Z. Wahrscheinlichkeitstheorie verw. Geb. 18, 73–84 (1971)


# The Accuracy of the Normal Approximation for Minimum Contrast Estimates

R. MICHEL* and J. PFANZAGL

## 1. Introduction

Theorems on the asymptotic behavior of an estimate are useless for practical purposes, because they give no information on the behavior of the estimate for any sample size, however large. The usual practice disregards this fact. It is for instance common to use the asymptotic variance for computing the error of the estimate or for comparing the accuracy of two estimates. This reflects the common belief that asymptotic assertions provide a good approximation for the behavior of estimates for even small sample sizes.

This belief is supported by numerical computations occasionally performed for particular estimates. There are, however, hardly general results available. Kiefer ([9], p. 141) has called this a "terrifically difficult problem". The only pertinent result known to the authors is that of Linnik and Mitrofanova ([10] and [11]). From their Theorem 1 one can derive that for maximum likelihood estimates $\vartheta_n$, $n \in \mathbb{N}$, of a shift parameter family $P_\vartheta$, $\vartheta \in \Theta$:

$$\left| P_\vartheta^{\mathbb{N}}\left\{ \underline{x} \in X^{\mathbb{N}} \colon \frac{\vartheta_n(\underline{x}) - \vartheta}{\beta(\vartheta)} n^{\frac{1}{2}} < t \right\} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left[ -\frac{r^2}{2} \right] dr \right| \leqq c n^{-\frac{1}{2}} \quad \text{for } |t| \leqq a_\vartheta .$$

Our result obtained below holds for arbitrary families of probability measures (fulfilling certain regularity conditions) and for all $t \in \mathbb{R}$. It does, however, not contain the result of Linnik and Mitrofanova as a special case, because we obtain an upper bound of the order $n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$ only instead of $n^{-\frac{1}{2}}$. The method of proof is entirely different[1].

## 2. The Main Result

Let $(X, \mathscr{A})$ be a measurable space and $P_\vartheta | \mathscr{A}$, $\vartheta \in \Theta$, a family of probability measures. In the following, $\Theta$ will be assumed to be an open interval (possibly $\Theta = \mathbb{R}$). Let $\Theta^c$ denote the closure of $\Theta$ in $\overline{\mathbb{R}} := [-\infty, +\infty]$ and $\mathscr{B}$ the pertaining Borel field over $\Theta^c$.

A family of $\mathscr{A}$-measurable functions $f_\vartheta \colon X \to \overline{\mathbb{R}}$, $\vartheta \in \Theta^c$, is a family of *contrast functions* for $\{P_\vartheta \colon \vartheta \in \Theta\}$, if $P_\vartheta(f_\tau)$ exists[2] for all $\vartheta \in \Theta$, $\tau \in \Theta^c$ and if

$$P_\vartheta(f_\vartheta) < P_\vartheta(f_\tau) \qquad \text{for all } \vartheta \in \Theta,\ \tau \in \Theta^c,\ \vartheta \neq \tau. \tag{1}$$

---

* The paper was written while this author was employed by a grant of the Deutsche Forschungsgemeinschaft.

[1] *Note Added in Proof.* In the meantime, one of the authors has been able to establish a bound of the order $n^{-\frac{1}{2}}$ (to appear: Metrika 17, Heft 2).

[2] This means that not $P_\vartheta(f_\tau^-) = P_\vartheta(f_\tau^+) = \infty$.

An *estimate* for the sample size $n$ is an $\mathscr{A}^{\mathbb{N}}$-measurable map $\vartheta_n$: $X^{\mathbb{N}} \to \overline{\mathbb{R}}$, which depends on $x_1, \ldots, x_n$ only.

A *minimum contrast (m.c.) estimate* for the sample size $n$ is an estimate $\vartheta_n$ for which $\vartheta_n(X^{\mathbb{N}}) \subset \Theta^c$ and

$$\sum_{i=1}^{n} f_{\vartheta_n(\underline{x})}(x_i) = \inf_{\vartheta \in \Theta^c} \sum_{i=1}^{n} f_\vartheta(x_i).$$

We remark that maximum likelihood estimates are particular m.c. estimates. Let $\mu$ be a $\sigma$-finite measure such that $P_\vartheta|\mathscr{A} \ll \mu|\mathscr{A}$, $\vartheta \in \Theta$, and $h_\vartheta$ a density of $P_\vartheta|\mathscr{A}$ relative to $\mu|\mathscr{A}$. Assume that there exist functions $\bar{h}_\vartheta$: $X \to \overline{\mathbb{R}}$, $\vartheta \in \Theta^c$, such that

(i) $\bar{h}_\vartheta = h_\vartheta$ for $\vartheta \in \Theta$;

(ii) $\vartheta \to \bar{h}_\vartheta(x)$ is continuous on $\Theta^c$ for every $x \in X$;

(iii) $P_\vartheta(\log \bar{h}_\tau) < P_\vartheta(\log h_\vartheta)$ for all $\vartheta \in \Theta$ and $\tau \in \Theta^c$, $\vartheta \neq \tau$.

(We remark that condition (iii) is necessarily fulfilled for all $\tau \in \Theta$ such that $P_\vartheta(\log h_\tau) < \infty$ $(\tau \neq \vartheta)$ if $\tau \neq \vartheta$ implies $P_\tau \neq P_\vartheta$ and if, furthermore, $-\infty < P_\vartheta(\log h_\vartheta)$ for all $\vartheta \in \Theta$.)

If conditions (i), (ii), (iii) are fulfilled, then $f_\vartheta := -\log \bar{h}_\vartheta$, $\vartheta \in \Theta^c$, is a family of contrast functions. The pertaining m.c. estimates are maximum likelihood estimates.

For more details as well as for general conditions on the existence and consistency of m.c. estimates see Pfanzagl [12].

