The Annals of Statistics
1981, Vol. 9, No. 5, 1107–1113

CONSISTENCY OF MAXIMUM LIKELIHOOD AND BAYES ESTIMATES

BY HELMUT STRASSER

University of Bayreuth, Germany

It is proved in the paper that every possible set of conditions which implies consistency of the maximum likelihood method also implies consistency of Bayes estimates for a large class of prior distributions. The method of the proof is based on a necessary and sufficient condition for the consistency of maximum likelihood estimates which is due to Perlman (1972). Perlman's result is improved as far as certain questions of measurability are concerned.

1. Introduction. Assume that $(X, \mathscr{A})$ is a measurable space and $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ is a family of probability measures $P_\theta \mid \mathscr{A}$. Assume further that $\mathscr{P}$ is dominated by a σ-finite measure $\mu \mid \mathscr{A}$ and let

$$h_\theta = \frac{dP_\theta}{d\mu}, \qquad \theta \in \Theta.$$

No conditions are imposed on the structure of the parameter space Θ.

Let $(\theta_n)_{n \in N}$ be a sequence of estimates of $\theta \in \Theta$ which is approximately a sequence of maximum likelihood estimates (AML) for the n-fold independent repetitions of $(X, \mathscr{A}, \mathscr{P})$, $n \in N$. The well-known paper of Wald (1949) contains a general set of sufficient conditions for consistency of such a sequence of AML-estimates. Wald's ideas are used and modified in subsequent papers of LeCam (1953) and Pfanzagl (1969). An important paper is Perlman's (1972) where it is shown that a suitable modification of Wald's approach even leads to conditions which simultaneously are necessary and sufficient for consistency of AML-estimates.

Several papers have tried to use Wald's approach in order to obtain consistency proofs for Bayes estimates. Notable examples are LeCam (1953), and Bickel and Yahav (1969). Essentially it turned out that any set of conditions for consistency of AML-estimates is also sufficient for consistency of Bayes estimates. This fact leads to the suggestion that the mere fact of consistency of AML-estimates could be used as a sufficient condition for consistency of Bayes estimates. If such a result can be proved then any set of sufficient conditions for consistency of AML-estimates which is known or will be stated in the future is automatically sufficient for consistency of Bayes estimates.

In the present paper we prove a result of this type. Assuming that every sequence of AML-estimates is consistent, we obtain from Perlman's theorem a necessary condition for which we show that it is sufficient for consistency of Bayes estimates. The assertion is valid under some regularity conditions on $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. These conditions are rather weak. In particular, it is important to note that they do not imply consistency of AML-estimates.

As indicated above the main result of the present paper consists of the fact that Perlman's criterion is sufficient for consistency of Bayes estimates. Taken literally, such an assertion is of theoretical value only, due to the technical spirit of Perlman's criterion. Nevertheless, the practical value of our result is the possibility to avoid Perlman's condition for the proof of consistency of Bayes estimates.

Consider e.g. a case where the original conditions of Wald are not satisfied but where

Received December 1977; revised December 1980.
AMS 1970 *subject classifications.* 62F10, 62E20, 62F15.
Key words and phrases. Maximum likelihood estimates, Bayes estimates, consistent estimation.

it is possible to prove consistency of AML-estimates by ad hoc methods. Such an example can be found in Kiefer and Wolfowitz (1956, page 904f). There the conditions for consistency of Bayes estimates of Bickel and Yahav (1969) also do not work since their integrability condition (A 2.9) is not satisfied. However, in view of our result the mere knowledge of consistency of AML-estimates implies consistency of Bayes estimates.

Conversely, there are well-known examples where AML-estimates are not consistent whereas Bayes estimates are; see, e.g. Schwartz, (1965, page 16f). In view of the main result of the present paper, there is little hope to obtain examples where AML-estimates are consistent and Bayes estimates are not. Therefore, if one tries to get a consistent method of estimation which works for a large class of statistical models then the Bayes method seems to be a better candidate than the maximum likelihood method. The success which is obtained in this direction by Schwartz (1965) supports such a suggestion.

