Z. Wahrscheinlichkeitstheorie verw. Gebiete
35, 269–282 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Asymptotic Properties of Posterior Distributions

Helmut Strasser

Mathematisches Institut, Universität Gießen, Arndtstraße 2, D-6300 Gießen,
Federal Republic of Germany

Let $(X, \mathscr{A})$ be a measurable space, $\Theta \subseteq \mathbb{R}$ an open interval and $P_\vartheta | \mathscr{A}$, $\vartheta \in \Theta$, a family of probability measures fulfilling certain regularity conditions. Let λ be a prior distribution on Θ and $R_{n,\underline{x}}$ be the posterior distribution for the sample size n given $\underline{x} \in X^n$. ϑ_n denotes the maximum likelihood estimate for the sample size n. It is shown that under certain regularity conditions on λ for every compact $K \subseteq \Theta$ there exists $c_K \geqq 0$ such that

$$\sup_{\vartheta \in K} P_\vartheta^n \{\underline{x} \in X^n | d(R_{n,\underline{x}}, Q_{n,\underline{x}}^\vartheta) \geqq c_K (\log n)^{1/2} n^{-1/2}\} = o(n^{-1/2}),$$

where $Q_{n,\underline{x}}^\vartheta$ is the Borel measure on Θ having Lebesgue density

$$\sqrt{\frac{n}{2\pi a(\vartheta)}} \exp\left[-\frac{n}{2} \frac{(\sigma - \vartheta_n(\underline{x}))^2}{a(\vartheta)}\right]$$

and $d(R_{n,\underline{x}}, Q_{n,\underline{x}}^\vartheta)$ denotes the variational distance between $R_{n,\underline{x}}$ and $Q_{n,\underline{x}}^\vartheta$.

1. Introduction

Let $(X, \mathscr{A})$ be a measurable space and $P_\vartheta | \mathscr{A}$, $\vartheta \in \Theta$, a family of probability measures. Θ denotes a parameter set. Let $\mathscr{B}$ be a σ-algebra on Θ and let λ be a prior distribution on $(\Theta, \mathscr{B})$. For every $n \in \mathbb{N}$ we define a probability measure R_n on $(X^n \times \Theta, \mathscr{A}^n \otimes \mathscr{B})$ by

$$R_n(A \times \Sigma) = \int_\Sigma P_\vartheta^n(A)\, \lambda(d\vartheta), \qquad A \in \mathscr{A}^n,\ \Sigma \in \mathscr{B},$$

and $\underline{x} \mapsto R_{n,\underline{x}}$, $\underline{x} \in X^n$, denotes a version of the conditional probability of R_n under the hypothesis $\mathscr{A}^n$. $R_{n,\underline{x}} | \mathscr{B}$ is called the posterior distribution given $\underline{x} \in X^n$.

Asymptotic properties of posterior distributions are closely related with the asymptotic behaviour of Bayes estimates. In 1949 Doob, [1], proved a general result concerning consistency of Bayes estimates of a real parameter relative to quadratic loss. His result was generalized in 1965 by Lorraine Schwartz, [11], and in 1973 by the author, [12] and [13]. Roughly speaking, for every prior distribution λ on a parameter set Θ any sequence of Bayes estimates is strongly consistent λ-almost everywhere, iff there exists a λ-almost exact estimate. The existence of a λ-almost exact estimate can be interpreted as the possibility to separate the values of the parameter by infinitely many observations. It was also shown by Lorraine Schwartz, [11], that for those points $\vartheta \in \Theta$, which can be separated from the complement U' of any of their open neighbourhoods U by a uniformly consistent test, the posterior probabilities $R_{n,\underline{x}}(U')$ converge to zero $P_\vartheta^{\mathbb{N}}$-almost everywhere. In a recent paper LeCam, [4], 1973, showed that under dimensional restrictions on Θ a similar result is true for a sequence (U_n) of neighbourhoods which decreases of the order $n^{-1/2}$ relative to the Hellinger distance. These results show that even in general cases for large sample size the mass of the posterior distribution concentrates in arbitrary small neighbourhoods of the true parameter value. Already in 1953 it was shown by LeCam, [3], that in sufficiently regular cases the variational distance between the posterior distribution and the normal distribution centered in $\vartheta_n(\underline{x})$ with variance $a(\vartheta)/n$ converges to zero $P_\vartheta^{\mathbb{N}}$-a.e. for every $\vartheta \in \Theta$, (cf. also Schmetterer, [10], pp. 391 ff.) It is one of the aims of the present paper to estimate the speed of convergence of the posterior distribution to the normal distribution. Previous results of this kind were given by van der Waerden, [14], and Johnson, [2], (cf. Remark 3).

2. Discussion of the Results

The results are stated in the framework of minimum contrast estimation. For the motivation of this approach confer Pfanzagl, [7].

