The Annals of Mathematical Statistics
1972, Vol. 43, No. 1, 74-83

EXISTENCE AND CONSISTENCY OF MODIFIED MINIMUM CONTRAST ESTIMATES

BY D. LANDERS

University of Cologne

It is the purpose of this paper to explore the efficiency of a modified definition for maximum likelihood estimates which depends on the whole equivalence class of densities only and not—as in the classical case—on the particular choice of versions. We prove the existence of measurable maximum likelihood estimates in the new sense for compact metrizable families of probability measures without any continuity assumption for the densities. For appropriate families of probability measures the modified asymptotic maximum likelihood estimates are exactly the strongly consistent estimates. The paper uses Huber's concept of minimum contrast estimates which covers maximum likelihood estimates as a special case.

Introduction. Maximum likelihood estimates in the usual sense are based on fixed versions of the densities. In [6] and [4] it was shown that the particular choice of the versions essentially influences existence as well as consistency of the maximum likelihood estimates.

It is the purpose of this paper to explore the efficiency of a modified definition[1] of maximum likelihood and asymptotic maximum likelihood estimates which depends on the whole equivalence class of densities only and not on the particular choice of versions any more. It turns out that this new definition has a number of advantages:

(i) For a compact metrizable family of probability measures measurable maximum likelihood estimates exist without any continuity assumption for the densities (Theorem 3.1), whereas measurable maximum likelihood estimates in the usual sense exist in general only under the assumption that the densities are upper-semicontinuous (see [6], page 253).

(ii) Under appropriate regularity conditions (e.g. exactly the conditions of the "classical" consistency theorem) the property of being a sequence of asymptotic maximum likelihood estimates is not only sufficient but also necessary for the strong consistency of the sequence of estimates (Theorem 3.2).

Example 3.3 shows that strongly consistent estimates (and hence asymptotic maximum likelihood estimates in our sense) are, however, not necessarily asymptotic maximum likelihood estimates in the usual sense.

(iii) The proofs for the main theorems are simpler than in the "classical" case since with modified maximum likelihood estimates no measurability

Received July 22, 1970.

[1] This definition was suggested to the author by J. Pfanzagl.

problems (such as $\sup h_C$ is measurable for all compact sets C, see [6]) arise.

(iv) The presented concept covers the "classical" concept in all important cases.

Since this creates no additional difficulties we use in Section 2 Huber's [3] concept of minimum contrast and asymptotic minimum contrast estimates. In Section 3 we apply the results of Section 2 to the case of maximum likelihood estimates. Section 4 collects auxiliary lemmas for Section 2 and Section 3.

For references concerning the "classical" theory of minimum contrast and maximum likelihood estimates we refer the reader to [6].

1. Preliminaries. Let $(T, \mathscr{U})$ be a topological space. The Borel-field $\sigma(\mathscr{U})$ is the σ-field, generated by $\mathscr{U}$. $\bar{A}$ denotes the closure of a set $A \subset T$. Let $(L, \leqq)$ be a complete lattice, i.e. a partially ordered set such that $\inf L_0$ and $\sup L_0$ exist for each $L_0 \subset L$. If $t \to g_t \in L$, $t \in T$, is any map, we write for every $S \subset T$: $\inf g_S := \inf\{g_t : t \in S\}$, $\sup g_S := \sup\{g_t : t \in S\}$. The map $t \to g_t \in L$, $t \in T$, is *lower semicontinuous* (l.s.c.) [*upper semicontinuous* (u.s.c.)], iff $g_t = \sup_{t \in U \in \mathscr{U}} \inf g_U [g_t = \inf_{t \in U \in \mathscr{U}} \sup g_U]$ for every $t \in T$. If L is the real line this is the usual definition of l.s.c. and u.s.c.

Let μ be a σ-finite measure on a σ-field $\mathscr{F}$ on X. Denote by $M(\mathscr{F}, \mu)$ the set of all μ-equivalence classes of $\mathscr{F}$-measurable functions on X with values in $[-\infty, +\infty]$. Elements of $M(\mathscr{F}, \mu)$ will be denoted by $\tilde{f}$. It is known that $M(\mathscr{F}, \mu)$ endowed with the natural ordering (i.e. $f \leqq g$ iff $f(x) \leqq g(x)$ μ-a.e.) is a complete lattice ([2] page 335). Furthermore each subfamily of $M(\mathscr{F}, \mu)$ contains a countable subfamily with the same 'infimum' and 'supremum' (*loc. cit.*). We remark that the infimum or supremum over a countable family, say M_0, can be taken pointwise, i.e. $f \in \tilde{f} \in M_0$ implies $\inf_{M_0} f \in \inf M_0$.

