conditions are satisfied by $f(x, \vartheta)$.

Received 20 January 1967.

[1] This forms a part of the author's dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in Statistics of Michigan State University. Prepared with the partial support of the Office of Naval Research, Contract Nonr-2587(02) and with the partial support of the National Science Foundation, GP-7454. Reproduction in whole or in part is permitted for any purpose of the United States Government.

(2.1) For each $\vartheta \neq \vartheta_0 \varepsilon [a, b]$, there exists a $\delta(\vartheta, \vartheta_0) > 0$ such that

$$E_{\vartheta_0} [\text{Sup} \{\log f(x, \varphi) - \log f(x, \vartheta_0) : |\varphi - \vartheta| \leqq \delta(\vartheta, \vartheta_0)\}]$$

is less than zero.

(2.2) For every ϑ, ϑ_0 in $[a, b]$, $\partial g(x, \vartheta)/\partial\vartheta$, $\partial^2 g(x, \vartheta)/\partial\vartheta^2$ exist,

$$E_{\vartheta_0}[|[\partial g(x, \vartheta)/\partial\vartheta]_{\vartheta=\vartheta_0}|] < \infty \quad \text{and} \quad E_{\vartheta_0}[|\partial^2 g(x, \vartheta)/\partial\vartheta^2|] \leqq K_1(\vartheta_0) < \infty,$$

for some constant $K_1(\vartheta_0)$ depending on ϑ_0.

(2.3) For every $\vartheta_0 \varepsilon [a, b]$, $E_{\vartheta_0}\{[\partial \log f(x, \vartheta)/\partial\vartheta]_{\vartheta=\vartheta_0}\} = 0$.

(2.4) For every $\vartheta_0 \varepsilon [a, b]$, $|f(x, \vartheta_0) - f(\vartheta_0, \vartheta_0)| \leqq K_2(\vartheta_0)|x - \vartheta_0|^\lambda$ for all $x \varepsilon [-A, A]$, where $K_2(\vartheta_0)$ is a constant depending on ϑ_0.

(2.5) $\beta(\vartheta)$ and $\gamma(\vartheta)$ are differentiable twice at all ϑ with bounded second derivatives.

3. Some asymptotic properties of MLE. Since our interest centers around obtaining the limiting distribution of the MLE of ϑ, we can assume, without loss of generality, that the true value of ϑ is $\vartheta_0 = 0$.

Let $K_1(\vartheta_0) = K_1$ and $K_2(\vartheta_0) = K_2$. Let x_i, $1 \leqq i \leqq n$, be independent observations from $f(x, \vartheta_0)$. Let $\hat{\vartheta}_n$ denote the MLE of ϑ_0.

LEMMA 3.1. *$\hat{\vartheta}_n$ is strongly consistent under the condition* (2.1).

PROOF. Let S_ϑ denote the interval $(\vartheta - \delta_\vartheta, \vartheta + \delta_\vartheta)$ where $\delta_\vartheta = \delta(\vartheta, 0)$ is given by (2.1). Let

$$L_K(\vartheta) = \textstyle\sum_{i=1}^K \log f(x_i, \vartheta). \tag{3.1}$$

Choose any $\eta > 0$ and define $\Omega = [a, b] \cap [\vartheta : |\vartheta| \geqq \eta]$. Since Ω is compact and $\bigcup\{S_\vartheta : \vartheta \varepsilon \Omega\} \supset \Omega$, there exists a finite set $\vartheta_1, \vartheta_2, \cdots, \vartheta_M$ in Ω such that $\bigcup_{i=1}^M S_i \supset \Omega$ where $S_i = S_{\vartheta_i}$.

It follows from (2.1) by the strong law of large numbers that for any $\epsilon > 0$ there exist integers $N(\vartheta_i, \epsilon)$ such that for $1 \leqq i \leqq M$, and $n \geqq \max_i N(\vartheta_i, \epsilon)$,

$$P_0[\textstyle\bigcup_{k \geqq n} \{\sum_{j=1}^K \{\sup_{\vartheta \varepsilon S_i} (\log f(x_j, \vartheta) - \log f(x_j, 0))\} < 0\}] > 1 - \epsilon/M. \tag{3.2}$$

Now

$$\begin{aligned} P_0[\textstyle\bigcup_{k \geqq n} \{|\hat{\vartheta}_k| \geqq \eta\}] & \\ &\leqq P_0[\textstyle\bigcup_{k \geqq n} \{\sup_{\vartheta \varepsilon \Omega} L_k(\vartheta) > L_k(0)\}] \\ &\leqq \textstyle\sum_{i=1}^M P_0[\bigcup_{k \geqq n} \{\sup_{\vartheta \varepsilon S_i} L_k(\vartheta) > L_k(0)\}] \\ &\leqq M \cdot \epsilon/M = \epsilon \quad \text{from} \quad (3.2). \end{aligned}$$

This establishes the strong consistency of $\hat{\vartheta}_n$. □

Let us now consider the log-likelihood ratio

$$L_n(\vartheta) - L_n(0) = \sum_{i=1}^{*n} [\epsilon(x_i, \vartheta)|x_i - \vartheta|^\lambda - \epsilon(x_i, 0)|x_i|^\lambda] + \sum_{i=1}^n [g(x_i, \vartheta) - g(x_i, 0)]$$

where $\sum^*$ denotes that the sum is extended over those x_i for which $|x_i| \leqq A$. We shall now prove some lemmas which lead to the calculation of $E_0[L_n(\vartheta) - L_n(0)]$, $\text{Var}_0 [L_n(\vartheta) - L_n(0)]$ and $\text{Var}_0 [L_n(\vartheta) - L_n(\varphi)]$.

