# A SIMPLE ANALYSIS OF THIRD ORDER EFFICIENCY OF ESTIMATES

BY

PETER J. BICKEL

F. GOETZE

W. R. VAN ZWET

TECHNICAL REPORT NO. 28

JANUARY 1984

DEPARTMENT OF STATISTICS

UNIVERSITY OF CALIFORNIA, BERKELEY

RESEARCH SUPPORTED IN PART BY

OFFICE OF NAVAL RESEARCH

CONTRACT N00014-80-C-0163

REVISED

# A simple analysis of third order efficiency of estimates

Peter J. Bickel[1]
University of California, Berkeley

F. Goetze[1]
University of Cologne

W. R. van Zwet[1]
University of Leiden

## 1. Introduction

During the last few years investigators in many countries have made advances in the higher order asymptotic optimality theory of statistical procedures. Perhaps the central contribution has been the discovery by Pfanzagl (1978a), 1979) and Chibisov (1974) that, under various regularity conditions, for smooth parametric models in the i.i.d. case first order efficiency of tests (and estimates) implies second order efficiency (to order $n^{-1/2}$) and the discovery by Pfanzagl (1978b)), Takeuchi and Akahira (1976), Ghosh and Subramanyam (1974) and Efron (1975) that maximum likelihood estimates are at least as good to third order $(n^{-1})$ as any competitors with the same bias to third order $(n^{-1})$. The last phenomenon is called, depending on the writer, second or third order efficiency. Presentations of these results have in general involved Edgeworth expansion and/or cumulant expansions, and it has not been clear to what extent the stringent conditions on the models and on the classes of procedures studied are needed for the validity of the results. In a previous paper, Bickel, Chibisov, and van Zwet (1981), we argued that the first order efficiency implies second order efficiency result for tests is a consequence of the Neyman-Pearson Lemma and the structure of the likelihood functions of the experiments considered. In this paper we study the third order property of maximum likelihood like estimates and again deduce it as a general feature of smooth likelihood functions. Our approach was suggested

[1]This research was supported in part by Office of Naval Research Contract N00014-80-C-0163.

by the discussion of L. Le Cam on Berkson (1980). The program we follow is to show:

1) That for a given loss function $W$ and suitable priors $\Pi$ the Bayes estimate is to third order of the form $\hat{\theta}_n + \frac{b(\hat{\theta}_n)}{n}$ where $\hat{\theta}_n$ is maximum likelihood (like).

2) It is possible to define a perturbed version $W_n$ of $W$ such that

a) $\hat{\theta}_n$ is itself third order Bayes with respect to $W_n$.

b) For estimates with the same bias to third order as $\hat{\theta}_n$, the Bayes risk under $W_n$ coincides with that under $W$.

From this it is easy to deduce the property. We illustrate this idea by applying it to the estimation of the mean $\theta$ of a normal distribution with variance 1 with quadratic loss and the usual notion of bias. In this context the idea corresponds to a Bayesian argument for the U.M.V.U. property of $\bar{X}$. Suppose $\theta \sim N(\mu,\tau^2)$.

Define a new loss function by

$$L(\theta,a) = (\theta-a)^2 + \frac{2\lambda}{1-\lambda}(\theta-\mu)(\theta-a)$$

where $\lambda = (n\tau^2+1)^{-1}$. The posterior density of $\theta$ is $N(\lambda\mu + (1-\lambda)\bar{X},\ \lambda\tau^2)$ so that the Bayes estimate of $\theta$ under $L$ is just

$$(1-\lambda)^{-1}E(\theta|\bar{X}) - \lambda(1-\lambda)^{-1}\mu = \bar{X}\ .$$

But then, for any estimate $T$,

$$E\{L(\theta,T)\} = E(T-\theta)^2 + \frac{2\lambda}{1-\lambda}\int(\theta-\mu)E_\theta(\theta-T)\pi(\theta)d\theta$$

where $\pi$ is the prior density and for unbiased estimates

$$E\{L(\theta,T)\} = E(T-\theta)^2$$

The optimality of $\bar{X}$ with respect to quadratic loss follows from the Bayesian optimality. We can proceed to derive optimality at individual $\theta$ by taking $\mu = \theta$ and letting $\tau^2 \to 0$. Note that this is essentially a Lagrange multiplier argument.

Our results are very close save for the more general setting to those of Ghosh, Sinha and Joshi (1982). They also obtain an expansion and apply it to the third order efficiency of maximum likelihood estimates. However their method still leaves them with limitations on the class of competitors to maximum likelihood that are being considered as well as their self-imposed restriction to quadratic loss functions. Their expansion and necessarily their application is also restricted to the case of independent identically distributed observations. Related results can be found in Takeuchi (1982).

## 2. The main results

Here are basic assumptions and notation.

Model: We observe $X^{(n)}$, a random element taking values in $\mathcal{X}_n$. The possible distributions of $X^{(n)}$ are $P_\theta^{(n)}$, $\theta \in \Theta$, an open interval. Suppose $P_\theta^{(n)} < \mu^{(n)}$, $\sigma$ finite. Let

$$f_n(\cdot,\theta) \triangleq \frac{dP_\theta^{(n)}}{d\mu^{(n)}} \tag{2.1}$$

$$\ell_n(\theta) \triangleq \log f_n(\cdot,\theta) .$$

We will typically drop the superscript $n$ for expectations and probabilities.

We postulate an expansion for $\ell_n(\theta+hn^{-1/2})$ around $\theta$ in powers of $hn^{-1/2}$ with coefficients which we can think of as derivatives of $\ell_n$ and a remainder. Our further assumptions on the model are framed in terms of stability conditions on the coefficients and bounds on the remainder. We write

$$(2.2)\quad \ell_n(\theta+hn^{-1/2}) = \ell_n(\theta) + n^{1/2}\bar{\ell}_n^{(1)}(\theta)h + \bar{\ell}_n^{(2)}(\theta)\frac{h^2}{2} + \sum_{i=1}^{j} n^{-i/2}\bar{\ell}_n^{(i+2)}(\theta)\frac{h^{i+2}}{(i+2)!} + \Delta_{nj}(\theta,h)h^{j+3}$$

$\bar{\ell}_n^{(i)}$ can be thought of as $n^{-1}\frac{\partial^i}{\partial\theta^i}\ell_n(\theta)$.

We postulate also non random functions $\lambda_i : \Theta \to R$ which approximate the $\bar{\ell}_n^{(i)}$ and write

$$\bar{\ell}_n^{(i)}(\theta) = \lambda_i(\theta) + \tilde{\Delta}_{ni}(\theta)$$

The $\lambda_i$ can be thought of as approximations to $E_\theta\bar{\ell}_n^{(i)}(\theta)$.

Loss Function: We are given a symmetric bowl-shaped function $W: R \to R^+$, which is bounded and satisfies

$$(2.3)\qquad W(t) = W(-t)\ ,\quad \forall t$$

$$(2.4)\qquad W(|t|)\ \text{non decreasing, non constant}\ .$$

Define the risk of an estimate $T_n: X_n \to R$

$$(2.5)\qquad R_W(\theta,T_n) = E_\theta W(n^{1/2}(T_n-\theta))\ .$$

Prior Distributions: We will consider prior distributions $\Pi$ on $\Theta$ with densities $\pi$ which are smooth.

