Ann. Inst. Statist. Math.
Vol. 55, No. 1, 95–110 (2003)

ASYMPTOTIC BOUNDS FOR ESTIMATORS WITHOUT LIMIT DISTRIBUTION

J. PFANZAGL

Mathematical Institute, University of Cologne, Weyertal 86-90, 50931 Cologne, Germany

(Received August 20, 2001; revised March 22, 2002)

Abstract. Let $\mathfrak{P}$ be a general family of probability measures, $\kappa : \mathfrak{P} \to \mathbb{R}$ a functional, and $N_{(0,\sigma^2(P))}$ the optimal limit distribution for regular estimator sequences of κ. On intervals symmetric about 0, the concentration of this optimal limit distribution can be surpassed by the asymptotic concentration of an arbitrary estimator sequence only for P in a "small" subset of $\mathfrak{P}$. For asymptotically median unbiased estimator sequences the same is true for arbitrary intervals containing 0. The emphasis of the paper is on "pointwise" conditions for $P \in \mathfrak{P}$, as opposed to conditions on shrinking neighbourhoods, and on "general" rather than parametric families.

Key words and phrases: Estimation, asymptotic theory, local uniformity, nonparametric theory, minimax bounds.

1. Introduction and summary

Bounds for the asymptotic concentration of estimator sequences in nonparametric families are the main object of the present paper. We start with a survey of what is known for parametric families.

Let $\{P_\vartheta : \vartheta \in \Theta\}$, $\Theta \subset \mathbb{R}^k$, be a family of probability measures fulfilling the usual regularity conditions (say a LAN-condition for every $\vartheta \in \Theta$). The problem is to estimate a "smooth" functional $\kappa : \Theta \to \mathbb{R}$ from a sample $(x_1, \dots, x_n)$, governed by P_ϑ^n. Let $\kappa^{(n)} : X^n \to \mathbb{R}$ denote an estimator.

As a consequence of the convolution theorem, for every $\vartheta \in \Theta$ there exists an optimal limit distribution, say Q_ϑ, such that

$$(1.1) \qquad \liminf_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \kappa(\vartheta)))dP_\vartheta^n \geq \int \ell dQ_\vartheta$$

for every subconvex loss function ℓ which is symmetric about 0. Relation (1.1) is true for estimator sequences $(\kappa^{(n)})_{n\in\mathbb{N}}$ which are regular (in the sense that $P^n_{\vartheta+c_n^{-1}a} \circ c_n(\kappa^{(n)} - \kappa(\vartheta + c_n^{-1}a))$, $n \in \mathbb{N}$, converges to the same limit distribution for every a in a neighbourhood of 0).

If our intention is to interpret a limit relation like (1.1) as "approximately true" for large samples, the assumption of a "regularly attainable limit distribution" comes in like a "deus ex machina". It seems therefore advisable to consider the approach via the convolution theorem as just a technical device for obtaining an asymptotic risk-bound, and to search for a role this bound could play in a wider context.

If relation (1.1) were true for every estimator sequence, this would be a convincing expression of the asymptotic optimality of Q_ϑ. Yet, straightforward examples show that there are always estimator sequences such that

$$\lim_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \kappa(\vartheta)))dP_\vartheta^n < \int \ell dQ_\vartheta \quad \text{for some} \quad \vartheta \in \Theta. \tag{1.2}$$

Many scholars consider, therefore, the following "local asymptotic minimax theorem" as an adequate expression of the optimality of Q_ϑ, i.e. the relation

$$\lim_{u\uparrow\infty} \liminf_{n\to\infty} \sup_{|a|\le u} \int \ell(c_n(\kappa^{(n)} - \kappa(P_{\vartheta+c_n^{-1}a})))dP^n_{\vartheta+c_n^{-1}a} \ge \int \ell dQ_\vartheta \tag{1.3}$$

for every $\vartheta \in \Theta$ and every estimator sequence $(\kappa^{(n)})_{n\in\mathbb{N}}$.

In cold fact, relation (1.3) is not very informative. It does not tell us that

(i) a relation like

$$\limsup_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \kappa(\vartheta)))dP_\vartheta^n < \int \ell dQ_\vartheta \tag{1.4}$$

can't possibly hold for every $\vartheta \in \Theta$, and that

(ii) terrible things are bound to happen in the neighbourhood of a point of superefficiency, namely: If (1.4) holds true for a certain ϑ, then the inequality in (1.3) is strict for this ϑ.

It requires a different approach to show that, for every estimator sequence, relation (1.4) can hold on a λ^k-null set only, i.e. that for every estimator sequence and every symmetric subconvex loss function,

$$\limsup_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \kappa(\vartheta)))dP_\vartheta^n \ge \int \ell dQ_\vartheta \quad \text{for} \quad \lambda^k\text{-a.a. } \vartheta \in \Theta. \tag{1.5}$$

Relation (1.5) is by necessity restricted to *symmetric* loss functions. If Q_ϑ is symmetric about 0 (the usual case is $Q_\vartheta = N_{(0,\sigma^2(\vartheta))}$), and $P_\vartheta^n \circ c_n(\kappa^{(n)} - \kappa(\vartheta)) \Rightarrow Q_\vartheta$, then the estimator sequence $\kappa^{(n)} + c_n^{-1}$, $n \in \mathbb{N}$, has an asymptotic risk smaller than $\int \ell dQ_\vartheta$ if ℓ is a loss function like $1 - 1_{[0,2]}$. It is only in the case of properly centered estimator sequences that asymmetric loss functions can be used for expressing optimality. In this connection, the appropriate concept of "properly centered" is asymptotic median unbiasedness. Whereas the existence of a limit distribution is something fictitious, median unbiasedness is a property which refers to every sample size.

It will be shown in Section 2 that for asymptotically median unbiased estimator sequences, relation (1.5) holds λ^k-a.e. for loss functions which are not necessarily symmetric, a result which was known up to now under the assumption that the asymptotic median unbiasedness holds locally uniformly.

The main purpose of the present paper is to obtain results in the spirit of (1.5) for more general families $\mathfrak{P}$ of probability measures. If $\mathfrak{P}$ is endowed with a suitable topology, then such assertions hold for every $P \in \mathfrak{P}$, except for a set of first category. Whether a set of first category can be considered as "small" (with the same force as sets of Lebesgue measure zero are considered as small) will be discussed in Section 2. Mind that the results for general families do not imply the results for parametric families: Borel sets of first category may be of positive Lebesgue measure.

2. The main results

In this section we present asymptotic bounds for the concentration of asymptotically median unbiased estimator sequences on arbitrary intervals containing 0, and for arbitrary estimator sequences on intervals symmetric about 0. These bounds hold for every probability measure in the given family, except for a "small" subset. The theorem refers to a general family $\mathfrak{P}$ of probability measures, endowed with a topology $\mathcal{U}$, alternatively to a parametric family $\{P_\vartheta : \vartheta \in \Theta\}$ with Θ an open subset of $\mathbb{R}^k$. For parametric families the exceptional set is of Lebesgue measure zero, in the general case it is of first category.