Θ is assumed to be an open interval of $\mathbb{R}$ and $\mathscr{B}$ denotes the Borel σ-algebra of Θ. A family of $\mathscr{A}$-measurable functions $f_\vartheta \colon X \mapsto \overline{\mathbb{R}}$, $\vartheta \in \overline{\Theta}$, (sometimes denoted by $x \mapsto f(x, \vartheta)$), is a family of contrast functions for $\{P_\vartheta \mid \vartheta \in \Theta\}$ if $P_\vartheta(f_\tau)$ exists for all $\vartheta \in \Theta$, $\tau \in \overline{\Theta}$, and if

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

A minimum contrast estimate for the sample size n is an $\mathscr{A}^n$-measurable function $\vartheta_n \colon X^n \to \overline{\Theta}$ satisfying

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

For those $\underline{x} \in X^n$ for which it is possible we define the probability measure

$$R_{n,\underline{x}}(\Sigma) = \frac{\int_\Sigma \exp\left(-\sum_{i=1}^n f_\sigma(x_i)\right) \lambda(d\sigma)}{\int_\Theta \exp\left(-\sum_{i=1}^n f_\sigma(x_i)\right) \lambda(d\sigma)}, \qquad \Sigma \in \mathscr{B}.$$

Lemma 1 gives conditions under which for every compact $K \subseteq \Theta$

$$\sup_{\vartheta \in K} P_\vartheta^n \left\{ \underline{x} \in X^n \,\middle|\, \sup_{\sigma \in \Theta} \exp\left(-\sum_{i=1}^n f(x_i, \sigma)\right) = \infty \right\} = O(n^{-1}).$$

Since in the following these conditions are always imposed we may restrict our attention to those $\underline{x} \in X^n$ for which $R_{n,\underline{x}}$ can be defined.

In the following we make assertions on the limit behaviour of sequences of events. Let $E_n(\underline{x}, \vartheta, K)$ be a statement depending on $n \in \mathbb{N}$, $\underline{x} \in X^n$, $\vartheta \in \Theta$ and $K \subseteq \Theta$. For notational convenience the assertion: "For every compact $K \subseteq \Theta$ there exists $c_K \geqq 0$ and $\alpha \in \mathbb{R}$ such that

$$\sup_{\vartheta \in K} P_\vartheta^n \{\underline{x} \in X^n \mid E_n(\underline{x}, \vartheta, K)\} \leqq c_K\, n^{-\alpha} \qquad \text{for all } n \in \mathbb{N}\text{"}$$

is abbreviated by

$$E_n(\underline{x}, \vartheta, K) \sim O_K(n^{-\alpha}).$$

If c_K can be replaced by a sequence $c_K^{(n)} \geqq 0$, $n \in \mathbb{N}$, satisfying $\lim c_K^{(n)} = 0$, we write

$$E_n(\underline{x}, \vartheta, K) \sim o_K(n^{-\alpha}).$$

The regularity conditions cited below are listed in the next paragraph which also contains several auxiliary results. In order to improve the readibility of the paper the proofs of the theorems are collected at the end of the paper.

The first result estimates the speed of convergence in Theorem 6.1 of Lorraine Schwartz, [11]. For every $\vartheta \in \Theta$ define $B_\delta(\vartheta)$ to be the open ball of radius δ around ϑ.

Theorem 1. *Assume that regularity conditions* (i), (ii), (iii)$'$, (vii) *and* (j) *are satisfied. Then for every $\delta > 0$ and every compact $K \subseteq \Theta$ there exists $\eta_K > 0$ such that*

$$R_{n,\underline{x}}(B_\delta(\vartheta)') > \exp(-\eta_K\, n) \sim O_K(n^{-1}).$$

The next theorem shows that a result similar to Theorem 1 is true when the radius of the neighbourhoods decreases of the order $(\log n)^{1/2}\, n^{-1/2}$. It is however necessary to replace the true parameter value ϑ by the sequence of minimum contrast estimates. Let

$$W_n^\vartheta(\underline{x}, s) = \{\sigma \in \Theta \mid n^{1/2} |\sigma - \vartheta(x)| / a(\vartheta)^{1/2} \leqq (s \log n)^{1/2}\},$$

if $s > 0$, $\vartheta \in \Theta$, $\underline{x} \in X^n$, $n \in \mathbb{N}$.

Theorem 2. *Assume that regularity conditions* (i)–(iv), (v)(b), (vi)–(viii) *and* (jj) *are satisfied. Then for every $r > 0$ and every compact $K \subseteq \Theta$ there exist $s_K \geqq 0$ and $c_K \geqq 0$*

$$R_{n,\underline{x}}(W_n^\vartheta(\underline{x}, s_K)') \geqq c_K\, n^{-r} \sim O_K(n^{-1}).$$

Remark 1. If condition (vi) is replaced by (vi)$'$ and if (jj) holds with $t = 1/2$ for every compact $K \subseteq \Theta$ then the assertion of Theorem 2 holds with $s_K = 1 + 2r$ and $O_K(n^{-1})$ replaced by $o_K(n^{-1/2})$. This can be proved with the aid of Lemma 5.

Remark 2. If condition (vi) is replaced by (vi)$'$ then the assertion of Theorem 2 even holds with $\vartheta_n(\underline{x})$ replaced by the true parameter ϑ and $O_K(n^{-1})$ by $o_K(n^{-1/2})$. This can be proved with the aid of Lemma 3 of Pfanzagl, [9].

Define

$$\tilde{Q}_{n,\underline{x}}^\vartheta(\Sigma) = \frac{\int_\Sigma \exp[-n(\sigma - \vartheta_n(\underline{x}))^2 / 2a(\vartheta)]\, \lambda(d\sigma)}{\int_\Theta \exp[-n(\sigma - \vartheta_n(\underline{x}))^2 / 2a(\vartheta)]\, \lambda(d\sigma)},$$

$\Sigma\in\mathscr{B}$, $\underline{x}\in X^n$, $n\in\mathbb{N}$, $\vartheta\in\Theta\cdot d(R_{n,\underline{x}},\tilde{Q}^{\vartheta}_{n,\underline{x}})$ denotes the variational distance of $R_{n,\underline{x}}$ and $\tilde{Q}^{\vartheta}_{n,\underline{x}}$.