Convention: Given random variables $R_n(\theta)$, $\theta \in \Theta$

$$R_n \overset{\Delta}{=} o(a_n,b_n) \Leftrightarrow P_{\theta_0}[|R_n(\theta_0)| \geq a_n\epsilon] = o(b_n)$$

$\forall\epsilon > 0$, uniformly for $\theta_0 \in K$ compact. Note that $R_n = o(a_n,b_n) \Leftrightarrow \exists\epsilon_n \downarrow 0 \ni P_{\theta_0}[|R_n(\theta_0)| \geq \epsilon_n a_n] = o(b_n)$ uniformly for $\theta_0 \in K$.

$$R_n \overset{\Delta}{=} O(a_n,b_n) \Leftrightarrow P_{\theta_0}[|R_n(\theta_0)| \geq a_n c_n] = o(b_n)$$

for all $c_n \uparrow \infty$ uniformly for $\theta_0 \in K$ compact.

Bias: We define bias side conditions as in Pfanzagl and Wefelmeyer (1978) through a function $d: R \to R$ which is assumed bounded, increasing and non constant. Define the $d$ bias of $T_n$

$$B(T_n,\theta) = E_\theta d(n^{1/2}(T_n-\theta)) .$$

We say $T_n$, $S_n$ have the same $d$ bias to third order if

$$B(T_n,\theta) = B(S_n,\theta) + o(n^{-1/2}) . \tag{2.6}$$

In particular if $d$ is the identity (which we exclude!) (2.6) says that the biases of $T_n$ and $S_n$ agree up to $o(n^{-1})$.

Estimate: We will distinguish an estimate $\hat{\theta}_n$ which can be thought of as the maximum likelihood estimate but will be specified by its asymptotic properties.

Assumptions:

$C_j$: For some $0 \leq \delta < \frac{1}{6(j+1)}$, $c_n \uparrow \infty$, $\epsilon > 0$,

(1) $\text{Sup}\{|\Delta_{nj}(\theta_0+\eta n^{-1/2},h)|: |\eta| \leq c_n n^\delta, |h| \leq c_n n^\delta\}$, considered as a process in $\theta_0$, $= o(n^{-j/2},n^{-j/2})$.

(2) $\text{Sup}\{|\tilde{\Delta}_{ni}(\theta_0+\eta n^{-1/2})|: |\eta| \leq c_n n^\delta\}$, considered as a process in $\theta_0$, $= o(n^{-\epsilon},n^{-j/2})$, $2 \leq i \leq j+2$,

(3) (a) $-\lambda_2(\theta) = I(\theta) > 0$, $\forall\theta$

(b) $\lambda_i(\cdot)$, $2 \leq i \leq j+2$ are continuous.

Clearly $C_2 \Rightarrow C_1 \Rightarrow C_0$.

Note that we can replace "$c_n \uparrow \infty$" by "every fixed c" and also that without loss of generality we can take $c_n = o(n^{\gamma})$, $\forall \gamma > 0$.

For $\delta$ as in assumption $C_j$,

$E_j$:

(1) $n^{1/2}\bar{\ell}_n^{(1)}(\hat{\theta}_n) = o(n^{-j/2-\delta}, n^{-j/2})$.

(2) (a) $n^{1/2}|\hat{\theta}_n - \theta_0| = O(1,1)$

(b) $= O(n^{\delta}, n^{-j/2})$.

(considered as a process in $\theta_0$).

Condition (1) says that $\hat{\theta}_n$ behaves like a root of the likelihood equation to order j. Condition (2a) corresponds to $\sqrt{n}$ consistency and (2b) to the usual bound for probabilities of intermediate deviations.

Our main result is

Theorem 1. (a) *If* $C_1$ *and* $E_1$ *hold, then*

$$\varliminf_n \sup_{|\theta-\theta_0|\le\epsilon} n^{1/2}\{R_W(\theta,T_n)-R_W(\theta,\hat{\theta}_n)\} \ge 0 \tag{2.7}$$

$\forall\theta_0 \in \Theta$, $\epsilon > 0$.

(b) *Suppose* $C_2$ *and* $E_2$ *hold and also that* $\lambda_2$, $\lambda_3$ *are continuously differentiable. Then if* $T_n$ *and* $\hat{\theta}_n$ *have the same* d *bias (in the sense of* (2.6))

$$\varliminf_n \sup_{|\theta-\theta_0|\le\epsilon} n\{R_W(\theta,T_n)-R_W(\theta,\hat{\theta}_n)\} \ge 0 \tag{2.8}$$

$\forall\theta_0 \in \Theta$, $\epsilon > 0$.

The corresponding statement for $j = 0$ is a version of the Hájek-Le Cam minimax theorem.

<u>Corollary 1</u>. *Suppose for both* $S_n = T_n$ *and* $S_n = \hat{\theta}_n$

$$(2.9) \qquad R_W(\theta,S_n) = \sum_{i=1}^{j} r_i(\theta)n^{-i/2} + o(n^{-j/2})$$

*uniformly on compacts where* $r_i$ *are continuous functions which depend on* $\{S_n\}_{n\geq 1}$ *but not on* $n$. *Then under the hypotheses of Theorem 1*

$$(2.10) \qquad \underline{\lim}_n \; n^{j/2}\{R_W(\theta,T_n)-R_W(\theta,\hat{\theta}_n)\} \geq 0$$

*for* $j = 1, 2$ *as appropriate.*

This corollary for $j = 1$ corresponds to the usual assertion of second order efficiency for the M.L.E. while $j = 2$ corresponds to the usual assertion of third order efficiency <u>after equating biases</u>.

Conditions $C_j$, $E_j$ and (2.9) follow from the assumptions of Pfanzagl (1976), Pfanzagl and Wefelmeyer (1978), Ghosh and Subramanyam (1974), Ghosh, Sinha and Wieand (1980) so that the conclusions of these authors can be subsumed under Theorem 1.

The theorem follows from a study of the structure of Bayes solutions. The priors $\Pi$ that we consider have densities $\pi$. Let $\gamma = \log \pi$, and suppose the following assumption holds.

$P_j$: $\Pi$ has compact support. For $\delta$ and $\epsilon$ as in assumption $C_j$ and
$0 \leq \beta < \frac{1}{2(j+1)} - \delta$, $\beta < \epsilon$,

$$(2.11) \qquad \Pi[\theta: \sup\{|\gamma^{(i)}(\theta+tn^{-1/2})|: |t| \leq c_n^2 n^{\delta}\} \geq n^{\beta}] = o(n^{-j/2})$$

$1 \leq i \leq j+1$, $c_n \uparrow \infty$ as given in assumption $C_j$. Here $\gamma^{(i)}(\theta)$ is the $i^{th}$ derivative of $\gamma$ for $\theta$ in the interior of the support of $\Pi$ and $\gamma^{(i)}(\theta) = \infty$ otherwise.

A class of examples of $\Pi$ satisfying $P_j$ are those with $(j+1)$ times continuously differentiable densities and $\pi^{(j+1)}(t) \sim c(t-b)^m$, $m > \frac{j(j+1)^2}{1 - 2\delta(j+1)} - (j+2)$ for $b$ a boundary point of support.