In his original paper Perlman considers AML-estimates regardless of whether they are measurable or not. Hence the sufficiency part of his proof yields consistency of a larger class of estimator sequences than usual. The necessity part, however, is based on a slightly overly strong assumption. We show by a minor modification of Perlman's proof that his condition is even necessary if only all sequences of measurable AML-estimates are consistent. Large parts of this proof are quite similar to Perlman's and therefore we only indicate what is really essential. A detailed version is available from the author.

Finally, we note that all results of the present paper can be generalized to the Markov case without any substantial difficulty. An elaboration of this remark also is available.

2. The results. Let us topologize the parameter set Θ by the distance $d(\sigma, \tau) = \| P_\sigma - P_\tau \|$, $\sigma \in \Theta$, $\tau \in \Theta$, where $\| \cdot \|$ denotes the variational norm for measures on $(X, \mathscr{A})$. At a first glance, the topology seems to be unusual for statistical purposes. But Landers and Rogge (1972) showed that under the regularity conditions which are usually considered in the theory of maximum likelihood estimation, the topologies considered there are equivalent to the topology defined here.

The functions $f(x, \theta) = -\log h_\theta(x)$, $x \in X$, $\theta \in \Theta$, are called likelihood contrast functions. Let us denote

$$f_n(\mathbf{x}, \theta) = \frac{1}{n} \sum_{i=1}^n f(x_i, \theta) \quad \text{if} \quad \mathbf{x} \in X^n, n \in N, \theta \in \Theta.$$

It is easy to see that for every $n \in N$ the stochastic process $\{f_n(\cdot, \theta)\}_{\theta \in \Theta}$ is uniformly continuous in μ-measure. If we assume that (Θ, d) is a separable metric space, then we may choose a separable version of $\{f_n(\cdot, \theta)\}_{\theta \in \Theta}$

It should be noted, however, that the separable version satisfies

$$f_n(\mathbf{x}, \theta) = \frac{1}{n} \sum_{i=1}^n f(x_i, \theta)$$

only μ^n—a.e., $\theta \in \Theta$. Another consequence of separability of (Θ, d) is that the function $(\mathbf{x}, \theta) \to f_n(\mathbf{x}, \theta)$, $(\mathbf{x}, \theta) \in X^n x \Theta$ can be chosen $\mathscr{A}^n \otimes \mathscr{B}$-measurable where $\mathscr{B}$ denotes the Borel-σ-field of (Θ, d).

We require some regularity conditions. It should be noted that conditions (2) and (3) may be omitted if the densities h_θ, $\theta \in \Theta$, are continuous functions of θ.

Regularity Conditions.

(1) The metric space (Θ, d) is separable.

(2) The functions $(f_n(\cdot, \theta))_{\theta \in \Theta}$, $n \in N$, are separable and $\mathscr{A}^n \otimes \mathscr{B}$-measurable random functions.

(3) The densities h_θ, $\theta \in \Theta$, are lower semicontinuous, i.e. $\limsup_{n\to\infty} h_{\theta_n} \leq h_\theta$ μ—a.e. if $\lim_{n\to\infty} d(\theta_n, \theta) = 0$.

(4) For every $\theta \in \Theta$ and $\sigma \in \Theta$ there is an open neighbourhood $U_{\theta,\sigma}$ of σ such that $P_\theta^n (\inf f_n(\cdot, U_{\theta,\sigma})) > -\infty$ for at least one $n \in N$.

The last regularity condition needs some comments. If the condition is required for $n = 1$, then it is equivalent to a local integrability condition which has been used extensively in dealing with maximum likelihood estimates. Our condition, which obviously is considerably weaker, has been introduced by Perlman (1972).

There are some technicalities with defining maximum likelihood estimates since the contrast functions may attain the values $+\infty$ and $-\infty$. For convenience, we use the mapping $\varphi:[-\infty, +\infty] \to [-1, +1]$ which satisfies $\varphi(x) = x/(1 + |x|)$. This is a strictly increasing homeomorphism.

