Revista de la
Unión Matemática Argentina
Volumen 27, 1974.

SEQUENTIAL ESTIMATION OF A TRUNCATION PARAMETER

Eduardo Warner de Weerth

1. INTRODUCTION.

Let X_1 be a random variable defined on a probability space $(\Omega, \mathcal{B}, P)$ and whose distribution belongs to the family $\{P_\theta : \theta \in \Theta\}$ where Θ is a finite or infinite interval of the real line R. We want to estimate the true value of the parameter θ on the basis of a series of independent observations $X_1, X_2, \ldots$ of the variable X_1. After taking each observation we decide, on the basis of the observations that we have already taken, whether to take still another observation or to stop and estimate. We suppose that it costs us c units to take an observation and that if we estimate θ by d when its true value is θ we lose $L(\theta,d)$. Our total loss then, upon taking n observations and deciding d is $L(\theta,d) + cn$.

A sequential decision procedure is a pair (δ,t) where t is the stopping rule which tells us, for each possible sequence of observations $x = (x_1,x_2,\ldots)$, when to stop, and $\delta = \{\delta_n : n = 1,2,\ldots\}$ is a sequence of terminal decision functions. The meaning of δ is as follows: for each n, if we stopped after taking n observations ($t(x) = n$) and have observed $x_1,x_2,\ldots,x_n$, we must decide $\theta = \delta_n(x_1,\ldots,x_n)$. If we use the procedure (δ,t), our loss as a function of $x \in R^\infty$ is given by:

$$L[\theta,\delta_{t(x)}(x)] + ct(x)$$

if θ is the true value of the parameter. The average loss is called the risk and is given by:

$$R(\delta,t,\theta) = E_\theta[L(\theta,\delta_t)] + cE_\theta(t)$$

The most desirable sequential procedure would be one that minimizes the risk uniformly in θ. Such a procedure, however, is unattainable in all but trivial cases.

In the Bayesian setup we assume that there is not one true value of the parameter but rather that θ is a random variable with a distribution Ψ. The variables $X_1, X_2, \ldots$ are then assumed to be independent and identically distributed given θ. This determines the joint distribution of $\theta, X_1, X_2, \ldots$. A procedure will be optimal in the Bayes sense (Bayes procedure) if it minimizes the Bayes risk: $\int_{\Theta} R(\delta, t, \theta)\Psi(d\theta)$.

Assuming that $P_\theta(dx_1) = f(x_1, \theta)\mu(dx_1)$, where μ is a σ-additive measure, and $\Psi(d\theta) = \psi(\theta)d\theta$, the Bayes risk for a sample of fixed size n is:

$$(1.1) \quad \int_{\Theta}\{\int_{R^n} L[\theta, \delta_n(x_1, \ldots, x_n)] f(x_1, \theta)\ldots f(x_n, \theta)\mu(dx_1)\ldots\mu(dx_n)\}.$$

$$.\ \psi(\theta)d\theta + cn =$$

$$= \int_{R^n}\{\int_{\Theta} L[\theta, \delta_n(x_1, \ldots, x_n)]\psi(\theta | x_1, \ldots, x_n)d\theta\}Q_n(dx_1, \ldots, dx_n) + cn$$

where $\psi(\theta | x_1, \ldots, x_n) = \dfrac{f(x_1, \theta)\ldots f(x_n, \theta)\psi(\theta)}{\int_{\Theta} f(x_1, \lambda)\ldots f(x_n, \lambda)\psi(\lambda)d\lambda}$ is the conditional distribution of θ given $X_1, \ldots, X_n$ and Q_n is the marginal distribution of $X_1, \ldots, X_n$.

Suppose that for each n there exists a measurable function of the observations, $\tilde{\theta}_n = \tilde{\theta}_n(X_1, \ldots, X_n)$, such that

$$(1.2) \quad \int_{\Theta} L(\theta, \tilde{\theta}_n)\psi(\theta | X_1, \ldots, X_n)d\theta = \min_d \int_{\Theta} L(\theta, d)\psi(\theta | X_1, \ldots, X_n)d\theta$$

From (1.1) and (1.2) it follows easily (cf.[1]) that the sequence of terminal decision functions of a Bayes procedure must be: $\{\tilde{\theta}_n\}$.