Theorem 1(a) follows immediately from

<u>Theorem 2</u>. *If* $\Pi$ *satisfies* $P_1$ *and* $C_1$, $E_1$ *hold, then*

$$\int R_W(\theta,\hat{\theta}_n)\Pi(d\theta) = \mathrm{Inf}_{T_n} \int R_W(\theta,T_n)\Pi(d\theta) + o(n^{-1/2}) . \tag{2.12}$$

Let $\pi_n(\cdot|x)$ be the posterior density of $n^{1/2}(\theta-\hat{\theta}_n)$ given $X^{(n)} = x$,

$$r_n(\Delta,x) = \int W(t-\Delta)\pi_n(t|x)dt , \tag{2.13}$$

the posterior risk incurred by action $\hat{\theta}_n + \Delta n^{-1/2}$, and

$$r_n(x) = \mathrm{Inf}_\Delta\ r_n(\Delta,x) \tag{2.14}$$

the Bayes posterior risk. So, (2.12) is equivalent to

$$E_\Pi(r_n(0,X^{(n)}) - r_n(X^{(n)})) = o(n^{-1/2})$$

where $E_\Pi$, $P_\Pi$ correspond to computation under $\int P_\theta^{(n)}\Pi(d\theta)$. Now,we can write, for $\pi(\hat{\theta}_n) > 0$,

$$\begin{aligned}\log \pi_n(t|x) = {} & \ell_n(\hat{\theta}_n+tn^{-1/2}) - \ell_n(\hat{\theta}_n) + \gamma(\hat{\theta}_n+tn^{-1/2}) \\ & - \gamma(\hat{\theta}_n) - \log N_n(x)\end{aligned} \tag{2.15}$$

where

$$N_n(x) \triangleq [f_n(x,\hat{\theta}_n)\pi(\hat{\theta}_n)]^{-1} \int f_n(x,\hat{\theta}_n+tn^{-1/2})\pi(\hat{\theta}_n+tn^{-1/2})dt . \tag{2.16}$$

We need an Edgeworth expansion on the posterior. Define a set $B_n \subset X^{(n)}$ by: $x \in B_n$ if and only if,

(i) $\mathrm{Sup}\{|\Delta_{nj}(\hat{\theta}_n,t)|: |t| \le c_n n^\delta\} \le \varepsilon_n n^{-j/2}$

(ii) $\bar{\ell}_n^{(2)}(\hat{\theta}_n) \ge -a$, $a > 0$,

(iii) $n^{1/2}|\bar{\ell}_n^{(1)}(\hat{\theta}_n)| < \varepsilon_n n^{-j/2-\delta}$

(iv) $|\bar{\ell}_n^{(i+2)}(\hat{\theta}_n) - \lambda_{i+2}(\hat{\theta}_n)| \le \varepsilon_n$, $1 \le i \le j$

(v) $\sup\{|\gamma^{(i)}(\hat{\theta}_n + tn^{-1/2})| : |t| \le c_n n^{\delta}\} \le n^{\beta}$, $1 \le i \le j+1$,

where $\varepsilon_n \downarrow 0$. Define for $x \in B_n$

$$\pi_{n0}(t|x) = N_n^{-1}(x)\exp\{\bar{\ell}_n^{(2)}(\hat{\theta}_n)\frac{t^2}{2}\} \tag{2.17}$$

$$\pi_{n1}(t|x) = \pi_{n0}(t|x)\left[1 + n^{-1/2}\left(\bar{\ell}_n^{(3)}(\hat{\theta}_n)\frac{t^3}{6} + \gamma^{(1)}(\hat{\theta}_n)t\right)\right] \tag{2.18}$$

More generally let

$$\pi_{nj}(t|x) = \pi_{n0}(t|x)(1 + \textstyle\sum_{i=1}^{j} n^{-i/2}A_i(t,x)) \tag{2.19}$$

where $A_i$ are defined as the coefficient of $n^{-i/2}$ in the formal expansion.

$$\begin{aligned} \exp \textstyle\sum_{i=1}^{\infty} \left[\frac{\bar{\ell}_n^{(i+2)}(\hat{\theta}_n)}{(i+2)!}\, t^{i+2} + \frac{\gamma^{(i)}(\hat{\theta}_n)}{i!} t^i\right] n^{-i/2} \\ = 1 + \textstyle\sum_{i=1}^{\infty} A_i(t,x)n^{-i/2} \end{aligned} \tag{2.20}$$

Define $\pi_{nj} = 0$ otherwise.

<u>Lemma 1</u>. *If* $C_j$, $E_j$, $P_j$ *hold, then*

$$E_{\Pi}[\int |\pi(t|X^{(n)}) - \pi_{nj}(t|X^{(n)})|dt] = o(n^{-j/2}) . \tag{2.21}$$

<u>Proof</u>. For $x \in B_n$, $|t| \le c_n n^{\delta}$ write

$$\begin{aligned} \pi_n(t|x) = \pi_{n0}(t|x)\exp\{&\textstyle\sum_{i=1}^{j} Q_i(t,x)n^{-i/2} \\ &+ \frac{(n^{-1/2}t)^{j+1}}{j!}\int_0^1 \gamma^{(j+1)}(\hat{\theta}_n + \lambda t n^{-1/2})(1-\lambda)^j d\lambda + n^{1/2}\bar{\ell}_n^{(1)}(\hat{\theta}_n)t + \Delta_{nj}(\hat{\theta}_n, t) \end{aligned}$$

where

$$Q_i(t,x) = \frac{\bar{\ell}_n^{(i+2)}(\hat{\theta}_n)}{(i+2)!}\, t^{i+2} + \frac{\gamma^{(i)}(\hat{\theta}_n)}{i!}\, t^i . \tag{2.22}$$

By construction of $B_n$ the last three terms in the exponent are $o(n^{-j/2})$ and for $n$ sufficiently large,

$$n^{-1/2}|Q_i(t,x)| \le 1$$

uniformly for $(x,t)$ as above. Therefore

(2.23) $$\pi_n(t|x) = \pi_{n0}(t|x)\exp\{\textstyle\sum_{i=1}^{j} Q_i(t,x)n^{-i/2}\}(1+o(n^{-j/2})) .$$

But, by standard arguments,

(2.24) $$\begin{aligned}&\pi_{no}(t|x)[\exp\{\textstyle\sum_{i=1}^{j} Q_i(t,x)n^{-i/2}\} - (1+\sum_{i=1}^{j} A_i(t,x)n^{-i/2})] \\ &\le \pi_{n0}(t|x)O(n^{-\frac{(j+1)}{2}}\max_i |Q_i(t,x)|^{\frac{j+1}{i}}) \\ &= O(\pi_{n0}(t|x)n^{-j/2}\max_i\{n^{-\frac{1}{2}+(i+2)(\frac{j+1}{i})\delta} + n^{-\frac{1}{2}+(\beta+i\delta)(\frac{j+1}{i})}\})\end{aligned}$$

uniformly as above.

Therefore

(2.25) $$\pi_n(t|x) = \pi_{nj}(t|x) + \pi_{n0}(t|x)o(n^{-j/2})$$

uniformly as above.

