Ann. Inst. Statist. Math.
38 (1986), Part A, 137-144

# A CHARACTERIZATION OF LIMITING DISTRIBUTIONS OF ESTIMATORS IN AN AUTOREGRESSIVE PROCESS

WEI-MIN HUANG

(Received May 21, 1984; revised Jan. 11, 1985)

## Summary

Let the random variables $X_1, X_2, \cdots, X_n$ be generated by the first-order autoregressive model $X_i=\theta X_{i-1}+e_i$ where $e_i$, $i=1, 2, \cdots, n$, are i.i.d. random variables with mean zero, variance $\sigma^2$, and with unspecified density function $g(\cdot)$. In the present paper we obtain a characterization of limiting distributions of nonparametric and parametric estimators of $\theta$ as well as a local asymptotic minimax bound of the risks of estimators.

## 1. Introduction

Let the random variables $X_1, X_2, \cdots, X_n$ be generated by the first-order autoregressive model

$$X_i=\theta X_{i-1}+e_i$$

where $e_i$, $i=1, 2, 3, \cdots, n$, are i.i.d. random variables with mean zero, variance $\sigma^2$, and with the density function $g(\cdot)$. We assume that $|\theta|<1$. In the present paper we obtain a characterization of limiting distributions of regular estimators of $\theta$ as well as a local asymptotic minimax bound of the risks of estimators of $\theta$.

Akahira [1] considered the parametric first-order autoregressive model and developed the asymptotic efficiency of the process within the class of all asymptotically median unbiased (AMU) estimators of $\theta$. Recently, Akahira and Takeuchi [2] studied higher order asymptotic efficiency of statistical estimators in parametric models, the results have been applied to the first-order autoregressive process. Also Kabaila [12] discussed the asymptotic efficiency of the AMU estimators for an autoregressive moving average process. The purpose of this paper is to investigate the asymptotic efficiency of the estimation of $\theta$ without

Key words and phrases: Autoregressive process, local likelihood ratio, adaptive estimation, asymptotic minimax bounds, representation theorem.

specifying the density function $g(\cdot)$. That is $\theta$ is the parameter of interest and the infinite dimensional parameter $g$ is a nuisance function. The density function $g$ is assumed to have certain regularity conditions and belongs to a class of density functions. The results provide a characterization of limiting distributions of estimators for $\theta$ as well as a local asymptotic minimax bound of the risks of estimators. The estimators are not restricted to AMU estimators and the efficiency result is more than to compare the limiting variances of estimators. A convolution type representation is derived for the limiting distributions of regular estimators for $\theta$, and the local asymptotic minimax bound leads to a comparison of asymptotic risks of arbitrary estimators for $\theta$.

Let us now look at the above AR(1) model from a different viewpoint. Suppose that we observe a sequence of numbers $x_1, x_2, \cdots, x_n$ which are generated by the deterministic AR(1) model, i.e. $x_i = \theta x_{i-1}$. We can easily recover the parameter $\theta$ through many different ways, for example, $\theta = x_i/x_{i-1}$ for $i=2, 3, \cdots, n$; $\theta = x_i x_{i-1}/x_i^2$ for $i=1, 2, \cdots, n$; $\theta = \sum_{i=1}^{n} x_i x_{i-1} \Big/ \sum_{n=1}^{n} x_i^2$; $\theta =$ median $\{x_i x_{i-1} | i=1, 2, \cdots, n\}$/median $\{x_i^2 | i=1, 2, \cdots, n\}$; or $\theta$ is defined to be the value minimizing $\sum_{i=1}^{n} \rho(x_i - \theta x_{i-1})$ where $\rho(t) \geqq 0$ for $t \in (-\infty, \infty)$, etc. We then can derive a class of estimators of $\theta$ by simply replacing $x_i$ with $X_i$ in all the cases. (Notice that $\sum X_i X_{i-1}/\sum X_i^2$ is the least squared estimator of $\theta$.) This intuitive way of producing estimators is of interest in itself but we will not pursue further here. Two interesting problems are the following: (1) Can we tell the difference among the above estimators when the density function $g$ is unspecified? (2) What is the information contained in the models of $g$ being fixed and specified and of $g$ being unspecified and belonging to a class of densities? These two problems are basically asking how well can we estimate $\theta$.