**Theorem.** *Assume that $\Theta \subset \mathbb{R}$ is an open interval. Let $\vartheta_n$, $n \in \mathbb{N}$, denote a sequence of minimum contrast estimates. Under the regularity conditions listed below, for every compact $K \subset \Theta$ there exists $c \in \mathbb{R}$ (depending on $K$) such that*

$$\left| P_\vartheta^{\mathbb{N}} \left\{ \underline{x} \in X^{\mathbb{N}} : \frac{\vartheta_n(\underline{x}) - \vartheta}{\beta(\vartheta)} n^{\frac{1}{2}} < t \right\} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left[-\frac{r^2}{2}\right] dr \right|$$
$$\leq c n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}} \quad \text{for all } t \in \mathbb{R},\ \vartheta \in K,\ n \in \mathbb{N},$$

*where*

$$\beta(\vartheta) := \frac{[P_\vartheta((f'(\cdot, \vartheta))^2)]^{\frac{1}{2}}}{P_\vartheta(f''(\cdot, \vartheta))}.$$

*Remark.* For maximum likelihood estimates the last formula reduces to

$$\beta(\vartheta) = \left[ P_\vartheta \left( \left( \frac{\partial}{\partial \vartheta} \log h_\vartheta \right)^2 \right) \right]^{-\frac{1}{2}}.$$

A straightforward application of Schwarz's inequality yields that under suitable regularity conditions

$$\left( \text{namely } P_\vartheta(f'(\cdot, \vartheta)) = 0,\ P_\vartheta(f''(\cdot, \vartheta)) > 0, \right.$$
$$\left. \text{and } \frac{\partial}{\partial \vartheta} \mu(f'(\cdot, \vartheta) h_\vartheta) = \mu \left( \frac{\partial}{\partial \vartheta} (f'(\cdot, \vartheta) h_\vartheta) \right) \right)$$
$$\left[ P_\vartheta \left( \left( \frac{\partial}{\partial \vartheta} \log h_\vartheta \right)^2 \right) \right]^{-\frac{1}{2}} \leq \frac{[P_\vartheta((f'(\cdot, \vartheta))^2)]^{\frac{1}{2}}}{P_\vartheta(f''(\cdot, \vartheta))}$$

for any family of contrast functions. This is in accord with the fact that maximum likelihood estimates are asymptotically efficient (see Bahadur [2], Schmetterer [13]).

The theorem in its present form has no immediate practical consequences, because it asserts the existence of a number $c$, but gives no method for determining its value for a given family of probability measures. Work in this direction by one of the authors is, however, in progress.

Let $X=\{0,1\}$, $\Theta=(0,1)$ and $P_\vartheta\{0\}=1-\vartheta$, $P_\vartheta\{1\}=\vartheta$, $\vartheta\in(0,1)$. Then $f_\vartheta(x)=(x-1)\log(1-\vartheta)-x\log\vartheta$, $x\in X$, $\vartheta\in[0,1]$ (where $0\infty=0$), is a family of contrast functions and $\vartheta_n(\underline{x})=n^{-1}\sum_{i=1}^{n} x_i$ is the pertaining sequence of m.c. estimates. For every compact subset $K\subset\Theta$ with $\frac{1}{2}\in K$ we obtain from Esseen ([4], p. 161) that:

$$\lim_{n\in\mathbb{N}}\left(\sup_{\vartheta\in K}\sup_{t\in\mathbb{R}} n^{\frac{1}{2}}\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \frac{\vartheta_n(\underline{x})-\vartheta}{\vartheta^{\frac{1}{2}}(1-\vartheta)^{\frac{1}{2}}}\, n^{\frac{1}{2}}<t\right\}\right.\right.$$
$$\left.\left.-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left(-\frac{r^2}{2}\right)dr\right|\right)\geq\frac{1}{\sqrt{2\pi}}.$$

Hence a better bound than $cn^{-\frac{1}{2}}$ cannot be achieved in general. The range between $cn^{-\frac{1}{2}}$ and $cn^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$ remains to be explored.

The theorem is proved under the following

*Regularity Conditions.*

(i) $\vartheta\to P_\vartheta$ is continuous in $\Theta$ with respect to the supremum-metric on $\{P_\vartheta:\vartheta\in\Theta\}$ (defined by the distance function $d(P,Q)=\sup_{A\in\mathscr{A}}|P(A)-Q(A)|$).

(ii) For each $x\in X$, $\vartheta\to f_\vartheta(x)$ is continuous in $\Theta^c$.

(iii) For every $\vartheta\in\Theta$ and every compact $K\subset\Theta$:

(a) $\sup_{\tau\in K} P_\tau(f_\vartheta^2)<\infty$.

(b) $f_\vartheta$ is uniformly integrable on $\{P_\tau:\tau\in K\}$.

($g$ is uniformly integrable on $\{P_\tau:\tau\in K\}$, if for every $\varepsilon>0$ there exists $a_\varepsilon>0$ such that $\sup_{\tau\in K} P_\tau(|g|\,1_{\{x:|g(x)|>a_\varepsilon\}})<\varepsilon$.)

(iv) For each $x\in X$, $\vartheta\to f_\vartheta(x)$ is twice differentiable in $\Theta$. With

$$f'(x,\vartheta):=\frac{\partial}{\partial\vartheta}f_\vartheta(x)\quad\text{and}\quad f''(x,\vartheta):=\frac{\partial^2}{\partial\vartheta^2}f_\vartheta(x)$$

we have for all $\vartheta\in\Theta$: $P_\vartheta(f'(\cdot,\vartheta))=0$.

(We remark that, because of (1), this condition and the condition $P_\vartheta(f''(\cdot,\vartheta))>0$ mentioned below essentially require that the order of differentiation and integration may be interchanged.)

(v) For every compact $K\subset\Theta$:

(a) $\inf_{\vartheta\in K} P_\vartheta((f'(\cdot,\vartheta))^2)>0$.

(b) $\inf_{\vartheta\in K} P_\vartheta\big((f''(\cdot,\vartheta)-P_\vartheta(f''(\cdot,\vartheta)))^2\big)>0.$

(c) $\inf_{\vartheta\in K} P_\vartheta(f''(\cdot,\vartheta))>0.$

(By (i) and (iv), these conditions are fulfilled if $P_\vartheta(f''(\cdot,\vartheta))>0$ and $P_\vartheta\big((f''(\cdot,\vartheta)-P_\vartheta(f''(\cdot,\vartheta))^2\big)>0$ for all $\vartheta\in\Theta$, and if $(f'(\cdot,\vartheta))^2$ and $(f''(\cdot,\vartheta))^2$ are uniformly integrable on $\{P_\tau\colon \tau\in K\}$.)