Let us define

$$\text{(3.3)} \qquad \Psi(x, \vartheta) = \epsilon(x, \vartheta)|x - \vartheta|^\lambda - \epsilon(x, 0)|x|^\lambda \quad \text{for} \quad |x| \leqq A$$
$$= 0 \quad \text{for} \quad |x| > A,$$

$$\text{(3.4)} \qquad \beta = \beta(0), \qquad \gamma = \gamma(0), \qquad f = f(0, 0),$$

$$\text{(3.5)} \qquad \Phi(x, \vartheta) = \lambda\vartheta\epsilon(x, 0) \operatorname{Sgn} x \, |x|^{\lambda-1} \quad \text{for} \quad |x| \leqq A$$
$$= 0 \quad \text{for} \quad |x| > A,$$

and

$$\text{(3.6)} \quad C = \Gamma(\lambda + 1)\Gamma(\tfrac{1}{2} - \lambda)[2^{2\lambda+1}\pi^{\frac{1}{2}}(2\lambda + 1)]^{-1}[\beta^2 + \gamma^2 - 2\beta\gamma \cos \pi\lambda].$$

Specifically, we shall prove the following results about the log-likelihood ratio $L_n(\vartheta) - L_n(0)$.

LEMMA 3.2.

$$\text{(3.7)} \qquad E_0[L_n(\vartheta) - L_n(0)] = -nCf|\vartheta|^{2\lambda+1}[1 + o(1)]$$

where $o(1)$ is in ϑ and in general for any $\vartheta \varepsilon [a, b]$,

$$\text{(3.8)} \qquad E_0[L_n(\vartheta) - L_n(0)] \leqq -nH|\vartheta|^{2\lambda+1}$$

where H is a constant independent of ϑ and n.

LEMMA 3.3.

$$\text{(3.9)} \qquad \text{Var}_0[L_n(\vartheta) - L_n(0)] = 2nCf|\vartheta|^{2\lambda+1}[1 + o(1)]$$

where $o(1)$ is in ϑ and in general for any ϑ and $\varphi \varepsilon [a, b]$

$$\text{(3.10)} \qquad \text{Var}_0[L_n(\vartheta) - L_n(\varphi)] \leqq nQ|\vartheta - \varphi|^{2\lambda+1}$$

where Q is a constant independent of ϑ, φ and n.

We shall now prove some results which lead to the above two lemmas.

LEMMA 3.4.

$$\text{(3.11)} \qquad E_0[\Psi(x, \vartheta) - \Psi(x, \varphi)]^2 = 2Cf|\vartheta - \varphi|^{2\lambda+1}[1 + o(1)]$$

as $|\vartheta| \to 0$ and $|\varphi| \to 0$ and in general for any ϑ and φ in $[a, b]$,

$$\text{(3.12)} \qquad E_0[\Psi(x, \vartheta) - \Psi(x, \varphi)]^2 \leqq B|\vartheta - \varphi|^{2\lambda+1}$$

where B is a constant independent of ϑ, φ.

PROOF. Let us assume without loss of generality that $\vartheta > \varphi$ and let $\eta = \vartheta - \varphi$. Now

$$(3.13)\quad E_0[\Psi(x, \vartheta) - \Psi(x, \varphi)]^2 = \int_{-A}^{A} \{\epsilon(x, \vartheta)|x - \vartheta|^\lambda - \epsilon(x, \varphi)|x - \varphi|^\lambda\}^2 f(x, 0)\, dx.$$

Let us now define

$$(3.14)\quad T_1 = \int_{-A}^{A} \{\epsilon(x, \vartheta)|x - \vartheta|^\lambda - \epsilon(x, \varphi)|x - \varphi|^\lambda]^2 f(0, 0)\, dx,$$

$$(3.15)\quad T_2 = \int_{-A}^{A} [\epsilon(x, \vartheta)|x - \vartheta|^\lambda - \epsilon(x, \varphi)|x - \varphi|^\lambda]^2 [f(x, 0) - f(0, 0)]\, dx,$$

$$(3.16)\quad T_3 = \int_{-A}^{A} [\epsilon(x - \eta, \varphi)|x - \vartheta|^\lambda - \epsilon(x, \varphi)|x - \varphi|^\lambda]^2 |x|^\lambda\, dx,$$

$$(3.17)\quad T_4 = \int_{-A}^{A} [\epsilon(x - \eta, \varphi) - \epsilon(x, \vartheta)]^2 |x - \vartheta|^{2\lambda} |x|^\lambda\, dx,$$

$$(3.18)\quad T_5 = \int_{-A}^{A} [\epsilon(x - \eta, \varphi)|x - \vartheta|^\lambda - \epsilon(x, \varphi)|x - \varphi|^\lambda]^2 f\, dx,$$

$$(3.19)\quad T_6 = \int_{-A}^{A} [\epsilon(x, \vartheta) - \epsilon(x - \eta, \varphi)]^2 |x - \vartheta|^{2\lambda} f(0, 0)\, dx.$$