Let $\mathbb{N}$ be the set of natural numbers. For each $n \in \mathbb{N}$ let μ^n on the product σ-field $\mathscr{F}^n$ be the Cartesian product of n identical components μ on $\mathscr{F}$. For any function $f: X \to [-\infty, +\infty]$ we denote by $f^{(n)}: X^n \to [-\infty, +\infty]$ the function defined by $f^{(n)}(x_1, \cdots, x_n) = n^{-1} \sum_{i=1}^n f(x_i)$, $(x_1, \cdots, x_n) \in X^n$, using the convention $\infty - \infty = -\infty$. Sometimes $f^{(n)}$ is considered in a natural way as function defined on the countable product space $X^{\mathbb{N}}$. If $\tilde{f} \in M(\mathscr{F}, \mu)$, denote by $\tilde{f}^{(n)} \in M(\mathscr{F}^n, \mu^n)$ the μ^n-equivalence class of functions on X^n containing the functions $f^{(n)}$, $f \in \tilde{f}$. For $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$, and $f_t \in \tilde{f}_t$, $t \in T$, we have—as introduced above—the following denotations:

$$\inf \tilde{f}_S^{(n)} = \inf\{\tilde{f}_t^{(n)} : t \in S\} \qquad \text{and} \qquad \inf f_S^{(n)} = \inf\{f_t^{(n)} : t \in S\},$$

where $S \subset T$.

The following notion is basic for our concept of minimum contrast estimates.

DEFINITION. Let $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$. A subset $T_0 \subset T$ is a separant for $\tilde{f}_t$,

$t \in T$, iff T_0 is countable and $\inf \tilde{f}_{U \cap T_0} = \inf \tilde{f}_U$ for all $U \in \mathscr{U}$. T_0 is *a strong separant* for $\tilde{f}_t$, $t \in T$, iff for every $n \in \mathbb{N}$, T_0 is *a separant* for the family $\tilde{f}_t^{(n)}$, $t \in T$, (i.e. $\inf \tilde{f}_{U \cap T_0}^{(n)} = \inf \tilde{f}_U^{(n)}$ for all $U \in \mathscr{U}, n \in \mathbb{N}$).

A family of versions $f_t \in \tilde{f}_t$, $t \in T$, is *separable* (see Doob (1953) page 52) iff there exist a countable set $T_0 \subset T$ and a μ-null set $X_0 \in \mathscr{F}$ such that $\inf f_{U \cap T_0}(x) = \inf f_U(x)$ for all $x \notin X_0$, $U \in \mathscr{U}$.

If $(T, \mathscr{U})$ has a countable base, for every family $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$, there exist a strong separant (Lemma 4.1) and a separable family of versions $f_t \in \tilde{f}_t$, $t \in T$ (see [1] page 57).

2. Minimum contrast estimates. Throughout this section let $(T, \mathscr{U})$ be a topological space, μ a σ-finite measure on a σ-field $\mathscr{F}$ and $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$.

DEFINITION 2.1. A function $\varphi : X \to T$ is a *minimum contrast* (m.c.) *estimate* for $\tilde{f}_t$, $t \in T$, iff there exist a separant $T_0 \subset T$, versions $f_t \in \tilde{f}_t$, $t \in T$, and a μ-null set $X_0 \in \mathscr{F}$ such that

$$\inf f_{U \cap T_0}(x) = \inf f_{T_0}(x)$$

for all $x \notin X_0$, $U \in \mathscr{U}$ with $\varphi(x) \in U$. φ_n is a m.c. *estimate for the sample size n*, iff φ_n is a m.c. estimate for the family $\tilde{f}_t^{(n)}$, $t \in T$.

If $(T, \mathscr{U})$ has a countable base, the definition of m.c. estimates is independent of the special separant T_0 and the particular choice of versions $f_t \in \tilde{f}_t$, $t \in T$ (see Lemma 4.2 (i)).

We remark that in contrast to the classical case we do not require a m.c. estimate to be measurable, because strong consistency of sequences of m.c. estimates can be proved without such measurability assumptions (see Theorem 2.8).

The following proposition shows that minimum contrast estimates in the sense of Huber [3] and Pfanzagl [6], derived from separable families of versions, are also minimum contrast estimates in our sense.

PROPOSITION 2.2. *If $(T, \mathscr{U})$ has a countable base and $f_t \in \tilde{f}_t$, $t \in T$, is a separable family of versions, then each function $\varphi : X \to T$ with $f_{\varphi(x)}(x) = \inf f_T(x)$, $x \in X$, is a* m.c. *estimate for $\tilde{f}_t$, $t \in T$.*

PROOF. As the family f_t, $t \in T$, is separable, there exists a countable $T_0 \subset T$ and a μ-null set $X_0 \in \mathscr{F}$ such that $\inf f_U(x) = \inf f_{U \cap T_0}(x)$ for all $x \notin X_0$, $U \in \mathscr{U}$. Then T_0 is a separant for $\tilde{f}_t$, $t \in T$, and $x \notin X_0$, $\varphi(x) \in U \in \mathscr{U}$ imply $\inf f_{U \cap T_0}(x) = \inf f_U(x) = \inf f_T(x) = \inf f_{T_0}(x)$ which proves the assertion.

THEOREM 2.3. *Let $(T, \mathscr{U})$ be compact metrizable, and $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$. Then $\mathscr{F}^n$, $\sigma(\mathscr{U})$-measurable minimum contrast estimates exist for every sample size $n \in \mathbb{N}$.*

PROOF. Let T_0 be a strong separant for $\tilde{f}_t$, $t \in T$, and let $f_t \in \tilde{f}_t$, $t \in T$. Apply now Lemma 4.5 for each $n \in \mathbb{N}$ to $f_t^{(n)}$, $t \in T$, and T_0.