The corresponding stopping rule is the one that minimizes:

$$(1.3) \quad E[Y_t + ct] = \int_{R^\infty}[Y_{t(x)}(x) + ct(x)]Q(dx)$$

where Q is the marginal distribution of $(X_1,X_2,\dots)$ and

$$(1.4)\qquad Y_n(X_1,\dots,X_n) = \int_{\Theta} L[\theta,\tilde{\theta}_n(X_1,\dots,X_n)]\,\psi(\theta|X_1,\dots,X_n)\,d\theta$$

is called the posterior (Bayes) risk after n observations.

Bickel and Yahav introduced in [2] the notion of asymptotically pointwise optimal (A.P.O.) stopping rules. A stopping rule t will be called pointwise optimal if $Q[\frac{Y_t+ct}{Y'_t+ct'} \leqslant 1] = 1$ for any other stopping rule t'.

Pointwise optimal rules exist only in essentially deterministic situations. One such case obtains when there exists a random variable V, $0 < V < \infty$ such that

$$(1.5)\qquad Y_n = \frac{V}{n}, \quad n = 1,2,\dots .$$

In this case it is easy to see that the stopping rule:

$$\text{"stop as soon as } \frac{V}{n(n+1)} \leqslant c\text{"}$$

is pointwise optimal.

In what follows we will call a function from the interval $(0,\infty)$ to the set T of all possible stopping rules for the sequence $\{Y_n\}$ also a stopping rule. Then, a stopping rule t(c) is called A.P.O. if

$$\lim_{c\to 0} \frac{Y_{t(c)}+ct(c)}{\inf\{Y_n+cn:\ n=1,2,\dots\}} = 1 \quad \text{a.s.}\quad Q$$

In their paper [3] Bickel and Yahav proved

THEOREM 1.1. *If*

(1.6) $Y_n > 0$ *a.s.* Q *for all* n

(1.7) $n^{\beta}Y_n \to V$ *a.s.* Q, *where* $\beta > 0$ *and* $0 < V < \infty$ *a.s.* Q

then the stopping rule $\tilde{t}(c)$: *stop for the first* n *such that*

$$(1.8)\qquad Y_n[1 - (\frac{n}{n+1})^{\beta}] \leqslant c$$

is A.P.O.

REMARK. The theorem as stated in [3] is somewhat more general. It is not hard to see that any rule s(c) such that

$$\frac{s(c)}{\tilde{t}(c)} \longrightarrow 1 \quad a.s. \quad Q$$

is also A.P.O. In particular, the rule t'(c) defined by: "stop for the first n such that $Y_n \frac{\beta}{n+1} \leq c$" is A.P.O. The rule t'(c) is obtained when we replace in (1.8) the expression $1 - (\frac{n}{n+1})^\beta$ by its first order approximation $\frac{\beta}{n+1}$.

The relation between the deterministic case (1.5) and (1.7) is clear.

Bickel and Yahav proved furthermore (cf. also [4]) that the relation (1.7) was fulfilled (in the case of quadratic loss, for $\beta=1$) under very general conditions that may be described roughly as those insuring the existence and asymptotic normality of maximun likelihood estimators.

It is the purpose of this work to extend the results of Bickel and Yahav to a case that clearly does not satisfy the above conditions, the case of the estimation of a truncation parameter. The model we consider is as follows.

Let h be a strictly positive and continuous function on the open interval $\Theta = (\theta_0, \infty)$. Here θ_0 can be finite or $-\infty$. For each $\theta \in \Theta$ we assume that

$$C(\theta) = \int_{\theta_0}^{\theta} h(x)dx < \infty$$

Then, for each $\theta \in \Theta$, the function $\frac{h(x)}{C(\theta)}$, $0 \leq x \leq \theta$, is the density function of a probability distribution P_θ concentrated on the interval (θ_0, θ), and we seek to estimate the truncation point θ sequentially.

In section 2 we prove that (1.7) (with $\beta = 2$) holds in this model for loss functions of the type $L(\theta,d) = B(\theta)(\theta-d)^2$ where $B(\theta)$ is a positive, continuous function of θ such that

$\int_\Theta (1+\theta^2)B(\theta)\psi(\theta)d\theta < \infty$. It then follows from the results of Bickel and Yahav that the stopping rules:

$\tilde{t}(c)$: stop for the first n such that

$$Y_n[1 - (\frac{n}{n+1})^2] \leq c$$

$t'(c)$: stop for the first n such that

$$Y_n \frac{2}{n+1} \leq c$$

are A.P.O.