By the same expansion $(j=0)$, for $x \in B_n$,

(2.26) $$N_n(x) \ge \int \exp\{-a\,\frac{s^2}{2} + o(1)\}ds \ge \varepsilon > 0$$

and

(2.27) $$\int \pi_{n0}(t|x)dt \le M$$

independent of $x$, $n$. By (2.25), (2.26) and (2.27) the lemma follows if

(2.28) $$P_\pi[n^{1/2}|\theta-\hat{\theta}_n| > c_n n^\delta] + P_\pi[X^{(n)} \notin B_n] = o(n^{-j/2}) ,$$

(2.29) $$E_{\Pi}\int_{\{|t|>c_n n^{\delta}\}} \pi_{nj}(t|X^{(n)})dt = o(n^{-j/2}) \ .$$

But

$$P_{\Pi}[n^{1/2}|\theta-\hat{\theta}_n| > c_n n^{\delta}] \leq \sup_K P_{\theta}[n^{1/2}|\theta-\hat{\theta}_n| > c_n n^{\delta}]$$
$$P_{\Pi}[X^{(n)} \notin B_n] \leq \sup_K P_{\theta}[X^{(n)} \notin B_n] \ .$$

Therefore if we take (say)

$$a = -\frac{1}{2} \inf_K \lambda_2(\theta)$$

it is easy to see that $C_j$, $P_j$ imply the existence of $\varepsilon_n$ such that (2.28) holds. A direct calculation yields (2.29) and the lemma. □

Lemma 2. *For* W *as specified, let* $\phi$ *be the standard normal density,*

$$A(\sigma^2) = \frac{1}{\sigma^5}\int W(t)(t^2-\sigma^2)\phi(\frac{t}{\sigma})dt \ .$$

*Then* $A(\sigma^2) > 0$ *and continuous.*

Proof. W is symmetric. $W(|t|)$ and $(t^2-\sigma^2)$ are both nondecreasing in $|t|$. By Chebyshev's theorem,

$$\sigma^4 A(\sigma^2) \geq \frac{1}{\sigma^2} \int W(t)\phi(\frac{t}{\sigma})dt \int (t^2-\sigma^2)\phi(\frac{t}{\sigma})dt = 0$$

with strict inequality unless W is constant. The lemma follows. □

Proof of Theorem 2. By Lemma 1

(2.30) $$E_{\Pi}(r_n(0,X^{(n)}) - r_n(X^{(n)}))$$
$$= E_{\Pi}\left(\int W(t)\pi_{n1}(t|X^{(n)})dt - \inf_{\Delta}\left(\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt\right)\right) + o(n^{-1/2}) \ .$$

Moreover,

$$\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt$$
$$= \int W(t-\Delta)\pi_{n0}(t|X^{(n)})dt + \int W(t-\Delta)\pi_{n0}(t|X^{(n)})n^{-1/2}\{\bar{\ell}_n^{(3)}(\hat{\theta}_n)\frac{t^3}{6} + \gamma^{(1)}(\hat{\theta}_n)t\}dt.$$

The integrand of the last integral is odd for $\Delta = 0$ and therefore this integral is $0(n^{-1/2+\beta}|\Delta|)$ as well as $0(n^{-1/2+\beta})$ on $B_n$. Since $\int W(t-\Delta)\pi_{n0}(t|X^{(n)})dt$ is increasing in $|\Delta|$ by Anderson's lemma, we see that $\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt$ can't assume its minimum as a function of $\Delta$ outside any fixed neighborhood of zero for sufficiently large $n$. If $\Delta = o(1)$ as $n \to \infty$, however, we have

$$\int W(t-\Delta)\pi_{n0}(t|X^{(n)})dt$$
$$= \int W(t)\pi_{n0}(t+\Delta|X^{(n)})dt$$
$$= \int W(t)\pi_{n0}(t|X^{(n)})dt + \frac{\Delta^2}{2}\int W(t)\{\bar{\ell}_n^{(2)}(\hat{\theta}_n) + t^2(\bar{\ell}_n^{(2)}(\hat{\theta}_n))^2\}\pi_{n0}(t|X^{(n)})dt$$
$$+ 0(|\Delta|^3) .$$

The coefficient of $\Delta^2$ in the second term is positive and bounded away from zero by Lemma 2, say $\geq \alpha > 0$. Hence, for $\Delta = o(1)$,

$$\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt = \int W(t-\Delta)\pi_{n0}(t|X^{(n)})dt + 0(n^{-1/2+\beta}|\Delta|)$$
$$\geq \int W(t)\pi_{n0}(t|X^{(n)})dt + \alpha\Delta^2 + 0(n^{-1/2+\beta}|\Delta| + |\Delta|^3) ,$$

and for sufficiently large $C > 0$, no minima of $\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt$ can occur for $|\Delta| \geq Cn^{-1/2+\beta}$. But for $|\Delta| < Cn^{-1/2+\beta}$ we have

$$\int W(t-\Delta)\pi_{n1}(t|X^{(n)})dt = \int W(t)\pi_{n0}(t|X^{(n)})dt + o(n^{-1/2})$$

as $1 - 2\beta > \frac{1}{2}$. The theorem follows by (2.30). □

To deal with third order efficiency we extend Theorem 2 as follows. Suppose $d$ defines bias as in (2.6).

Let $c: \Theta \to R$. Define

$$W_n(\theta,a) = W(n^{1/2}(\theta-a)) + h(n^{-1/2}c(\theta))d(n^{1/2}(\theta-a)) \tag{2.31}$$

where $h(t) = t$, $|t| \leq 1$ and 0 otherwise. $W_n$ is an asymmetric perturbation of W. Assume

Q: (1) c is differentiable on $\Theta$ and

(2) $\Pi[\theta: \sup\{|c^{(i)}(\theta+tn^{-1/2})|: |t| \leq c_n^2 n^{\delta}\} \leq n^{\alpha}] = 1 - o(n^{-1})$, $0 \leq i \leq 2$

for a prior $\Pi$, $\delta$ as in $C_2$, $c_n \uparrow \infty$ given in $C_2$, $\alpha < \frac{1}{2} - 2\beta$, $\beta$ as in $P_2$.

Define

$$b_W(\theta) = v(\theta)A^{-1}(I^{-1}(\theta)) \tag{2.32}$$

where

$$v(\theta) = \{(\gamma^{(1)}(\theta)I(\theta) - \frac{\lambda_3(\theta)}{2})\int s^2 W(s)\phi(s,I^{-1}(\theta))ds + \frac{\lambda_3(\theta)I(\theta)}{6}\int s^4 W(s)\phi(s,I^{-1}(\theta))ds - \gamma^{(1)}(\theta)\int W(s)\phi(s,I^{-1}(\theta))ds\}$$

and $\phi(\cdot,\sigma^2)$ is the $N(0,\sigma^2)$ density, and

$$b_{W_n}(\theta) = b_W(\theta) + c(\theta)D(\theta)I(\theta)A^{-1}(I^{-1}(\theta)) \tag{2.33}$$

where

$$D(\theta) = \int d(v)v\phi(v,I^{-1}(\theta))dv . \tag{2.34}$$

Note that $D(\theta) > 0$ by Chebyshev's theorem since d is nondecreasing, non constant.

Theorem 3. *If* $\Pi$ *satisfies* $P_2$ *and* $C_2$, $E_2$ *hold,*

(2.35) $$\int R_W(\theta, \hat\theta_n + b_W(\hat\theta_n)n^{-1})\Pi(d\theta) = \operatorname*{Inf}_{T_n} \int R_W(\theta, T_n)\Pi(d\theta) + o(n^{-1}) .$$

If $c$ satisfies $Q$ then (2.35) holds with $W$ replaced by $W_n$.

In words, $\hat\theta_n + b_W(\hat\theta_n)n^{-1}$ is Bayes to third order under $W$. The formula for the correction $b_W n^{-1}$ is unimportant. The main point is that it is a function of $\hat\theta_n$ only and that the corresponding additional correction for $W_n$ is linear in $c$.