Now we consider sequences of asymptotic minimum contrast estimates (for short: a.m.c. sequences). We remind the reader that a sequence $\varphi_n : X^{\mathbb{N}} \to T$, $n \in \mathbb{N}$, is a 'classical' a.m.c. sequence for versions $f_t \in \tilde{f}_t$, $t \in T$, iff for all $x \in X^{\mathbb{N}}$

$$\text{(2.4)} \qquad \lim_{n \in \mathbb{N}}(\exp f^{(n)}_{\varphi_n(x)}(x) - \exp \inf f_T{}^{(n)}(x)) = 0$$

(see [6] page 251). Equivalently this means

$$\text{(2.5)} \qquad \limsup_{n \in \mathbb{N}_0} f^{(n)}_{\varphi_n(x)}(x) = \limsup_{n \in \mathbb{N}_0} \inf f_T{}^{(n)}(x)$$

for all $x \in X^{\mathbb{N}}$ and all subsequences $\mathbb{N}_0 \subset \mathbb{N}$.

To obtain a concept which is independent of the special choice of versions $f_t \in \tilde{f}_t$, $t \in T$, we do not consider the values of the versions pertaining to $\varphi_n(x)$ any more, we are only interested in their behavior in the neighborhoods of accumulation points of $(\varphi_n(x))_{n \in \mathbb{N}}$. This leads to the following modification of (2.5):

Definition 2.6. Let $P \ll \mu$ be a probability measure on $\mathscr{F}$. A sequence $\varphi_n : X^{\mathbb{N}} \to T$, $n \in \mathbb{N}$, is a P-a.m.c. *sequence* for $\tilde{f}_t$, $t \in T$, iff there exist a strong separant T_0, versions $f_t \in \tilde{f}_t$, $t \in T$, and a $P^{\mathbb{N}}$-null set $X_0 \in \mathscr{F}^{\mathbb{N}}$ such that

$$\limsup_{n \in \mathbb{N}_0} \inf f^{(n)}_{U \cap T_0}(x) = \limsup_{n \in \mathbb{N}_0} \inf f^{(n)}_{T_0}(x)$$

if $x \notin X_0$, $U \in \mathscr{U}$ and $\varphi_n(x) \in V$ for all $n \in \mathbb{N}_0$ and some $V \in \mathscr{U}$ with $\bar{V} \subset U$. If $(T, \mathscr{U})$ has a countable base this definition is independent of the special strong separant T_0 and the special choice of versions $f_t \in \tilde{f}_t$, $t \in T$ (see Lemma 4.2 (ii)).

Proposition 2.7. *If $(T, \mathscr{U})$ has a countable base and if φ_n is* m.c. *estimate (for $\tilde{f}_t$, $t \in T$) for each sample size $n \in \mathbb{N}$, then $(\varphi_n)_{n \in \mathbb{N}}$ is a P-*a.m.c. *sequence for every probability measure $P \ll \mu$ on $\mathscr{F}$.*

Proof. Let T_0 be a strong separant for $\tilde{f}_t$, $t \in T$, and $f_t \in \tilde{f}_t$, $t \in T$. As φ_n is m.c. estimate for $\tilde{f}_t{}^{(n)}$, $t \in T$, there exists a μ^n-null set $A_n \in \mathscr{F}^n$ such that for all $x \in X^{\mathbb{N}}$ which are not element of the cylinder set over A_n, say $X_n \in \mathscr{F}^{\mathbb{N}}$, we have $\inf f^{(n)}_{U \cap T_0}(x) = \inf f^{(n)}_{T_0}(x)$ if $\varphi_n(x) \in U \in \mathscr{U}$. Then $X_0 = \cup_{n \in \mathbb{N}} X_n$ is a $P^{\mathbb{N}}$-null set for each probability measure $P \ll \mu$, and for each $n \in \mathbb{N}$ we have $\inf f^{(n)}_{U \cap T_0}(x) = \inf f^{(n)}_{T_0}(x)$ if $x \notin X_0$ and $\varphi_n(x) \in U \in \mathscr{U}$. This immediately implies the assertion.

The 'if part' of the following theorem is closely related to Theorem 1.12 of Pfanzagl [6]. As Example 3.3 shows, we prove, however, strong consistency for a larger class of estimates. Moreover our proof is simpler than that of Pfanzagl because no measurability problems such as those found there arise.