In section 3 we show, using a theorem of Bickel and Yahav [3], that the rules $\tilde{t}(c)$ and $t'(c)$ are also asymptotically optimal in the sense of Kiefer and Sacks [5].

It is interesting to remark that in our development the function $\frac{h(\theta)}{C(\theta)}$ plays a role analogous to that played by the Fisher information (or information matrix) in the work of Bickel and Yahav.

It is clear, furthermore, that if θ_o is finite we can always assume that $\theta_o = 0$.

Finally, we wish to remark that all our results, obtained for a distribution truncated above, can easily be translated into the corresponding result for a distribution truncated below.

2. A LIMIT THEOREM.

We turn now to the truncation parameter model described in section 1 and use the same notation as there. We assume that the loss function is

$$L(\theta,d) = B(\theta)(\theta-d)^2 \tag{2.1}$$

where $B(\theta)$ is a positive and continuous function of θ. The prior density $\psi(\theta)$ is a positive continuous and bounded function of θ and we assume that $B(\theta)$ and $\psi(\theta)$ are such that

$$\int_{\theta_o}^{\infty} B(\theta)(1+\theta^2)\psi(\theta)d\theta < \infty \quad \text{and} \quad \int_{\theta_o}^{\infty} \theta^2 \psi(\theta)d\theta < \infty \tag{2.2}$$

We denote by $\tilde{\theta}_n$ the Bayes estimator based on n observations and we will now show that it exists and compute its value. The conditional distribution of θ given $X_1,\ldots,X_n$ has the density

$$(2.3)\qquad \psi(\theta|x_1,\ldots,x_n) = (\psi(\theta)C^{-n}(\theta)I_{[\theta\ge\hat{\theta}_n]})/\int_{\hat{\theta}_n}^{\infty}\psi(\lambda)C^{-n}(\lambda)d\lambda$$

where I_A denotes the indicator of the set A and $\hat{\theta}_n = \max(x_1,\ldots,x_n)$ is as before the M.L.E. based on n observations. It follows from (1.1) that the Bayes estimator is the one that attains the minimun in the expression

$$Y_n = \inf_d \int_{\theta_o}^{\infty} B(\theta)(\theta-d)^2\psi(\theta|x_1,\ldots,x_n)d\theta$$

From

$$Y_n = \inf_d \int_{\hat{\theta}_n}^{\infty} B(\theta)(\theta-d)^2\psi(\theta)C^{-n}(\theta)d\theta/\int_{\hat{\theta}_n}^{\infty}\psi(\lambda)C^{-n}(\lambda)d\lambda =$$

$$= \inf_d \int_{\hat{\theta}_n}^{\infty}(\theta-d)^2(B(\theta)\psi(\theta)C^{-n}(\theta)\,/\int_{\hat{\theta}_n}^{\infty}B(\lambda)\psi(\lambda)C^{-n}(\lambda)d\lambda)d\theta.$$

$$\cdot\frac{\int_{\hat{\theta}_n}^{\infty}B(\lambda)\psi(\lambda)C^{-n}(\lambda)d\lambda}{\int_{\hat{\theta}_n}^{\infty}\psi(\lambda)C^{-n}(\lambda)d\lambda}$$

it is clear that

$$(2.4)\qquad \tilde{\theta}_n = \int_{\hat{\theta}_n}^{\infty}\theta(B(\theta)\psi(\theta)C^{-n}(\theta)/\int_{\hat{\theta}_n}^{\infty}B(\lambda)\psi(\lambda)C^{-n}(\lambda)d\lambda)d\theta$$

and therefore

$$(2.5)\qquad Y_n = \int_{\theta_o}^{\infty}B(\theta)(\theta-\tilde{\theta}_n)^2\psi(\theta|x_1,\ldots,x_n)d\theta$$

We now prove

LEMMA 2.1. $\hat{\theta}_n \to \theta$ *a.s.* P_θ *for every* $\theta\in\Theta$.

Proof. $P_\theta[\bigcap_{n>N}\{|\theta-\hat{\theta}_n|<\varepsilon\}] = P_\theta\{\hat{\theta}_N > \theta-\varepsilon\} = 1 - P_\theta\{\hat{\theta}_N \le \theta-\varepsilon\}$

But

$$P_\theta\{\hat\theta_N \leqslant \theta-\varepsilon\} = \left(\frac{C(\theta-\varepsilon)}{C(\theta)}\right)^N$$

Since the term on the right goes to 0 as $N \longrightarrow \infty$ the lemma is proved.