The basic assumption: For every $P \in \mathfrak{P}$ we are given a family of sequences $(P_{n,u})_{n\in\mathbb{N}}$, $u \in \mathbb{R}$, fulfilling a condition slightly weaker than LAN (see (4.2) and (4.8)) which determines the variance $\sigma^2(P)$ in the concentration bound $N_{(0,\sigma^2(P))}$. For parametric families, the sequence $(P_{n,u})_{n\in\mathbb{N}}$ pertaining to P_ϑ is assumed to be $P_{\vartheta+c_n^{-1}u}$.

Though this does not enter in the following theorem, the interest is in "narrow" bounds which, in this case, means "large" $\sigma(P)$ resulting from "least favourable" sequences.

DEFINITION 2.1.
An estimator sequence $(\kappa^{(n)})_{n\in\mathbb{N}}$ is asymptotically median unbiased for κ at P if

$$\liminf_{n\to\infty} P^n\{\kappa^{(n)} \le \kappa(P)\} \ge 1/2 \tag{2.1'}$$

and

$$\liminf_{n\to\infty} P^n\{\kappa^{(n)} \ge \kappa(P)\} \ge 1/2. \tag{2.1''}$$

Equivalently,

$$\limsup_{n\to\infty} P^n\{\kappa^{(n)} < \kappa(P)\} \le 1/2 \le \liminf_{n\to\infty} P^n\{\kappa^{(n)} \le \kappa(P)\}.$$

We hope that the reader will not be confused by our endeavour to combine 4 theorems (parametric / nonparametric; median unbiased / arbitrary) in one.

THEOREM 2.1. (i) *If $(\kappa^{(n)})_{n\in\mathbb{N}}$ is asymptotically median unbiased for every $P \in \mathfrak{P}$, then there exists an exceptional set $\mathfrak{P}_+$ and a subsequence $\mathbb{N}_0$ such that, for $P \notin \mathfrak{P}_+$,*

$$\limsup_{n\in\mathbb{N}_0} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} \le N_{(0,\sigma^2(P))}(I) \tag{2.2}$$

for arbitrary intervals I containing 0.

(ii) *If $(\kappa^{(n)})_{n\in\mathbb{N}}$ is an arbitrary estimator sequence, then for every subsequence $\mathbb{N}_0$ there exists an exceptional set $\mathfrak{P}_+$ such that $P \notin \mathfrak{P}_+$ implies*

$$\liminf_{n\in\mathbb{N}_0} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} \le N_{(0,\sigma^2(P))}(I) \tag{2.3}$$

for intervals I symmetric about 0.

For general families, the exceptional set $\mathfrak{P}_+$ is of first category, for parametric families of Lebesgue-measure zero.

REGULARITY CONDITIONS. For every $P \in \mathfrak{P}$ there exists a family of sequences $(P_{n,u})_{n\in\mathbb{N}}$ fulfilling condition (4.2) in case (i) and conditions (4.8), (4.9) in case (ii).

a) General families: $\mathfrak{P}$ is endowed with a topology, at least as fine as the topology of the sup-distance, such that κ is continuous and the sequences $(P_{n,u})_{n\in\mathbb{N}}$ from Propositions 4.1 and 4.2, respectively, converge to P. Then relation (2.2) holds with $\mathbb{N}_0 = \mathbb{N}$. Relation (2.3) holds if σ is, in addition, upper semicontinuous.

b) Parametric families: $\mathfrak{P} = \{P_\vartheta : \vartheta \in \Theta\}$, Θ an open subset of $\mathbb{R}^k$. The functions $\vartheta \to \kappa(P_\vartheta)$ and $\vartheta \to P_\vartheta(A)$, $A \in \mathcal{A}$, are measurable. The conditions (4.2) and (4.8), (4.9), respectively, hold with $P_{n,u}$ replaced by $P_{\vartheta+c_n^{-1}u}$.

Observe that convergence to a limit distribution is not required in Theorem 2.1. If $P^n \circ c_n(\kappa^{(n)} - \kappa(P)) \Rightarrow Q_P$, we obtain from (2.3) that $Q_P(I) \le N_{(0,\sigma^2(P))}(I)$ for every interval symmetric about 0. If $(\kappa^{(n)})_{n\in\mathbb{N}}$ is asymptotically median unbiased, Q_P has median 0, and we obtain from (2.2) that $Q_P(I) \le N_{(0,\sigma^2(P))}(I)$ for arbitrary intervals containing 0. This is, however, not the sharpest result which could be obtained in this case. In fact, Q_P is in the spread order equivalent or inferior to $N_{(0,\sigma^2(P))}$ (see Pfanzagl (2002), Section 6).

It would be preferable, of course, to have relation (2.3) with $\liminf_{n\in\mathbb{N}_0}$ replaced by $\limsup_{n\to\infty}$. Yet, this is impossible. There are always estimator sequences $(\kappa^{(n)})_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} = 1$ for every $P \in \mathfrak{P}$ and every nondegenerate interval I. (Hint: Let $r_n \in \mathbb{R}$, $n \in \mathbb{N}$, be a countable subset of $\mathbb{R}$ such that 0 is an accumulation point of $c_n(r_n - r)$, $n \in \mathbb{N}$, for every $r \in \mathbb{R}$, and define $\kappa^{(n)}(x) = r_n$ for every $x \in X^n$. See van der Vaart (1997), Example, p. 407.)

An estimator sequence fulfilling $\limsup_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} > N_{(0,\sigma^2(P))}(I)$ is of no use if the "superefficiency" occurs for different P along different subsequences. What would count for possible applications is the existence of a subsequence $\mathbb{N}_0$ (from which one could choose the sample size) such that

$$\liminf_{n\in\mathbb{N}_0} P^n\{(\hat{\kappa}^{(n)} - \kappa(P)) \in I\} > N_{(0,\sigma^2(P))}(I) \tag{2.4}$$

for P in a large subset of $\mathfrak{P}$. This is, however, impossible: Whatever the subsequence $\mathbb{N}_0$, Theorem 2.1 (ii) implies that the set of those P for which (2.4) holds true, is "small" (in the sense of being of first category or of Lebesgue-measure 0, respectively).

2.1 *The topology on $\mathfrak{P}$*

Not much has been said so far about the topology on $\mathfrak{P}$. The interpretation of a set of first category as "negligible" is easier to justify if the topology $\mathcal{U}$ is derived from a metric ϱ under which $\mathfrak{P}$ is complete. In this case, any set of first category has an empty interior. Equivalently: Every nonempty open subset of $\mathfrak{P}$ contains elements which are not in the exceptional set of first category.

A metric which renders a particular family $\mathfrak{P}$ complete has to be invented, taking details of $\mathfrak{P}$ into account. (See Pfanzagl (2002), Section 8, for examples.) Since we use Corollary 5.1, the metric ϱ has to be "stronger" than the sup-metric. The sequences $(P_{n,u})_{n\in\mathbb{N}}$ converge to P with respect to the sup-metric as a consequence of conditions (4.2) or (4.8). In using a metric ϱ stronger than d, one has to make sure that the sequences $(P_{n,u})_{n\in\mathbb{N}}$ are chosen such that they converge to P also with respect to ϱ, and that κ is ϱ-continuous (and σ upper semicontinuous).