Theorem 2.8. *Let μ be a σ-finite measure on a σ-field $\mathscr{F}$ and $P \ll \mu$ a probability measure on $\mathscr{F}$. Let $(T, \mathscr{U})$ be a compact metrizable space, $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$, and $S \subset T$ dense in T. Assume that*

(i) *$t \to \tilde{f}_t$, $t \in T$, is* l.s.c. *If for some $t_0 \in S$*

(ii) $P(\tilde{f}_{t_0}) < P(\tilde{f}_t)$ *for all* $t \neq t_0$, $t \in T$,

(iii) $P(\inf \tilde{f}_C) > -\infty$ *for all compact sets* $C \subset T$ *with* $t_0 \notin C$, *then a sequence of estimates* $\varphi_n : X^{\mathbb{N}} \to S$, $n \in \mathbb{N}$, *converges* $P^{\mathbb{N}}$-a.e. *to* t_0 *if and only if* $(\varphi_n)_{n \in \mathbb{N}}$ *is a* P-a.m.c. *sequence for* $\tilde{f}_t$, $t \in S$.

PROOF. Let $f_t \in \tilde{f}_t$, $t \in T$. Then, according to the strong law of large numbers, there exists a $P^{\mathbb{N}}$-null set $X_1 \in \mathscr{F}^{\mathbb{N}}$, such that for all $x \notin X_1$ and all $S_1 \subset T$ with $t_0 \in S_1$

$$\limsup_{n \in \mathbb{N}} \inf f_{S_1}^{(n)}(x) \leqq \lim_{n \in \mathbb{N}} f_{t_0}^{(n)}(x) = P(f_{t_0}) \,. \tag{2.9}$$

Let $U_k \in \mathscr{U}$, $k \in \mathbb{N}$, be a base for the neighborhood system of t_0 such that $U_{k+1} \subset \bar{U}_{k+1} \subset U_k$, $k \in \mathbb{N}$. By Lemma 4.1 there exists a countable $T_0 \subset T$, $t_0 \in T_0$, which contains a strong separant for all the families $\tilde{f}_t$, $t \in Z$, where $Z = S, T, T - U_k, S - U_k$, $k \in \mathbb{N}$. As for each $k \in \mathbb{N}$ the assumptions of Lemma 4.3 are fulfilled for $T - U_k$, there exists a $P^{\mathbb{N}}$-null set $X_2 \in \mathscr{F}^{\mathbb{N}}$ such that for all $x \notin X_2$ and all $k \in \mathbb{N}$

$$P(f_{t_0}) < \inf_{t \in T - U_k} P(f_t) = \lim_{n \in \mathbb{N}} \inf f_{(T-U_k) \cap T_0}^{(n)}(x) \tag{2.10}$$

where the first inequality follows from (i) and the fact that a l.s.c. function attains its infimum on a compact set.

(a) Assume that $(\varphi_n)_{n \in \mathbb{N}}$ is a P-a.m.c. sequence for $\tilde{f}_t$, $t \in S$. As $S_0 = S \cap T_0$ is a strong separant for $\tilde{f}_t$, $t \in S$, there exists a $P^{\mathbb{N}}$-null set $X_3 \in \mathscr{F}^{\mathbb{N}}$ such that for all $x \notin X_3$, $\mathbb{N}_0 \subset \mathbb{N}$ with $\varphi_n(x) \in V \subset \bar{V} \subset U$ for all $n \in \mathbb{N}_0 (U, V \in \mathscr{U})$

$$\limsup_{n \in \mathbb{N}_0} \inf f_{U \cap S_0}^{(n)}(x) = \limsup_{n \in \mathbb{N}_0} \inf f_{S_0}^{(n)}(x) \,. \tag{2.11}$$

Let $X_0 = X_1 \cup X_2 \cup X_3$. Then $P^{\mathbb{N}}(X_0) = 0$. If $x \notin X_0$ and $\varphi_n(x) \in T - \bar{U}_{k-1} \subset T - U_{k-1} \subset T - \bar{U}_k$ for infinitely many $n \in \mathbb{N}$, say $\mathbb{N}_0$, and some $k > 1$, we have by (2.11)

$$\limsup_{n \in \mathbb{N}_0} \inf f_{(T-\bar{U}_k) \cap S_0}^{(n)}(x) = \limsup_{n \in \mathbb{N}_0} \inf f_{S_0}^{(n)}(x) \,.$$

As $T - \bar{U}_k \subset T - U_{k+1}$ and $S_0 \subset T_0$, we have by (2.10)

$$P(f_{t_0}) < \limsup_{n \in \mathbb{N}_0} \inf f_{S_0}^{(n)}(x) \,.$$

This, however, contradicts (2.9). Hence for all $x \notin X_0 (\varphi_n(x))_{n \in \mathbb{N}}$ is eventually in each $\bar{U}_k$, $k \in \mathbb{N}$, whence $\lim_{n \in \mathbb{N}} \varphi_n(x) = t_0$ $P^{\mathbb{N}}$-a.e.

(b) Assume conversely that the sequence of functions $\varphi_n : X^{\mathbb{N}} \to T$, $n \in \mathbb{N}$, converges $P^{\mathbb{N}}$-a.e. to t_0. Hence there exists a $P^{\mathbb{N}}$-null set $X_3' \in \mathscr{F}^{\mathbb{N}}$ such that $\lim_{n \in \mathbb{N}} \varphi_n(x) = t_0$ for all $x \notin X_3'$.