In what follows we assume throughout that θ is fixed and that we are dealing with a fixed sequence of observations for which $\hat\theta_n \longrightarrow \theta$.

THEOREM 2.2. *Let* $\nu_n(s) = (C(\hat\theta_n)/C(sn^{-1}+\hat\theta_n))^n$. *Then*

$$(2.6)\quad \int_0^\infty (1+s^2)\left|B(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)\nu_n(s)-B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}\right|ds \to 0$$

as $n \longrightarrow \infty$.

Proof. We can write the integral in (2.6) as the sum of the following two integrals, where δ is a positive number to be determined later.

$$(2.7)\quad \int_0^{n\delta} (1+s^2)\left|B(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)\nu_n(s)-B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}\right|ds$$

$$(2.8)\quad \int_{n\delta}^\infty (1+s^2)\left|B(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)\nu_n(s)-B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}\right|ds$$

We first consider (2.7)

$$(2.9)\qquad \nu_n(s) = e^{-n[\log C(sn^{-1}+\hat\theta_n)-\log C(\hat\theta_n)]} = e^{-\frac{h(\xi)}{C(\xi)}s}$$

where $\hat\theta_n < \xi < sn^{-1}+\hat\theta_n$.

If $0 \leqslant s \leqslant n\delta$ and $n > N(\delta)$, we get

$$\theta-\delta < \hat\theta_n < \xi < \hat\theta_n+\delta \leqslant \theta+\delta$$

Therefore

$$(2.10)\qquad \nu_n(s) \leqslant e^{-ws} \text{ where } w = \inf\left\{\frac{h(x)}{C(x)} : |x-\theta| < \delta\right\}$$

It is clear, furthermore, that we can assume δ sufficiently small and N sufficiently large so that, in addition

(2.11) $B(sn^{-1}+\hat{\theta}_n) < 1 + B(\theta)$ for $0 \leq s \leq n\delta$ and $n > N$.

It then follows that if K is the upper bound of ψ,

$$(2.12)\quad \begin{aligned} &(1+s^2)\,|B(sn^{-1}+\hat{\theta}_n)\psi(sn^{-1}+\hat{\theta}_n)\nu_n(s)-B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}|\,I_{[0\leq s\leq n\delta]} \\ &\leq (1+s^2)[(1+B(\theta))Ke^{-ws} + B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}]\,I_{[0\leq s]} \end{aligned}$$

Since it is clear from (2.9) that

$$\nu_n(s) \longrightarrow e^{-\frac{h(\theta)}{C(\theta)}s}$$

it follows from (2.12) and the dominated convergence theorem that the integral (2.7) goes to 0 as $n \longrightarrow \infty$.

We now consider integral (2.8). From (2.9) it is clear that

$$\sup\{\nu_n(s):\ s \geq n\delta\} = e^{-n\delta\frac{h(\xi)}{C(\xi)}} \leq e^{-n\delta w}$$

where w is as before.

Therefore, putting $s = n(\theta-\hat{\theta}_n)$

$$(2.13)\quad \int_{n\delta}^{\infty}(1+s^2)\,|B(sn^{-1}+\hat{\theta}_n)\psi(sn^{-1}+\hat{\theta}_n)\nu_n(s)-B(\theta)\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}|\,ds \leq$$

$$\leq e^{-n\delta w}\int_{n\delta}^{\infty}(1+s^2)B(sn^{-1}+\hat{\theta}_n)\psi(sn^{-1}+\hat{\theta}_n)ds \;+$$

$$+\; B(\theta)\psi(\theta)\int_{n\delta}^{\infty}(1+s^2)e^{-\frac{h(\theta)}{C(\theta)}s}ds \;=$$

$$= ne^{-n\delta w}\int_{\delta+\hat{\theta}_n}^{\infty}[1+n^2(\theta-\hat{\theta}_n)^2]B(\theta)\psi(\theta)d\theta \;+$$

$$+\; B(\theta)\psi(\theta)\int_{n\delta}^{\infty}(1+s^2)e^{-\frac{h(\theta)}{C(\theta)}s}ds$$

Due to (2.2) this last expression goes to 0 as $n \longrightarrow \infty$ and the proof of the theorem is thus concluded.