Since ϑ belongs to a finite interval $[a, b]$ and since $\beta(\vartheta)$ and $\gamma(\vartheta)$ have bounded derivatives, it follows that

$$(3.20)\quad T_4 = \eta^2 O(1), \qquad T_6 = \eta^2 O(1).$$

Let us now consider T_5. We have

$$\begin{aligned} T_5 &= f\eta^{2\lambda+1} \int_{-A/\eta}^{A/\eta} [\epsilon(\eta y - \eta, \varphi)|y - \vartheta/\eta|^\lambda - \epsilon(\eta y, \varphi)|y - \varphi/\eta|^\lambda]^2\, dy \\ &= f\eta^{2\lambda+1} \int_{B_1}^{B_2} [\epsilon(\eta z + \varphi - \eta, \varphi)|z - 1|^\lambda - \epsilon(\eta z + \varphi, \varphi)|z|^\lambda]^2\, dz \\ &\qquad \text{where } B_1 = -(A + \varphi)/\eta \quad \text{and} \quad B_2 = (A - \varphi)/\eta, \\ &= f\eta^{2\lambda+1} \int_{-\infty}^{\infty} [h(z - 1)|z - 1|^\lambda - h(z)|z|^\lambda]^2\, dz \\ &\quad - f\eta^{2\lambda+1} \int_{-\infty}^{B_1} [h(z - 1)|z - 1|^\lambda - h(z)|z|^\lambda]^2\, dz \\ &\quad - f\eta^{2\lambda+1} \int_{B_2}^{\infty} [h(z - 1)|z - 1|^\lambda - h(z)|z|^\lambda]^2\, dz \end{aligned}$$

where

$$\begin{aligned} h(z) &= \beta(\varphi) \quad \text{if} \quad z < 0 \\ &= \gamma(\varphi) \quad \text{if} \quad z \geqq 0. \end{aligned}$$

It was shown in Prakasa Rao [4] that

$$\int_{-\infty}^{\infty} [h(z - 1)|z - 1|^\lambda - h(z)|z|^\lambda]^2\, dz = 2C(\varphi)$$

where

$$C(\varphi) = \Gamma(\lambda + 1)\Gamma(\tfrac{1}{2} - \lambda)[2^{2\lambda+1}\pi^{\frac{1}{2}}(2\lambda + 1)]^{-1} \cdot [\beta^2(\varphi) + \gamma^2(\varphi) - 2\beta(\varphi)\gamma(\varphi) \cos \pi\lambda].$$

Therefore,

$$(3.21)\quad T_5 = 2C(\varphi)f\eta^{2\lambda+1} - f\eta^{2\lambda+1}\int_{-\infty}^{B_1}[h(z-1)|z-1|^\lambda - h(z)|z|^\lambda]^2\,dz - f\eta^{2\lambda+1}\int_{B_2}^{\infty}[h(z-1)|z-1|^\lambda - h(z)|z|^\lambda]^2\,dz.$$

As a consequence of (3.21), we find that for any ϑ and φ in $[a, b]$,

$$(3.22)\qquad T_5 = \eta^{2\lambda+1}O(1)$$

and as $\vartheta \to 0$, and $\varphi \to 0$,

$$(3.23)\qquad T_5 = 2Cf\eta^{2\lambda+1}[1 + o(1)].$$

In a similar way, it can be shown that for any ϑ and φ in $[a, b]$,

$$(3.24)\qquad T_3 = \eta^{2\lambda+1}O(1),$$

and as $\vartheta \to 0$, and $\varphi \to 0$, by bounded convergence theorem, that

$$(3.25)\qquad T_3 = \eta^{2\lambda+1}o(1).$$

Since $|T_1^{\frac{1}{2}} - T_5^{\frac{1}{2}}| \leqq T_6^{\frac{1}{2}}$, (3.20) and (3.22) imply that

$$(3.26)\qquad T_1 = T_5 + \eta^{\lambda+3/2}O(1) + \eta^2 O(1).$$

Let us now consider the general case when ϑ and φ are any two numbers in $[a, b]$. Now combining all the previous results, we have

$$(3.27)\qquad \begin{aligned} E_0[\Psi(x,\vartheta) - \Psi(x,\varphi)]^2 &= T_1 + T_2 \\ &\leqq T_5 + \eta^{\lambda+3/2}O(1) + \eta^2 O(1) + 2K_2(T_3 + T_4) \\ &\leqq \eta^{2\lambda+1}[\eta^{1-2\lambda}O(1) + \eta^{\frac{1}{2}-\lambda}O(1) + O(1)] \\ &= \eta^{2\lambda+1}O(1) \end{aligned}$$

since $0 < \lambda < \frac{1}{2}$ and η is bounded. This establishes (3.12). Let us now suppose that ϑ and φ approach zero. Now

$$\begin{aligned} E_0[\Psi(x,\vartheta) - \Psi(x,\varphi)]^2 &= T_1 + T_2 \\ &= T_5 + \eta^{\lambda+3/2}O(1) + \eta^2 O(1) + \eta^{2\lambda+1}o(1) \\ &= 2Cf\eta^{2\lambda+1}[1 + o(1)] \end{aligned}$$

since $0 < \lambda < \frac{1}{2}$. This establishes (3.11). □

LEMMA 3.5. *For any* $\vartheta \in [a, b]$,

$$(3.28)\quad E_0[\Psi(x,\vartheta) + \Phi(x,\vartheta)] = -Cf|\vartheta|^{2\lambda+1}[1 + o(1)] + \int_{-A}^{A}[\epsilon(x,\vartheta) - \epsilon(x-\vartheta, 0)]|x-\vartheta|^\lambda f(x, 0)\,dx.$$

Proof of this lemma is omitted for the reason that it runs along the same lines as that of the previous lemma and that it is too lengthy. Details can be found in Prakasa Rao [4].