Proof of Theorem 3. We have

$$\begin{aligned}
&\int W(t-\Delta)\pi_{n2}(t|X^{(n)})dt \\
&\quad = \int W(t-\Delta)\pi_{n0}(t|X^{(n)})dt \\
&\qquad + n^{-1/2}\Delta\int W(t)\pi_{n0}(t|X^{(n)})[\bar\ell_n^{(2)}(\hat\theta_n) + \{\bar\ell_n^{(3)}(\hat\theta_n)\frac{t^3}{6} + \gamma^{(1)}(\hat\theta_n)t\} \\
&\qquad\qquad + \bar\ell_n^{(3)}(\hat\theta_n)\frac{t^2}{2} + \gamma^{(1)}(\hat\theta_n)]dt \\
&\qquad + \psi_n(X^{(n)}) + O(n^{-1/2}|\Delta|^3 + n^{-1+2\beta}\Delta^2) ,
\end{aligned}$$

where $\psi$ is a function independent of $\Delta$. Arguing as in the proof of Theorem 2, we find that we can restrict attention to $|\Delta| \le Cn^{-1/2+\beta}$. By assumption $C_2(2)$ we may replace $\bar\ell_n^{(i)}(\hat\theta_n)$ by $\lambda_i(\hat\theta_n)$ for $i = 2,3$ and obtain

$$\begin{aligned}
&\int W(t-\Delta)\pi_{n2}(t|X^{(n)})dt \\
&\quad = \int W(t)\pi_{n0}(t|X^{(n)})dt + \frac{\Delta^2}{2}A(I^{-1}(\hat\theta_n)) - n^{-1/2}\Delta v(\hat\theta_n) + \psi_n(X^{(n)}) + o(n^{-1}) .
\end{aligned}$$

Claim (2.35) follows. Its extension to $W_n$ follows similarly. $\square$

We can now complete the proof of Theorem 1, part (b). Choose a prior $\Pi$ satisfying $P_2$, $\int|\pi^{(1)}(\theta)|d\theta < \infty$. Define the function $c(\theta)$ by

$$(2.38) \qquad \begin{aligned} c(\theta) &= -b_W(\theta)A(I^{-1}(\theta))I^{-1}(\theta)D^{-1}(\theta) \\ &= -v(\theta)I^{-1}(\theta)D^{-1}(\theta) \ . \end{aligned}$$

Note that by assumption $I^{-1}(\theta)D^{-1}(\theta)$ is continuously differentiable while

$$\Pi[\sup\{|v^{(j)}(\theta_0+tn^{-1/2})|: |t| \le c_n n^{\delta}\} \le n^{\beta}] = 1 - o(n^{-1})$$

for $0 \le j \le 2$ by $P_2$. So the conditions of Theorem 3 are satisfied for the choice of $\Pi$ and $c$. Moreover, by construction $b_{W_n}(\theta) = 0$. So,

$$(2.39) \qquad \int R_{W_n}(\theta,T_n)\Pi(d\theta) \ge \int R_{W_n}(\theta,\hat\theta_n)\Pi(d\theta) + o(n^{-1}) \ .$$

But for any $T_n$,

$$\begin{aligned} &\int R_{W_n}(\theta,T_n)\Pi(d\theta) \\ &\quad = \int R_W(\theta,T_n)\Pi(d\theta) + \int B(n^{1/2}(T_n-\theta))h(n^{-1/2}c(\theta))\Pi(d\theta) + o(n^{-1}) \ . \end{aligned}$$

Therefore if $\hat\theta_n$ and $T_n$ have the same bias to third order,

$$(2.40) \qquad \begin{aligned} &\int [R_{W_n}(\theta,T_n) - R_{W_n}(\theta,\hat\theta_n)]\Pi(d\theta) \\ &\quad = \int [R_W(\theta,T_n) - R_W(\theta,\hat\theta_n)]\Pi(d\theta) \\ &\quad\quad + n^{-1/2}\int [B(n^{1/2}(T_n-\theta) - B(n^{1/2}(\hat\theta_n-\theta))]c(\theta)\Pi(d\theta) \\ &\quad\quad + O\Big(n^{-1/2}\int_{\{|c(\theta)|>n^{1/2}\}} |c(\theta)|\Pi(d\theta)\Big) + o(n^{-1}) \ . \end{aligned}$$

But $\int|c(\theta)|\Pi(d\theta) < \infty$ since $\int|\pi^{(1)}(\theta)|d\theta < \infty$. Part (b) of Theorem 1 follows from (2.39) and (2.40).

Extensions:

(1) If $c(\cdot)$ is bounded or more generally satisfies $Q_1$ on compacts then the assertions of Theorem 1 hold with $\hat{\theta}_n$ replaced by $\hat{\theta}_n + \frac{c(\hat{\theta}_n)}{n}$. Of course, the competitors admitted under the bias equivalence condition depend on $c$.

(2) Theorems 1-3 can be straightforwardly extended to the multiparameter case. With $\theta = (\theta_1,\ldots,\theta_k)$ let $\ell_n^{(j)}(\theta)$ be the $j^{th}$ differential with respect to $\theta$ thought of as a j-linear function on $R^k$ or equivalently as a point in $R^{k^j}$.

$$\ell_n^{(j)}(\theta)(t_1,\ldots,t_j) = \sum\left\{\frac{\partial^j \ell_n(\theta)}{\partial\theta_{i_1}\cdots\partial\theta_{i_j}} t_{1i_1}\cdots t_{ji_j} : i_1,\ldots,i_j \in\{1,\ldots,k\}\right\}$$

where $t_q = (t_{q1},\ldots,t_{qk})$, $1 \leq q \leq j$ and write $\ell_n^{(j)}(\theta)t^j$ for $\ell_n^{(j)}(\theta)(t,t,\ldots,t)$. With this convention reinterpret (2.2) for $\theta$ a vector. The loss function $W: R^k \rightarrow R^+$ is assumed to be bounded and

(a) $W(t) = W(-t)$ for all $t$

(b) $\{t: W(t) \leq w\}$ is convex for all $w$.

(c) For every $\lambda \in R^k$, $W(\lambda t)$ is non constant in $t \in R$.

If $\phi(\cdot,\Sigma)$ is the k-variate normal density with positive definite covariance matrix $\Sigma$, define the matrix $A(\Sigma)$ by

(2.41) $$A(\Sigma) = \Sigma^{-1}\int W(t)(t^T t - \Sigma)\phi(t,\Sigma)dt\ \Sigma^{-1} .$$

Conditions (a)-(b) on $W$ guarantee that $A(\Sigma)$ is nonnegative definite, cf. Lemma 5.8, Pfanzagl and Wefelmeyer (1978). A more careful argument shows that (c) implies that $A$ is positive definite.

Risk is defined by (2.5). The scalar bias function $d$ is replaced by the vector $d: R^k \rightarrow R^k$ and the scalar $D(\theta)$ by the matrix

(2.42) $$D(\theta) = \int d^T(s)s\phi(s,I^{-1}(\theta))ds .$$

We require that $D$ be nonsingular for each $\theta$. Conditions $C_j$ and $E_j$ need to be changed only to the extent that absolute values become vector norms and $I(\theta) > 0$ becomes $I(\theta)$ positive definite. If we interpret $\gamma^{(i)}$ as the $i^{th}$ differential of the log of prior density $\pi$ defined on an open ball in $R^k \subset \Theta$ then $P_j$ also need only be modified by substituting vector norms for absolute values. If we add the assumption that $D(\theta)$ is nonsingular to the reinterpreted $C_j$, $E_j$, $P_j$, Theorems 1-3 carry over to the multiparameter case without change in proof provided that we interpret $A$ as a matrix and define the vector $v(\theta)$ by

$$\begin{aligned} v_j(\theta) = {} & \sum_{a,b} \left(\gamma_a^{(1)}(\theta) I_{bj}(\theta) - \frac{\lambda_{abj}(\theta)}{2}\right) \int s_a s_b W(s)\phi(s, I^{-1}(\theta))\,ds \\ & + \sum_{a,b,c,d} \frac{\lambda_{abc}(\theta)}{6} I_{dj}(\theta) \int s_a s_b s_c s_d W(s)\phi(s, I^{-1}(\theta))\,ds \\ & - \gamma_j^{(1)}(\theta) \int W(s)\phi(s, I^{-1}(\theta))\,ds \end{aligned}$$

where $\lambda_{abc}(\theta)$ are the components of $\lambda^{(3)}(\theta)$ and subscripts denote elements of vectors and matrices.