We state for later use

COROLLARY 2.3.

$$(2.14)\qquad \int_0^\infty sB(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)\nu_n(s)ds \longrightarrow B(\theta)\psi(\theta)\left(\frac{C(\theta)}{h(\theta)}\right)^2$$

$$(2.15)\qquad \int_0^\infty B(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)\nu_n(s)ds \longrightarrow B(\theta)\psi(\theta)\frac{C(\theta)}{h(\theta)}$$

Proof. Immediate.

THEOREM 2.4. $n(\tilde\theta_n-\hat\theta_n) \longrightarrow \frac{C(\theta)}{h(\theta)}$ *a.s.* P_θ *as* $n \longrightarrow \infty$.

Proof. It follows from (2.4) that

$$n(\tilde\theta_n-\hat\theta_n) = \int_{\hat\theta_n}^\infty n(\theta-\hat\theta_n)B(\theta)\psi(\theta)C^{-n}(\theta)d\theta \Big/ \int_{\hat\theta_n}^\infty B(\lambda)\psi(\lambda)C^{-n}(\lambda)d\lambda$$

and making the substitution $s = n(\theta-\hat\theta_n)$ we get

$$n(\tilde\theta_n-\hat\theta_n) = \frac{\int_0^\infty sB(sn^{-1}+\hat\theta_n)\psi(sn^{-1}+\hat\theta_n)C^n(\hat\theta_n)C^{-n}(sn^{-1}+\hat\theta_n)ds}{\int_0^\infty B(vn^{-1}+\hat\theta_n)\psi(vn^{-1}+\hat\theta_n)C^n(\hat\theta_n)C^{-n}(vn^{-1}+\hat\theta_n)dv}$$

The theorem now follows from a direct application of Corollary 2.3.

THEOREM 2.5. $n^2Y_n \longrightarrow B(\theta)\left(\frac{C(\theta)}{h(\theta)}\right)^2$ *a.s.* P_θ *as* $n \longrightarrow \infty$.

Proof. Putting $s = n(\theta-\hat\theta_n)$

$$n^2Y_n = \int_{\hat\theta_n}^\infty B(\theta)(\theta-\tilde\theta_n)^2\psi(\theta|x_1,\ldots,x_n)d\theta =$$

$$= \int_{\hat\theta_n}^\infty B(\theta)[n(\theta-\hat\theta_n)-n(\tilde\theta_n-\hat\theta_n)]^2\psi(\theta|x_1,\ldots,x_n)d\theta =$$

$$= \frac{\int_0^\infty B(sn^{-1}+\hat\theta_n)[s-n(\tilde\theta_n-\hat\theta_n)]^2\psi(sn^{-1}+\hat\theta_n)\nu_n(s)ds}{\int_0^\infty \psi(vn^{-1}+\hat\theta_n)\nu_n(v)dv}$$

By Theorem 2.4. for $n > N$ sufficiently large

$$n(\tilde{\theta}_n - \hat{\theta}_n) \leq \frac{C(\theta)}{h(\theta)} + 1$$

and

$$[s - n(\tilde{\theta}_n - \hat{\theta}_n)]^2 \leq (\frac{C(\theta)}{h(\theta)} + 1)^2 + s^2$$

Therefore

$$[s - n(\tilde{\theta}_n - \hat{\theta}_n)]^2 B(sn^{-1} + \hat{\theta}_n)\psi(sn^{-1} + \hat{\theta}_n)\nu_n(s) \leq$$

$$\leq [(\frac{C(\theta)}{h(\theta)} + 1)^2 + s^2] B(sn^{-1} + \hat{\theta}_n)\psi(sn^{-1} + \hat{\theta}_n)\nu_n(s)$$

and it follows from Theorem 2.2, the dominated convergence theorem (see Loève [6] p.162) and (2.15) (with $B \equiv 1$), that

$$n^2 Y_n \longrightarrow \psi(\theta)B(\theta)\int_0^\infty (s - \frac{C(\theta)}{h(\theta)})^2 e^{-\frac{h(\theta)}{C(\theta)}s} ds / \psi(\theta)\frac{C(\theta)}{h(\theta)} = B(\theta)(\frac{C(\theta)}{h(\theta)})^2$$

as was to be proved.