PROOF OF LEMMA 3.2. Let us assume that $\vartheta > 0$ without loss of generality.

We have

$$
\begin{aligned}
E_0[L_n(\vartheta) - L_n(0)] &= nE_0[\log f(x, \vartheta) - \log f(x, 0)] \\
&= nE_0[\Psi(x, \vartheta)] + nE_0[g(x, \vartheta) - g(x, 0)] \qquad (3.29)\\
&= nE_0[\Psi(x, \vartheta)] + n\vartheta E_0[g'(x, 0)] \\
&\quad + n\vartheta^2 E_0[\textstyle\int_0^1 (1-t) g''(x, \vartheta t)\, dt]
\end{aligned}
$$

where $g'(x, \vartheta) = \partial g(x, \vartheta)/\partial\vartheta$ and $g''(x, \vartheta) = \partial^2 g(x, \vartheta)/\partial\vartheta^2$.

Let

$$
(3.30) \quad T_7 = E_0[\Psi(x, \vartheta) + \Phi(x, \vartheta)] - \int_{-A}^{A} [\epsilon(x, \vartheta) - \epsilon(x - \vartheta, 0)]|x - \vartheta|^{\lambda} f(x, 0)\, dx,
$$

and

$$
(3.31) \quad T_8 = E_0[\vartheta g'(x, 0) - \Phi(x, \vartheta)] + \int_{-A}^{A} [\epsilon(x, \vartheta) - \epsilon(x - \vartheta, 0)]\, |x - \vartheta|^{\lambda} f(x, 0)\, dx.
$$

By the conditions given in Section 2,

$$
(3.32) \qquad E_0[L_n(\vartheta) - L_n(0)] = nT_7 + nT_8 + n\vartheta^2 O(1).
$$

Since $E_0[\partial \log f(x, \vartheta)/\partial\vartheta|_{\vartheta=0}] = 0$,

$$
\int_{-A}^{A} \{\epsilon'(x, 0)|x|^{\lambda} - \lambda\epsilon(x, 0) \operatorname{Sgn} x\, |x|^{\lambda-1}\} f(x, 0)\, dx + \int_{-\infty}^{\infty} g'(x, 0) f(x, 0)\, dx = 0.
$$

This implies that

$$
(3.33) \qquad E_0[\vartheta g'(x, 0) - \Phi(x, \vartheta)] = -\vartheta \int_{-A}^{A} \epsilon'(x, 0)|x|^{\lambda} f(x, 0)\, dx.
$$

Combining (3.31) and (3.33), we have

$$
T_8 = \int_{-A}^{A} \{[\epsilon(x, \vartheta) - \epsilon(x - \vartheta, 0)]|x - \vartheta|^{\lambda} - \vartheta\epsilon'(x, 0)|x|^{\lambda}\} f(x, 0)\, dx
$$

which implies that

$$
(3.34) \qquad T_8 = \vartheta^2 O(1)
$$

since ϑ is bounded and $\beta(\vartheta)$, $\gamma(\vartheta)$ have bounded second derivatives.

(3.32), (3.34) together with Lemma 3.5 establish (3.7). (3.8) can be established using (3.7) under the condition 2.1. □

PROOF OF LEMMA 3.3. Let us first compute $\operatorname{Var}_0[L_n(\vartheta) - L_n(0)]$ for ϑ approaching zero. Now

$$
\begin{aligned}
\operatorname{Var}_0[L_n(\vartheta) - L_n(0)] &= n \operatorname{Var}_0[\log f(x, \vartheta) - \log f(x, 0)] \\
&= nE_0[\log f(x, \vartheta) - \log f(x, 0)]^2 \\
&\quad - n\{E_0[\log f(x, \vartheta) - \log f(x, 0)]\}^2.
\end{aligned}
$$

It is easily seen from Lemmas 3.4 and 3.5 and conditions (2.2)–(2.5) that as $\vartheta \to 0$,

$$E_0[\log f(x, \vartheta) - \log f(x, 0)]^2 = 2Cf|\vartheta|^{2\lambda+1}[1 + o(1)],$$

and

$$E_0[\log f(x, \vartheta) - \log f(x, 0)] = -Cf|\vartheta|^{2\lambda+1}[1 + o(1)].$$

Therefore we have for $\vartheta \to 0$,

$$\mathrm{Var}_0[L_n(\vartheta) - L_n(0)] = 2nCf|\vartheta|^{2\lambda+1}[1 + o(1)].$$