(vi) For every compact $K\subset\Theta$:

(a) $\sup_{\vartheta\in K} P_\vartheta(|f'(\cdot,\vartheta)|^3)<\infty.$

(b) $\sup_{\vartheta\in K} P_\vartheta(|f''(\cdot,\vartheta)|^3)<\infty.$

(vii) For every $\vartheta\in\Theta^c$ there exists a neighborhood $U_\vartheta$ of $\vartheta$ such that

(a) for every neighborhood $U$ of $\vartheta$, $U\subset U_\vartheta$, and every compact $K\subset\Theta$

$$\sup_{\vartheta\in K} P_\vartheta\big((\inf_{\sigma\in U} f_\sigma)^2\big)<\infty.$$

(b) for every compact $K\subset\Theta$ the function $\inf_{\sigma\in U_\vartheta} f_\sigma$ is uniformly integrable from below on $\{P_\tau\colon \tau\in K\}$.

($g$ is uniformly integrable from below on $\{P_\tau\colon \tau\in K\}$, if for every $\varepsilon>0$ there exists $a_\varepsilon>0$ such that $\inf_{\tau\in K} P_\tau(g\,1_{\{x\colon g(x)<-a_\varepsilon\}})>-\varepsilon$.)

(viii) For every $\vartheta\in\Theta$ there exists an open neighborhood $V_\vartheta$ of $\vartheta$ and a measurable function $k_\vartheta\colon X\to\overline{\mathbb{R}}$ such that $\sup_{\tau\in K} P_\tau(k_\vartheta^2)<\infty$ for every compact $K\subset\Theta$ and

$$|f''(x,\tau')-f''(x,\tau)|\leqq|\tau'-\tau|\,k_\vartheta(x)\qquad \text{for all } \tau',\tau\in V_\vartheta,\ x\in X.$$

The following notations will be used:

$$\Phi(t):=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\exp\left[-\frac{r^2}{2}\right]dr,$$

$$u(x,\vartheta):=-\big[P_\vartheta((f'(\cdot,\vartheta))^2)\big]^{-\frac12} f'(x,\vartheta),$$

$$a(\vartheta):=\big[P_\vartheta(f''(\cdot,\vartheta))\big]^{-1}.$$

The letter $c$ will be used to denote a generic constant. In this way we avoid the useless distinction between a great number of different constants.

*Proof of the Theorem.* Let $d$ be the distance between $K$ and $\overline{\mathbb{R}}-\Theta$. If $\vartheta\in K$ and $|\vartheta_n(\underline{x})-\vartheta|<d$, the m.c. estimate $\vartheta_n$ fulfills

$$\sum_{i=1}^{n} f'(x_i,\vartheta_n(\underline{x}))=0.$$

Starting from the expansion

$$\sum_{i=1}^{n} f'(x_i,\vartheta_n(\underline{x}))=\sum_{i=1}^{n} f'(x_i,\vartheta)+(\vartheta_n(\underline{x})-\vartheta)\sum_{i=1}^{n} f''(x_i,\hat\vartheta_n(\underline{x},\vartheta)), \tag{2}$$

$|\hat\vartheta_n(\underline{x},\vartheta)-\vartheta|\leqq|\vartheta_n(\underline{x})-\vartheta|$, we therefore obtain:

$$n^{\frac12}\,\frac{\vartheta_n(\underline{x})-\vartheta}{\beta(\vartheta)}\left[n^{-1}a(\vartheta)\sum_{i=1}^{n} f''(x_i,\hat\vartheta_n(\underline{x},\vartheta))\right]=n^{-\frac12}\sum_{i=1}^{n} u(x_i,\vartheta). \tag{3}$$

Condition (iv) implies $P_\vartheta(u(\cdot,\vartheta))=0$. Furthermore, $P_\vartheta((u(\cdot,\vartheta))^2)=1$.

In the following we shall apply Lemma 1 for $Z=X^{\mathbb{N}}$, $Q=P_\vartheta^{\mathbb{N}}$, $r=an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$ (where $a$ is the constant occurring in Lemma 7),

$$f\colon \underline{x}\to n^{-\frac{1}{2}}\sum_{i=1}^{n} u(x_i,\vartheta), \qquad g\colon \underline{x}\to n^{-1}a(\vartheta)\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta)).$$

We remark that $g$ is measurable because of (2).

Since

$$n^{-1}a(\vartheta)\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))>0$$

implies

$$\frac{\vartheta_n(\underline{x})-\vartheta}{\beta(\vartheta)}\,n^{\frac{1}{2}}<t$$

iff

$$n^{-\frac{1}{2}}\sum_{i=1}^{n} u(x_i,\vartheta)<t n^{-1}a(\vartheta)\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta)),$$

we obtain from Lemma 1 for all $t\in\mathbb{R}$, $\vartheta\in K$, $n\in\mathbb{N}$:

$$\begin{aligned}&\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}\colon \frac{\vartheta_n(\underline{x})-\vartheta}{\beta(\vartheta)}\,n^{\frac{1}{2}}<t\right\}-\Phi(t)\right|\\ &\quad\leqq P_\vartheta^{\mathbb{N}}\{\underline{x}\in X^{\mathbb{N}}\colon |\vartheta_n(\underline{x})-\vartheta|\geqq d\}\\ &\qquad+\sup_{s\in\mathbb{R}}\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}\colon n^{-\frac{1}{2}}\sum_{i=1}^{n}u(x_i,\vartheta)<s\right\}-\Phi(s)\right|\\ &\qquad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}\colon \left|n^{-1}\sum_{i=1}^{n}\left(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1\right)\right|\geqq an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\\ &\qquad+an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}.\end{aligned}$$

(Hint: $n^{-1}a(\vartheta)\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))\leqq 0$ implies

$$\left|n^{-1}\sum_{i=1}^{n}\left(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1\right)\right|\geqq an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}},$$

except $an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}>1$; in the latter case the inequality is trivial.)