Notice that the (semi)continuity of σ is not continuity along the sequences $(P_{n,u})_{n\in\mathbb{N}}$. It is continuity of σ as a function of P. $\sigma(P)$ is determined by the family of paths converging to P. Since these paths are chosen for each P separately, one could ask where the continuity of σ should come from. There is a natural answer to this question: If one chooses for every $P \in \mathfrak{P}$ the least favourable path, this connects the values of $\sigma(P)$ for different P, and the resulting minimal variances are, in the usual instances, continuous functions of P.

Even if the metric ϱ should be artificial, on occasion: If superefficiency is impossible on ϱ-open subsets of $\mathfrak{P}$, it is a fortiori impossible on d-open subsets, a property which is easier to interpret.

2.2 *Results in terms of loss functions*

Since most results available in the literature are formulated in terms of loss functions, we mention that relation (2.2) implies immediately that

$$\liminf_{n\in\mathbb{N}_0} \int \ell(c_n(\kappa^{(n)} - \kappa(P)))dP^n \geq \int \ell dN_{(0,\sigma^2(P))} \quad \text{for} \quad P \notin \mathfrak{P}_+ \tag{2.5}$$

for not necessarily symmetric loss functions. In the general case, this holds with $\mathbb{N}_0 = \mathbb{N}$; for parametric families, it seems preferable to replace $\liminf_{n\in\mathbb{N}_0}$ by $\limsup_{n\to\infty}$.

As against that, relation (2.3) does not lend itself to a conversion into loss functions. This is, in the author's opinion, not a big disadvantage. Yet, it should be mentioned that such a result is available for $\Theta \subset \mathbb{R}$, namely

$$\limsup_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \vartheta))dP_\vartheta^n \geq \int \ell dN_{(0,\sigma^2(\vartheta))} \tag{2.6}$$

for every symmetric subconvex loss function up to a set of Lebesgue measure 0. That the exceptional set in relation (2.6) is of Lebesgue measure 0 and of first category was already claimed in Le Cam ((1953), p. 292). A proof for "Lebesgue measure zero" follows from his Corollary 8.1, p. 314. Le Cam's remarks following Theorem 4a, p. 296, indicate that the exceptional set is of first category. Not all readers are willing to accept Le Cam's proofs in this paper (so, for instance, Wolfowitz (1965), p. 249), which Le Cam himself later calls "rather incorrect" (see Le Cam (1974), p. 254).

Strictly speaking, Le Cam's assertion is more restrictive than (2.6). In harmony with his definition of "superefficiency" (see Definition 4, p. 283) he just says: "If (2.6) holds with $\leq$ for every $\vartheta \in \Theta$, then it holds with $=$ except for a set of Lebesgue measure 0". This assertion follows by a Bayesian argument which was, at this time, obviously in the air. Wolfowitz ((1953), p. 116) gives an informal proof of about the same result. A mathematically contestable Bayesian result, from which the Lebesgue-a.a. version could easily be obtained, is Proposition 1 in Strasser (1978), p. 37. Explicitly, the Lebesgue-a.a. version occurs in van der Vaart (1997), p. 407.

2.3 *A side result for finite sample sizes*

Relation (2.3) implies that

$$\liminf_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} \leq N_{(0,\sigma^2(P))}(I)$$

for symmetric intervals I (up to an exceptional set of first category / Lebesgue measure zero). Being merely an asymptotic assertion, this is compatible with

$$P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} > N_{(0,\sigma^2(P))}(I) \text{ for every } P \in \mathfrak{P} \text{ and every } n \in \mathbb{N}. \tag{2.7}$$

Straightforward examples show that this does, in fact, occur. (Take $\mathfrak{P} = \{N_{(0,\sigma^2)} : \sigma^2 > 0\}$ and $\kappa^{(n)}(x_1, \ldots, x_n) = (n+2)^{-1} \sum_{\nu=1}^n x_\nu^2$.)

Relation (2.3) was derived from Proposition 4.2, a slightly stronger result which can be used to show that (2.7) cannot hold for every P in some neighbourhood of a point of asymptotic superefficiency (i.e. where (2.3) holds with $\leq$ replaced by $>$).

PROPOSITION 2.1. *Assume the conditions of Proposition* 4.2, *and* $\lim_{n\to\infty} \sigma^2(P_{n,u}) = \sigma^2(P)$ *for every* $P \in \mathfrak{P}$.

Then relation (2.7) *implies*

$$\lim_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in I\} = N_{(0,\sigma^2(P))}(I) \quad \textit{for every} \quad P \in \mathfrak{P}.$$

Mind that there is no exceptional set in $\mathfrak{P}$.

PROOF. Relation (2.7), with $P_{n,u}$ in place of P, reads

$$P^n_{n,u}\{c_n(\kappa^{(n)} - \kappa(P_{n,u})) \in I\} > N_{(0,\sigma^2(P_{n,u}))}(I) \quad \text{for} \quad u \in \mathbb{R} \text{ and } n \in \mathbb{N}.$$

Since σ^2 is continuous, this implies

$$\liminf_{n\to\infty} P^n_{n,u}\{c_n(\kappa^{(n)} - \kappa(P_{n,u})) \in I\} \geq N_{(0,\sigma^2(P_0))}(I),$$

and the assertion follows from Proposition 4.2. □

In terms of loss functions, Proposition 2.1 asserts that

$$\int \ell(c_n(\kappa^{(n)} - \kappa(P)))dP^n < \int \ell dN_{(0,\sigma^2(P))} \quad \text{for} \quad P \in \mathfrak{P} \text{ and } n \in \mathbb{N}$$

implies

$$\lim_{n\to\infty} \int \ell(c_n(\kappa^{(n)} - \kappa(P)))dP^n = \int \ell dN_{(0,\sigma^2(P))} \quad \text{for} \quad P \in \mathfrak{P},$$

for 1-dimensional functionals and bounded, symmetric loss functions. The example of Stein's shrinkage estimator for the mean vector μ in the family $\{N_{(\mu,I_3)} : \mu \in \mathbb{R}^3\}$ shows that this is not so in general. The optimal limit distribution, attained by the sample mean, is $N_{(0,I_3)}$. Stein's estimator $\kappa^{(n)} : X^n \to \mathbb{R}^3$ fulfills for the loss function $\ell(u) = \|u\|^2$ the relations

$$\int \ell(n^{1/2}(\kappa^{(n)} - \mu))dN^n_{(\mu,I_3)} < \int \ell dN_{(0,I_3)} \quad \text{for every} \quad \mu \in \mathbb{R}^3 \textit{ and } n \in \mathbb{N},$$

and

$$\lim_{n\to\infty} \int \ell(n^{1/2}(\kappa^{(n)} - \mu))dN^n_{(\mu,I_3)} < \int \ell dN_{(0,I_3)} \quad \text{for} \quad \mu = 0.$$

3. Proofs of the main results

The proofs of the various versions of Theorem 2.1 (median unbiased / general; parametric / nonparametric) follow the same pattern: (i) If a certain asymptotic property of an estimator sequence holds for every $P \in \mathfrak{P}_0$, then it holds locally uniformly on $\mathfrak{P}_0$, except for a "small" subset. (ii) If this asymptotic property holds locally uniformly at some P, then one may derive asymptotic bounds for the concentration of this estimator sequence at P by the usual techniques.