DEFINITION 2.1. Assume that regularity conditions (1) and (2) are satisfied. A sequence of $(\mathscr{A}^n, \mathscr{B})$-measurable estimates $T_n:X^n \to \Theta$ is a sequence of approximate maximum likelihood estimates (AML-estimates) at $\theta \in \Theta$ if

$$\limsup_{n\to\infty} \{\varphi\circ f_n(\mathbf{x}, T_n(\mathbf{x})) - \inf_{\sigma\in\Theta} \varphi\circ f_n(t, \sigma)\} \leq 0 \; P_\theta^N\text{—a.e.}$$

It is clear that every sequence of maximal likelihood estimates is a sequence of AML-estimates. In contrast to exact maximum likelihood estimates which need not exist, there are always AML-estimates if regularity condition (1) is satisfied. This is proved by a standard argument.

The following lemma improves the necessity part of Perlman's Theorem 4.1 in that we consider only measurable estimates.

LEMMA 2.2. *Assume that regularity conditions* (1) *and* (2) *are satisfied. Let $\theta \in M \subseteq \Theta$ where M is closed. Then the following assertions are equivalent*:

(1) *Every sequence of AML-estimates (T_n) at θ satisfies*

$$P_\theta^N \liminf_{n\to\infty} \{T_n \in M\} = 1.$$

(2) *There is some open set $U \supseteq M$ such that*

$$\lim_{n\to\infty} (\inf \varphi\circ f_n(\cdot, \Theta\backslash M) - \inf_{\sigma\in U} \varphi\circ f_n(\cdot, \sigma)) > 0 \qquad P_\theta^N\text{—a.e.}$$

(3) *For every sequence of open sets $U_i \supseteq M$, $i \in N$,*

$$\lim_{i\in N} \lim_{n\to\infty}(\inf \varphi\circ f_n(\cdot, \Theta\backslash M) - \inf_{\sigma\in U} \varphi\circ f_n(\cdot, \sigma)) > 0 \qquad P_\theta^N\text{—a.e.}$$

PROOF. It is clear that (3) $\Rightarrow$ (2). The proof of (2) $\Rightarrow$ (1) is essentially Wald's (1949). Thus it remains to show that (1) $\Rightarrow$ (3). This is done in Section 3 where the proofs of the paper are collected. □

Using the martingale techniques of Perlman, in particular proving an analog to his Theorem 2.2, we obtain a strengthened version of the preceding lemma.

THEOREM 2.3. *Assume that regularity conditions* (1)–(4) *are satisfied. Let $\theta \in M \subseteq \Theta$ where M is compact. Then the following assertions are equivalent*:

(i) *Every sequence of AML-estimates (T_n) at θ satisfies*

$$P_\theta^N \lim_{n\to\infty} \inf\{T_n \in M\} = 1.$$

(ii)
$$\lim_{n\to\infty} \inf_{\sigma\in\Theta\backslash M} f_n(\cdot, \sigma) > P_\theta(f_\theta) \qquad P_\theta^N\text{—a.e.}$$

PROOF. The proof follows from Lemma 2.2 by arguments which are similar to those of Perlman proving his Theorem 2.2. The main steps are indicated in paragraph 3. □

The fact that (ii) implies (i) is already contained in LeCam (1953) and elaborated by Pfanzagl (1969). It will be shown in the following that (ii) implies consistency of Bayes estimates.

Let $\pi \mid \mathscr{B}$ be a probability measure which serves as prior distribution. We require some conditions on π.

CONDITIONS ON PRIOR MEASURES.
(5) For every $\theta \in \Theta$ and $\epsilon > 0$

$$\pi\{\sigma \in \Theta : P_\theta(f_\sigma) < P_\theta(f_\theta) + \epsilon\} > 0.$$

(6) For every $\theta \in \Theta$ there is some $n_\theta \in N$ such that

$$P_\theta^n\left\{\mathbf{x} \in X^n : \int \prod_{i=1}^n h_\sigma(x_i)\pi\,(d\sigma) < \infty\right\} = 1 \qquad \text{if} \quad n \geq n_\theta.$$

If $\sigma \to P_\theta(f_\sigma)$ is continuous and open sets of (Θ, d) have positive π-measure, then condition (5) is satisfied. Condition (6) has been used previously by Bickel and Yahav (1969).