Let us now compute $\mathrm{Var}_0[L_n(\vartheta) - L_n(\varphi)]$ for any ϑ, φ in $[a, b]$. Now by Lemma 3.4,

$$\begin{aligned}\mathrm{Var}_0[L_n(\vartheta) - L_n(\varphi)] &\leqq nE_0[\log f(x, \vartheta) - \log f(x, \varphi)]^2 \\ &\leqq 2nE_0[\Psi(x, \vartheta) - \Psi(x, \varphi)]^2 \\ &\qquad + 2nE_0[g(x, \vartheta) - g(x, \varphi)]^2 \\ &\leqq 2n[B|\vartheta - \varphi|^{2\lambda+1} + |\vartheta - \varphi|^2 O(1)] \\ &\leqq nQ|\vartheta - \varphi|^{2\lambda+1}\end{aligned}$$

for some constant Q since $0 < \lambda < \frac{1}{2}$ and ϑ, φ belong to a finite interval. □

We shall now prove a theorem by means of Lemmas 3.2, 3.3 which enables us to conclude that the probability, that the log-likelihood ratio $L_n(\vartheta) - L_n(0)$ attains its maximum outside the interval $[-\tau n^{-\rho}, \tau n^{-\rho}]$, approaches zero for sufficiently large n as $|\tau| \to \infty$ where $\rho = (1 + 2\lambda)^{-1}$. More precisely,

THEOREM 3.6. *There exists* $\eta > 0$ *such that*

$$\lim_{\tau\to\infty} \limsup_n P_0[\sup_{|\vartheta|>\tau n^{-\rho}}\{M_n(\vartheta)/n|\vartheta|^{2\lambda+1}\} \geqq -\eta] = 0 \tag{3.35}$$

where

$$M_n(\vartheta) = L_n(\vartheta) - L_n(0). \tag{3.36}$$

PROOF. Since $M_n(\vartheta)$ is continuous in ϑ, it is sufficient to prove that there exists an $\eta > 0$ such that

$$\lim_{\tau\to\infty} \limsup_n P_0[\sup_{|\vartheta_{ijk}|>\tau n^{-\rho}}\{M_n(\vartheta_{ijk})/n|\vartheta_{ijk}|^{2\lambda+1}\} \geqq -\eta] = 0. \tag{3.37}$$

where the set $\{\vartheta_{ijk}\}$ is dense in the set $\{\vartheta: |\vartheta| > \tau n^{-\rho}\}$.

Let $\vartheta_{ijk} = \tau n^{-\rho} 2^{i+k2^{-j}}$ for $i \geqq 0$, $j \geqq 0$, and $0 \leqq k < 2^j$. Obviously the set $\{\vartheta_{ijk}\}$ is dense in $\{\vartheta: |\vartheta| > \tau n^{-\rho}\}$.

Let $\zeta = 2\lambda + 1$. Let us now define

$$T_n(\vartheta_{ijk}) = M_n(\vartheta_{ijk}) - E_0[M_n(\vartheta_{ijk})] - nH\vartheta_{i00}^{\zeta} \tag{3.38}$$

where H is given by the Lemma 3.2. We observe that

$$\text{(3.39)}\quad \begin{aligned} &\text{(i) } E_0[T_n(\vartheta_{i00})] = -H\tau^{\zeta}2^{i\zeta}, \\ &\text{(ii) } E_0[T_n(\vartheta_{i,j,2k+1}) - T_n(\vartheta_{i,j-1,k})] = 0, \qquad \text{for} \quad 0 \leqq k < 2^{j-1} - 1. \\ &\text{(iii) } \mathrm{Var}_0[T_n(\vartheta_{i00})] = \mathrm{Var}_0[M_n(\vartheta_{i00})] \leqq Q\tau^{\zeta}2^{i\zeta}, \\ &\text{(iv) } \mathrm{Var}_0[T_n(\vartheta_{i,j,2k+1}) - T_n(\vartheta_{i,j-1,k})] \leqq Q\tau^{\zeta}(2\log 2)^{\zeta}2^{(i-j)\zeta}. \end{aligned}$$

The last two inequalities are obtained from Lemma 3.3 after some manipulation. We now observe that for any $0 < \eta < H$,

$$\text{(3.40)}\quad P_0[\sup_{\vartheta_{ijk} > \tau n^{-\rho}}\{M_n(\vartheta_{ijk})/n\vartheta_{ijk}^{\zeta}\} \geqq -\eta] \leqq P_0[\sup_{\vartheta_{ijk} > \tau n^{-\rho}}\{T_n(\vartheta_{ijk})/n\vartheta_{i00}^{\zeta}\} \geqq -\eta].$$

Let $0 < \xi < H$, $p_j = 2^{-\lambda j/2}\delta$ where δ is chosen so that $0 < \delta < \xi(2^{\lambda/2} - 1)$ and let $\eta = \xi - \delta(2^{\lambda/2} - 1)^{-1}$. We notice that

$$\begin{aligned} P_0[\sup_{\vartheta_{ijk} > \tau n^{-\rho}}&\{T_n(\vartheta_{ijk})/n\vartheta_{i00}^{\zeta}\} \geqq -\eta] \\ &\leqq \textstyle\sum_{i=0}^{\infty} P_0[T_n(\vartheta_{i00}) \geqq -n\xi\vartheta_{i00}^{\zeta}] \\ &\quad + \textstyle\sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\sum_{k=0}^{2^{j-1}-1} P_0\{T_n(\vartheta_{i,j,2k+1}) - T_n(\vartheta_{i,j-1,k}) \geqq np_j\vartheta_{i00}^{\zeta}\} \\ &\leqq Q\tau^{-\zeta}(1 - 2^{-\zeta})^{-1}[(H - \xi)^{-2} + \delta^{-2}(2\log 2)^{\zeta}(2^{\lambda+1} - 2)^{-1}]. \end{aligned}$$