Theorem 3. *Assume that regularity conditions* (i)–(v), (vi)′, (vii), (viii) *and* (jj) *are satisfied. Then for every compact $K\subseteq\Theta$ there exists $c_K\geqq 0$ such that*

$$d(R_{n,\underline{x}},\tilde{Q}^{\vartheta}_{n,\underline{x}})\geqq c_K(\log n)^{3/2}\,n^{-1/2}\sim o_K(n^{-1/2}).$$

If stronger regularity conditions on λ are imposed the assertion of Theorem 3 can be improved considerably. Let $Q^{\vartheta}_{n,\underline{x}}$ be the measure on Θ having Lebesgue density

$$\sqrt{n/2\pi\,a(\vartheta)}\exp\left[-n(\sigma-\vartheta_n(\underline{x}))^2/2a(\vartheta)\right],\qquad \sigma\in\Theta.$$

Theorem 4. *Assume that regularity conditions* (i)–(v), (vi)′, (vii), (viii) *and* (jjj) *are satisfied. Then for every compact $K\subseteq\Theta$ there exists $c_K\geqq 0$ such that*

$$d(R_{n,\underline{x}},Q^{\vartheta}_{n,\underline{x}})\geqq c_K(\log n)^{1/2}\,n^{-1/2}\sim o_K(n^{-1/2}).$$

In Theorems 1 and 2 conditions on second moments and in Theorems 3 and 4 conditions on third moments are involved. The respective estimates of speed of convergence are $O_K(n^{-1})$ and $o_K(n^{-1/2})$. Speed of convergence of higher order can be achieved by modifying the regularity conditions (towards existence of higher moments) and applying Lemmas 1 and 2 of Pfanzagl, [9].

Remark 3. Theorem 4 is related to a result of Johnson, [2]. Assume that $P_\vartheta\ll\mu$ for every $\vartheta\in\Theta$, where μ is a σ-finite measure on $\mathscr{A}$, and let $h_\vartheta\in dP_\vartheta/d\mu$, $\vartheta\in\Theta$. Let $F^{\vartheta}_{n,\underline{x}}$ be the distribution function of the measure induced by $R_{n,\underline{x}}$ and

$$T^{\vartheta}_{n,\underline{x}}:\sigma\mapsto n^{1/2}(\sigma-\vartheta_n(\underline{x}))\,b_{n,\underline{x}}(\vartheta_n(\underline{x})),\qquad \sigma\in\mathbb{R},$$

where

$$b_{n,\underline{x}}(\vartheta)=\left[-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\tau^2}\log h_\tau(x_i)\bigg|_{\tau=\vartheta}\right]^{1/2},\qquad \vartheta\in\Theta.$$

Then under certain regularity conditions (involving fourth continuous partial derivatives of $\log h_\vartheta(x)$ with respect to ϑ and second derivatives of the density of λ) Theorem 2 in [2] implies that

$$\limsup_{n\in\mathbb{N}}\sup_{t\in\mathbb{R}}|n^{1/2}(F^{\vartheta}_{n,\underline{x}}(t)-\Phi(t))-\varphi(t)(a_{\underline{x}}t^2+b_{\underline{x}})|=0\quad P^{\mathbb{N}}_{\vartheta}\text{-a.e.}$$

for every $\vartheta\in\Theta$. The constants $a_{\underline{x}}$ and $b_{\underline{x}}$ depend on $\underline{x}\in X^n$ through $\vartheta_n(\underline{x})$. If in the definition of $F^{\vartheta}_{n,\underline{x}}$ the term $b_{n,\underline{x}}(\vartheta_n(\underline{x}))$ is replaced by $a(\vartheta)^{-1/2}$, where

$$a(\vartheta)=-P_\vartheta\left[\frac{\partial^2}{\partial\tau^2}\log h_\tau\bigg|_{\tau=\vartheta}\right]^{-1},$$

then Johnson's result leads to the conjecture that for the so modified distributions $F^{\vartheta}_{n,\underline{x}}$ the $P^{\mathbb{N}}_{\vartheta}$-probability of deviations

$$\sup_{t\in\mathbb{R}} n^{1/2}\,|F^{\vartheta}_{n,\underline{x}}(t)-\Phi(t)|\geqq c(\log n)^{1/2}$$

converges to zero for every $\vartheta\in\Theta$ (Φ denotes the distribution function of the standard normal distribution). Theorem 4 proves this assertion even for the variational distance of $F^{\vartheta}_{n,x}$ and Φ, and establishes an estimate for the speed of convergence. The imposed regularity conditions do neither involve differentiability assumptions on a density of λ nor higher partial derivatives of $\log h_{\vartheta}$ than of second order.

3. Auxiliary Results

Regularity conditions. (i) $\vartheta \mapsto P_{\vartheta}$ is continuous in Θ with respect to the supremum metric.

(ii) For each $x \in X$, $\vartheta \mapsto f_{\vartheta}(x)$ is continuous on $\bar{\Theta}$.

(iii) For every $\vartheta \in \Theta$ there exists a neighbourhood W_{ϑ} of ϑ such that $\sup_{\tau \in W_{\vartheta}} P_{\tau}(f_{\vartheta}^2) < \infty$.