As the regularity conditions (v)(a), (vi)(a) are fulfilled, the Berry-Esseen theorem (see e.g. Feller [6], p. 515, Theorem 1) yields the existence of a constant $c''$ such that

$$\sup_{s\in\mathbb{R}}\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}\colon n^{-\frac{1}{2}}\sum_{i=1}^{n}u(x_i,\vartheta)<s\right\}-\Phi(s)\right|\leqq c''n^{-\frac{1}{2}} \qquad \text{for all } \vartheta\in K,\ n\in\mathbb{N}.$$

Together with Lemmas 4 and 7 this implies

$$\begin{aligned}&\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}\colon \frac{\vartheta_n(\underline{x})-\vartheta}{\beta(\vartheta)}\,n^{\frac{1}{2}}<t\right\}-\Phi(t)\right|\\ &\qquad\leqq c'n^{-1}+c''n^{-\frac{1}{2}}+c'''n^{-\frac{1}{2}}+an^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\end{aligned}$$

for all $t\in\mathbb{R}$, $\vartheta\in K$, and $n\in\mathbb{N}$, where $c'$, $c''$, $c'''$ are real numbers depending on $K$.

## 3. A Few Lemmas

**Lemma 1.** *Let* $(Z, \mathscr{C}, Q)$ *be a probability space and* $f\colon Z \to \mathbb{R}$, $g\colon Z \to \mathbb{R}$ $\mathscr{C}$*-measurable functions. Then:*

$$|Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)|$$
$$\leqq \sup_{s\in\mathbb{R}} |Q\{z\in Z: f(z)<s\}-\Phi(s)|+Q\{z\in Z: |g(z)-1|>r\}+r \quad \textit{for all } t\in\mathbb{R},\ r>0.$$

*Proof.* As the assertion is trivial for $r\geqq 1$, we shall assume $r\in(0,1)$ in the following.

(i) *Lower bound* for $Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)$:

(a) $t\geqq 0$:

$f(z)<t(1-r)$ implies $f(z)<t\,g(z)$ or $g(z)<1-r$. Hence

$$Q\{z\in Z: f(z)<t(1-r)\}\leqq Q\{z\in Z: f(z)<t\,g(z)\}+Q\{z\in Z: |g(z)-1|>r\}. \tag{4}$$

Furthermore,

$$\left|\Phi(t)-\Phi\big(t(1-r)\big)\right|\leqq \min\left(\tfrac{1}{2}, |t|\,r(2\pi)^{-\frac{1}{2}}\exp\left[-\tfrac{1}{2}t^2(1-r)^2\right]\right)$$
$$\leqq \min\left(\tfrac{1}{2}, (2\pi e)^{-\frac{1}{2}}\,r(1-r)^{-1}\right)\leqq r.$$

Together with (4) this implies:

$$Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)\geqq -\left|Q\{z\in Z: f(z)<t(1-r)\}-\Phi\big(t(1-r)\big)\right|$$
$$-Q\{z\in Z: |g(z)-1|>r\}-r.$$

(b) $t<0$:

$f(z)<t(1+r)$ implies $f(z)<t\,g(z)$ or $g(z)>1+r$. Hence

$$Q\{z\in Z: f(z)<t(1+r)\}\leqq Q\{z\in Z: f(z)<t\,g(z)\}+Q\{z\in Z: |g(z)-1|>r\}. \tag{5}$$

Furthermore,

$$\left|\Phi\big(t(1+r)\big)-\Phi(t)\right|\leqq |t|\,r(2\pi)^{-\frac{1}{2}}\exp\left[-\tfrac{1}{2}t^2\right]\leqq r(2\pi e)^{-\frac{1}{2}}\leqq r.$$

Together with (5) this implies:

$$Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)\geqq -\left|Q\{z\in Z: f(z)<t(1+r)\}-\Phi\big(t(1+r)\big)\right|$$
$$-Q\{z\in Z: |g(z)-1|>r\}-r.$$

Regardless of whether $t\geqq 0$ or $t<0$, we therefore obtain the following lower bound:

$$Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)\geqq -\sup_{s\in\mathbb{R}}|Q\{z\in Z: f(z)<s\}-\Phi(s)|$$
$$-Q\{z\in Z: |g(z)-1|>r\}-r.$$

(ii) Similarly as in (i) we obtain the following upper bound:

$$Q\{z\in Z: f(z)<t\,g(z)\}-\Phi(t)\leqq \sup_{s\in\mathbb{R}}|Q\{z\in Z: f(z)<s\}-\Phi(s)|$$
$$+Q\{z\in Z: |g(z)-1|>r\}+r.$$

(Hint: For $t\geqq 0$ use that $f(z)<t\,g(z)$ implies $f(z)<t(1+r)$ or $g(z)>1+r$; for $t<0$ use that $f(z)<t\,g(z)$ implies $f(z)<t(1-r)$ or $g(z)<1-r$.)

**Lemma 2.** *Let $(X, \mathscr{A})$ be a measurable space, $\mathfrak{P}|\mathscr{A}$ a family of probability measures and $g\colon X \to \mathbb{R}$ an $\mathscr{A}$-measurable function which is uniformly integrable from below [above] on $\mathfrak{P}$. Then $P \to P(g)$ is lower [upper] semicontinuous on $\mathfrak{P}$.*

*Proof.* We shall restrict ourselves to prove lower semicontinuity at $P_0$ for the case $P_0(g) < \infty$. The case $P_0(g) = \infty$ requires obvious modifications.

(i) At first we shall show that $P \to P(g)$ is lower semicontinuous if $g$ is nonnegative. For every $\varepsilon > 0$ there exists $b_\varepsilon$ such that with $B_\varepsilon := \{x \in X\colon g(x) \leqq b_\varepsilon\}$, $P_0(g 1_{B_\varepsilon}) \geqq P_0(g) - \varepsilon/2$. This implies $P(g) \geqq P(g 1_{B_\varepsilon}) \geqq P_0(g 1_{B_\varepsilon}) - 2 b_\varepsilon d(P, P_0)$. Hence $d(P, P_0) \leqq \varepsilon/4 b_\varepsilon$ implies $P(g) \geqq P_0(g) - \varepsilon$.