Let $X_0' = X_1 \cup X_2 \cup X_3'$. Then $P^{\mathbb{N}}(X_0') = 0$ and we have by (2.9) and (2.10) for all $x \notin X_0'$, $\mathbb{N}_0 \subset \mathbb{N}$, $k \in \mathbb{N}$:

$$\limsup_{n \in \mathbb{N}_0} \inf f_{S_0 \cap U_k}^{(n)}(x) \leqq P(f_{t_0}) < \lim_{n \in \mathbb{N}_0} \inf f_{(T-U_k) \cap S_0}^{(n)}(x)$$

and therefore

$$\limsup_{n\in \mathrm{N}_0} \inf f_{S_0}^{(n)}(x) = \limsup_{n\in \mathrm{N}_0} \inf f_{S_0\cap U_k}^{(n)}(x) .$$

Since $\lim_{n\in \mathrm{N}} \varphi_n(x) = t_0$ for all $x \notin X_0'$, since S_0 is a strong separant for $\tilde f_t$, $t\in S$, and since $U_k \cap S$, $k \in \mathbb{N}$, is a base for the neighborhood system of t_0 in S, this implies that $(\varphi_n)_{n\in \mathrm{N}}$ is a P-a.m.c. sequence for $\tilde f_t$, $t \in S$.

3. Maximum likelihood estimates. In this section we shall apply the results obtained in Section 2 to maximum likelihood estimation. Let P be a family of probability measures endowed with a topology $\mathscr{U}$ and dominated by a σ-finite measure μ on the σ-field $\mathscr{F}$. For each $P \in \mathsf{P}$ let $\tilde h_P$ be the equivalence class of densities of P with respect to μ.

A function $\varphi_n : X^n \to \mathsf{P}$ is a *maximum likelihood estimate* for P at sample size n, iff φ_n is m.c. estimate at sample size n for the family $\tilde f_P = -\log \tilde h_P$, $P \in \mathsf{P}$. A sequence of functions $\varphi_n : X^{\mathrm{N}} \to \mathsf{P}$, $n \in \mathbb{N}$, is an *asymptotic maximum likelihood* (a.m.l.) sequence for P, iff for every $P_0 \in \mathsf{P}$ the sequence φ_n, $n \in \mathbb{N}$, is P_0-a.m.c. sequence for $\tilde f_P = -\log \tilde h_P$, $P \in \mathsf{P}$.

A sequence of functions $\varphi_n : X^{\mathrm{N}} \to \mathsf{P}$, $n \in \mathbb{N}$, is *strongly consistent* for P, iff for every $P \in \mathsf{P} : \lim_{n\in \mathrm{N}} \varphi_n(x) = P$ P^{N}-a.e.

THEOREM 3.1. *If* P *is endowed with a compact metrizable topology* $\mathscr{U}$, *for each* $n \in \mathbb{N}$ *there exists an* $\mathscr{F}^n$, $\sigma(\mathscr{U})$*-measurable maximum likelihood estimate for* P *at sample size* n.

PROOF. Apply Theorem 2.3 to $T = \mathsf{P}$ and $\tilde f_t = -\log \tilde h_t$, $t \in T$.

THEOREM 3.2. *Let* P *be endowed with a compact metrizable topology and assume that* $P \to \tilde h_P$, $P \in \mathsf{P}$, *is upper semicontinuous. If for every* $P \in \mathsf{P}$

(i) $P(\log \tilde h_P) > -\infty$,

(ii) $P(\log \sup \tilde h_C) < +\infty$ *for each compact* $C \subset \mathsf{P}$ *with* $P \notin C$, *then a sequence of functions* $\varphi_n : X^{\mathrm{N}} \to \mathsf{P}$, $n \in \mathbb{N}$, *is strongly consistent for* P *if and only if it is a sequence of asymptotic maximum likelihood estimates for* P.

PROOF. As by (i) and (ii) $P(\log \tilde h_P) > -\infty$ and $P(\log \tilde h_Q) < \infty$ if $Q \neq P$, we have for all $P, Q \in \mathsf{P}$ with $P \neq Q$

$$P(\log \tilde h_Q) - P(\log \tilde h_P) < \log P(\tilde h_Q/\tilde h_P) \leqq \log \mu(\tilde h_Q) = 0 ,$$

and hence $P(\log \tilde h_Q) < P(\log \tilde h_P)$. Therefore we may apply Theorem 2.8 for each $P_0 \in \mathsf{P}$ to $S = T = \mathsf{P}$, $P = P_0$, $t_0 = P_0$ and the family $\tilde f_P = -\log \tilde h_P$, $P \in \mathsf{P}$. This implies the assertion.

We remark that—similar as Theorem 2.6 of Pfanzagl [6]—Theorem 3.2 can be formulated with a compact metric space $T \supset \mathsf{P}$, where for $t \in T - \mathsf{P}$ an equivalence class $\tilde h_t \geqq 0$ with $\mu(\tilde h_t) \leqq 1$ is given. Then we have in addition to assume that $t \to \tilde h_t$, $t \in T$, is u.s.c. and that 3.2 (ii) holds with compact $C \subset T$

instead of $C \subset \mathsf{P}$. In this case the assertion follows from Theorem 2.8., applied to T, $S = \mathsf{P}$ and $\tilde{f}_t = -\log \tilde{h}_t$, $t \in T$.