(iv) For each $x \in X$, $\vartheta \mapsto f_{\vartheta}(x)$ is twice differentiable in Θ. 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 $\vartheta \in \Theta$: $P_{\vartheta}(f'(., \vartheta)) = 0$.

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

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

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

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

(a) $\sup_{\vartheta \in K} P_{\vartheta}((f'(., \vartheta))^2) < \infty$,

(b) $\sup_{\vartheta \in K} P_{\vartheta}((f''(., \vartheta))^2) < \infty$.

(vii) For every $\vartheta \in \bar{\Theta}$ there exists a neighbourhood U_{ϑ} of ϑ such that for every neighbourhood U of ϑ, $U \subseteq U_{\vartheta}$, and every compact $K \subseteq \Theta$, $\sup_{\tau \in K} P_{\tau}((\inf_{\sigma \in U} f_{\sigma})^2) < \infty$.

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

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

Conditions for the prior distribution.

(j) For every $\delta > 0$ and every compact $K \subseteq \Theta$

$$\inf_{\vartheta \in K} \lambda \{\sigma \in \Theta \,|\, |\sigma - \vartheta| < \delta\} > 0.$$

(jj) For every compact $K \subseteq \Theta$ there exists $t \geqq 0$ (depending on K) such that

$$\liminf_{n \in \mathbb{N}} \inf_{\vartheta \in K} n^t \lambda \{\sigma \in \Theta \,|\, |\sigma - \vartheta| < n^{-1/2}\} > 0.$$

(jjj) λ has a continuous, positive density p on Θ with respect to the Lebesgue measure satisfying the following condition: For every compact $K \subseteq \Theta$ there exist constants $d_K > 0$, $c_K > 0$, such that $\sigma \in \Theta$, $\vartheta \in K$, and $|\sigma - \vartheta| \leqq d_K$ imply

$$\left| \frac{p(\sigma)}{p(\vartheta)} - 1 \right| \leqq c_K \, |\sigma - \vartheta|.$$

Obviously we have (jjj) $\Rightarrow$ (jj) $\Rightarrow$ (j).

Remark 4. Simple Taylor expansion arguments show that conditions (iii), (iv), (vi) and (viii) imply the following condition:

(iii)′ For every $\vartheta \in \Theta$ there exists a neighbourhood W_ϑ of ϑ such that

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

In order to avoid differentiability assumptions where it is not necessary we sometimes propose condition (iii)′ instead of conditions (iii), (iv), (vi) and (viii).

Remark 5. In Theorems 3 and 4 it is necessary to replace condition (vi) by

(vi)′ For every compact $K \subseteq \Theta$

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

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

The proof of the following Lemma 1 is strongly inspired by the proof of Lemma 4 in Michel and Pfanzagl, [5]. Conversely, Lemma 4 in Michel and Pfanzagl, [5], is an easy consequence of Lemma 1.

Lemma 1. *Let conditions* (i)–(iii) *and* (vii) *be satisfied. Then for every* $\delta > 0$ *and every compact* $K \subseteq \Theta$ *there exists* $\varepsilon_K > 0$ *such that*

$$\inf_{|\sigma - \vartheta| \geqq \delta} \frac{1}{n} \sum_{i=1}^n f(x_i, \sigma) \leqq P_\vartheta(f_\vartheta) + \varepsilon_K \sim O_K(n^{-1})$$

Proof. Let $C = \{(\vartheta, \tau) \in K \times \overline{\Theta} \,|\, |\vartheta - \tau| \geqq \delta\}$. Obviously, $(\vartheta, \tau) \in C$ implies $P_\vartheta(f_\vartheta) < P_\vartheta(f_\tau)$. Because of conditions (ii) and (vii)(b), Lemma 3 of Michel and Pfanzagl, [5], yields the existence of an open neighbourhood $U_{(\vartheta,\tau)} \subseteq U_\tau$ of τ such that

$$P_\vartheta(f_\vartheta) < P_\vartheta(\inf f(., U_{(\vartheta,\tau)})).$$

Because of conditions (i), (iii) and (vii), Lemma 2 of Michel and Pfanzagl, [5], implies that $\gamma \mapsto P_\gamma(f_\vartheta)$ is continuous and $\gamma \mapsto P_\gamma\,(\inf f(., U_{(\vartheta,\tau)}))$ is lower semicontinuous. Hence there exists a compact neighbourhood $C_{(\vartheta,\tau)} \subseteq \Theta$ of ϑ such that

$$P_\gamma(f_\vartheta) < P_\gamma(\inf f(., U_{(\vartheta,\tau)})) \quad \text{if } \gamma \in C_{(\vartheta,\tau)}.$$

Since $\{\mathring{C}_{(\vartheta,\tau)} \times U_{(\vartheta,\tau)} \,|\, (\vartheta, \tau) \in C\}$ is an open cover of C, there exists a finite subcover determined by $(\vartheta_j, \tau_j) \in C$, $j = 1, \ldots, m$. Let $C_j = C_{(\vartheta_j,\tau_j)}$ and $U_j = U_{(\vartheta_j,\tau_j)}$. Then $\sigma \in U_j$ implies

$$f(x, \sigma) \geqq \inf f(x, U_j), \qquad x \in X,$$

and $\vartheta \in C_j$ implies

$$P_\vartheta(f_\vartheta) < P_\vartheta(f_{\vartheta_j}) < P_\vartheta(\inf f(., U_j)).$$