The results also carry over directly to the estimation of a subvector $(\theta_1,\ldots,\theta_p)$, $p < k$ with an appropriately redefined loss function.

(3) Given any estimate $T_n$, define recursively

$$T_n^{(0)} = T_n \tag{2.43}$$
$$T_n^{(i+1)} = T_n^{(i)} - \frac{\bar{\ell}_n^{(1)}(T_n^{(i)})}{\bar{\ell}_n^{(2)}(T_n^{(i)})}$$

That is, $T_n^{(j)}$ is defined by taking $j$ Newton-Rapson steps in the solution of $\bar{\ell}_n^{(1)}(\theta) = 0$ starting from $T_n$.

Theorem 4: *Suppose* $C_j$ *holds and in addition* $\ell_n$ *is* $j+1$ *times continuously differentiable with* $\ell_n^{(j)}$ *being its derivatives, and (as a process in* $\theta_0$*)*

$$\bar{\ell}_n^{(1)}(\theta_0) = o(n^{-\frac{1}{2}+\delta}, n^{-\frac{j}{2}}) . \tag{2.44}$$

*Then if* $T_n$ *satisfies* $E_j(2)$, $T_n^{(j+1)}$ *satisfies* $E_j(1)$ *and* $E_j(2)$.

Proof. We shall argue by induction for $i = 0,\ldots,j$ that

$$\bar{\ell}_n^{(1)}(T_n^{(i+1)}) = o(n^{-\frac{i}{2}-\delta}, n^{-\frac{j}{2}}) , \tag{2.45}$$

$$|T_n^{(i+1)} - \theta| = o(n^{\delta-\frac{1}{2}}, n^{-\frac{j}{2}}) . \tag{2.46}$$

Note first that, by (2.43),

$$\bar{\ell}_n^{(1)}(T_n^{(i+1)}) = \frac{\bar{\ell}_n^{(3)}(T_n^*)}{2}\left\{\frac{\bar{\ell}_n^{(1)}}{\bar{\ell}_n^{(2)}}(T_n^{(i)})\right\}^2 \tag{2.47}$$

where $|T_n^*-T_n^{(i)}| \le |T_n^{(i+1)}-T_n^{(i)}|$; also,

$$(2.48)\qquad \frac{\mathcal{L}_n^{(1)}}{\mathcal{L}_n^{(2)}}(T_n) = \frac{\bar{\ell}_n^{(1)}}{\bar{\ell}_n^{(2)}}(\theta) + \left\{1 - \frac{\bar{\ell}_n^{(1)}(\theta^*)\bar{\ell}_n^{(3)}(\theta^*)}{[\bar{\ell}_n^{(2)}(\theta^*)]^2}\right\}(T_n-\theta)$$

with $|\theta^*-\theta| \leq |T_n - \theta|$. From $C_j$ and (2.44),

$$(2.49)\qquad T_n - T_n^{(1)} = \frac{\bar{\ell}_n^{(1)}(T_n)}{\bar{\ell}_n^{(2)}(T_n)} = o(n^{-\frac{1}{2}+\delta}, n^{-\frac{j}{2}}) \ .$$

So, by (2.47),

$$(2.50)\qquad n^{\frac{1}{2}}\bar{\ell}_n^{(1)}(T_n^{(1)}) = o(n^{-\frac{1}{2}+2\delta}, n^{-\frac{j}{2}}) = o(n^{-\delta}, n^{-\frac{j}{2}})$$

for $\delta < \frac{1}{6}$. Case $i=0$ now follows. If the claim holds for $i$, then by (2.47) and induction,

$$\begin{aligned}\bar{\ell}_n^{(1)}(T_n^{(i+2)}) &= o(n^{-i-2\delta}, n^{-j/2}) \\ &= o(n^{-\frac{(i+1)}{2}-\delta}, n^{-\frac{j}{2}}) \ .\end{aligned}$$

Since

$$T_n^{(i+2)} - T_n^{(i+1)} = -\frac{\bar{\ell}_n^{(1)}}{\bar{\ell}_n^{(2)}}(T_n^{(i+1)}) = o(n^{-\frac{i}{2}-\delta}, n^{-\frac{j}{2}}) \ ,$$

the induction and result follow. □

## 3. Examples of situations in which the regularity conditions hold

The IID Case: Consider the following conditions.

$I_j(1)$: $\ell_1$ is differentiable to order $(j+3)$ and

$$E_\theta \sup\{|\ell_1^{(j+3)}(\theta')|^{j'}: \theta' \in K\} \leq M(K,K') < \infty$$

for $\theta \in K' \supset K$ arbitrary compacts, and $j' = j \vee 2$. $I_j(1)$ may be replaced by the condition

$$\sup E_\theta\{|\ell_1^{(j+4)}(\theta')|^{j'}: \theta' \in K\} \leq M(K,K') < \infty$$

$I_j(2)$: $$E_\theta|\ell_1^{(i)}(\theta)|^{j'+\delta}$$

bounded for $\theta \in K$ compact, $2 \leq i \leq j+2$.

Define

$I_j(3)$: $$\lambda_i(\theta) = E_\theta \ell_1^{(i)}(\theta) \ .$$

Under $I_j(1)$, $\theta \to \lambda_i(\theta)$ are continuous, $1 \leq i < j+2$. $\theta \to E_\theta[\ell_1^{(1)}]^2(\theta)$ is positive.

It is easy to see that $I_j(1) \Rightarrow C_j(1)$, $I_j(2) \Rightarrow C_j(2)$ and $I_j(1) - I_j(3)$ $\Rightarrow C_j(3)$ and

$$\lambda_1(\theta) = 0$$
$$I(\theta) = E_\theta[\ell_1^{(1)}]^2(\theta)$$

It is also easy to see that the minimum distance estimate $T_n$ constructed by Le Cam (1969), pp. 103-107 satisfies $E_j(2)$ provided that $I_j(1)$-$I_j(3)$ hold and the parameter is identifiable. We need only remark that if $F_n$ is the empirical distribution function and $F_\theta$ the true,

$$P_\theta[n^{1/2}|T_n-\theta| \geq c_n n^\delta] \sim P_\theta[\sup_x n^{1/2}\|\hat{F}_n - F_\theta\| \geq \Omega(c_n n^\delta)] = o(n^{-\alpha}) \quad \forall\alpha > 0$$

by the well-known Dvoretzky-Kiefer-Wolfowitz inequality. If we now require that, in addition to $I_j(1)$-$I_j(3)$,

$I_j(4)$: $$E_\theta|\ell_1^{(1)}(\theta)|^{j'+2}$$

is bounded, uniformly on compacts, then (2.44) holds by Bhattacharya and Ranga Rao (1976) p. 178. Thus Theorem 4 yields $\hat{\theta}_n$ which are suitable. There are many alternative possibilities for $\hat{\theta}_n$ including the construction of Pfanzagl and Wefelmeyer (1978), Bayes estimates and of course MLE's obeying $E_j(2)$. In any case, $I_j(1)$-$I_j(4)$ guarantee our theorems. All of these conditions save for $I_j(4)$ are implied by the conditions of Ghosh and Subramanyan (and Pfanzagl and Wefelmeyer). But $I_j(4)$ is only used in verifying $E_j$, a condition which is easily seen to be satisfied for $\hat{\theta}$, the M.L.E., under the Ghosh-Subramanyan conditions.