According to Theorem 3.2 each sequence of 'classical' a.m.l. estimates for P which turns out to be strongly consistent is a sequence of a.m.l. estimates for P, if P fulfills the assumptions of Theorem 3.2. The following example shows that even for a compact metric P with uniformly bounded upper semicontinous densities there exist large classes of a.m.l. (and hence strongly consistent) sequences for P, which cannot be obtained as classical a.m.l. estimates from a suitable family of versions of the densities. Example 3.3 improves an example of Lüpsen [5].

EXAMPLE 3.3. Let $X = (0, 2)$, $\mathscr{F}$ the Borel-field on X and μ the Lebesgue measure on $\mathscr{F}$. For each $t \in T = [0, 1]$ let P_t be the probability measure on $\mathscr{F}$ with density $h_t = 1_{(t,t+1)}$ with respect to μ; where 1_A denotes the indicator function of the set $A \subset X$. Let $\mathsf{P} = \{P_t : t \in T\}$, and let $\mathscr{U}$ be the topology on P induced by the 1—1 map $t \to P_t$, $t \in T$, and the usual topology on $T = [0, 1]$. Then the assumptions of Theorem 3.2 are fulfilled for $(\mathsf{P}, \mathscr{U})$. Let

$$X_0 = \{(x_i)_{i \in \mathbb{N}} \in X^{\mathbb{N}} : \sup(x_1, \cdots, x_n) - \inf(x_1, \cdots, x_n) < 1 \quad \text{for all} \quad n \geqq 2\} .$$

Then $P_t^{\mathbb{N}}(X_0) = 1$ for all $t \in T$. Let $T_0 = \{t_i : i \in \mathbb{N}\}$ be dense in T. Then we have for all $(x_i)_{i \in \mathbb{N}} \in X_0$ and all $n \geqq 2$

$$\sup_{t \in T} \prod_{i=1}^{n} h_t(x_i) = \sup_{t \in T_0} \prod_{i=1}^{n} h_t(x_i) = 1. \tag{3.4}$$

Now define for all $x = (x_i)_{i \in \mathbb{N}} \in X^{\mathbb{N}}$ and all $n \in \mathbb{N}$: $\varphi_n(x) = 1$ iff $\inf(x_1, \cdots, x_n) \geqq 1$, and $\varphi_n(x) = t_k$ otherwise, where k is the smallest index $i \in \mathbb{N}$ such that $t_i \in (\inf(x_1, \cdots, x_n), \inf(x_1, \cdots, x_n) + 1/n)$. It is easy to see that $\varphi_n : X^{\mathbb{N}} \to T$ is measurable for each $n \in \mathbb{N}$. As for all $t \in T$ there exists $X_t \in \mathscr{F}^{\mathbb{N}}$ with $P_t^{\mathbb{N}}(X_t) = 1$ such that $\lim_{n \in \mathbb{N}} \inf(x_1, \cdots, x_n) = t$ for all $(x_i)_{i \in \mathbb{N}} \in X_t$, the sequence $(\varphi_n)_{n \in \mathbb{N}}$ is strongly consistent for P and hence an a.m.l. sequence for P (Theorem 3.2).

Now we shall show that for no choice of versions $(\varphi_n)_{n \in \mathbb{N}}$ is an a.m.l. sequence in the usual sense. Let $h_t' \in \tilde{h}_t$, $t \in T$. As T_0 is countable there exists $X_1 \in \mathscr{F}$ with $\mu(X - X_1) = 0$ such that

$$h_t(x) = h_t'(x) \quad \text{for all} \quad x \in X_1, \ t \in T_0 . \tag{3.5}$$

Let $X_t = X_t' \cap X_0 \cap X_1^{\mathbb{N}}$, $t \in T$. As $P_t(X_1) = 1$ we have $P_t^{\mathbb{N}}(X_1^{\mathbb{N}}) = 1$ and hence $P_t^{\mathbb{N}}(X_t) = 1$ for all $t \in T$. Let $f_t = -\log h_t$, $f_t' = -\log h_t'$, $t \in T$. Since φ_n assumes its values only in $T_0 \cup \{1\}$, we obtain by (3.5) for all $t \in [0, 1)$ and all $x = (x_i)_{i \in \mathbb{N}} \in X_t$:

$$\frac{1}{n} \sum_{i=1}^{n} f'_{\varphi_n(x)}(x_i) = \frac{1}{n} \sum_{i=1}^{n} f_{\varphi_n(x)}(x_i) = +\infty \tag{3.6}$$

for all sufficiently large $n \in \mathbb{N}$.

By (3.4) and (3.5) we have for all $(x_i)_{i\in\mathbb{N}} \in X_0 \cap X_1{}^{\mathbb{N}}$ and all $n \geqq 2$

$$\text{(3.7)} \qquad \inf \frac{1}{n} \sum_{i=1}^{n} f_t'(x_i) \leqq \inf \frac{1}{n} \sum_{i=1}^{n} f_t(x_i) = 0 .$$

According to (3.6) and (3.7) relation 2.5 is violated for all $x \in X_t$, $t \in [0, 1)$ for both the families f_t, $t \in T$, and f_t', $t \in T$. Hence $(\varphi_n)_{n\in\mathbb{N}}$ is not an a.m.l. sequence in the usual sense.