Bayes estimates are consistent if the posterior distributions concentrate around the true parameter value. The following definition makes things precise.

DEFINITION 2.4. Assume that regularity condition (6) is satisfied. For every $n \geq n_\theta$, $B \in \mathscr{B}$ and $\mathbf{x} \in X^n$ let

$$F_{n,\mathbf{x}}(B) = \frac{\int_B \prod_{i=1}^n h_\sigma(x_i)\pi\,(d\sigma)}{\int_\Theta \prod_{i=1}^n h_\sigma(x_i)\pi\,(d\sigma)}$$

if the denominator is positive, and $F_{n,\mathbf{x}}(B) = 0$ otherwise. The function $B \to F_{n,\mathbf{x}}(B)$, $B \in \mathscr{B}$, is called posterior distribution based on the prior measure π.

Now we are in a position to state our main result.

THEOREM 2.5. *Assume that regularity conditions* (1)–(6) *are satisfied. Let* $\theta \in M \subseteq \Theta$ *where* M *is compact. If every sequence of AML-estimates* (T_n) *satisfies*

(a) $$P_\theta^N \lim_{n\to\infty} \inf\{T_n \in M\} = 1,$$

then

(b) $$P_\theta^N\{\lim_{n\to\infty} \inf F_{n,\mathbf{x}}(M) = 1\} = 1.$$

PROOF. The proof is given in Section 3, which follows. □

Let us indicate what the preceding assertion means in terms of consistency. To this end we assume that (Θ, d) is a locally compact space. If every sequence of AML-estimates is (strongly) consistent then (a) is satisfied for every compact neighbourhood M_θ of any $\theta \in \Theta$. Hence (b) is also satisfied for every compact neighbourhood M_θ of any $\theta \in \Theta$ and hence even for every neighbourhood. This implies that the posterior distributions concentrate around θ with P_θ^N-probability one.

3. Proofs.

PROOF OF LEMMA 2.2. We have to show that (1) ⇒ (3). Assume that (3) is not satisfied for a sequence (U_i) of open sets $U_i \supseteq M$, $i \in N$. We will construct a sequence of AML-estimates (T_n), for which

$$P_\theta^N \lim_{n\in N}\{T_n \in M\} < 1.$$

Define

$$A: = \{\inf_{i\in N} \lim_{n\to\infty} (\inf_{\sigma\in\Theta\setminus M} \varphi\circ f_n(\cdot, \sigma) - \inf_{\sigma\in U_i} \varphi\circ f_n(\cdot, \sigma)) \le 0\}$$

and

$$A_{n,k}: = \cup_{i=1}^{\infty} \left\{\inf_{\sigma\in\Theta\setminus M} \varphi\circ f_n(\cdot, \sigma) - \inf_{\sigma\in U_i} \varphi\circ f_n(\cdot, \sigma) < \frac{1}{k}\right\}, \qquad n \in N, \qquad k \in N.$$

According to our assumption we have $P_\theta^N(A) = \alpha > 0$. It is easy to see that $A \subseteq \cap_{m=1}^{\infty} \cup_{n=m}^{\infty} A_{n,k}$ for every $k \in N$. It follows that there is a strictly increasing sequence $(N(k))_{k\in N}$ such that $N(k) \uparrow \infty$ and

$$P_\theta^N \cup_{n=N(k)}^{N(k+1)-1} (A_{n,k} \cup A) \ge \alpha\left(1 - \frac{1}{2^{k+1}}\right), \qquad k \in N,$$

and therefore

$$P_\theta^N \cap_{k=1}^{\infty} \cup_{n=N(k)}^{N(k+1)-1} (A_{n,k} \cap A) \ge \frac{\alpha}{2} > 0.$$