For every $(\vartheta, \sigma) \in C$ there exists $j \in \{1, \ldots, m\}$ such that $\vartheta \in C_j$ and $\sigma \in U_j$. This implies

$$\frac{1}{n} \sum_{i=1}^n \inf f(x_i, U_j) - P_\vartheta(f_{\vartheta_j}) \leqq \frac{1}{n} \sum_{i=1}^n f(x_i, \sigma) - P_\vartheta(f_\vartheta).$$

Since $\gamma \mapsto P_\gamma(\inf f(.,U_j)) - P_\gamma(f_{\vartheta_j})$ is lower semicontinuous and positive on C_j, it follows that

$$a_j = \inf_{\gamma \in C_j} (P_\gamma(\inf f(.,U_j)) - P_\gamma(f_{\vartheta_j})) > 0, \qquad j = 1, \ldots, m.$$

Hence we obtain that

$$\inf_{|\sigma - \vartheta| \geqq \delta} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) - P_\vartheta(f_\vartheta) \leqq \varepsilon_K$$

implies

$$\frac{1}{n} \sum_{i=1}^{n} \inf f(x_i, U_j) - P_\vartheta(\inf f(.,U_j)) \leqq \varepsilon_K - a_j$$

for at least one $j \in \{1, \ldots, m\}$. Choosing $\varepsilon_K < \min_{1 \leqq j \leqq m} a_j$ and applying Čebyšev's inequality proves the assertion. □

Lemma 2. *Let conditions* (i), (ii) *and* (iii)′ *be satisfied. Then for every* $\varepsilon > 0$ *and every compact* $K \subseteq \Theta$ *there exists* $\delta_K > 0$ *such that*

$$\sup_{|\sigma - \vartheta| < \delta_K} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) \geqq P_\vartheta(f_\vartheta) + \varepsilon \sim O_K(n^{-1}).$$

Proof. Conditions (ii) and (iii)′ imply by Lemma 3 of Michel and Pfanzagl, [5], that for every $\vartheta \in K$ there exists an open neighbourhood V_ϑ of ϑ, $V_\vartheta \subseteq W_\vartheta$, such that

$$P_\vartheta(\sup f(.,V_\vartheta)) < P_\vartheta(f_\vartheta) + \frac{\varepsilon}{2}.$$

Lemma 2 of Michel and Pfanzagl, [5], and condition (iii)′ imply that $\sigma \mapsto P_\sigma(\sup f(.,V_\vartheta))$ is continuous on W_ϑ. Lemma 5 of Michel and Pfanzagl, [6], and condition (iii)′ imply that $\sigma \mapsto P_\sigma(f_\sigma)$ is lower semicontinuous. Therefore

$$O_\vartheta = \{\sigma \in V_\vartheta \mid P_\sigma(\sup f(.,V_\vartheta)) < P_\sigma(f_\sigma) + \varepsilon/2\}$$

is an open neighbourhood of ϑ. Since K is compact there exist $\vartheta_1, \ldots, \vartheta_m$ such that the sets $O_j = O_{\vartheta_j}$, $1 \leqq j \leqq m$, cover K. Let $\delta_K > 0$ be such that for every $\vartheta \in K$ the δ_K-neighbourhood of ϑ is contained in at least one O_j, $1 \leqq j \leqq m$. Then for every $\vartheta \in K$

$$\sup_{|\sigma - \vartheta| < \delta_K} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) \geqq P_\vartheta(f_\vartheta) + \varepsilon$$

implies

$$\frac{1}{n} \sum_{i=1}^{n} \sup f(x_i, O_j) \geqq P_\vartheta(\sup f(.,O_j)) + \varepsilon/2$$

for a certain $j \in \{1, \ldots, m\}$. Hence for every $\vartheta \in K$ there exists O_j, $1 \leqq j \leqq m$, such that

$$P_\vartheta^n \left\{ \underline{x} \in X^n \,\middle|\, \sup_{|\sigma - \vartheta| < \delta_K} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) \geqq P_\vartheta(f_\vartheta) + \varepsilon \right\}$$

$$\leqq \sup_{\tau \in O_j} P_\tau^n \left\{ \underline{x} \in X^n \,\middle|\, \frac{1}{n} \sum_{i=1}^{n} \sup f(x_i, O_j) \geqq P_\tau(\sup f(.,O_j)) + \varepsilon/2 \right\}. \quad \square$$

Lemma 3. *Conditions* (vi)(b) *and* (viii) *imply that for every compact* $K \subseteq \Theta$ *there exist constants* $b_K > 0$, $d_K > 0$, *and for every* $\vartheta \in K$, $n \in \mathbb{N}$ *a set* $A_{n,\vartheta K} \in \mathscr{A}^n$ *such that*

(a) $\sup_{\vartheta \in K} P_\vartheta^{\mathbb{N}}(X^{\mathbb{N}} \setminus A_{n,\vartheta,K}) = O(n^{-1})$,

(b) $\underline{x} \in A_{n,\vartheta,K}$, $\tau \in \Theta$ and $|\tau - \vartheta| \leq d_K$ *imply*

$$\left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \tau) - \frac{1}{n}\sum_{i=1}^{n} f''(x_i, \vartheta)\right| \leq |\tau - \vartheta|\, b_K.$$