Recently Begun, Hall, Huang, and Wellner [3] studied the asymptotic efficiencies of estimators for general i.i.d. parametric-nonparametric models. Their work answers problems similar to (2) above in the i.i.d. cases (with or without censored data). They developed asymptotic lower bounds for estimation of the parametric and nonparametric components of mixed models. The results are based on work of LeCam [13], Hajek [7], [8], Beran [5], and Millar [15]. Inagaki [11] has also independently developed a similar convolution representation as in Hajek [7]. Representation theorems and asymptotic minimax theorems have also been established for a variety of parametric and completely nonparametric problems, see, for example, Levit [14], Beran [5], Millar [15], Ibragimov and Hasminskii [10], Pfanzagl [16], and Wellner [17]. The current paper is dealing with a mixed parametric-nonparametric (semiparametric) model with dependent observations. An important feature of semiparametric models such as AR(1) is that the parameter of interest $\theta$ can be interpreted in many ways as it is discussed in the previous paragraph; see Bickel [6] and Begun, Hall, Huang and Wellner [3] for more examples in semiparametric models. A surprising property of the AR(1) model may lead to a possible construction of adaptive estimators for $\theta$ which perform as well asymptotically as when we know the density function even without the knowledge of the density function. See Bickel [6] or Begun, Hall, Huang and Wellner [3] for a definition of adaptive estimators. Under certain regularity conditions, Beran [4] constructed a class of partially adaptive estimators of $\theta$ whose asymptotic performance will dominate that of the least squares estimators for all $g$ of interest.

For notational simplicity we only consider AR(1) model, the current results can be extended to AR($p$) model, $p \geqq 1$, as long as similar conditions as in Lemma 1.2 are imposed.

## 2. Conditioning and local asymptotic normality

Suppose that $e_1, e_2, \cdots, e_n$ are i.i.d. random variables with density function $g \in \mathcal{G}$ and $\mathcal{G}$ is the collection of all mean zero densities with respect to a sigma-finite measure $\mu$ on the real line. Let $L^2(\mu)$ denote the usual $L^2$-space of square-integrable function and let $\langle \cdot, \cdot \rangle_\mu$ and $\|\cdot\|_\mu$ denote the usual inner product and norm in $L^2(\mu)$. Thus $g^{1/2} \in L^2(\mu)$.

The random variables $X_1, X_2, \cdots, X_n$ generated by the first-order autoregressive model have joint density, by conditioning,

$$f(X; \theta, g) = f_1(X_1; \theta, g) \cdot \prod_{j=2}^{n} f_j(X_j \mid X_1, \cdots, X_{j-1}, \theta, g) \tag{1.1}$$
$$= g(X_1 - \theta X_0) \cdot \prod_{j=2}^{n} g(X_j - \theta X_{j-1})$$
$$= \prod_{j=1}^{n} g(e_j)$$

where $|\theta| < 1$, $g \in \mathcal{G}$, and $X_0 \equiv 0$. Hence the log-likelihood function is a sum of i.i.d. unobservable random variables.

Now let $\Theta(h)$ denote the collection of all sequences $\{\theta_n\}_{n \geq 1}$ such that $|\theta_n| < 1$ and $|n^{1/2}(\theta_n - \theta) - h| \to 0$ as $n \to \infty$ where $h \in R^1$, and let $\Theta \equiv U\{\Theta(h); h \in R^1\}$. Similarly, let $C(g, \beta)$ denote the collection of all sequences $\{g_n\}_{n \geq 1}$ with each $g_n \in \mathcal{G}$ such that $\|n^{1/2}(g_n^{1/2} - g^{1/2}) - \beta\|_\mu \to 0$ as $n \to \infty$ where $\beta \in L^2(\mu)$ and $\langle \beta, g^{1/2} \rangle_\mu = 0$ necessarily.