Independent Observations: Let $f_{kn}$ denote the density of $X_k$, $\ell_{kn}$ its log likelihood, etc. Assume $\ell_{kn}$ is (j+3) times differentiable and let $\ell_n^{(i)} = \sum_{k=1}^n \ell_{kn}^{(i)}$ be the $i^{th}$ derivative of $\ell_n$. Conditions $I_j(1)$, $I_j(2)$ generalize straightforwardly.

$I_j'(1)$: $$\frac{1}{n}\sum_{k=1}^n E_\theta\{\sup|\ell_{kn}^{(j+3)}(\theta')|^{j'} : \theta' \in K\} \leq M(K,K') < \infty$$

for $\theta \in K' \supset K$ both compact independent of $n$.

$I_j'(2)$: $$\frac{1}{n}\sum_{k=1}^n E_\theta|\ell_{kn}^{(i)}(\theta)|^{j'+\delta}$$

bounded for $\theta \in K$ independent of $n$, $2 \leq i \leq j+2$.

$I_j(3)$ becomes

$I_j'(3)$: $$\theta \to \frac{1}{n}\sum_{k=1}^n E_\theta \ell_{kn}^{(i)}(\theta) \quad \text{continuous,}$$
$$\frac{1}{n}\sum_{k=1}^n E_\theta \ell_{kn}^{(i)}(\theta) \to \lambda_i(\theta) \quad \text{uniformly on compacts,}$$
$$\theta \to \frac{1}{n}\sum_{k=1}^n E_\theta[\ell_{kn}^{(1)}(\theta)]^2 \quad \text{continuous.}$$

The existence of (Bayes) estimates satisfying $E_j(2)$ follows from Theorem III2.2 of Ibragimov-Hasminskii (1980) if in addition to $I_j'(1)$, $I_j'(3)$, we require as a replacement for identifiability,

$$\sum_{k=1}^n \int (f_{kn}^{1/2}(x,\theta+sn^{-1/2}) - f_{kn}^{1/2}(x,\theta))^2 \mu(dx) \geq c \min(|s|^\beta, |s|^2)$$

for some $\beta > 0$, $c < \infty$ independent of $s$, $\theta$. Theorem 4 then yields estimates satisfying $E_j(1)$, $E_j(2)$ provided that we have

$I_j'(4)$: $$\frac{1}{n}\sum_{k=1}^{n} E_\theta|\ell_{kn}^{(1)}(\theta)|^{j'+2}$$

bounded, independent of $n$, on compacts.

Markov Processes: For simplicity we consider Markov chains with starting density $f(x_1,\theta)$ and transition densities $f(x_k,x_{k+1},\theta)$ with respect to a $\sigma$-finite measure $\mu$ on $X$. Following Billingsley (1961) assume the existence of a unique stationary distribution $S_\theta(dx)$ such that for each $x \in X$

$$(3.1) \qquad P_\theta(\cdot|x) << S_\theta$$

where $$P_\theta(A|x) = \int_A f(x,y,\theta)\mu(dy) .$$

Also assume that the Markov chain is aperiodic. Condition 3.1 holds for a discrete state space provided that for each $\theta$ the chain is irreducible and positive recurrent. Assume $\ell(x,\theta) = \log f(x,\theta)$, $\ell(x,y,\theta) = \log f(x,y,\theta)$ are (j+3) times continuously differentiable and

$M_j(1)$: $E_\theta\{\sup_K[|\ell^{(j+3)}(X_1,\theta')|^{j'+\delta} + |\ell^{(j+3)}(X_1,X_2,\theta')|^{j'+\delta}]\}$
uniformly bounded for $\theta \in K' \supset K$

$M_j(2)$: $E_\theta[|\ell^{(i)}(X_1,\theta)|^{j'+\delta} + |\ell^{(i)}(X_1,X_2,\theta)|^{j'+\delta}]$
uniformly bounded for $\theta \in K$

$M_j(1)$, $M_j(2)$ and boundedness on compacts of $\lambda_i$ below imply $C_j(1)$, $C_j(2)$. To see this, suppose without loss of generality that the initial distribution is stationary and write, for example,

$$(3.2) \quad \bar{\ell}_n^{(i)}(X^{(n)},\theta) = n^{-1}\sum_{k=1}^{n-1}[\ell^{(i)}(X_k,X_{k+1},\theta) - \lambda_i(\theta)] + \frac{n-1}{n}\lambda_i(\theta) + n^{-1}\ell^{(i)}(X_1,\theta)$$

where $\lambda_i(\theta) = E_\theta\ell^{(i)}(X_1,X_2,\theta)$.

Since $(X_1,X_2),(X_2,X_3),\ldots$ is a stationary mixing sequence with exponential rate (see Doob (1953) p. 221, (7.1)), we can apply the moment bounds for sums of mean 0 functions of such variables, see e.g. Ibragimov and Linnik (1971) Lemma 18.5.2.

To get $C_j(3)$, having defined $\lambda_i$ above, we require

$M_j(3)$: $\lambda_i$ continuous

$\theta \to E_\theta \ell^2(X_1,X_2,\theta)$ continuous and positive

The existence of estimates satisfying $E_j(2)$ follows as in the i.i.d. case, using a Dvoretzky-Kiefer-Wolfowitz inequality for the empirical distribution function of $\phi$-mixing random variables (Sen (1974) Theorem 3.2). Theorem 4 is applicable if also

$M_j(4)$: $E_\theta|\ell^{(1)}(X_1,X_2,\theta)|^{j'+3}$ is bounded on compacts .

These results require the application of Theorem 2.11, Götze-Hipp (1982) which guarantee that the moderate deviation estimates of Bhattacharya and Ranga Rao continue to hold in this situation. The conditions of Theorem 2.11 are guaranteed by (3.1) since the chain is then strongly mixing with exponential rate.

These conditions and situations are given as samples only. More general classes of dependent situations to which these conclusions apply may be obtained, for instance by modifying the conditions in Basawa and Rao (1980) Section 10.3.

## References


BASAWA, I. V. and RAO, Prakasa (1980). *Statistical Inference for Stochastic Processes.* Academic Press, New York.

BERKSON, J. (1980). Minimum chi-square not maximum likelihood (with discussion). *Ann. Statist.* **8**, 457-487.

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximations and Asymptotic Expansions.* J. Wiley, New York.

BICKEL, P. J., CHIBISOV, D. M. and VAN ZWET, W. R. (1981). On efficiency of first and second order. *International Statistical Review* **49**, 169-175.

BILLINGSLEY, P. (1961). *Statistical Inference for Markov Processes.* The University of Chicago Press, Chicago.

BURNAŞEV, M. V. (1982). Investigation of Second Order Properties of Statistical Estimators in a Scheme of Independent Observations. Mathematics U.S.S.R. *Izvestija* **18**, 439-467.

CHIBISOV, D. M. (1974). Asymptotic expansions for some asymptotically optimal tests. In *Proc. of the Prague Symp. on Asymptotic Statistics*, vol. **2**, (J. Hájek, ed.), 37-68. Charles University, Prague.