Proof. Confer Lemma 5 of Pfanzagl, [9]. □

Lemma 4. *Let conditions* (iv), (vi)(b) *and* (viii) *be satisfied. Then for every* $\varepsilon > 0$ *and every compact* $K \subseteq \Theta$ *there exists a* $\delta_K > 0$ *such that*

$$\sup_{|\sigma - \vartheta| \leq \delta_K} \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \sigma) - P_\vartheta(f''(\cdot, \vartheta))\right| > \varepsilon \sim O_K(n^{-1}).$$

Proof. Choose b_K and d_K according to Lemma 3. For every $\underline{x} \in A_{n,\vartheta,K}$, $\delta > 0$, $\delta < d_K$, we have

$$\sup_{|\sigma - \vartheta| \leq \delta} \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \sigma) - P_\vartheta(f''(\cdot, \vartheta))\right| \leq \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right| + \delta b_K.$$

According to Čebyšev's inequality condition (vi)(b) implies

$$\left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right| \geq \frac{\varepsilon}{2} \sim O_K(n^{-1}).$$

Choosing $\delta \leq \frac{\varepsilon}{2 b_K}$ proves the assertion. □

Lemma 5. *Let conditions* (i)–(v), (vi)′, (vii) *and* (viii) *be satisfied. Then for every compact* $K \subseteq \Theta$ *and every* $s > 0$ *there exists* $c_K \geq 0$ *such that*

$$\sup_{\sigma \in W_n^\vartheta(\underline{x}, s)} \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \sigma) - P_\vartheta(f''(\cdot, \vartheta))\right| \geq c_K (\log n)^{1/2} n^{-1/2} \sim o_K(n^{-1/2})$$

Proof. Choose d_K, b_K according to Lemma 3. Then

$$\sup_{\sigma \in W_n^\vartheta(\underline{x}, s)} \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \sigma) - P_\vartheta(f''(\cdot, \vartheta))\right| \leq \left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right|$$
$$+ b_K(|\vartheta - \vartheta_n(\underline{x})| + (s\, a(\vartheta))^{1/2} (\log n)^{1/2} n^{-1/2}).$$

Lemma 1 of Pfanzagl, [9], implies that there exists $c_K \geq 0$ such that

$$\left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i, \vartheta) - P_\vartheta(f''(\cdot, \vartheta))\right| \geq \frac{c_K}{2} (\log n)^{1/2} n^{-1/2} \sim o_K(n^{-1/2}).$$

Lemma 3 of Pfanzagl, [9], implies that for sufficiently large $c_K \geq 0$

$$|\vartheta_n(\underline{x}) - \vartheta| \geq \left(\frac{c_K}{2 b_K} - (s\, a(\vartheta))^{1/2}\right) (\log n)^{1/2} n^{-1/2} \sim o_K(n^{-1/2}). \quad \square$$

Lemma 6. *Let conditions* (i)–(iii) *and* (vii) *be satisfied. For a Borel set* $\Sigma \subseteq \mathbb{R}$ *define*

$$V_n^\vartheta(\underline{x}, \Sigma) := \{\sigma \in \Theta \mid n^{1/2}(\sigma - \vartheta_n(\underline{x}))/a(\vartheta)^{1/2} \in \Sigma\}, \quad \underline{x} \in X^n, \ n \in \mathbb{N}.$$

Then for every $s > 0$, *every sequence* $(\gamma_n) \subseteq \mathbb{R}$, $|\gamma_n| \leqq 1/2$, $n \in \mathbb{N}$, *and for every compact* $K \subseteq \Theta$ *there exists* $c_K \geqq 0$ *such that*

$$\sup_{\Sigma \in \mathscr{B}} \Big| \int_{V_n^\vartheta(\underline{x}, \Sigma) \cap W_n^\vartheta(\underline{x}, s)} \exp[-n(1+\gamma_n)(\sigma - \vartheta_n(\underline{x}))^2 / 2a(\vartheta)]\, d\sigma$$
$$- a(\vartheta)^{1/2} n^{-1/2} \int_\Sigma \exp(-\tau^2/2)\, d\tau \Big| \geqq c_K |\gamma_n| n^{-1/2} \sim O_K(n^{-1}).$$

Proof. Let $\rho_K > 0$ such that $\{\sigma \in \mathbb{R} \mid \operatorname{dist}(\sigma, K) \leqq \rho_K\}$ is a subset of Θ. There exists n_K such that $n \geqq n_K$ implies for every $\vartheta \in K$

$$(s\, a(\vartheta))^{1/2} (\log n)^{1/2} n^{-1/2} \leqq \rho_K / 2.$$

Since

$$|\vartheta_n(\underline{x}) - \vartheta| \geqq \rho_K/2 \sim O_K(n^{-1})$$

we may restrict our attention to those $\underline{x} \in X^n$ for which $n \geqq n_K$ implies

$$\{\sigma \in \mathbb{R} \mid -(s \log n)^{1/2} \leqq n^{1/2}(\sigma - \vartheta_n(\underline{x}))/a(\vartheta)^{1/2} \leqq (s \log n)^{1/2}\} \subseteq \Theta.$$

Let $n \geqq n_K$. A simple substitution yields

$$\int_{V_n^\vartheta(\underline{x}, \Sigma) \cap W_n^\vartheta(\underline{x}, s)} \exp[-n(1+\gamma_n)(\sigma - \vartheta_n(\underline{x}))^2 / 2a(\vartheta)]\, d\sigma$$
$$= a(\vartheta)^{1/2} n^{-1/2} \int_{\Sigma \cap \{t^2 \leqq s \log n\}} \exp(-\tau^2(1+\gamma_n)/2)\, d\tau$$

for every $\Sigma \in \mathscr{B}$. Thus the proof is finished since for every $\gamma \in \mathbb{R}$, $|\gamma| \leqq 1/2$, there exists $c > 0$ such that

$$\sup_{\Sigma \in \mathscr{B}} \Big| \int_{\Sigma \cap \{t^2 \leqq s \log n\}} \exp[-\tau^2(1+\gamma)/2]\, d\tau - \int_\Sigma \exp[-\tau^2/2]\, d\tau \Big| \leqq c|\gamma|. \quad \square$$