For $f_n \equiv f(X; \theta_n, g_n)$ and $f \equiv f(X; \theta, g)$ where $|\theta| < 1$, $\theta_n \in \Theta$, $g \in \mathcal{G}$, and $g_n \in C(g) \equiv U\ C(g, \beta)$ and the union is taking over all $\beta$ such that $\beta \in L^2(\mu)$ and $\langle \beta, g^{1/2} \rangle_\mu = 0$, define the local-likelihood ratio $L_n$, whenever the right side is finite, by

(1.2) $$L_n=\log\{[f(X;\theta_n,g_n)/f(X;\theta,g)]\}\ .$$

Hence we have

$$L_n=\log\{[f(X;\theta_n,g_n)/f(X;\theta_n,g)]\}+\log\{[f(X;\theta_n,g)/f(X;\theta,g)]\}$$
$$=L_n^{(1)}+L_n^{(2)}\ ,\qquad \text{say}\ .$$

The following Lemma is a consequence of (1.1) and arguments of LeCam [13] and Beran [5].

LEMMA 1.1. *For $\theta_n\in\Theta$ and $g_n\in C(g,\beta)$, we have*

$$L_n^{(1)}=2n^{-1/2}\sum_{j=1}^n\beta(X_j-\theta X_{j-1})g^{-1/2}(X_j-\theta X_{j-1})-\frac{1}{2}\sigma_1^2+o_p(1)$$

*where $\sigma_1^2=4\|\beta\|_\mu^2$. Thus, under $P_f$,*

$$L_n^{(1)}\xrightarrow{d}N\left(-\frac{1}{2}\sigma_1^2,\sigma_1^2\right)\qquad \textit{as } n\to\infty\ ,$$

*and the sequences $\{f(X;\theta_n,g_n)\}$ and $\{f(X;\theta_n,g)\}$ are contiguous. The sequences of probability measures $\{Q_n\}$ and $\{P_n\}$ (or their corresponding sequences of densities with respect to a dominating measure) on $\mathcal{A}_n$ are said to be contiguous if for any $\mathcal{A}_n$-measurable random variables $Y_n$, $Y_n\to 0$ in $Q_n$-probability if and only if $Y_n\to 0$ in $P_n$-probability.*

The following lemma is a result due to Akahira [1], and the technical assumptions were also introduced there. The lemma is restated in terms of $L_2$-framework.

LEMMA 1.2. (*Akahira* [1]) *Under the assumptions*

(i) *$g(t)$ is positive for all $t$ on $R^1$ and $\lim_{t\to\pm\infty}g(t)=0$.*

(ii) *$g$ is three times differentiable and $\lim_{t\to\pm\infty}g'(t)=0$.*

(iii) *$d^2\log g(t)/dt^2$ is a bounded function and $E(|e_i|^4)<\infty$.*

(iv) *For each $\theta$, $|\theta|<1$ and $\theta_n\in\Theta(h)$ the following hold:*

(a) $\lim_{n\to\infty}n^{-3/2}\sum_{j=1}^n E_{\theta_n}[|X_{j-1}|^3\cdot\sup_{\eta\in A_j}|b'(e_j+\eta)|]=0$;

(b) $\lim_{n\to\infty}n^{-3}E_{\theta_n}\left\{\left[\sum_{j=1}^n|X_{j-1}|^3\cdot\sup_{\eta\in A_j}|b'(e_j+\eta)|\right]^2\right\}=0,$

*where $b(t)=-d^2\log g(t)/dt^2$, $A_j=\{\eta\,|\,0<|\eta|<hn^{-1/2}|X_{j-1}|\}$.*