It follows from Theorem 2.5 and Theorem 1.1 that the stopping rules $\tilde{t}(c)$ and $t'(c)$ defined in section 1 are A.P.O. in our truncation parameter model with loss functions of the type $L(\theta,d) = B(\theta)(\theta-d)^2$ and subject to the stated conditions on ψ and B.

3. ASYMPTOTIC OPTIMALITY.

Following Kiefer and Sacks [5], we say that a stopping rule t(c) is asymptotically optimal if

$$(3.1) \quad \lim_{c \to 0} \sup [E(Y_{t(c)} + ct(c)) / \inf\{E(Y_{s(c)} + cs(c)) : s(c) \in T\}] \leq 1$$

where T is the set of all stopping rules.

Then we have

THEOREM 3.1. (Bickel and Yahav [3]). *Under the conditions of Theorem 1.1 and if*

$$(3.2) \quad \sup_n n^\beta E(Y_n) < \infty$$

then the stopping rules $\tilde{t}(c)$ *and* $t'(c)$ *are asymptotically optimal.*

The following corollary is an immediate consequence of Theorem 3.1, upon applying (1.1) and the definition of Bayes risk (1.4).

COROLLARY 3.2. (Bickel and Yahav [3]). *If the conditions of Theorem 1.1 hold and if there exists a sequence of estimates* δ_n *such that*

$$\sup_n n^{\beta} \int_{\Theta} E_{\theta}[L(\theta,\delta_n)]\psi(\theta)d\theta < \infty$$

then the rules $\tilde{t}(c)$ *and* $t'(c)$ *are asymptotically optimal.*

We now apply this corollary to our truncation parameter model and prove

THEOREM 3.3. *In the truncation parameter model with loss function* $L(\theta,d) = B(\theta)(\theta-d)^2$, *if the conditions of section 2 are satisfied and furthermore*

$$\frac{h(\theta)}{C(\theta)} \geqslant a > 0 \quad \text{for every } \theta \tag{3.3}$$

then the rules $\tilde{t}(c)$ *and* $t'(c)$ *are asymptotically optimal.*

Proof. Integrating by parts,

$$E_{\theta}[(\theta-\hat{\theta}_n)^2] = \int_{\theta_o}^{\theta} (\theta-x)^2 d\left(\frac{C(x)}{C(\theta)}\right)^n = \lim_{A\to\theta_o} \int_A^{\theta} (\theta-x)^2 d\left(\frac{C(x)}{C(\theta)}\right)^n = \tag{3.4}$$

$$= \lim_{A\to\theta_o} \left[-(\theta-A)^2\left(\frac{C(A)}{C(\theta)}\right)^n + 2\int_A^{\theta}(\theta-x)\left(\frac{C(x)}{C(\theta)}\right)^n dx\right] \leqslant$$

$$\leqslant \lim_{A\to\theta_o} 2\int_A^{\theta}(\theta-x)\left(\frac{C(x)}{C(\theta)}\right)^n dx = 2\int_{\theta_o}^{\theta}(\theta-x)\left(\frac{C(x)}{C(\theta)}\right)^n dx$$

Making the substitution: $x = \theta+sn^{-1}$, we get

$$\int_{\theta_o}^{\theta}(\theta-x)\left(\frac{C(x)}{C(\theta)}\right)^n dx = n^{-2}\int_{n(\theta_o-\theta)}^{0}(-s)\left(\frac{C(\theta+sn^{-1})}{C(\theta)}\right)^n ds \leqslant \tag{3.5}$$

$$\leqslant n^{-2}\int_{n(\theta_o-\theta)}^{0}(-s)e^{as}ds < n^{-2}\int_{-\infty}^{0}(-s)e^{as}ds$$

because by (3.3)

$$\left(\frac{C(\theta+sn^{-1})}{C(\theta)}\right)^n = e^{n[\log C(\theta+sn^{-1})-\log C(\theta)]} = e^{\frac{h(\xi)}{C(\xi)}s} \leqslant e^{as}$$

for every $n(\theta_o-\theta) \leqslant s \leqslant 0$ $(\theta+sn^{-1} < \xi < \theta)$.

From (3.4) and (3.5) it follows that

$$n^2 B(\theta)E_\theta[(\theta-\hat{\theta}_n)^2] \leqslant KB(\theta)$$

where K is a constant.