4. Proofs

Proof of Theorem 1. Let $\delta > 0$ and choose $\varepsilon_K > 0$ according to Lemma 1. Let $\eta_K < \varepsilon_K$, $\eta_K > 0$. It is easy to see that for every $\delta_1 > 0$

$$a = \inf_{\vartheta \in K} (\log \lambda\{\sigma \in \Theta \,|\, |\sigma - \vartheta| < \delta_1\} - \log \lambda(\Theta)) > -\infty.$$

The inequality

$$R_{n,\underline{x}}\{\sigma \in \Theta \,|\, |\sigma - \vartheta| \geqq \delta\} > \exp(-\eta_K n)$$

is equivalent with

$$\frac{1}{n} \log \int_{|\sigma - \vartheta| \geqq \delta} \exp\left[-\sum_{i=1}^n f(x_i, \sigma)\right] \lambda(d\sigma)$$
$$-\frac{1}{n} \log \int_\Theta \exp\left[-\sum_{i=1}^n f(x_i, \sigma)\right] \lambda(d\sigma) > -\eta_K$$

which implies for arbitrary $\delta_1 > 0$ and $\vartheta \in K$ that

$$\sup_{|\sigma-\vartheta|<\delta_1} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) - \inf_{|\sigma-\vartheta| \geqq \delta} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) > -\eta_K + \frac{a}{n}.$$

From Lemma 1 we obtain

$$-P_\vartheta(f_\vartheta) - \varepsilon_K \leqq - \inf_{|\sigma-\vartheta| \geqq \delta} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) \sim O_K(n^{-1}).$$

Therefore we have only to show that

$$\sup_{|\sigma-\vartheta|<\delta_1} \frac{1}{n} \sum_{i=1}^{n} f(x_i, \sigma) - P_\vartheta(f_\vartheta) > \varepsilon_K - \eta_K + \frac{a}{n} \sim O_K(n^{-1}).$$

But this assertion follows from Lemma 2 choosing δ_1 sufficiently small. □

Lemma 4 of Michel and Pfanzagl, [5], implies that under regularity conditions (i)–(iii) and (vii)

$$\vartheta_n(\underline{x}) \notin \Theta \sim O_K(n^{-1}).$$

Hence we may assume $\vartheta_n(\underline{x}) \in \Theta$. In the following we use repeatedly a Taylor expansion argument which runs as follows: For every $\sigma \in \Theta$

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

where $|\vartheta_n(\underline{x}) - \hat{\vartheta}_n(\underline{x}, \sigma)| \leqq |\vartheta_n(\underline{x}) - \sigma|$. For notational convenience define

$$A_n(\underline{x}, \sigma) = \exp\left[-\frac{n}{2} (\sigma - \vartheta_n(\underline{x}))^2 \frac{1}{n} \sum_{i=1}^{n} f''(x_i, \hat{\vartheta}_n(\underline{x}, \sigma))\right],$$

$\underline{x} \in X^n$, $n \in \mathbb{N}$. It follows that for every $\Sigma \in \mathscr{B}$

$$R_{n,\underline{x}}(\Sigma) = \frac{\int_\Sigma A_n(\underline{x}, \sigma)\, \lambda(d\sigma)}{\int_\Theta A_n(\underline{x}, \sigma)\, \lambda(d\sigma)}.$$

Easy computations show that Theorem 1 implies that for every $\delta > 0$ there exists $\eta_K > 0$ such that

$$\sup_{\Sigma \in \mathscr{B}} \left| R_{n,\underline{x}}(\Sigma) - \frac{\int_{\Sigma \cap B_\delta(\vartheta)} A_n(\underline{x}, \sigma)\, \lambda(d\sigma)}{\int_{B_\delta(\vartheta)} A_n(\underline{x}, \sigma)\, \lambda(d\sigma)} \right| > \exp(-\eta_K\, n) \sim O_K(n^{-1}).$$

Proof of Theorem 2. Lemma 4 of Michel and Pfanzagl, [5], and Theorem 1 imply that for every $\delta > 0$ there exists $\eta_K > 0$ such that

$$R_{n,\underline{x}}\{\sigma \in \Theta \,|\, |\sigma - \vartheta_n(\underline{x})| \geqq \delta\} > \exp(-\eta_K\, n) \sim O_K(n^{-1}).$$

Hence it is sufficient to prove that for every $r > 0$ and every compact $K \subseteq \Theta$ there exist $s_K > 0$, c_K and $\delta_K > 0$ such that

$$R_{n,\underline{x}}\{\sigma \in \Theta \,|\, \delta_K > |\sigma - \vartheta_n(\underline{x})| > (s_K\, a(\vartheta) \log n)^{1/2}\, n^{-1/2}\} \geqq c_K\, n^{-r} \sim O_K(n^{-1}).$$

It follows from Lemma 4 that for every $\varepsilon>0$ there exists $\delta_K>0$ such that

$$\sup_{|\sigma-\vartheta_n(\underline{x})|<\delta_K}\left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\sigma))-P_\vartheta(f''(\cdot,\vartheta))\right|>\varepsilon\sim O_K(n^{-1}),$$

(use Lemma 4 of Michel and Pfanzagl, [5]). The set of all $\underline{x}\in X^n$ such that the inequality on the left hand is true will be denoted by $M_n^{(1)}(\varepsilon,\delta_K,\vartheta)$.