(v) $E[|g'(e_j)/g(e_j)|^4]<\infty$,

*then*

$$L_n^{(2)}=2n^{-1/2}\sum_{j=1}^n hX_{j-1}\cdot\rho(X_j-\theta X_{j-1})g^{-1/2}(X_j-\theta X_{j-1})-\frac{1}{2}\sigma_2^2+o_p(1)$$

*where $\sigma_2^2=4h^2\sigma^2\|\rho\|_\mu^2/(1-\theta^2)$ and $\rho(\cdot)=-2^{-1}g'(\cdot)g^{-1/2}(\cdot)\in L^2(\mu)$. Thus, under $P_f$*

$$L_n^{(2)} \xrightarrow{d} N\left(-\frac{1}{2}\sigma_2^2, \sigma_2^2\right) \qquad as\ n \to \infty ,$$

*and the sequences* $\{f(\boldsymbol{X}; \theta_n, g)\}$ *and* $\{f(\boldsymbol{X}; \theta, g)\}$ *are contiguous.*

LEMMA 1.3. (*Orthogonality of score functions*). *The score functions* $S_\theta^{(j)} \equiv X_{j-1} \cdot \rho(X_j - \theta X_{j-1}) g^{-1/2}(X_j - \theta X_{j-1})$ *and* $S_\beta^{(j)} \equiv \beta(X_j - \theta X_{j-1}) g^{-1/2}(X_j - \theta X_{j-1})$ *of* $\theta$ *and* $g$ (*in direction* $\beta$, $\beta \in L^2(\mu)$), *respectively, satisfy for every pair* $(i, j)$, $i \neq j$, $1 \leqq i, j \leqq n$,

(i) $\mathrm{Cov}\,[S_\theta^{(j)}, S_\theta^{(i)}] = 0$,
(ii) $\mathrm{Cov}\,[S_\beta^{(j)}, S_\beta^{(i)}] = 0$, *and*
(iii) $\mathrm{Cov}\,[S_\theta^{(i)}, S_\beta^{(j)}] = 0$, *this holds for every* $(i, j)$.

PROOF. These follow from $\langle g^{1/2}, \beta \rangle_\mu = 0$, mean $(g) = 0$, and the following identity, for $i < j$,

$$\begin{aligned}\mathrm{Cov}\,(S_i, S_j) = {} & E\{\mathrm{Cov}\,(S_i, S_j \mid X_1, X_2, \cdots, X_{j-1})\} \\ & + \mathrm{Cov}\,\{E(S_i \mid X_1, X_2, \cdots, X_{j-1}), E(S_j \mid X_1, X_2, \cdots, X_{j-1})\}\end{aligned}$$

where $S_i$ and $S_j$ are arbitrary random variables depending on $(X_1, \cdots, X_i)$ and $(X_1, \cdots, X_j)$ respectively.

Asymptotic normality of the local log-likelihood ratio follows from Lemmas 1.1, 1.2 and 1.3, we have

PROPOSITION 1.1. (*Local asymptotic normality*). *Suppose that* $\theta_n \in \Theta$, $g_n \in C(g)$, *and* $f \equiv f(\boldsymbol{X}; \theta, g)$, $f_n \equiv f(\boldsymbol{X}; \theta_n, g_n)$, *then* $L_n$ *in* (1.2) *satisfy for every* $\varepsilon > 0$,

$$P_f\left\{\left| L_n - 2n^{-1/2} \sum_{j=1}^{n} \alpha_j(\boldsymbol{X}) g^{-1/2}(X_j - \theta X_{j-1}) + \frac{1}{2}\delta^2 \right| > \varepsilon\right\} \to 0 \qquad as\ n \to \infty , \tag{1.3}$$

*where* $\alpha_j(\boldsymbol{X}) = hX_{j-1} \cdot \rho(X_j - \theta X_{j-1}) + \beta(X_j - \theta X_{j-1})$, $\delta^2 = 4(h^2\sigma^2 \|\rho\|_\mu^2 + \|\beta\|_\mu^2) = \sigma_1^2 + (1-\theta^2)\sigma_2^2$. *Thus, under* $P_f$,

$$L_n \xrightarrow{d} N(-2^{-1}\delta^2, \delta^2) \qquad as\ n \to \infty ,$$

*and the sequences* $f(\boldsymbol{X}; \theta_n, g_n)$ *and* $f(\boldsymbol{X}; \theta, g)$ *are contiguous.*