This inequality is obtained by Chebyshev's inequality using (3.39) after some manipulation. Therefore (3.40) implies that

$$\lim_{\tau\to\infty} \limsup_n P_0[\sup_{\vartheta_{ijk} > \tau n^{-\rho}}\{M_n(\vartheta_{ijk})/n\vartheta_{ijk}^{\zeta}\} \geqq -\eta] = 0.$$

In an analogous way, one can establish that

$$\lim_{\tau\to\infty} \limsup_n P_0[\sup_{\vartheta_{ijk} < -\tau n^{-\rho}}\{M_n(\vartheta_{ijk})/n\vartheta_{ijk}^{\zeta}\} \geqq -\eta] = 0.$$

These two results together establish (3.37) which in turn proves the theorem.

4. Reduction to a problem in stochastic processes. We shall now reduce the problem of determining the asymptotic distribution of the MLE $\hat{\vartheta}_n$ to the problem of determining the distribution of the location of the maximum of a Gaussian process.

In view of Theorem 3.6, the log-likelihood ratio $M_n(\vartheta)$ has a maximum in the interval $[-\tau n^{-\rho}, \tau n^{-\rho}]$ with probability approaching one for large τ. For any such $\tau > 0$, let

$$\text{(4.1)}\qquad X_n(t) = M_n(tn^{-\rho}) \quad \text{for} \quad t \in [-\tau, \tau]$$

and let X be the continuous normal process on $[-\tau, \tau]$ with

$$\text{(4.2)}\qquad \begin{aligned} &\text{(i)} \qquad E[X(t)] = -Cf|t|^{\zeta}, \\ &\text{(ii) } \mathrm{Cov}[X(t_1), X(t_2)] = Cf[|t_1|^{\zeta} + |t_2|^{\zeta} - |t_1 - t_2|^{\zeta}]. \end{aligned}$$

Let

$$\text{(4.3)}\qquad A_n(t) = X_n(t) - E[X_n(t)],$$

and

$$A(t) = X(t) - E[X(t)]. \tag{4.4}$$

LEMMA 4.1. *For any* $t \in [-\tau, \tau]$, $A_n(t)$ *is asymptotically normal with mean* 0 *and Variance* $2Cf|t|^\zeta$.

PROOF. Notice that $A_n(t) = B_n(t) + C_n(t)$, where

$$B_n(t) = \sum_{i=1}^n \Psi(X_i, \vartheta) - nE_0[\Psi(X, \vartheta)], \tag{4.5}$$

$$C_n(t) = \sum_{i=1}^n \{g(X_i, \vartheta) - g(X_i, 0)\} - nE_0[g(X, \vartheta) - g(X, 0)], \tag{4.6}$$

and $\vartheta = tn^{-\rho}$. It is easily seen that $E_0[C_n(t)] = 0$ and $\text{Var}_0[C_n(t)] \to 0$ as $n \to \infty$ since $0 < \lambda < \frac{1}{2}$. In other words, $C_n(t)$ converges to zero in probability as $n \to \infty$. Let $F_n(x)$ be the distribution function of $B_n(t)$ and let $\Phi(x)$ be the distribution function of standard normal random variable. By the normal approximation theorem in Loève [3], it follows that

$$|F_n(x) - \Phi(x)| \leqq nC_0\{\text{Var}_0[B_n(t)]\}^{-3/2}E_0|Y - E(Y)|^3$$

where C_0 is an universal constant and $Y = \Psi(X, \vartheta)$. It is known that $\text{Var}_0[B_n(t)] = 2Cf|t|^\zeta[1 + o(1)]$ by Lemma 3.3 and it can be easily shown in a similar way that

$$E_0|Y - E(Y)|^3 \leqq C_1|\vartheta|^{3\lambda+1}$$

for some constant C_1 independent of ϑ. Therefore,

$$|F_n(x) - \Phi(x)| \leqq C_0C_1|t|^{3\lambda+1}\{2Cf|t|^\zeta(1 + o(1))\}^{-3/2}n^{-\lambda/1+2\lambda}.$$

Since the right hand term tends to zero as $n \to \infty$, we have established that $B_n(t)$ is asymptotically normal with mean 0 and Variance $2Cf|t|^\zeta$. Since $C_n(t)$ converges to zero in probability, Slutsky's theorem implies the required result. □

REMARK. It can be shown in a similar manner by normal approximation theorem that for any collection a_i, $1 \leqq i \leqq k$, and $t_1, t_2, \cdots, t_k$ in the interval $[-\tau, \tau]$, $\sum_{i=1}^k a_i A_n(t_i)$ is asymptotically normal with mean 0 and Variance

$$Cf[\sum_{i=1}^k \sum_{j=1}^k a_i a_j\{|t_i|^\zeta + |t_j|^\zeta - |t_i - t_j|^\zeta\}].$$

The next theorem shows that the processes A_n on $[-\tau, \tau]$ satisfy an equicontinuity condition.