Choose $k(n) \in N$ in such a way that $N(k(n)) \le n < N(k(n) + 1)$, $n \in N$. Note that $A_{n,k(n)} \in \mathscr{A}^n$, $n \in N$. If $\mathbf{x} \in A_{n,k(n)}$ choose $T_n(\mathbf{x}) \in \Theta\setminus M$ such that

$$\varphi\circ f_n(\mathbf{x}, T_n(\mathbf{x})) - \inf_{\sigma\in\Theta\setminus M} \varphi\circ f_n(\mathbf{x}, \sigma) \le \frac{1}{k(n)},$$

and if $\mathbf{x} \notin A_{n,k(n)}$ choose $T_n(\mathbf{x}) \in \Theta$ such that

$$\varphi\circ f_n(\mathbf{x}, T_n(\mathbf{x})) - \inf_{\sigma\in\Theta} \varphi\circ f_n(\mathbf{x}, \sigma) \le \frac{1}{n}.$$

Since (Θ, d) is separable this can be done in such a way that T_n is $(\mathscr{A}^n, \mathscr{B})$-measurable. (T_n) is a sequence of AML-estimates at θ since for $\mathbf{x} \in A_{n,k(n)}$

$$\begin{aligned}\varphi\circ f_n(\mathbf{x}, T_n(\mathbf{x})) &\le \inf_{\sigma\in\Theta\setminus M} \varphi\circ f_n(\mathbf{x}, \sigma) - \frac{1}{k(n)} \\ &\le \min\left\{\inf_{\sigma\in\Theta\setminus M} \varphi\circ f_n(\mathbf{x}, \sigma), \sup_{i\in N} \inf_{\sigma\in U_i} \varphi\circ f_n(\mathbf{x}, \sigma) + \frac{1}{k(n)}\right\} + \frac{1}{k(n)} \\ &\le \inf_{\sigma\in\Theta} \varphi\circ f_n(\mathbf{x}, \sigma) + \frac{2}{k(n)}.\end{aligned}$$

But the sequence (T_n) leads to the desired contradiction since

$$A_{n,k(n)} \subseteq \{T_n \notin M\}, \qquad n \in N,$$

and

$$\begin{aligned}\cap_{m=1}^{\infty} \cup_{n=m}^{\infty} A_{n,k(n)} &= \cap_{k=1}^{\infty} \cup_{n=N(k)+1}^{\infty} A_{n,k(n)} \supseteq \cap_{k=1}^{\infty} \cup_{n=N(k)}^{N(k+1)-1} A_{n,k(n)} \\ &= \cap_{k=1}^{\infty} \cup_{n=N(k)}^{N(k+1)-1} A_{n,k}. \ \square\end{aligned}$$

PROOF OF THEOREM 2.3. Let us try to simplify the assertions (2) and (3) by computing the limits

$$\lim_{n\to\infty} \inf_{\sigma\in U} \varphi\circ f_n(\cdot, \sigma), \qquad U \subseteq \Theta.$$

First we consider an open subset $U \subseteq \Theta$ such that for some $n_0 \in N$

$$P_\theta^N(\inf_{\sigma\in U} f_{n_0}(\cdot, \sigma)) > -\infty.$$

Noting that $(\inf_{\sigma\in U} f_n(\cdot, \sigma))_{n\in N}$ is a reversed supermartingale for the symmetric σ-fields $\mathscr{T}_n \subseteq \mathscr{A}^n$, $n \in N$, we obtain from the Hewitt-Savage zero-one law that

$$\lim_{n\to\infty} \inf_{\sigma\in U} f_n(\cdot, \sigma) = \sup_{n\in N} P_\theta^N(\inf_{\sigma\in U} f_n(\cdot, \sigma)), \ P_\theta^N\text{—a.e.}$$