EFRON, B. (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.* **3**, 1189-1242.

GHOSH, J. K., SINHA, B. K. and JOSHI, S. M. (1982). Expansions for posterior probability and integrated Bayes risk. *Proc. III Purdue Symp. on Decision Theory and Related Topics*, vol. II, J. Berger and S. Gupta, Eds., Academic Press, New York.

GHOSH, J. K., SINHA, B. K. and WIEAND, H. S. (1980). Second order efficiency of the M.L.E. with respect to any bowl-shaped loss function. *Ann. Statist.* **8**, 506-521.

GHOSH, J. K. and SUBRAMANYAM, K. (1974). Second-order efficiency of maximum likelihood estimators. *Sankhyā Ser. A* **36**, 325-358.

GÖTZE, F. and HIPP, C. (1982). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Warscheinlichkeitstheorie verw. Gebiete* **64**, 211-239.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1980). *Statistical Estimation--Asymptotic Theory.* Springer Verlag, New York.

IBRAGIMOV, I. A. and LINNIK, Y. (1971). *Independent and stationary sequences of random variables.* Wolters-Nordhoff Publishing, Groningen.

LE CAM, L. (1969). *Théorie asymptotique de la decision statistique.* Presses de l'Université de Montréal.

PFANZAGL, J. (1979). First order efficiency implies second order efficiency. In *Contributions to Statistics*, J. Jureckova, ed., 167-196. Academia, Prague.

PFANZAGL, J. and WEFELMEYER, W. (1978a). A third-order optimum property of the maximum likelihood estimator. *J. Multivariate Anal.* **8**, 1-29.

PFANZAGL, J. and WEFELMEYER, W. (1978b). An asymptotically complete class of tests. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **45**, 49-72.

SEN, P. K. (1974). Weak convergence of multidimensional empirical processes for stationary $\phi$-mixing processes. *Ann. Prob.* **2**, 147-154.

TAKEUCHI, K. (1982). Higher order asymptotic efficiency of estimators in decision procedures. *Proc. III Purdue Symp. on Decision Theory and Related Topics*, vol. II, J. Berger and S. Gupta, Eds., Academic Press, New York.

TAKEUCHI, K. and AKAHIRA, M. (1976). On the second order asymptotic efficiency of estimators. *Proc. 3rd Japan-USSR Symp. on Probability Theory*. Lecture Notes in Mathematics 550, pp. 604-638. Springer Verlag, New York.

TAKEUCHI, K. and AKAHIRA, M. (1978). Asymptotic optimality of the generalized Bayes estimator. *Reports of the University of Electro Communication* **28**, 37-45.

# TECHNICAL REPORTS

## Statistics Department

## University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review. Statistical Science, 1988, Vol.3 No. 2 239-271.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984, 579-611.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist. 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data. Scandinavian J. Statist., 1988, 15, 1-23.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data. Biometrika, 1987, 74, 799-808.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (Nov 1986, revised Aug 1988). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, 23, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model. J. Amer. Statist. Assoc., 1988, 83, 744-749.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (Aug 1987, revised Oct 1988). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Achieving information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (Oct. 1987, revised Mar. 1988, Oct. 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. To appear in Statistics a Guide to the Unknown.

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. DUBINS, L.E. and SCHWARZ, G. (December 1987). A sharp inequality for martingales and stopping-times.

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (Mar. 1989). a) A study of second- and third-order spectral procedures and maximum likelihood identification of a bilinear system. b) Some statistical aspects of NMR spectroscopy, *Actas del 2° congreso lantinoamericano de probabilidad y estadistica matematica*, Caracas, 1985.

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (Feb. 1988, revised Nov. 1988). Randomization in the two-armed bandit problem.

144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.

145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.

146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.

147. STARK, P.B. (March 1988). Strict bounds and applications.

148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.

149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.

150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.

151. NOLAN, D. (March 1988). Limit theorems for a random convex set.

152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.

153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.

154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.

155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.

156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.

157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.

158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.

159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.

160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.

161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).

162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.

163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform. (Revised by Tech Report No. 180).

164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.

165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.

166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.

167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.

168. LE CAM, L. (August 1988). Maximum likelihood an introduction.

169. BREIMAN, L. (Aug.1988, revised Feb. 1989). Submodel selection and evaluation in regression I. The X-fixed case and little bootstrap.

170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.

171. STONE, C.J. (September 1988). Large-sample inference for logspline models.

172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

173. MILLAR, P.W. (October 1988). Optimal estimation in the non-parametric multiplicative intensity model.

174. YOR, M. (October 1988). Interwinings of Bessel processes.

175. ROJO, J. (October 1988). On the concept of tail-heaviness.

176. ABRAHAMS, D.M. and RIZZARDI, F. (September 1988). BLSS - The Berkeley interactive statistical system: An overview.

177. MILLAR, P.W. (October 1988). Gamma-funnels in the domain of a probability, with statistical implications.

178. DONOHO, D.L. and LIU, R.C. (October 1988). Hardest one-dimensional subfamilies.

179. DONOHO, D.L. and STARK, P.B. (October 1988). Recovery of sparse signals from data missing low frequencies.

180. FREEDMAN, D.A. and PITMAN, J.A. (Nov. 1988). A measure which is singular and uniformly locally uniform. (Revision of Tech Report No. 163).

181. DOKSUM, K.A. and HOYLAND, ARNLJOT (Nov. 1988, revised Jan. 1989). A model for step-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution.

182. DALANG, R.C., MORTON, A. and WILLINGER, W. (November 1988). Equivalent martingale measures and no-arbitrage in stochastic securities market models.

183. BERAN, R. (November 1988). Calibrating prediction regions.

184. BARLOW, M.T., PITMAN, J. and YOR, M. (Feb. 1989). On Walsh's Brownian Motions.

185. DALANG, R.C. and WALSH, J.B. (Dec. 1988). Almost-equivalence of the germ-field Markov property and the sharp Markov property of the Brownian sheet.

186. HESSE, C.H. (Dec. 1988). Level-Crossing of integrated Ornstein-Uhlenbeck processes

187. NEVEU, J. and PITMAN, J.W. (Feb. 1989). Renewal property of the extrema and tree property of the excursion of a one-dimensional brownian motion.

188. NEVEU, J. and PITMAN, J.W. (Feb. 1989). The branching process in a brownian excursion.

189. PITMAN, J.W. and YOR, M. (Mar. 1989). Some extensions of the arcsine law.

190. STARK, P.B. (Dec. 1988). Duality and discretization in linear inverse problems.

191. LEHMANN, E.L. and SCHOLZ, F.W. (Jan. 1989). Ancillarity.

192. PEMANTLE, R. (Feb. 1989). A time-dependent version of Pólya's urn.

193. PEMANTLE, R. (Feb. 1989). Nonconvergence to unstable points in urn models and stochastic approximations.

194. PEMANTLE, R. (Feb. 1989). When are touchpoints limits for generalized Pólya urns.

195. PEMANTLE, R. (Feb. 1989). Random walk in a random environment and first-passage percolation on trees.

196. BARLOW, M., PITMAN, J. and YOR, M. (Feb. 1989). Une extension multidimensionnelle de la loi de l'arc sinus.

197. BREIMAN, L. and SPECTOR, P. (Mar. 1989). Submodel selection and evaluation in regression — the X-random case.

198. BREIMAN, L., TSUR, Y. and ZEMEL, A. (Mar. 1989). A simple estimation procedure for censored regression models with known error distribution.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.