THEOREM 4.2. *For any* t_1, t_2 in $[-\tau, \tau]$,

$$E_0|A_n(t_1) - A_n(t_2)|^2 \leqq Q|t_1 - t_2|^\zeta$$

where Q *is a constant independent of* n, t_1, t_2 *and* $\zeta = 2\lambda + 1$.

PROOF. This theorem is an immediate consequence of Lemma 3.3, since

$$\begin{aligned} E_0|A_n(t_1) - A_n(t_2)|^2 &= \text{Var}_0[M_n(\vartheta_1) - M_n(\vartheta_2)] \\ &\leqq nQ|\vartheta_1 - \vartheta_2|^\zeta \\ &= Q|t_1 - t_2|^\zeta. \end{aligned}$$

□

We shall now state a theorem due to Prohorov connected with convergence of distributions of stochastic processes on $C[a, b]$, where $C[a, b]$ denotes the space of continuous functions on $[a, b]$. Let us endow $C[a, b]$ with topology generated by supremum norm. We refer to Sethuraman [7] for a proof of the theorem.

THEOREM 4.3. *Let X_n be a sequence of stochastic processes on $C[a, b]$ and X be another process on $C[a, b]$ such that*

(i) *for any $t_i \varepsilon [a, b]$, $1 \leqq i \leqq k$, the joint distribution of $[X_n(t_1), \cdots, X_n(t_k)]$ converges weakly to the joint distribution of $[X(t_1), \cdots, X(t_k)]$ and*

(ii) *there exists constants $A_1 > 0$, $A_2 > 0$, $A_3 > 0$ independent of n, t_1, and t_2 such that*

$$E|X_n(t_1) - X_n(t_2)|^{A_1} < A_3|t_1 - t_2|^{1+A_2}.$$

Let μ_n and μ be the distributions induced by X_n and X respectively on the σ-field of Borel subsets of $C[a, b]$. Then μ_n converges to μ weakly. In other words X_n converges in distribution to X.

In view of Lemma 4.1, Theorem 4.2, and the remarks made at the end of Lemma 4.1, follows that the processes A_n on $[-\tau, \tau]$ converge in distribution to the process A on $[-\tau, \tau]$ by Theorem 4.3 since the trajectories of the processes A_n and A belong to $C[-\tau, \tau]$. Since $E[X_n(t)]$ converges to $E[X(t)]$ uniformly for $t \varepsilon [-\tau, \tau]$, the following theorem can be obtained by an extension of Slutsky's theorem generalized to processes (Rubin [6]).

THEOREM 4.4. *The processes X_n converge in distribution to the process X.*

For any $x \varepsilon C[-\tau, \tau]$, let us denote by $g(x)$ the value of t for which $x(t)$ is maximum over $[-\tau, \tau]$. It is easily seen that $x_n \to x$ and x has a unique maximum imply that $g(x_n) \to g(x)$. Since the process X is continuous, the set of discontinuities of g has measure zero with respect to the measure induced by the process X on the space $C[-\tau, \tau]$. Therefore, it follows from a theorem of Rubin [6], that $g(X_n)$ converges in law to $g(x)$. Hence we have the following theorem.

THEOREM 4.5. *The distribution of the location of the maximum of log-likelihood ratio $M_n(\vartheta)$ over $[-\tau n^{-\rho}, \tau n^{-\rho}]$ converges weakly to the distribution of the location of the maximum of the non-stationary Gaussian process X over $[-\tau, \tau]$ defined in (4.2).*

The next theorem proves that the Gaussian process X over $(-\infty, \infty)$, with mean function and covariance function given by (4.2), has its maximum in a finite interval $[-\tau, \tau]$ with probability approaching one as $|\tau| \to \infty$. More precisely,

THEOREM 4.6.

$$P[\limsup_{|\tau|\to+\infty} \{X(\tau)/Cf|\tau|^{\zeta}\} \leqq -1] = 1.$$

PROOF. Define

$$\text{(i)}\quad A(\tau) = X(\tau) + Cf|\tau|^{\zeta},$$

$$\text{(4.7)}\qquad \text{(ii)}\quad Z_0 = \sup\{A(\tau) : 1 \leqq \tau \leqq 2\},$$

$$\text{(iii)}\quad Z_n = \sup\{A(\tau) : 2^n \leqq \tau \leqq 2^{n+1}\},$$

$$\text{(iv)}\quad U = \sup\{|A(\tau) - A(1)| : 1 \leqq \tau \leqq 2\}.$$

Since the process A is separable, Z_0, Z_n and U are random variables. Since A is a continuous process with probability one,

$$U = \sup\{T_j : j \geqq 0\} \text{ a.s.} \tag{4.8}$$

where

$$T_j = \sup\{|A((s-1)2^{-j}) - A(s2^{-j})| : 2^j + 1 \leqq s \leqq 2^{j+1}\}. \tag{4.9}$$

Now for any $a > 0$, $0 < r < 1$,

$$\begin{aligned} P[T_j > ar^j] &\leqq \textstyle\sum_{s=2^j+1}^{2^{j+1}} P\{|A((s-1)2^{-j}) - A(s2^{-j})| > ar^j\} \\ &= (2^{j+1} - 2^j)P\{|A(1) - A(1+2^{-j})| > ar^j\} \\ &\leqq Cfa^{-2}r^{-2j}2^{1-2\lambda j}. \end{aligned}$$