To make the proof of Theorem 2.1 as simple as possible, we isolate the abstract core of our argument. Let X be an arbitrary set, and f_n, g_n, $n = 0, 1, 2, \dots$ functions from X to $\mathbb{R}$.

CONDITION A. For every $x \in X$ there exists a set $\Xi_A(x)$ of sequences $x_n \in X$, $n \in \mathbb{N}$, with the following property: If

$$\liminf_{n\to\infty} f_n(x) \geq f_0(x) \quad \text{for} \quad x \in X_0 \subset X, \tag{3.1}$$

then there exists an exceptional set $X_+ \subset X_0$ and a subsequence $\mathbb{N}_0$ such that for $x \in X_0 - X_+$

$$\liminf_{n\in\mathbb{N}_0} f_n(x_n) \geq f_0(x) \quad \text{for} \quad (x_n)_{n\in\mathbb{N}} \in \Xi_A(x). \tag{3.2}$$

CONDITION B. For every $x \in X$ there exists a set $\Xi_B(x)$ of sequences $x_n \in X$, $n \in \mathbb{N}$, with the following property: For every $x \in X$ and every subsequence $\mathbb{N}_0$,

$$\liminf_{n\in\mathbb{N}_0} f_n(x_n) \geq f_0(x) \quad \text{for} \quad (x_n)_{n\in\mathbb{N}} \in \Xi_B(x), \tag{3.3}$$

implies

$$\limsup_{n\in\mathbb{N}_0} g_n(x) \leq g_0(x). \tag{3.4}$$

THEOREM 3.1. *Assume Conditions* A *and* B, *with* $\Xi_B(x) \subset \Xi_A(x)$ *for every* $x \in X$. *Let* X_0 *denote the set of all* $x \in X$ *for which* (3.1) *holds true. Let* X_+ *denote the exceptional set and* $\mathbb{N}_0$ *the subsequence from Condition* A. *Then* (3.4) *holds for* $x \in X_0 - X_+$.

Addendum. If $g_n = f_n$, $n = 0, 1, 2, \dots$, then

$$\liminf_{n\to\infty} f_n(x) \leq f_0(x) \quad \text{for} \quad x \in X_+^c. \tag{3.5}$$

PROOF. Theorem 3.1 is an immediate consequence of Conditions A and B. The Addendum follows from

$$\liminf_{n\to\infty} f_n(x) < f_0(x) \quad \text{for} \quad x \in X_0^c \text{ and}$$

$$\liminf_{n\to\infty} f_n(x) \leq \limsup_{n\in\mathbb{N}_0} f_n(x) \leq f_0(x) \quad \text{for} \quad x \in X_0 - X_+. \qquad \square$$

Theorem 2.1 now follows by specializing Theorem 3.1. Since $N_{(0,\sigma^2)}$ is nonatomic, it suffices to prove (2.2) and (2.3) for closed intervals.

a) General families $\mathfrak{P}$

Let $f_n(P) := P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in B\}$ for $n \in \mathbb{N}$, and $f_0(P) := 1/2$ in case (i) of Theorem 2.1, $f_0(P) = N_{(0,\sigma^2(P))}(B)$ in case (ii) of Theorem 2.1.

If B is closed, f_n is for $n \in \mathbb{N}$ upper semicontinuous by Corollary 5.1. Since f_0 is lower semicontinuous (for this we need upper semicontinuity of σ in case (ii) of Theorem 2.1), Proposition 5.2 implies Condition A for $X = P$ with $\mathbb{N}_0 = \mathbb{N}$, $\mathfrak{P}_+$ a set of first category, and $\Xi_A(P)$ the set of all sequences $(P_n)_{n\in\mathbb{N}} \to P$ $(\mathcal{U})$.

(ai) If $(\kappa^{(n)})_{n\in\mathbb{N}}$ is asymptotically median unbiased, Condition A holds for $f_n(P) := P^n\{\kappa^{(n)} \geq \kappa(P)\}$ and $f_0(P) := 1/2$. Proposition 4.1 implies Condition B for $X = P$ with $g_n(P) := P^n\{c_n(\kappa^{(n)} - \kappa(P)) \leq t\}$ for $n \in \mathbb{N}$ and $g_0(P) := N_{(0,\sigma^2(P))}(-\infty, t]$, with $\Xi_B(P)$ consisting of the sequences $(P_{n,u})_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$.

Provided $(P_{n,u})_{n\in\mathbb{N}}$ converges to P in the topology $\mathcal{U}$, Theorem 3.1 implies

$$\limsup_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \leq t\} \leq N_{(0,\sigma^2(P))}(-\infty, t]$$
$$\text{for every } P \notin \mathfrak{P}'_+ \text{ and every } t > 0.$$

The corresponding argument using $f_n(P) = P^n\{\kappa^{(n)} \leq \kappa(P)\}$ yields

$$\limsup_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \geq -t\} \leq N_{(0,\sigma^2(P))}[-t, \infty)$$
$$\text{for every } P \notin \mathfrak{P}''_+ \text{ and every } t > 0.$$

Hence (2.2) follows with $\mathbb{N}_0 = \mathbb{N}$ for $P \notin \mathfrak{P}'_+ \cup \mathfrak{P}''_+$.

(aii) For arbitrary estimator sequences, we apply the Addendum to Theorem 3.1 with $f_n(P) := P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in [-t, t]\}$ for $n \in \mathbb{N}$ and $f_0(P) := N_{(0,\sigma^2(P))}[-t, t]$. Condition A is fulfilled with an exceptional set of first category which now depends on t, say $\mathfrak{P}_t$. By Proposition 4.2, Condition B holds with $f_n = g_n$ for $n = 0, 1, 2, \ldots$ for $\Xi_B(P)$ consisting of the sequences $(P_{n,u})_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$.

Provided $(P_{n,u})_{n\in\mathbb{N}}$ converges to P in the topology $\mathcal{U}$, the Addendum to Theorem 3.1 implies

$$(3.6) \qquad \liminf_{n\to\infty} P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in [-t, t]\} \leq N_{(0,\sigma^2(P))}[-t, t] \qquad \text{for every} \quad P \in \mathfrak{P}_t.$$

It is now easy to see that (3.6) holds simultaneously for all $t > 0$ if $P \notin \cup\{\mathfrak{P}_t : t \in \mathbb{Q} \cap (0, \infty)\}$, again a set of first category.