Then

$$n^2\int_{\theta_o}^{\infty} B(\theta)E_\theta[(\theta-\hat{\theta}_n)^2]\psi(\theta)d\theta \leqslant K\int_{\theta_o}^{\infty} B(\theta)\psi(\theta)d\theta < \infty$$

by (2.2).

The theorem now follows from Corollary 3.2.

4. REMARK.

An important corollary can be obtained from Theorem 2.2. Putting $B \equiv 1$ in (2.6) and disregarding the s^2 we obtain:

$$(4.1)\quad \int_0^{\infty}\left|\psi(sn^{-1}+\hat{\theta}_n)\nu_n(s) - \psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}\right|ds \to 0 \text{ as } n \to \infty$$

If we now divide by

$$(4.2)\quad \int_0^{\infty}\psi(vn^{-1}+\hat{\theta}_n)\nu_n(v)dv = A_n$$

we get

$$(4.3)\quad \int_0^{\infty}\left|\frac{\psi(sn^{-1}+\hat{\theta}_n)C^{-n}(sn^{-1}+\hat{\theta}_n)}{\int_0^{\infty}\psi(vn^{-1}+\hat{\theta}_n)C^{-n}(vn^{-1}+\hat{\theta}_n)dv} - \frac{\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}}{A_n}\right|ds =$$

$$= \int_0^{\infty}\left|\psi(sn^{-1}+\hat{\theta}_n|X_1,\ldots,X_n) - \frac{\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}}{A_n}\right|ds \to 0 \text{ as } n \to \infty$$

Furthermore

$$(4.4)\quad \int_0^\infty \left|\psi(sn^{-1}+\hat{\theta}_n|X_1,\ldots,X_n) - \frac{h(\theta)}{C(\theta)}\, e^{-\frac{h(\theta)}{C(\theta)}s}\right| ds \leqslant$$

$$\leqslant \int_0^\infty \left|\psi(sn^{-1}+\hat{\theta}_n|X_1,\ldots,X_n) - \frac{\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}}{A_n}\right| ds +$$

$$+ \int_0^\infty \left|\frac{\psi(\theta)e^{-\frac{h(\theta)}{C(\theta)}s}}{A_n} - \frac{h(\theta)}{C(\theta)}\, e^{-\frac{h(\theta)}{C(\theta)}s}\right| ds \to 0 \quad \text{as} \quad n \to \infty$$

because of (4.3) and (2.15).

The limit theorem for the posterior distribution embodied in (4.4) is an analogue of the Bernstein-Von Mises theorem (cf. Bickel and Yahav [4], Theorem 2.2).

ACKNOWLEDGEMENT. This research is part of my Ph. D. Thesis completed at the University of California, Berkeley, under the guidance of Prof. Peter J. Bickel. The problem was suggested as well by Prof. Bickel to whom I wish to express my deepest gratitude.

REFERENCES

[1] ARROW,K., BLACKWELL,D. and GIRSHICK,M.A., *Bayes and minimax solutions of sequential decision problems*, Econometrica 17, pp. 213-244 (1949).

[2] BICKEL,P.J. and YAHAV, J.A., *Asymptotically pointwise optimal procedures in sequential analysis*, Proc. Fifth Berk. Symp. Math. Stat. Prob. (1965).

[3] BICKEL, P.J. and YAHAV,J.A., *Asymptotically optimal Bayes and minimax procedures in sequential estimation*, Ann. Math. Statist. 39, pp. 442-456 (1968).

[4] BICKEL,P.J. and YAHAV,J.A., *Some contributions to the asymptotic theory of Bayes solutions*, Zeit. Wahr. verw. Geb. 11, pp. 257-276 (1969).

[5] KIEFER,J. and SACKS,J., *Asymptotically optimal sequential inference and design*, Ann. Math. Statist. 34, pp. 705-750 (1963).

[6] LOÈVE,M., *Probability theory*, 3rd ed., D. Van Nostrand Co. Inc. (1963).

Universidad Nacional del Sur
Bahía Blanca, Argentina

Recibido en marzo de 1974

En el curso de la impresión del presente ejemplar de la Revista de la UMA hemos sufrido la infortunada pérdida de nuestros compañeros de trabajo y queridos amigos Evelio T. Oklander y Ricardo Alemany, a cuya memoria dedicamos este trabajo.