4. Auxiliary Lemmas. In this section we collect lemmas which are auxiliary for the results in Section 2 and Section 3. Throughout this section let μ be a σ-finite measure on a σ-field $\mathscr{F}$ over X.

LEMMA 4.1. *Let $(T, \mathscr{U})$ be a topological space with countable base and $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$. Then for each countable family $\mathscr{C}$ of subsets of T there exists a countable $T_0 \subset T$ which contains for every $C \in \mathscr{C}$ a strong separant for the family $\tilde{f}_t$, $t \in C$.*

PROOF. As each subset of T with the relative topology has a countable base, for each $C \in \mathscr{C}$, $n \in \mathbb{N}$, there exists a separant, say $T(C, n) \subset C$, for the family $\tilde{f}_t^{(n)}$, $t \in C$ (see [6], Corollary 3.2). Then $T_0 = \cup\, \{T(C, n) : C \in \mathscr{C}, n \in \mathbb{N}\}$ is countable and fulfills the assertion.

LEMMA 4.2. *Let $(T, \mathscr{U})$ be a topological space with countable base, $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$ and f_t, $g_t \in \tilde{f}_t$, $t \in T$.*

(i) *If T_0, $T_1 \subset T$ are separants, then there exists a μ-null set $X_0 \in \mathscr{F}$ such that for all $x \notin X_0$ and all $U \in \mathscr{U}$*

$$\inf f_{U\cap T_0}(x) = \inf g_{U\cap T_1}(x) .$$

(ii) *If T_0, $T_1 \subset T$ are strong separants, then there exists $X_0 \in \mathscr{F}^{\mathbb{N}}$, which is $P^{\mathbb{N}}$-null set for each probability measure $P \ll \mu$ on $\mathscr{F}$, such that for all $x \notin X_0$, $U \in \mathscr{U}$, $n \in \mathbb{N}$*

$$\inf f^{(n)}_{U\cap T_0}(x) = \inf g^{(n)}_{U\cap T_1}(x) .$$

PROOF. We only prove (ii). The proof for (i) is similar. As T_0, T_1 are strong separants we have for all $n \in \mathbb{N}$, $U \in \mathscr{U}$

$$\inf \tilde{f}^{(n)}_{U\cap T_0} = \inf \tilde{f}^{(n)}_{U\cap T_1} .$$

Hence for each $U \in \mathscr{U}$, $n \in \mathbb{N}$, there exist a μ^n-null set $A_{U,n} \in \mathscr{F}^{\,n}$ such that for all $x \in X^{\mathbb{N}}$, which are not element of the cylinder set over $A_{U,n}$, say $X_{U,n} \in \mathscr{F}^{\mathbb{N}}$, we have

$$\inf f^{(n)}_{U\cap T_0}(x) = \inf g^{(n)}_{U\cap T_1}(x) .$$

Let $\mathscr{U}_0$ be a countable base for $\mathscr{U}$ and define $X_0 = \cup\, \{X_{U,n} : U \in \mathscr{U}_0, n \in \mathbb{N}\}$. If $P \ll \mu$ is a probability measure on $\mathscr{F}$ then $P^{\mathbb{N}}(X_{U,n}) = 0$ for all $U \in \mathscr{U}$,

$n \in \mathbb{N}$, whence $P^{\mathbb{N}}(X_0) = 0$. As $\inf f^{(n)}_{U \cap T_0}(x) = \inf g^{(n)}_{U \cap T_1}(x)$ for all $U \in \mathscr{U}_0$, $x \notin X_0$, $n \in \mathbb{N}$, and each $U \in \mathscr{U}$ is (countable) union of elements of $\mathscr{U}_0$, this implies the assertion.

The following lemma is a slight modification of Lemma 3.11 of Pfanzagl [6]:

LEMMA 4.3. *Let $(T, \mathscr{U})$ be a compact metric space and $\tilde{f}_t \in M(\mathscr{F}, \mu)$, $t \in T$. Assume that*

(i) *$t \to \tilde{f}_t$, $t \in T$, is* l.s.c.

Let furthermore $P \ll \mu$ be a probability measure on $\mathscr{F}$ such that

(ii) $P(\inf \tilde{f}_T) > -\infty$.

Then

(a) *$t \to P(\tilde{f}_t)$, $t \in T$, is* l.s.c.

(b) *If $f_t \in \tilde{f}_t$, $t \in T$, and T_0 is a strong separant for $\tilde{f}_t$, $t \in T$, then*

$$\inf_{t \in T} P(\tilde{f}_t) = \lim_{n \in \mathbb{N}} \inf f^{(n)}_{T_0}(x)\ P^{\mathbb{N}}\text{-a.e.}$$

PROOF. The proof for (a) and for the relation

$$\inf_{t \in T} P(f_t) \leqq \lim\inf_{n \in \mathbb{N}} \inf f^{(n)}_{T_0}(x)\ P^{\mathbb{N}}\text{-a.e.} \tag{4.4}$$

runs analogously to the proof of the corresponding assertions in Lemma (3.11) of Pfanzagl [6].