Consider a fixed element $\tau \in \Theta$. Lower semicontinuity implies for every sequence of open neighbourhoods U_k of τ, $U_k \downarrow \{\tau\}$, $U_k \subseteq U$, $k \in \mathbb{N}$, that

$$\lim_{k\to\infty} P_\theta^n(\inf_{\sigma\in U_k} f_n(\cdot, \sigma)) = P_\theta(f_\tau).$$

To prove it one has to choose $(f_\sigma)_{\sigma\in\Theta}$ in such a way that it has lower semicontinuous paths μ—a.e. This is possible in view of a result of Pfanzagl (1969). Thus we obtain that for every compact $M \subseteq \Theta$ and every $\epsilon > 0$ there is an open set $U_\epsilon \subseteq M$ such that

$$-\infty < \sup_{n\in\mathbb{N}} P_\theta^{\mathbb{N}}(\inf_{\sigma\in U_\epsilon} f_n(\cdot, \sigma)) \le \inf_{\sigma\in M} P_\theta(f_\sigma) \le \sup_{n\in\mathbb{N}} P_\theta^{\mathbb{N}}(\inf_{\sigma\in U_\epsilon} f_n(\cdot, \sigma)) + \epsilon.$$

Now it is easy to obtain the assertion of Theorem 2.3 from Lemma 2.2. □

Proof of Theorem 2.5. We will use the following inequality: If $A \in \mathscr{A}$, $P(A) > 0$, and $f \ge 0$ is $\mathscr{A}$-measurable then

$$\frac{1}{n}\log \int f^n\, dP \ge \log \frac{1}{P(A)} \int_A f\, dP + \frac{1}{n}\log P(A)$$

which follows from Hölder's inequality

$$\int 1_A f\, dP \le \left(\int f^n\, dP\right)^{1/n} \left(\int 1_A\, dP\right)^{1-1/n}.$$

As a first step we show that

$$(7) \qquad P_\theta^{\mathbb{N}}\left\{\lim_{n\to\infty}\sup F_{n,\mathbf{x}}(M') \ne \lim_{n\to\infty}\sup \frac{\int_{M'} \exp(-nf_n(\mathbf{x}, \sigma))\pi\,(d\sigma)}{\int_\Theta \exp(-nf_n(\mathbf{x}, \sigma))\pi\,(d\sigma)}\right\} = 0.$$

To this end it is sufficient to show that for some $n_0 \in \mathbb{N}$

$$\int_\Theta \exp(-nf_n(\mathbf{x}, \sigma))\pi\,(d\sigma) > 0 \qquad \text{if} \quad n \ge n_0, \quad P_\theta^{\mathbb{N}}\text{—a.e.}$$

Let $\epsilon > 0$ and define $M_\epsilon = \{\sigma \in \Theta : P_\theta(f_\sigma) < P_\theta(f_\sigma) + \epsilon\}$. It follows from Fatou's lemma that

$$\begin{aligned}
&\lim_{n\to\infty}\inf \frac{1}{n}\log \int_\Theta \exp(-nf_n(\mathbf{x}, \sigma))\pi\,(d\sigma) \\
&\quad\ge \lim_{n\to\infty}\inf \log \frac{1}{\pi(M_\epsilon)} \int_{M_\epsilon} \exp(-f_n(\mathbf{x}, \sigma))\pi\,(d\sigma) + \lim_{n\to\infty} \frac{1}{n}\log \pi(M_\epsilon) \\
&\quad\ge \log \frac{1}{\pi(M_\epsilon)} \int_{M_\epsilon} \exp(-P_\theta(f_\sigma))\pi\,(d\sigma) \ge -P_\theta(f_\theta) - \epsilon \qquad P_\theta^{\mathbb{N}}\text{—a.e.}
\end{aligned}$$

This proves (7).