(ii) Now we shall show that $P \to P(g)$ is lower semicontinuous for functions $g$ which are uniformly integrable from below. By assumption, for every $\varepsilon > 0$ there exists $a_\varepsilon > 0$ such that, with $A_\varepsilon := \{x \in X\colon g(x) < -a_\varepsilon\}$:

$$P(g 1_{A_\varepsilon}) > -\varepsilon/2 \qquad \text{for all } P \in \mathfrak{P}.$$

Therefore,

$$\begin{aligned} P(g) - P_0(g) &= P(g 1_{A_\varepsilon}) - P_0(g 1_{A_\varepsilon}) + P(g 1_{\bar{A}_\varepsilon}) - P_0(g 1_{\bar{A}_\varepsilon}) \\ &\geqq -\varepsilon/2 + P(g 1_{\bar{A}_\varepsilon}) - P_0(g 1_{\bar{A}_\varepsilon}). \end{aligned}$$

Furthermore,

$$P(g 1_{\bar{A}_\varepsilon}) - P_0(g 1_{\bar{A}_\varepsilon}) \geqq P((g + a_\varepsilon) 1_{\bar{A}_\varepsilon}) - P_0((g + a_\varepsilon) 1_{\bar{A}_\varepsilon}) - a_\varepsilon d(P, P_0).$$

As $(g + a_\varepsilon) 1_{\bar{A}_\varepsilon}$ is nonnegative, the proof can be concluded by using the result of (i).

**Lemma 3.** *Let $(X, \mathscr{A}, P)$ be a probability space and $T$ a Hausdorff space fulfilling the first axiom of countability. Assume that $g_t\colon X \to \overline{\mathbb{R}}$, $t \in T$, is a family of $\mathscr{A}$-measurable functions such that for $t_0 \in T$:*

(i) *for every $x \in X$, $t \to g_t(x)$ is lower semicontinuous in $t_0$,*

(ii) $\inf\limits_{t \in U} g_t$ *is $\mathscr{A}$-measurable for every open neighborhood $U$ of $t_0$,*

(iii) $P(\inf\limits_{t \in T} g_t) > -\infty$.

*Then $\sup\limits_{U \in \mathscr{U}_0} P(\inf\limits_{t \in U} g_t) = P(g_{t_0})$, where $\mathscr{U}_0$ denotes the system of all open neighborhoods of $t_0$.*

*(From this we easily conclude that $t \to P(g_t)$ is lower semicontinuous in $t_0$, i.e., $\sup\limits_{U \in \mathscr{U}_0} \inf\limits_{t \in U} P(g_t) = P(g_{t_0})$.)*

*Proof.* Let $U_n \in \mathscr{U}_0$ be such that $U_{n+1} \subset U_n$ and $\bigcap\limits_{n=1}^{\infty} U_n = \{t_0\}$. By (i) we have $\lim\limits_{n \in \mathbb{N}} \inf\limits_{t \in U_n} g_t = g_{t_0}$. By (ii) the monotone convergence theorem is applicable and yields $\lim\limits_{n \in \mathbb{N}} P(\inf\limits_{t \in U_n} g_t) = P(g_{t_0})$. From this the assertion follows easily.

The following lemma is related to results of Wald ([14], p. 599, Theorem 1), Brillinger ([3], p. 576, Theorem), and Kaufmann ([7], p. 168, Theorem 4.1).

**Lemma 4.** *Let $\vartheta_n$, $n \in \mathbb{N}$, be a sequence of m.c. estimates. Assume that regularity conditions* (i), (ii), (iii), *and* (iv) *are fulfilled. Then for every $\varepsilon > 0$ and every compact $K \subset \Theta$ there exists $c \in \mathbb{R}$ (depending on $\varepsilon$ and $K$) such that*

$$P_\vartheta^{\mathbb{N}}\{\underline{x} \in X^{\mathbb{N}}\colon |\vartheta_n(\underline{x}) - \vartheta| \geqq \varepsilon\} \leqq c n^{-1} \qquad \textit{for all } \vartheta \in K,\ n \in \mathbb{N}.$$

*Proof.* Let $C:=\{(\vartheta,\tau)\in K\times\Theta^c: |\vartheta-\tau|\geqq\varepsilon\}$. W.l.g. $C\neq\emptyset$. $(\vartheta,\tau)\in C$ implies $\vartheta\neq\tau$ and therefore by (1) $P_\vartheta(f_\vartheta)<P_\vartheta(f_\tau)$. Because of conditions (ii) and (vii)(b), Lemma 3 may be applied for $P=P_\vartheta$, $T=U_\tau$, $g_t=f_t$, $t_0=\tau$ to obtain the existence of an open neighborhood $U_{(\vartheta,\tau)}\subset U_\tau$ of $\tau$ such that $P_\vartheta(f_\vartheta)<P_\vartheta(g_{(\vartheta,\tau)})$, where $g_{(\vartheta,\tau)}:=\inf\{f_\sigma: \sigma\in U_{(\vartheta,\tau)}\}$. Because of conditions (i), (iii)(b), and (vii)(b), Lemma 2 implies that the map $\delta\to P_\delta(f_\vartheta)$ is continuous and the map $\delta\to P_\delta(g_{(\vartheta,\tau)})$ is lower semicontinuous. Hence there exists a compact neighborhood $C_{(\vartheta,\tau)}\subset\Theta$ of $\vartheta$ such that $P_\delta(f_\vartheta)<P_\delta(g_{(\vartheta,\tau)})$ for all $\delta\in C_{(\vartheta,\tau)}$.

Let $A^\circ$ denote the interior of $A$. As $\{C^\circ_{(\vartheta,\tau)}\times U_{(\vartheta,\tau)}: (\vartheta,\tau)\in C\}$ covers the compact set $C$ (recall that $\Theta^c$ is the closure of $\Theta$ in $\mathbb{R}$), there exists a finite subcover determined by $(\vartheta_j,\tau_j)\in C$, $j=1,\dots,m$, say. For notational convenience, let $C_j:=C_{(\vartheta_j,\tau_j)}$, $U_j:=U_{(\vartheta_j,\tau_j)}$ and $g_j:=g_{(\vartheta_j,\tau_j)}$.