This proves relation (2.3) with $\mathbb{N}$ in place of $\mathbb{N}_0$. Starting these considerations ab ovo with a subsequence $\mathbb{N}_0$, relation (2.3) follows.

b) Parametric families

It is now convenient to consider f_n, g_n as functions on Θ, i.e. $f_n(\vartheta) := P_\vartheta^n\{c_n(\kappa^{(n)} - \kappa(P_\vartheta)) \in B\}$ for $n \in \mathbb{N}$, and $f_0(\vartheta) = 1/2$ in case (i) of Theorem 2.1, $f_0(\vartheta) = N_{(0,\sigma^2(P_\vartheta))}(B)$ in case (ii) of Theorem 2.1. Conditions (4.2) and (4.8), (4.9) are now understood with $P_{n,u}$ replaced by $\vartheta + c_n^{-1}u$. With $\vartheta \to P_\vartheta(A)$, $\vartheta \to \kappa(P_\vartheta)$ (and $\vartheta \to \sigma(P_\vartheta)$ in case (ii) of Theorem 2.1) measurable, the functions f_n, $n = 0, 1, 2, \ldots$ are measurable. The Addendum to Proposition 5.1 implies Condition A for $X = \Theta$ with a subsequence $\mathbb{N}_0$, a set Θ_+ of λ^k-measure 0, and $\Xi_A(\vartheta)$ the set of all sequences $(\vartheta + c_n^{-1}u)_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$.

(bi) If $(\kappa^{(n)})_{n\in\mathbb{N}}$ is asymptotically median unbiased, Condition A holds for $X = \Theta$ with $f_n(\vartheta) := P_\vartheta^n\{\kappa^{(n)} \geq \kappa(P_\vartheta)\}$ and $f_0(\vartheta) = 1/2$. Proposition 4.1 implies Condition B with $g_n(\vartheta) := P_\vartheta^n\{c_n(\kappa^{(n)} - \kappa(P_\vartheta)) \leq t\}$ for $n \in \mathbb{N}$ and $g_0(\vartheta) := N_{(0,\sigma^2(P_\vartheta))}(-\infty, t]$, with $\Xi_B(\vartheta)$ consisting of all sequences $(\vartheta + c_n^{-1}u)_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$. The proof is now concluded as under (ai).

(bii) For arbitrary estimator sequences, we apply the Addendum to Theorem 3.1 with $f_n(\vartheta) := P_\vartheta^n\{c_n(\kappa^{(n)} - \kappa(P_0)) \in [-t,t]\}$ for $n \in \mathbb{N}$ and $f_0(\vartheta) = N_{(0,\sigma^2(\vartheta))}[-t,t]$. Condition A is fulfilled with a subsequence $\mathbb{N}_t$, a set Θ_t of λ^k-measure 0, and $\Xi_A(\vartheta)$ the set of all sequences $(\vartheta + c_n^{-1}u)_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$. By Proposition 4.2, Condition B holds with $g_n = f_n$ for $n = 0,1,2,\dots$ for $\Xi_B(\vartheta)$ consisting of all sequences $(\vartheta + c_n^{-1}u)_{n\in\mathbb{N}}$, $u \in \mathbb{Q}$. The proof is now concluded as under (aii). (Notice that the subsequence which enters through Condition A drops out.)

4. Auxiliary results on concentration bounds

This section contains asymptotic bounds for the concentration of estimator sequences. The bound at a given $P \in \mathfrak{P}$ is based on the performance of the estimator sequence along sequences $(P_{n,u})_{n\in\mathbb{N}}$, $u \in \mathbb{R}$, fulfilling certain LAN-type regularity conditions.

The constitutive condition which connects the local properties of $\mathfrak{P}$ with local properties of the functional κ is

$$c_n(\kappa(P_{n,u}) - \kappa(P_0)) \to u\sigma(P_0) \quad \text{for} \quad u \in \mathbb{R}. \tag{4.1}$$

In the following, $(\kappa^{(n)})_{n\in\mathbb{N}}$ is an arbitrary estimator sequence for κ.

Using a suitable asymptotic version of the Neyman-Pearson Lemma (see e.g. Lemma 8.2.15, p. 275, in Pfanzagl (1994)) we obtain the following.

PROPOSITION 4.1. *Assume that*

$$P_0^n \circ \log dP_{n,u}^n/dP_0^n \Rightarrow N_{(-u^2/2,u^2)} \quad \textit{for} \quad u \in \mathbb{R}. \tag{4.2}$$

Then the following is true.

If for every $u \in \mathbb{Q}$,

$$\liminf_{n\to\infty} P_{n,u}^n\{\kappa^{(n)} \le \kappa(P_{n,u})\} \ge 1/2 \tag{4.3$'$}$$

and

$$\liminf_{n\to\infty} P_{n,u}^n\{\kappa^{(n)} \ge \kappa(P_{n,u})\} \ge 1/2, \tag{4.3$''$}$$

then

$$\limsup_{n\to\infty} P_0^n\{c_n(\kappa^{(n)} - \kappa(P_0)) \in [-t',t'']\} \le N_{(0,\sigma^2(P_0))}[-t',t''] \tag{4.4}$$

for all $t', t'' \ge 0$.

PROOF. (i) Assume that $C_n \in \mathcal{A}$, $n \in \mathbb{N}$, fulfills $\liminf_{n\to\infty} P_{n,u}^n(C_n) \ge 1/2$. Considering C_n as a critical region for the test problem $P_0^n : P_{n,u}^n$, we obtain from (4.2) by the Neyman-Pearson Lemma that

$$\liminf_{n\to\infty} P_0^n(C_n) \ge N_{(0,1)}(u,\infty). \tag{4.5}$$

(ii) Given $t > 0$, let $u > t$ be arbitrary. From (4.5), applied with $C_n = \{\kappa^{(n)} \ge \kappa(P_{n,u})\}$, we obtain

$$\liminf_{n\to\infty} P_0^n\{\kappa^{(n)} \ge \kappa(P_{n,u})\} \ge N_{(0,1)}(u,\infty). \tag{4.6}$$

By (4.1), we have $c_n(\kappa(P_{n,u}) - \kappa(P_0)) > t\sigma(P_0)$ for n large. This implies

$$\{c_n(\kappa^{(n)} - \kappa(P_0)) > t\sigma(P_0)\} \supset \{\kappa^{(n)} \geq \kappa(P_{n,u})\}.$$

Together with (4.6) this implies

$$\liminf_{n\to\infty} P_0^n\{c_n(\kappa^{(n)} - \kappa(P_0)) > t\sigma(P_0)\} \geq N_{(0,1)}(u,\infty). \tag{4.7}$$

Since $u > t$ was arbitrary, relation (4.7) holds with u replaced by t.