Furthermore, we have (denoting $(\mathbf{x})_{n_\theta} := (x_{n_\theta+1}, x_{n_\theta+2}, \cdots)$)

$$\begin{aligned}
&\lim_{n\to\infty}\sup \frac{1}{n}\log \int_{M'} \exp(-nf_n(\mathbf{x}, \sigma))\pi\,(d\sigma) \\
&\quad= \lim_{n\to\infty}\sup \frac{1}{n}\log \int_{M'} \exp\left(-\textstyle\sum_{i=n_\theta+1}^{n} f(x_i, \sigma)\right)\exp\left(-\textstyle\sum_{i=1}^{n} f(x_i, \sigma)\right)\pi\,(d\sigma)
\end{aligned}$$

$$\leq \lim_{n\to\infty} \sup\left(-\frac{1}{n} \inf_{\sigma\in M'} \sum_{i=n_\theta+1}^{n} f(x_i, \sigma)\right) + \lim_{n\to\infty} \frac{1}{n} \log \int \prod_{i=1}^{n} h_\sigma(x_i)\pi\,(d\sigma)$$

$$= -\lim_{n\to\infty} \inf \inf_{\sigma\in M'}, f_n((\mathbf{x})_{n_\theta}, M') \qquad P_\theta^N\text{—a.e.}$$

Putting terms together, we obtain

$$\lim_{n\to\infty} \sup \frac{1}{n} \log F_{n,\mathbf{x}}(M') \leq -\lim_{n\to\infty} \inf_{\sigma\in M'} f_n((\mathbf{x})_{n_\theta}, \sigma) + P_\theta(f_\theta) + \epsilon \qquad P_\theta^N\text{—a.e.}$$

Let

$$A_r = \left\{\lim_{n\to\infty} \inf_{\sigma\in M'} f_n((\mathbf{x})_{n_\theta}, M') \geq P_\theta(f_\theta) + \frac{1}{r}\right\}, \qquad r \in \mathbb{N}.$$

Since $P_\theta^N | \mathscr{A}^N$ is strictly stationary, Theorem 2.3 implies

$$P_\theta^N(\cup_{r=1}^{\infty} A_r) = 1.$$

Now the assertion follows from

$$P_\theta^N \cup_{r=1}^{\infty} \left\{\lim_{n\to\infty} \sup \frac{1}{n} \log F_{n,\mathbf{x}}(M') \leq -\frac{1}{2r}\right\}$$

$$= P_\theta^N \cup_{r=1}^{\infty} \left\{\lim_{N\to\infty} \sup \frac{1}{n} \log F_{n,\mathbf{x}}(M') \leq -\frac{1}{r} + \frac{1}{2r}\right\}$$

$$\geq P_\theta^N \cup_{r=1}^{\infty} \left\{-\lim_{n\to\infty} \inf_{\sigma\in M'} f_n((\mathbf{x})_{n_\theta}, \sigma) + P_\theta(f_\theta) + \frac{1}{2r} \leq -\frac{1}{r} + \frac{1}{2r}\right\}$$

$$\geq P_\theta^N \cup_{r=1}^{\infty} A_r = 1. \ \square$$

REFERENCES

BICKEL, P. J. and YAHAV, J.A. (1969). Some contributions to the asymptotic theory of Bayes solutions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **11** 257–276.

KIEFER, J. and WOLFOWITZ, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Ann. Math. Statist.* **27** 887–906.

LANDERS, D. and ROGGE, L. (1972). Characterization of the topologies used in the theory of maximum likelihood estimation. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **21** 197–200.

LECAM, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. California Publ. Statist.* **1**, 277–330.

PERLMAN, M. D. (1972). On the strong consistency of approximate maximum likelihood estimators. *Proc. VIIth Berkeley Symp. Prob. Math. Statist.*, 263–281.

PFANZAGL, J. (1969). On the measurability and consistency of minimum contrast estimates. *Metrika* **14** 249–272.

SCHWARTZ, L. (1965). On Bayes procedures. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **4** 10–26.

WALD, A. (1949). Note on the consistency of the maximum likelihood estimates. *Ann. Math. Statist.* **20** 595–601.

BOX 3008
MATHEMATICAL INSTITUTE
UNIVERSITY OF BAYREUTH
D-8580 BAYREUTH, WEST GERMANY