As $\vartheta_n(\underline{x})\in U_j$ implies

$$n^{-1}\sum_{i=1}^n g_j(x_i)\leqq n^{-1}\sum_{i=1}^n f_{\vartheta_j}(x_i)$$

by definition of the m.c. estimate, we obtain that

$$n^{-1}\sum_{i=1}^n (g_j(x_i)-P_\delta(g_j))\leqq -\tfrac{1}{2}(P_\delta(g_j)-P_\delta(f_{\vartheta_j}))$$

or

$$n^{-1}\sum_{i=1}^n (f_{\vartheta_j}(x_i)-P_\delta(f_{\vartheta_j}))\geqq \tfrac{1}{2}(P_\delta(g_j)-P_\delta(f_{\vartheta_j})).$$

As $\delta\to P_\delta(g_j)-P_\delta(f_{\vartheta_j})$ is lower semicontinuous and positive on $C_j$, we have

$$a_j:=\inf_{\delta\in C_j}(P_\delta(g_j)-P_\delta(f_{\vartheta_j}))>0,\qquad j=1,\dots,m.$$

If $\vartheta\in K$ and $|\vartheta_n(\underline{x})-\vartheta|\geqq\varepsilon$, then $(\vartheta,\vartheta_n(\underline{x}))\in C$ and therefore $(\vartheta,\vartheta_n(\underline{x}))\in C_j\times U_j$ for some $j\in\{1,\dots,m\}$. This implies

$$\begin{aligned}&P_\vartheta^{\mathbb{N}}\{\underline{x}\in X^{\mathbb{N}}: |\vartheta_n(\underline{x})-\vartheta|\geqq\varepsilon\}\\ &\leqq\sum_{j=1}^m P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: n^{-1}\sum_{i=1}^n(g_j(x_i)-P_\vartheta(g_j))\leqq -a_j/2\right\}\\ &\quad+\sum_{j=1}^m P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: n^{-1}\sum_{i=1}^n(f_{\vartheta_j}(x_i)-P_\vartheta(f_{\vartheta_j}))\geqq a_j/2\right\}.\end{aligned}$$

As conditions (iii)(a) and (vii)(a) are fulfilled, the assertion now follows from Chebychev's inequality.

The reader will note that ideas related to Bahadur ([1], p. 248, Lemma 5.3) are used in the proofs of the following lemmas.

**Lemma 5.** *Let $\vartheta_n$, $n\in\mathbb{N}$, be a sequence of m.c. estimates. Assume that regularity conditions* (i)–(iv), (v)(c), (vi)(b), (vii), *and* (viii) *are fulfilled. Then for every compact subset $K\subset\Theta$ there exists $c\in\mathbb{R}$ (depending on $K$) such that*

$$P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-1}\sum_{i=1}^n(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1)\right|\geqq\tfrac{1}{2}\right\}\leqq cn^{-1}$$

*for all $\vartheta\in K$, $n\in\mathbb{N}$ and every sequence $\hat{\vartheta}_n$, $n\in\mathbb{N}$, with the following properties:*

(a) $\hat{\vartheta}_n$: $X^{\mathbb{N}} \times \Theta \to \Theta^c$ *depends on* $x_1, \ldots, x_n$ *only;*

(b) $|\hat{\vartheta}_n(\underline{x}, \vartheta) - \vartheta| \leqq |\vartheta_n(\underline{x}) - \vartheta|$ *for all* $\underline{x} \in X^{\mathbb{N}}$, $\vartheta \in \Theta$;

(c) $\underline{x} \to \sum_{i=1}^{n} f''(x_i, \hat{\vartheta}_n(\underline{x}, \vartheta))$ *is* $\mathscr{A}^{\mathbb{N}}$*-measurable.*

*Proof.* Condition (viii) implies that for every compact $K \subset \Theta$ there exist $e > 0$ and $k$: $X \to \overline{\mathbb{R}}$ (both depending on $K$) such that $\vartheta \in K$, $\vartheta' \in \Theta^c$, and $|\vartheta' - \vartheta| < e$ imply $|f''(x, \vartheta') - f''(x, \vartheta)| \leqq |\vartheta' - \vartheta| \, k(x)$ for all $x \in X$, and such that $\sup_{\vartheta \in K} P_\vartheta(k^2) < \infty$.

To see this let $V_\vartheta$ and $k_\vartheta$ be given by condition (viii). As $K$ is compact, there exist $\vartheta_1, \ldots, \vartheta_m \in K$ such that $K \subset \bigcup_{j=1}^{m} V_{\vartheta_j}$. Let $k := \max\{k_{\vartheta_j}: j = 1, \ldots, m\}$. By the uniform cover theorem (see e.g. Kelley [8], p. 199, Theorem 33) there exists $e > 0$ such that for every $\vartheta \in K$ there exists $j$ such that $|\vartheta' - \vartheta| < e$ implies $\vartheta' \in V_{\vartheta_j}$, whence

$$|f''(x, \vartheta') - f''(x, \vartheta)| \leqq |\vartheta' - \vartheta| \, k_{\vartheta_j}(x) \leqq |\vartheta' - \vartheta| \, k(x).$$

The relations $\vartheta \in K$ and $|\vartheta_n(\underline{x}) - \vartheta| < e$ therefore imply

$$\left| n^{-1} \sum_{i=1}^{n} \left(a(\vartheta) f''(x_i, \hat{\vartheta}_n(\underline{x}, \vartheta)) - 1\right) \right| \leqq a \left| n^{-1} \sum_{i=1}^{n} \left(f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right) \right| + a \, |\vartheta_n(\underline{x}) - \vartheta| \, n^{-1} \sum_{i=1}^{n} k(x_i), \tag{6}$$

where $a := \sup_{\vartheta \in K} a(\vartheta)$.