(iii) The same argument, applied to $C_n = \{\kappa^{(n)} \leq \kappa(P_{n,u})\}$ with $-u < -t$ yields

$$\liminf_{n\to\infty} P_0^n\{c_n(\kappa^{(n)} - \kappa(P_0)) < -t\sigma(P_0)\} \geq N_{(0,1)}(-\infty,-t).$$

Both relations together imply (4.4). □

PROPOSITION 4.2. *Let $u_n, v_n \in \mathbb{R}$, $n \in \mathbb{N}$, be convergent sequences with $v := \lim_{n\to\infty} v_n$. Assume that*

$$P_{n,u_n}^n \circ \log dP_{n,u_n+v_n}^n / dP_{u_n}^n \Rightarrow N_{(-v^2/2,v^2)} \tag{4.8}$$

and that

$$u \to \kappa(P_{n,u}) \text{ is continuous on } \mathbb{R}. \tag{4.9}$$

Let $t > 0$ be fixed. If

$$\liminf_{n\to\infty} P_{n,u}^n\{c_n(\kappa^{(n)} - \kappa(P_{n,u})) \in [-t,t]\} \geq N_{(0,\sigma^2(P_0))}[-t,t] \quad \text{for every } u \in \mathbb{Q}, \tag{4.10}$$

then

$$\lim_{n\to\infty} P_0^n\{c_n(\kappa^{(n)} - \kappa(P_0)) \in [-t,t]\} = N_{(0,\sigma^2(P_0))}[-t,t]. \tag{4.11}$$

Proposition 4.2 is a modification of Theorem 8.6.3 in Pfanzagl ((1994), p. 298). We give the proof in extenso, since the proof given in l.c. is certainly less than optimal.

The implication from (4.10) to (4.11) is equivalent to the following.

$$\limsup_{n\to\infty} P_0^n\{c_n(\kappa^{(n)} - \kappa(P_0)) \in [-t,t]\} > N_{(0,\sigma^2(P_0))}[-t,t] \tag{4.12}$$

implies

$$\liminf_{n\to\infty} P_{n,u}^n\{c_n(\kappa^{(n)} - \kappa(P_{n,u})) \in [-t,t]\} < N_{(0,\sigma^2(P_0))}[-t,t] \quad \text{for some } u \in \mathbb{Q}. \tag{4.13}$$

Though technically less convenient, this expresses more clearly that superefficiency of $(\kappa^{(n)}))_{n\in\mathbb{N}}$ at P_0 brings about unwelcome properties of this estimator sequence in the neighbourhood of P_0.

PROOF. We shall prove that (4.12) implies (4.13). Since P_0 and t are fixed, they will be omitted if there is no danger of confusion.

Since $u \to c_n(\kappa(P_{n,u}) - \kappa(P_0))$ is continuous, relation (4.1) implies that for arbitrary $n, k \in \mathbb{N}$ there exists $u_{n,k}$ such that

(4.14) $$c_n(\kappa(P_{n,u_{n,k}}) - \kappa(P_0)) \in (2kt + (k-1)/n, 2kt + k/n)$$

and

(4.15) $$\lim_{n\to\infty} u_{n,k} = 2kt.$$

To simplify our notations, let $Q_{n,k} := P_{n,u_{n,k}}$. Moreover, let

(4.16′) $$A^-_{n,k} = \{\kappa^{(n)} < \kappa(Q_{n,k}) - c_n^{-1}t\}$$

(4.16″) $$A^+_{n,k} = \{\kappa^{(n)} > \kappa(Q_{n,k}) + c_n^{-1}t\}.$$

The relation $c_n(\kappa(Q_{n,k}) - \kappa(Q_{n,k-1})) > 2t$ implies

(4.17) $$(A^-_{n,k})^c \subset A^+_{n,k-1}.$$

Moreover, let

(4.18) $$\alpha^\pm_{n,k} = Q^n_{n,k}(A^\pm_{n,k}).$$

With these notations, relation (4.12) may be rewritten as

(4.19) $$\liminf_{n\to\infty}(\alpha^-_{n,0} + \alpha^+_{n,0}) < 2\Phi(-t/\sigma).$$

Let

(4.20) $$\alpha^+_k := \liminf_{n\to\infty} \alpha^+_{n,k}.$$

W.l.g. we may assume that

(4.21) $$\alpha^+_0 < \Phi(-t/\sigma).$$

Considering $A^+_{n,k-1}$ as a critical region for testing $Q^n_{n,k-1} : Q^n_{n,k}$ at level $\alpha^+_{n,k-1}$, we obtain from a suitable asymptotic version of the Neyman-Pearson Lemma that

(4.22) $$\liminf_{n\to\infty} Q^n_{n,k}(A^+_{n,k-1}) \le \Phi(\Phi^{-1}(\alpha^+_{k-1}) + 2t/\sigma).$$

Together with (4.17), evaluated by $Q^n_{n,k}$, we obtain

(4.23) $$\limsup_{n\to\infty} \alpha^-_{n,k} \ge \Phi(-\Phi^{-1}(\alpha^+_{k-1}) - 2t/\sigma).$$

Now we shall prove that

$$\limsup_{n\to\infty}(\alpha^-_{n,k_0} + \alpha^+_{n,k_0}) > 2\Phi(-t/\sigma) \quad \text{for some} \quad k_0 \in \mathbb{N}.$$

This is relation (4.13) with u_{n,k_0} in place of u. Because of (4.15), this implies (4.13) with $u = 2k_0t$.

Assume that, on the contrary,

$$\limsup_{n\to\infty}(\alpha_{n,k}^{-}+\alpha_{n,k}^{+})\le 2\Phi(-t/\sigma)\quad \text{for every}\quad k\in\mathbb{N}. \tag{4.24}$$

Since

$$\limsup_{n\to\infty}(\alpha_{n,k}^{-}+\alpha_{n,k}^{+})\ge \alpha_k^{+}+\limsup_{n\to\infty}\alpha_{n,k}^{-},$$

we obtain from (4.23) and (4.24) that

$$\alpha_k^{+}\le 2\Phi(-t/\sigma)-\Phi(-\Phi^{-1}(\alpha_{k-1}^{+})-2t/\sigma). \tag{4.25}$$

With $\delta_k := \Phi(-t/\sigma)-\alpha_k^{+}$, relation (4.25) may be rewritten as

$$\delta_k\ge \Phi(-\Phi^{-1}(\Phi(-t/\sigma)-\delta_{k-1})-2t/\sigma)-\Phi(-t/\sigma). \tag{4.26}$$

If $0<\delta_{k-1}<\Phi(-t/\sigma)$, relation (4.26) implies

$$\delta_k>\delta_{k-1}+\Delta\delta_{k-1}^2, \tag{4.27}$$

with $\Delta>0$ not depending on k. (By Lemma 8.6.23 in Pfanzagl (1994), p. 302, relation (4.27) holds with $\Delta = t/\sigma\varphi(t/\sigma)$.)