## 3. Representation theorem and local asymptotic lower bound

In this section we develop asymptotic lower bounds for estimation of $\theta$ in the presence of the unknown nuisance parameter $g$. If the density function $g$ of $e_j$ were known, $g = g_0$ say, then the result of Hájek [7] or Inagaki [11] together with Lemma 1.2 guarantee that any estimator of $\theta$ has a limiting distribution more dispersed than $N(0, 1/I_0)$

where $I_0=4\sigma^2\|\rho\|_\mu^2/(1-\theta^2)$ is the usual parametric Fisher information for $\theta$ but in an non-i.i.d. model. Akahira [1] considered the class of all asymptotically median unbiased (AMU) estimators of $\theta$ and derived a lower bound of the asymptotic distributions of AMU estimators; the limiting distribution is again $N(0, 1/I_0)$, the argument was based on the Neyman-Pearson fundamental lemma and it is different from the current investigation. It was also shown that the least squares estimator $\sum_{j=1}^{n} X_{j-1} \cdot X_j \Big/ \sum_{j=1}^{n} X_{j-1}^2$ of $\theta$ is asymptotically efficient if and only if the density function $g_0$ is a normal density function.

In the present paper we develop a representation theorem of the limiting distributions of regular estimators of $\theta$, as well as a local asymptotic minimax lower bound of the risks of arbitrary estimators of $\theta$. Both lower bounds are derived without specifying the nuisance function $g$. Suprisingly, the lower bounds turn out to be the same as when $g$ is specified. It then provides a necessary condition of constructing adaptive estimators (see Bickel [6] or Begun, Hall, Huang, and Wellner [3], for example) of $\theta$. Beran [4] provided a sufficient condition for constructing partially adaptive estimators by estimating the 'score function' $\rho(x)$ of $\theta$ of a one-step linearized estimators of the least squares estimators. The orthogonality property of score functions $S_\theta^{(j)}$ and $S_\beta^{(j)}$ in Lemma 1.3 suggests that we may be able to construct fully adaptive estimators of $\theta$.

We say that an estimator $\hat{\theta}_n$ of $\theta$ is regular at $(\theta, g)\in(-1, 1)\times\mathcal{G}$ if, for every sequence $\{(\theta_n, g_n), n\geqq 1\}\in\Theta\times C(g)$, the distribution of $n^{1/2}(\hat{\theta}_n-\theta_n)$ (under $f(X; \theta_n, g_n)$) converges weakly to a law $\mathcal{L}=\mathcal{L}(\theta, g)$ which depends only on $(\theta, g)$ but not on the particular sequence $\{(\theta_n, g_n), n\geqq 1\}$. Thus $\mathcal{L}$ does not depend on $h$ or $\beta$.

The proof of the following theorem is slightly different from the approach taken in Begun, Hall, Huang and Wellner [3] for i.i.d. cases; a directional approach is used in the current proof, this has also been used in Huang [9].

THEOREM 2.1. *Suppose that $\hat{\theta}_n$ is a regular estimator of $\theta$ with limiting law $\mathcal{L}$ and that the assumptions of Lemma 1.2 and Proposition 1.1 hold. Then $\mathcal{L}$ may be represented as the convolution of a $N(0, 1/I_0)$ distribution with $\mathcal{L}_1=\mathcal{L}_1(\theta, g)$, a distribution depending only on $(\theta, g)$, where $I_0=4\sigma^2\|\rho\|_\mu^2/(1-\theta^2)$.*