Let $k_0 := \sup_{\vartheta \in K} P_\vartheta(k^2)$. If $k_0 = 0$, the assertion is trivial. Hence we may assume $k_0 > 0$ in the following. Then

$$\vartheta \in K, \quad |\vartheta_n(\underline{x}) - \vartheta| < e' := \min\left(e, \frac{1}{8 a k_0}\right),$$

and

$$\left| n^{-1} \sum_{i=1}^{n} \left(a(\vartheta) f''(x_i, \hat{\vartheta}_n(\underline{x}, \vartheta)) - 1\right) \right| \geqq \tfrac{1}{2}$$

imply

$$\left| n^{-1} \sum_{i=1}^{n} \left(f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right) \right| \geqq \frac{1}{4a} \quad \text{or} \quad n^{-1} \sum_{i=1}^{n} k(x_i) \geqq 2 k_0.$$

The second inequality, in turn, implies

$$n^{-1} \sum_{i=1}^{n} \left(k(x_i) - P_\vartheta(k)\right) \geqq k_0.$$

Hence for $\vartheta \in K$:

$$\begin{aligned} P_\vartheta^{\mathbb{N}} &\left\{ \underline{x} \in X^{\mathbb{N}}: \left| n^{-1} \sum_{i=1}^{n} \left(a(\vartheta) \; f''(x_i, \hat{\vartheta}_n(\underline{x}, \vartheta)) - 1\right) \right| \geqq \tfrac{1}{2} \right\} \\ &\leqq P_\vartheta^{\mathbb{N}} \{ \underline{x} \in X^{\mathbb{N}}: |\vartheta_n(\underline{x}) - \vartheta| \geqq e' \} \\ &\quad + P_\vartheta^{\mathbb{N}} \left\{ \underline{x} \in X^{\mathbb{N}}: \left| n^{-1} \sum_{i=1}^{n} \left(f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right) \right| \geqq \frac{1}{4a} \right\} \\ &\quad + P_\vartheta^{\mathbb{N}} \left\{ \underline{x} \in X^{\mathbb{N}}: n^{-1} \sum_{i=1}^{n} \left(k(x_i) - P_\vartheta(k)\right) \geqq k_0 \right\}. \end{aligned}$$

Because of conditions (v)(c), (vi)(b), and (viii), Chebychev's inequality may be applied to show that the second and the third term have upper bounds of the type $cn^{-1}$. The proof can now be concluded by an application of Lemma 4.

**Lemma 6.** *Let $\vartheta_n$, $n\in\mathbb{N}$, be a sequence of m.c. estimates. Assume that regularity conditions* (i)–(iv), (v)(a), (v)(c), *and* (vi)–(viii) *are fulfilled. Then for every compact $K\subset\Theta$ there exists $c\in\mathbb{R}$ (depending on $K$) such that*

$$P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \frac{|\vartheta_n(\underline{x})-\vartheta|}{\beta(\vartheta)}\geqq 2n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\leqq cn^{-\frac{1}{2}} \quad \text{for all } \vartheta\in K,\ n\in\mathbb{N}.$$

*Proof.* Let $d$ be the distance between $K$ and $\overline{\mathbb{R}}-\Theta$. Using (3) we obtain for $\vartheta\in K$:

$$\begin{aligned}
&P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \frac{|\vartheta_n(\underline{x})-\vartheta|}{\beta(\vartheta)}\geqq 2n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\\
&\quad\leqq P_\vartheta^{\mathbb{N}}\{\underline{x}\in X^{\mathbb{N}}: |\vartheta_n(\underline{x})-\vartheta|\geqq d\}\\
&\qquad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-\frac{1}{2}}\sum_{i=1}^n u(x_i,\vartheta)\right|\geqq(\log n)^{\frac{1}{2}}\right\}\\
&\qquad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-1}\sum_{i=1}^n a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))\right|\leqq\tfrac{1}{2}\right\}.
\end{aligned}$$

We have

$$\begin{aligned}
&P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-\frac{1}{2}}\sum_{i=1}^n u(x_i,\vartheta)\right|\geqq(\log n)^{\frac{1}{2}}\right\}\\
&\quad\leqq \sup_{\tau>0}\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-\frac{1}{2}}\sum_{i=1}^n u(x_i,\vartheta)\right|\geqq\tau\right\}-2\Phi(-\tau)\right|\\
&\qquad+2\Phi(-(\log n)^{\frac{1}{2}})\leqq c'n^{-\frac{1}{2}}+c''n^{-\frac{1}{2}},
\end{aligned}$$

as

$$\sup_{\tau>0}\left|P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-\frac{1}{2}}\sum_{i=1}^n u(x_i,\vartheta)\right|\geqq\tau\right\}-2\Phi(-\tau)\right|\leqq c'n^{-\frac{1}{2}}$$

by the Berry-Esseen theorem (which is applicable by virtue of conditions (iv), (v)(a), and (vi)(a)), and as $2\Phi(-(\log n)^{\frac{1}{2}})\leqq c''n^{-\frac{1}{2}}$ by Feller ([5], p. 166, Lemma 2).

As

$$\left|n^{-1}\sum_{i=1}^n a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))\right|\leqq\tfrac{1}{2}$$

implies

$$\left|n^{-1}\sum_{i=1}^n (a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1)\right|\geqq\tfrac{1}{2},$$

the assertion follows from Lemmas 4 and 5.

**Lemma 7.** *Let $\vartheta_n$, $n\in\mathbb{N}$, be a sequence of m.c. estimates. Assume that regularity conditions* (i)–(viii) *are fulfilled. Then for every compact $K\subset\Theta$ there exist $a, c\in\mathbb{R}$*

*(both depending on $K$) such that*

$$P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-1}\sum_{i=1}^{n}\left(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1\right)\right|\geqq a n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\leqq c n^{-\frac{1}{2}}$$

*for all $\vartheta\in K$, $n\in\mathbb{N}$, and for every sequence $\hat{\vartheta}_n$, $n\in\mathbb{N}$, with the properties* (a), (b), (c) *stated in Lemma* 5.

*Proof.* The reader will note that the following proof is a suitable modification of the proof of Lemma 5.