As the function $t \to P(\tilde{f}_t)$, $t \in T$, is l.s.c. by (a), it attains its infimum for some $t_0 \in T$. Since $\inf \tilde{f}^{(n)}_{T_0} \leqq \tilde{f}^{(n)}_{t_0}$ for each $n \in \mathbb{N}$, we have by the strong law of large numbers $\lim\sup_{n \in \mathbb{N}} \inf f^{(n)}_{T_0}(x) \leqq \lim\sup_{n \in \mathbb{N}} f^{(n)}_{t_0}(x) = P(\tilde{f}_{t_0}) = \inf_{t \in T} P(\tilde{f}_t)$ $P^{\mathbb{N}}$-a.e. Together with (4.4) this implies (b).

LEMMA 4.5. *Let $(T, \mathscr{U})$ be a compact metric space and let $f_t \colon X \to [-\infty, +\infty]$, $t \in T$, be a family of $\mathscr{F}$-measurable functions. Then for every non void $T_0 \subset T$ there exists a $\mathscr{F}$, $\sigma(\mathscr{U})$-measurable function $\varphi \colon X \to T$ such that $\varphi(x) \in U \in \mathscr{U}$ implies* $\inf f_{T_0}(x) = \inf f_{U \cap T_0}(x)$.

PROOF. For each $x \in X$ let $Q_x = \{t \in T \colon \inf f_{U \cap T_0}(x) = \inf f_{T_0}(x) \text{ if } t \in U \in \mathscr{U}\}$. In order to apply Theorem 4.1 of Sion (1960)—which admits a measurable choice $\varphi(x) \in Q_x$, $x \in X$—we prove the following:

(a) $Q_x \neq \emptyset$ for each $x \in X$

Let $t_n \in T_0$, $n \in \mathbb{N}$, be such that $\lim_{n \in \mathbb{N}} f_{t_n}(x) = \inf f_{T_0}(x)$. Then every accumulation point of t_n, $n \in \mathbb{N}$, belongs to Q_x. As a compact metric space is sequentially compact this implies $Q_x \neq \emptyset$.

(b) Q_x is closed for each $x \in X$

Let t be limit point of the sequence $t_n \in Q_x$, $n \in \mathbb{N}$. Then $t \in U \in \mathscr{U}$ implies $t_n \in U$ for some $n \in \mathbb{N}$; hence $\inf f_{U \cap T_0}(x) = \inf f_{T_0}(x)$, whence $t \in Q_x$.

(c) $\{x \in X \colon Q_x \cap C = \emptyset\} \in \mathscr{F}$ for each compact $C \subset T$

Let $\mathscr{U}_0$ be a countable base of $(T, \mathscr{U})$ which is closed under finite unions. We shall prove

$$\{x \colon Q_x \cap C = \emptyset\} = \cup \{\{x \colon \inf f_{U \cap T_0}(x) > \inf f_{T_0}(x)\} \colon C \subset U \in \mathscr{U}_0\}. \tag{4.6}$$

The inclusion "$\supset$" in (4.6) is trivial. Assume, conversely, $Q_x \cap C = \emptyset$. Then for every $t \in C$ there exists $U_t \in \mathscr{U}_0$, $t \in U_t$, with $\inf f_{U \cap T_0}(x) > \inf f_{T_0}(x)$. As $\{U_t : t \in C\}$ is an open cover of the compact set C, there exists a finite subcover, say $U_{t_1}, \cdots, U_{t_n}$. Let $U = \{\cup_{t_i} : i = 1, \cdots, n\}$. Then $C \subset U \in \mathscr{U}_0$ and $\inf f_{U \cap T0}(x) > \inf f_{T_0}(x)$. This proves (4.6). As T_0 is countable and each f_t, $t \in T$, is measurable, (4.6) implies (c).

Because of (a), (b), (c), Theorem 4.1 of Sion (1960) is applicable and we obtain the existence of an $\mathscr{F}$, $\sigma(\mathscr{U})$ measurable function $\varphi : X \to T$ such that $\varphi(x) \in Q_x$ for all $x \in X$. This function has the asserted properties.

REFERENCES

[1] Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[2] Dunford, N. and Schwartz, J. (1964). *Linear Operators I.* Interscience, New York.

[3] Huber, P. J. (1966). The behavior of maximum likelihood estimates under non-standard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 221–233.

[4] Landers, D. (1968). Existenz und Konsistenz von Maximum Likelihood Schätzen. Ph. D. Dissertation, University of Cologne.

[5] Lüpsen, H. (1970). Existenz und Konsistenz von Minimum Kontrast Schätzen zu Familien von Funktionenäquivalenzklassen. Diplomarbeit. University of Cologne.

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

[7] Sion, M. (1969). On uniformization of sets in topological spaces. *Trans. Amer. Math. Soc.* **96** 237–245.