*Equivalently, if $T$, $Z_0$, and $W$ denote random variables with laws $\mathcal{L}$, $N(0, 1/I_0)$, and $\mathcal{L}_1$, respectively.*

$$T\overset{d}{=}Z_0+W \tag{2.1}$$

*where $Z_0$ and $W$ are independent.*

PROOF. For any particular 'direction' $\beta \in L_2(\mu)$, $\|\beta\|_\mu=1$, $\langle\beta, f^{1/2}\rangle=0$, let $g_{(\tau)}^{1/2}=(1-\tau^2)^{1/2}g^{1/2}+\tau\beta$ where $\tau$ is real and small. The sequence of probability measures $P_n(\cdot\,; \theta, g_{(\tau)}^{1/2}, \beta)$ determined by $f(X; \theta, g_{(\tau)})$ is essentially a two-parameter $(\theta, \tau)$ family. Note that $g_n^{1/2} \in C(g, h_1\beta)$ implies $g_n^{1/2}=(1-n^{-1}h_1^2)^{1/2}g^{1/2}+n^{-1/2}h_1\beta+o(n^{-1/2})$ and hence $P_n(\cdot\,; \theta, g_n^{1/2}, \beta)=P_n(\cdot\,; \theta, g_{(\tau_n)}^{1/2}+o(n^{-1/2}), \beta)$ at $\tau=0$ where $\tau_n=\tau+n^{-1/2}h_1$. The Hájek's [7], [8] theorems apply to this two parameter family at $(\theta, 0)$, it follows from Lemma 1.3 and Proposition 1.1 that, for all $(\theta_n, g_n) \in \Theta\times C(g, h_1\beta)$ and any regular estimator $\hat{\theta}_n$ of $\theta$, (2.1) holds. Note that Lemma 1.3 implies the convolution representation (2.1) is free of the chosen direction $\beta$. It completes the proof.

THEOREM 2.2. *Suppose that the assumptions of Lemma* 1.2 *and Proposition* 1.1 *hold, and that* $l(\cdot)$ *is a loss function satisfies the Hájek's conditions*:

(i) $l(x)=l(|x|)$, (ii) $l(x)\leqq l(y)$ *if* $|x|\leqq|y|$, *and*

(iii) $\int_{-\infty}^{\infty} l(x)\exp(-2^{-1}\lambda x^2)dx<\infty$ *for all* $\lambda>0$ *and* $l(0)=0$. *Then with* $f_n\equiv f(X; \theta_n, g_n)$, $(\theta_n, g_n) \in \Theta\times C(g)$,

$$\liminf_{n\to\infty}\ \inf_{\hat{\theta}_n}\ \sup_{(\theta_n, g_n)} E_{f_n} l[n^{1/2}(\hat{\theta}_n-\theta)]\geqq E[l(Z_0)]\ , \qquad \text{where } Z_0\sim N(0, 1/I_0)\ , \tag{2.2}$$

$I_0=4\sigma^2\|\rho\|_\mu^2/(1-\theta^2)$. *Furthermore, if the equality holds for a nonconstant* $l(\cdot)$, *then under* $P_f$,

$$\left[n^{1/2}(\hat{\theta}_n-\theta)-2n^{-1/2}\sum_{j=1}^{n} X_{j-1}\cdot\rho(X_j-\theta X_{j-1})g^{-1/2}(X_j-\theta X_{j-1})\right]\to 0 \quad \text{as } n\to\infty\ . \tag{2.3}$$

PROOF.

$$\begin{aligned}
&\liminf_{n\to\infty}\ \inf_{\hat{\theta}_n}\ \sup_{(\theta_n, g_n)} E_{f_n} l[n^{1/2}(\hat{\theta}_n-\theta)]\\
&\quad\equiv \liminf_{n\to\infty}\ \inf_{\hat{\theta}_n}\ \sup_{(\theta_n, g_n)\in\Theta\times C(g)} E_{f_n} l[n^{1/2}(\hat{\theta}_n-\theta)]\\
&\quad\geqq \liminf_{n\to\infty}\ \inf_{\hat{\theta}_n}\ \sup_{\theta_n\in\Theta} E_{f(X;\theta_n, g)} l[n^{1/2}(\hat{\theta}_n-\theta)]
\end{aligned}$$