Let

$$a:=5\max\left\{\sup_{\vartheta\in K}a(\vartheta)\left[P_\vartheta\left(\left(f''(\cdot,\vartheta)-P_\vartheta(f''(\cdot,\vartheta))\right)^2\right)\right]^{\frac{1}{2}},\ \sup_{\vartheta\in K}\beta(\vartheta)\,a(\vartheta)\,P_\vartheta(k)\right\}.$$

Starting from (6) we obtain that $\vartheta\in K$, $|\vartheta_n(\underline{x})-\vartheta|<e$, and

$$\left|n^{-1}\sum_{i=1}^{n}\left(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1\right)\right|\geqq a n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$$

imply

$$a(\vartheta)\left|n^{-1}\sum_{i=1}^{n}\left(f''(x_i,\vartheta)-P_\vartheta(f''(\cdot,\vartheta))\right)\right|\geqq\frac{a}{5}n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$$

or

$$\frac{|\vartheta_n(\underline{x})-\vartheta|}{\beta(\vartheta)}n^{-1}\sum_{i=1}^{n}m(x_i,\vartheta)\geqq\frac{4a}{5}n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}$$

with $m(x,\vartheta):=\beta(\vartheta)\,a(\vartheta)\,k(x)$.

The second inequality, in turn, implies $n^{-1}\sum_{i=1}^{n}m(x_i,\vartheta)\geqq\frac{2}{5}a$ and therefore

$$n^{-1}\sum_{i=1}^{n}\left(m(x_i,\vartheta)-P_\vartheta(m(\cdot,\vartheta))\right)\geqq\frac{a}{5},$$

or

$$\frac{|\vartheta_n(\underline{x})-\vartheta|}{\beta(\vartheta)}\geqq 2n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}.$$

Hence for $\vartheta\in K$:

$$\begin{aligned}
&P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-1}\sum_{i=1}^{n}\left(a(\vartheta)f''(x_i,\hat{\vartheta}_n(\underline{x},\vartheta))-1\right)\right|\geqq a n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\\
&\quad\leqq P_\vartheta^{\mathbb{N}}\{\underline{x}\in X^{\mathbb{N}}: |\vartheta_n(\underline{x})-\vartheta|\geqq e\}\\
&\quad\quad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: a(\vartheta)\left|n^{-1}\sum_{i=1}^{n}\left(f''(x_i,\vartheta)-P_\vartheta(f''(\cdot,\vartheta))\right)\right|\geqq\frac{a}{5}n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}\\
&\quad\quad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: n^{-1}\sum_{i=1}^{n}\left(m(x_i,\vartheta)-P_\vartheta(m(\cdot,\vartheta))\right)\geqq\frac{a}{5}\right\}\\
&\quad\quad+P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \frac{|\vartheta_n(\underline{x})-\vartheta|}{\beta(\vartheta)}\geqq 2n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}.
\end{aligned}$$

As conditions (v) (b) and (vi) (b) are fulfilled, the Berry-Esseen theorem applies and yields the existence of a number $c'$ such that

$$P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: a(\vartheta)\left|n^{-1}\sum_{i=1}^{n}(f''(x_i,\vartheta)-P_\vartheta(f''(\cdot,\vartheta)))\right|\geqq\frac{a}{5}n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\right\}$$

$$\leqq P_\vartheta^{\mathbb{N}}\left\{\underline{x}\in X^{\mathbb{N}}: \left|n^{-\frac{1}{2}}\frac{\sum_{i=1}^{n}(f''(x_i,\vartheta)-P_\vartheta(f''(\cdot,\vartheta)))}{[P_\vartheta((f''(\cdot,\vartheta)-P_\vartheta(f''(\cdot,\vartheta)))^2)]^{\frac{1}{2}}}\right|\geqq(\log n)^{\frac{1}{2}}\right\}$$

$$\leqq c'n^{-\frac{1}{2}}+2\Phi\left(-(\log n)^{\frac{1}{2}}\right)\leqq cn^{-\frac{1}{2}}.$$

This inequality together with Lemmas 4 and 6 implies the assertion.

## References

1. Bahadur, R. R.: On the asymptotic efficiency of tests and estimates. Sankhyā **22**, 229–252 (1960).
2. – A note on Fisher's bound for asymptotic variance. Ann. math. Statistics **35**, 1545–1552 (1964).
3. Brillinger, D.: A note on the rate of convergence of a mean. Biometrika **49**, 574–576 (1962).
4. Esseen, C. G.: A moment inequality with an application to the central limit theorem. Skand. Aktuarietidskr. **39**, 160–170 (1956).
5. Feller, W.: An introduction to probability theory and its applications, Vol. I. New York: Wiley 1957.
6. – An introduction to probability theory and its applications, Vol. II. New York: Wiley 1966.
7. Kaufmann, S.: Asymptotic efficiency of the maximum likelihood estimator. Ann. Inst. Statist. Math. **18**, 155–178 (1966).
8. Kelley, J. L.: General topology. New York: Van Nostrand 1955.
9. Kiefer, J. C.: Statistical inference. In: The future of statistics (ed. by D. G. Watts), 139–142. New York: Academic Press 1968.
10. Linnik, Yu. V., Mitrofanova, N. M.: On the asymptotic distribution of maximum likelihood estimates (in Russian). Doklady Akademie Nauk SSSR **149**, 518–520 (1963).
11. – Some asymptotic expansions for the distribution of the maximum likelihood estimate. Sankhyā **27** A, 73–82 (1965).
12. Pfanzagl, J.: On the measurability and consistency of minimum contrast estimates. Metrika **14**, 249–272 (1969).
13. Schmetterer, L.: On the asymptotic efficiency of estimates. In: Research papers in statistics; Festschrift for J. Neyman (ed. by F. N. David), 301–317. London: Wiley 1966.
14. Wald, A.: Note on the consistency of the maximum likelihood estimator. Ann. math. Statistics **20**, 595–601 (1949).

Dipl.-Math. R. Michel and Prof. Dr. J. Pfanzagl
Mathematisches Institut der Universität zu Köln
BRD-5000 Köln-Lindenthal, Weyertal 90
Germany

*(Received December 23, 1969)*