Recall that $\delta_0>0$ (by (4.21)). If (4.24) holds true for every $k=1,\dots,K$, we obtain from (4.27) that

$$\delta_K>\delta_0+K\Delta\delta_0^2. \tag{4.28}$$

Since $\delta_k\le\Phi(-t/\sigma)$ for every $k\in\mathbb{N}$, relation (4.24) must be violated for some k smaller than $\delta_0^{-2}\Delta^{-1}\Phi(-t/\sigma)$. □

Remark 4.1. Condition (4.8) follows from LAN. Assume that

$$\log\frac{dP_{n,u_n}^n}{dP_0^n}-\left(u\Delta_n-\frac{u^2}{2}\right)\to 0\ (P_0^n)\quad \text{if}\quad u_n\to u>0 \tag{4.29}$$

and

$$P_0^n\circ\Delta_n\Rightarrow N_{(0,1)}. \tag{4.30}$$

This condition implies that for every bounded sequence $(u_n)_{n\in\mathbb{N}}$ the sequences $(P_{n,u_n}^n)_{n\in\mathbb{N}}$ and $(P_0^n)_{n\in\mathbb{N}}$ are mutually contiguous. From this and Le Cam's 1st Lemma one obtains that

$$P_{n,u_n}^n\circ\Delta_n\Rightarrow N_{(u,1)}. \tag{4.31}$$

(For a convenient reference see Witting and Müller-Funk (1995), p. 311, Korollar 6.1.24 or Pfanzagl (1994), p. 217, Addendum to Corollary 6.7.11.)

Since (4.29) implies

$$\log\frac{dP_{n,u_n+v_n}^n}{dP_{n,u_n}^n}-\left(v\Delta_n-\frac{1}{2}(v^2+2uv)\right)\to 0\ (P_0^{(n)}),$$

relation (4.8) follows from (4.31).

5. Auxiliary results on locally uniform convergence

The proofs in Section 3 make use of the fact that a relation like $\liminf_{n\to\infty} f_n(x) \geq f_0(x)$ for $x \in X$ holds in some sense locally uniformly on a "large" subset of X. For the application to parametric families it suffices to have such a result for $X = \mathbb{R}^k$. This is Proposition 5.1. For the application to general families of probability measures, Proposition 5.2 presents results under topological conditions.

The following proposition is a slight generalization of Bahadur's Lemma: The assertion refers to lim inf rather than lim, and to a subset $X_0 \in \mathbb{B}^k$ (rather than $\mathbb{R}^k$ itself).

PROPOSITION 5.1. *Let $f_n : \mathbb{R}^k \to \mathbb{R}$, $n = 0, 1, 2, \ldots$ be measurable functions such that*

$$\liminf_{n\to\infty} f_n(x) \geq f_0(x) \quad \text{for} \quad x \in X_0 \in \mathbb{B}^k. \tag{5.1}$$

Then for every sequence $(u_n)_{n\in\mathbb{N}} \to 0$ there exists a λ^k-null set $X_+ \subset X_0$ and a subsequence $\mathbb{N}_0$ such that

$$\liminf_{n\in\mathbb{N}_0} f_n(x+u_n) \geq f_0(x) \quad \text{for} \quad x \in X_0 - X_+. \tag{5.2}$$

Addendum. Relation (5.2) holds simultaneously for any countable family of sequences $(u_n)_{n\in\mathbb{N}} \to 0$.

PROOF. (i) Let $N_{(0,I)} \mid \mathbb{B}^k$ denote the normal distribution with mean-vector 0 and covariance matrix I. Recall that for every function $f \in \mathcal{L}_1(\mathbb{R}^k, \mathbb{B}^k, N_{(0,I)})$ and every sequence $(u_n)_{n\in\mathbb{N}} \to 0$,

$$\lim_{n\to\infty} \int |f(x+u_n) - f(x)| N_{(0,I)}(dx) \to 0. \tag{5.3}$$

(Hint: Use Hewitt and Stromberg (1965), p. 199, Theorem 13.24.)

(ii) Let $g_n := f_n - f_0$. Using the decomposition $g_n = g_n^+ - g_n^-$, we obtain from (5.1) that $\lim_{n\to\infty} g_n^-(x) = 0$ for $x \in X_0$. Given a sequence $(u_n)_{n\in\mathbb{N}} \to 0$, let $X_n := X_0 + u_n$. We have

$$\begin{aligned} N_{(0,I)}\{x \in X_0 : g_n^-(x+u_n) > \varepsilon\} &= N_{(u_n,I)}\{x \in X_n : g_n^-(x) > \varepsilon\} \\ = N_{(0,I)}\{x \in X_n : g_n^-(x) > \varepsilon\} + o(n^0) &= N_{(0,I)}\{x \in X_0 : g_n^-(x) > \varepsilon\} = o(n^0). \end{aligned}$$

(Hint: $N_{(0,I)}(X_0 \triangle X_n) \to 0$ as a consequence of (5.3), applied with $f = 1_{X_0}$.)

Since the sequence $g_n^-(\cdot + u_n)$, $n \in \mathbb{N}$, converges to 0 in $N_{(0,I)}$-measure restricted to X_0, there exists (see e.g. Hewitt and Stromberg (1965), p. 156, Theorem 11.26) a subsequence $\mathbb{N}_0$ and a subset $X' \subset X_0$ with $N_{(0,I)}(X') = 0$ such that $\lim_{n\in\mathbb{N}_0} g_n^-(x+u_n) = 0$ for $x \in X_0 - X'$. Since $g_n^+(x+u_n) \geq 0$ for $x \in \mathbb{R}^k$, this implies

$$\liminf_{n\in\mathbb{N}_0} g_n(x+u_n) \geq 0 \quad \text{for} \quad x \in X_0 - X'. \tag{5.4}$$

(iii) By (5.3), applied with $f = f_0$ and $\mathbb{N}_0$ in place of $\mathbb{N}$, there exists $\mathbb{N}_1 \subset \mathbb{N}_0$ and a $N_{(0,I)}$-null set X'' such that $\lim_{n\in\mathbb{N}_1} f_0(x+u_n) = f_0(x)$ for $x \notin X''$. Together with (5.4) this implies (5.2) with $X_+ = X' \cup (X_0 \cap X'')$ and $\mathbb{N}_1$ in place of $\mathbb{N}_0$. □

The following proposition generalizes Lemma 9.1 in Pfanzagl (2002).

PROPOSITION 5.2. *Let $(X,\mathcal{U})$ be a Hausdorff space. For $n \in \mathbb{N}$ let $f_n : X \to \mathbb{R}$ be upper semicontinuous, and $f_0 : X \to \mathbb{R}$ lower semicontinuous.*

If

$$\liminf_{n\to\infty} f_n(x) \geq f_0(x) \quad \text{for} \quad x \in X_0, \tag{5.5}$$

then there exists a set of first category $X_+ \subset X_0$ such that for every $x_0 \in X_0 - X_+$

$$\liminf_{n\to\infty} f_n(x_n) \geq f_0(x_0) \text{ for every sequence } (x_n)_{n\in\mathbb{N}} \to x_0. \tag{5.6}$$

PROOF. (i) For $n, m \in \mathbb{N}$ let $A_{n,m} \subset X$ be closed. Then

$$\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m} - \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ}$$

is of first category.

If $x_0 \in (\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ})^c$, there exists m_0 such that $x_0 \in \bigcap_{n=1}^{\infty} (A_{n,m_0}^{\circ})^c \subset (A_{n,m_0}^{\circ})^c$ for every $n \in \mathbb{N}$.