Therefore,

$$\begin{aligned} P[U > a(1-r)^{-1}] &\leqq \textstyle\sum_{j=0}^{\infty} P[T_j > ar^j] \\ &\leqq 2Cfa^{-2}(1 - r^{-2}2^{-2\lambda})^{-1}. \end{aligned}$$

It can easily be seen from here that there exists a constant $D > 0$ such that

$$P[U > a] \leqq Da^{-2} \tag{4.10}$$

for any $a > 0$, which implies that $E(U) < \infty$. Since $|Z_0| \leqq |A(1)| + U$, it follows that $E|Z_0| < \infty$. Let $E|Z_0| = \alpha$. Now for any $\epsilon > 0$,

$$\begin{aligned} \textstyle\sum_{n=0}^{\infty} P[Z_n > \epsilon 2^{n\zeta}] &= \textstyle\sum_{n=0}^{\infty} P[Z_0 > \epsilon 2^{n\zeta/2}] \\ &\leqq \alpha\epsilon^{-1}(1 - 2^{-\zeta/2})^{-1} < \infty, \end{aligned}$$

and hence by the Borel-Cantelli lemma

$$P[Z_n > \epsilon 2^{n\zeta} \text{ i.o.}] = 0$$

for every $\epsilon > 0$. In other words $P[\limsup Z_n 2^{-n\zeta} \leqq 0] = 1$. Since $A(\tau)\tau^{-\zeta} \leqq Z_n 2^{-n\zeta}$ provided $2^n \leqq \tau \leqq 2^{n+1}$, we have

$$P[\limsup_{\tau\to+\infty} A(\tau)\tau^{-\zeta} \leqq 0] = 1.$$

It is easily seen from this result together with a similar result as $\tau \to -\infty$ that

$$P[\limsup_{|\tau|\to+\infty} \{X(\tau)/Cf\,|\tau|^{\zeta}\} \leqq -1] = 1. \qquad \square$$

5. Asymptotic distribution of MLE. We have the following final theorem from Theorems 3.6, 4.5 and 4.6.

THEOREM 5.1. *Consider the family of densities* $f(x, \vartheta)$ *given by* (1.1)–(1.4) *satisfying the regularity conditions* (2.1)–(2.5). *Let* $\hat{\vartheta}_n$ *denote a* MLE *of* ϑ *based on* n *independent observations. Let* ϑ_0 *denote the true value of* ϑ. *Then* $n^{1/1+2\lambda}[\hat{\vartheta}_n - \vartheta_0]$ *has a limiting distribution and it is the distribution of the location of the maximum of the nonstationary Gaussian process* X *with*

$$E[X(\tau)] = -C(\vartheta_0)f(\vartheta_0, \vartheta_0)|\tau|^{2\lambda+1},$$

and

$$\mathrm{Cov}[X(\tau_1), X(\tau_2)] = C(\vartheta_0)f(\vartheta_0, \vartheta_0)[|\tau_1|^{2\lambda+1} + |\tau_2|^{2\lambda+1} - |\tau_1 - \tau_2|^{2\lambda+1}]$$

where

$$C(\vartheta_0) = \Gamma(\lambda + 1)\Gamma(\tfrac{1}{2} - \lambda)[2^{2\lambda+1}\pi^{\frac{1}{2}}(2\lambda + 1)]^{-1}[\beta^2(\vartheta_0) + \gamma^2(\vartheta_0) - 2\beta(\vartheta_0)\gamma(\vartheta_0)\cos \pi\lambda].$$

In other words, the MLE $\hat{\vartheta}_n$ is hyper-efficient estimator of ϑ_0 when $0 < \lambda < \frac{1}{2}$.

REMARK. It can be shown by analogous methods that Bayes estimators for ϑ, for smooth prior densities, are also hyper-efficient and asymptotically the Bayes estimation of ϑ is equivalent to the estimation of the location of center of non-stationarity. The problem of hyper-efficiency of MLE when the cusp is of order $\frac{1}{2}$ is being investigated and we hope that the MLE is hyper-efficient.

Acknowledgment. This work was done under the supervision and guidance of Professor Herman Rubin of Michigan State University. It is a pleasure to express my gratitude to Professor Herman Rubin for suggesting the problem and for his encouragement.

REFERENCES

[1] CHERNOFF, H. and RUBIN, H. (1955). The estimation of the location of a discontinuity in density. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 19–37. Univ. of California.

[2] DANIELS, H. (1960). The asymptotic efficiency of a maximum likelihood estimator. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 151–163. Univ. of California.

[3] LOÈVE, M. (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

[4] PRAKASA RAO, B. L. S. (1966). Asymptotic distributions in some non-regular statistical problems. Technical report No. 9, Department of Statistics and Probability, Mich. State Univ.

[5] RUBIN, H. (1960). The estimation of discontinuities in multivariate densities, and related problems in stochastic processes. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 563–574. Univ. of California.

[6] RUBIN, H. (1963). Unpublished lecture notes, Michigan State University.

[7] SETHURAMAN, J. (1965). Limit theorems for stochastic processes. Technical report No. 10, Department of Statistics, Stanford Univ.