which, by Lemma 1.2 and Proposition 1.1 and the parametric model local asymptotic minimax theorem of Hájek [8], is $\geqq E[l(Z_0)]$; and moreover, equality in the final step holds only if, under $P_f$,

$$\left[n^{1/2}(\hat{\theta}_n-\theta)-2n^{-1/2}\sum_{j=1}^{n} X_{j-1}\cdot\rho(X_j-\theta X_{j-1})g^{-1/2}(X_j-\theta X_{j-1})\right]\to 0$$

that is, the limiting law of $n^{1/2}(\hat{\theta}_n-\theta)$ is $N(0, [4\sigma^2\|\rho\|_\mu^2/(1-\theta^2)]^{-1})$. We have essentially guessed the lower bound in the above inequality from Theorem 2.1 and local asymptotic minimax theorem of Hájek [8].

## Acknowledgments

This paper is related to part of the author's doctoral dissertation. The author is greatly indebted to his advisors, Professors W. J. Hall and Jon A. Wellner. I wish to thank Professor M. Akahira for valuable comments. My thanks also go to the editor and the referee for their helpful comments.

LEHIGH UNIVERSITY

## REFERENCES

[1] Akahira, M. (1976). On the asymptotic efficiency of estimators in an autoregressive process, *Ann. Inst. Statist. Math.*, 28, 35-48.

[2] Akahira, M. and Takeuchi, K. (1981). Asymptotic efficiency of statistical estimators: Concepts and higher order asymptotic efficiency, *Lecture Notes in Statistics*, ♯7, Springer-Verlag.

[3] Begun, Janet M., Hall, W. J., Huang, Wei-Min and Wellner, Jon A. (1983). Information and asymptotic efficiency in parametric-nonparametric models, *Ann. Statist.*, 11, 432-452.

[4] Beran, R. (1976). Adaptive estimate for autoregressive process, *Ann. Inst. Statist. Math.*, 28, 77-89.

[5] Beran, R. (1977). Robust location estimates, *Ann. Statist.*, 5, 431-444.

[6] Bickel, P. J. (1982). On adaptive estimation, *Ann. Statist.*, **10**, 647-671.

[7] Hájek, J. (1970). A characterization of limiting distributions of regular estimates, *Zeit. Wahrscheinlichkeitsth.*, 14, 323-330.

[8] Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation, *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **1**, 175-194, UC at Berkley.

[9] Huang, Wei-Min (1982). Parameter estimation when there are nuisance functions, *Technical Report 82/08*, University of Rochester, Rochester, New York.

[10] Ibragimov, I. A. and Hasminskii, R. Z. (1981). *Statistical Estimation: Asymptotic Theory*, Springer-Verlag, New York.

[11] Inagaki, N. (1970). On the limiting distribution of a sequence of estimators with uniformity property, *Ann. Inst. Statist. Math.*, 22, No. 1, 1-13.

[12] Kabaila, P. (1983). On the asymptotic efficiency of estimators of the parameters of an ARMA process, *J. Time Series Analysis*, 4, 37-47.

[13] LeCam, L. (1969). *Theorie Asymptotique de la Decision Statistique*, Les Press de l'Universite de Montreal.

[14] Levit, B. Ya (1975). On the efficiency of a class of non-parametric estimates, *Theory Probab. Appl.*, **20**, 723-740.

[15] Millar, P. W. (1979). Asymptotic minimax theorems for the sample distribution function, *Zeit. Wahrscheinlichkeitsth.*, 48, 233-252.

[16] Pfanzagl, J. (1982). Contributions to a general asymptotic statistical theory, *Lecture Notes in Statistics*, ♯**13**, Springer-Verlag.

[17] Wellner, Jon A. (1982). Asymptotic optimality of the product-limit estimator, *Ann. Statist.*, **10**, 595-602.