If $x_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}$, then $x_0 \in \bigcup_{n=1}^{\infty} A_{n,m_0}$; hence there exists n_0, such that $x_0 \in A_{n_0,m_0}$. Therefore, $x_0 \in A_{n_0,m_0} - A_{n_0,m_0}^{\circ}$, which implies

$$\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m} - \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ} \subset \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} (A_{n,m} - A_{n,m}^{\circ}).$$

Since $A_{n,m} - A_{n,m}^{\circ}$ is a closed set with empty interior, this proves the assertion under (i).

(ii) First we prove the assertion for $f_0(x) = 0$. Let now

$$A_{n,m} := \bigcap_{k=n}^{\infty} \{x \in X : f_k(x) \geq -1/m\}.$$

Since f_k is upper semicontinuous, the sets $\{x \in X : f_k(x) \geq -1/m\}$ are closed. This, in turn, implies that $A_{n,m}$ is closed.

We have

$$\liminf_{n\to\infty} f_n(x_0) \geq 0 \quad \text{iff} \quad x_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}.$$

Moreover, $x_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ}$ implies

$$\liminf_{n\to\infty} f_n(x_n) \geq 0 \quad \text{if} \quad (x_n) \underset{n\in\mathbb{N}}{\longrightarrow} x_0. \tag{5.7}$$

If $x_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ}$, then for every m there is n_m such that $x_0 \in A_{n_m,m}^{\circ} =: V_m$. The set V_m is open, and $\inf_{x\in V_m} f_k(x) \geq -1/m$ for $k \geq n_m$. Since $(x_k)_{k\in\mathbb{N}} \to x_0$ implies $x_k \in V_m$ for $k \geq n'_m$, we have $f_k(x_k) \geq -1/m$ for $k \geq \max\{n_m, n'_m\}$. Hence (5.7) is true for every $x_0 \in X_0$, except for $x_0 \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m} - \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m}^{\circ}$, a set of first category.

This proves the assertion with $f_0(x) = 0$. Applying this assertion for $f_n - f_0$ (which is upper semicontinuous, since f_0 is lower semicontinuous), we obtain from (5.7) for $x_0 \in X_0 - X_+$

$$\liminf_{n\to\infty}(f_n(x_n) - f_0(x_n)) \geq 0 \text{ for every sequence } (x_n)_{n\in\mathbb{N}} \to x_0.$$

Since f_0 is lower semicontinuous, we have $\liminf_{n\to\infty} f_0(x_n) \geq f_0(x_0)$, and (5.6) follows. □

The application of Proposition 5.2 in Section 3 requires the semicontinuity of functions like $f_n(P) = P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in B\}$. This is guaranteed by Corollary 5.1.

LEMMA 5.1. *Let $\mathfrak{Q}$ denote the family of all probability measures $Q \mid \mathbb{B}$, endowed with the topology $\mathcal{U}_d$ induced by the sup-distance. Let $\mathcal{V}$ denote the topology of the Euclidean distance on $\mathbb{R}$. Let $B \in \mathbb{B}$ be fixed.*

Then the map $(Q, t) \to Q(B+t)$ is lower [upper] semicontinuous with respect to the product topology $\mathcal{U}_d \times \mathcal{V}$ if B is open [closed].

PROOF. It suffices to show that $d(Q_0, Q_m) \to 0$ and $t_m \to t_0$, $m \in \mathbb{N}$, implies

$$\liminf_{m\to\infty} Q_m(B + t_m) \geq Q_0(B + t_0) \tag{5.8}$$

if B is open. W.l.g. we assume $t_0 = 0$. If B is open, $x \in B$ implies $x \in B + t$ if t is sufficiently small. Hence $B \cap (B + t_m)^c \downarrow \emptyset$ if $t_m \to 0$. Since $Q_0(B + t_m) \geq Q_0(B) - Q_0(B \cap (B + t_m)^c)$, this implies $\liminf_{n\to\infty} Q_0(B + t_m) \geq Q_0(B)$. Since

$$|Q_0(B + t_m) - Q_m(B + t_m)| \leq d(Q_0, Q_m),$$

relation (5.8) follows. This proves the assertion for B open. The assertion for closed B follows since $Q(B) = 1 - Q(B^c)$. □

COROLLARY 5.1. *Let $\mathcal{U} \supset \mathcal{U}_d$ be a topology on $\mathfrak{P}$. If $\kappa : \mathfrak{P} \to \mathbb{R}$ is $\mathcal{U}$-continuous, then, for every $n \in \mathbb{N}$, $P \to P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in B\}$ is lower [upper] semicontinuous with respect to $\mathcal{U}$ if B is open [closed].*

PROOF. If the maps $P \to Q_P$ and $P \to t(P)$ are $(\mathcal{U}, \mathcal{U}_d)$ and $(\mathcal{U}, \mathcal{V})$-continuous, respectively, then $P \to Q_P(B + t(P))$ is lower semicontinuous by Lemma 5.1 for B open. Applied with $P^n \circ c_n\kappa^{(n)}$ and $c_n\kappa(P)$ in place of Q_P and $t(P)$, this yields the assertion. □

Remark 5.1. Notice the analogy and the distinction between Corollary 5.1 and Alexandrov's theorem, asserting that

$$\liminf_{m\to\infty} Q_m(B) \geq Q_0(B) \text{ for open subsets } B, \text{if } Q_m \Rightarrow Q_0.$$

Since $P^n\{c_n(\kappa^{(n)} - \kappa(P)) \in B\}$ depends on P through P^n and through $\kappa(P)$, we need the stronger condition $d(P_0, P_m) \to 0$ instead of $P_m \Rightarrow P_0$.

Acknowledgements

The author wishes to thank L. Mattner for bringing an error in an earlier version of Lemma 5.1 to his attention, and a referee for suggesting improvements of the presentation.

References

Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis*, Springer, New York.

Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates, *University of California Publications in Statistics*, **1**, 277–330.

Le Cam, L. (1974). Notes on asymptotic methods in statistical decision theory, Research Memo., Centre de Recherches Mathématiques, Univ. de Montreal.

Pfanzagl, J. (with the assistance of R. Hamböker) (1994). *Parametric Statistical Theory*, de Gruyter, Berlin.

Pfanzagl, J. (2002). Asymptotic optimality of estimator sequences: "Pointwise" versus "Locally uniform", *Math. Methods Statist.*, **11**, 69–97.

Strasser, H. (1978). Global asymptotic properties of risk functions in estimation, *Z. Wahrsch. Verw. Gebiete*, **45**, 35–48.

van der Vaart, A. W. (1997). Superefficiency, *Festschrift for Lucien Le Cam*, Research Papers in Probability and Statistics (eds. D. Pollard, E. Torgersen and G. L. Yang), 397–410, Springer, New York.

Witting, H. and Müller-Funk, U. (1995). *Mathematische Statistik II*, Teubner, Stuttgart.

Wolfowitz, J. (1953). The method of maximum likelihood and the Wald theory decision functions, *Indag. Math.*, **15**, 114–119.

Wolfowitz, J. (1965). Asymptotic efficiency of the maximum likelihood estimator, *Theory Prob. Appl.*, **X**, 247–260.