Let $s>0$ and define

$$S_n^\vartheta(\underline{x},s)=\{\sigma\in\Theta\mid \delta_K>|\sigma-\vartheta_n(\underline{x})|>(s\,a(\vartheta)\log n)^{1/2}\,n^{-1/2}\}$$

and

$$V_n(\underline{x})=\{\sigma\in\Theta\mid |\sigma-\vartheta_n(\underline{x})|<n^{-1/2}\}.$$

It follows that $\underline{x}\notin M_n^{(1)}(\varepsilon,\delta_K,\vartheta)$ and $n\geqq n_K$ imply

$$\begin{aligned}R_{n,\underline{x}}^\vartheta(S_n^\vartheta(\underline{x},s))&=\frac{\int_{S_n^\vartheta(\underline{x},s)} A_n(\underline{x},\sigma)\,\lambda(d\sigma)}{\int_\Theta A_n(\underline{x},\sigma)\,\lambda(d\sigma)}\\&\leqq\frac{\int_{S_n^\vartheta(\underline{x},s)}\exp\left[-\frac{n}{2}(\vartheta_n(\underline{x})-\sigma)^2(a(\vartheta)^{-1}-\varepsilon)\right]\lambda(d\sigma)}{\int_{V_n(\underline{x})}\exp\left[-\frac{n}{2}(\vartheta_n(\underline{x})-\sigma)^2(a(\vartheta)^{-1}+\varepsilon)\right]\lambda(d\sigma)}\\&\leqq\frac{\exp\left[-\frac{1}{2}(\log n)\,s(1-\varepsilon\,a(\vartheta))\right]}{\lambda(V_n(\underline{x}))\exp\left[-\frac{1}{2}(a(\vartheta)^{-1}+\varepsilon)\right]}.\end{aligned}$$

Let $\rho>0$ be such that $K_1=\{\sigma\in\Theta\mid \mathrm{dist}(\sigma,K)\leqq\rho\}$ is a compact subset of Θ. Let $M_n^{(2)}(\rho,\vartheta)=\{\underline{x}\in X^n\mid |\vartheta_n(\underline{x})-\vartheta|>\rho\}$. Then $\sup_{\vartheta\in k}P_\vartheta^n(M_n^{(2)}(\rho,\vartheta))\leqq c_K\,n^{-1}$. $\underline{x}\notin M_n^{(2)}(\rho,\vartheta)$ implies

$$\liminf_{n\in\mathbb{N}} n^t\,\lambda(V_n(\underline{x}))=\liminf_{n\in\mathbb{N}} n^t\inf_{\vartheta\in K_1}\lambda\{\sigma\in\Theta\mid |\sigma-\vartheta|<n^{-1/2}\}>0$$

according to condition (jj). Thus it follows that there exists $c_K\geqq 0$ with

$$R_{n,\underline{x}}(S_n^\vartheta(\underline{x},s))>c_K\,n^{(2t-s(1-\varepsilon a(\vartheta)))/2}\sim O_K(n^{-1}).$$

It is obvious that for every $r>0$, $s>0$ can be chosen such that the assertion of the theorem holds. □

Easy computations show that Theorem 2 implies that for every $r>0$ and every compact $K\subseteq\Theta$ there exist $s_K>0$ and $c_K\geqq 0$ such that

$$\sup_{\Sigma\in\mathscr{B}}\left|R_{n,\underline{x}}(\Sigma)-\frac{\int_{\Sigma\cap W_n^\vartheta(\underline{x},s_K)} A_n(\underline{x},\sigma)\,\lambda(d\sigma)}{\int_{W_n^\vartheta(\underline{x},s_K)} A_n(\underline{x},\sigma)\,\lambda(d\sigma)}\right|>c_K\,n^{-r}\sim O_K(n^{-1}).$$

The proof of the next theorem is related to the proof of Satz 3.11 in Schmetterer, [10]. In the following we denote by $M_n^{(3)}(s,c,\vartheta)$ the set of all $\underline{x}\in X^n$ such that

$$\sup_{\sigma\in W_n^\vartheta(\underline{x},s)}\left|\frac{1}{n}\sum_{i=1}^{n} f''(x_i,\hat{\vartheta}_n(\underline{x},\sigma))-P_\vartheta(f''(\cdot,\vartheta))\right|\geqq c(\log n)^{1/2}\,n^{-1/2}.$$

It follows from Lemma 5 that for every compact $K \subseteq \Theta$ there exists $c_K > 0$ such that

$$\underline{x} \in M_n^{(3)}(s, c_K, \vartheta) \sim O_K(n^{-1/2}).$$

Proof of Theorem 3. For simplicity of notations we introduce the following abbreviation

$$B_n^\vartheta(\underline{x}, \sigma) = \exp(-n(\sigma - \vartheta_n(\underline{x}))^2 / 2a(\vartheta)).$$

According to Theorem 2 there exists $s > 0$ such that

$$\left| \frac{\int\limits_{W_n^\vartheta(\underline{x}, s_K)} A_n(\underline{x}, \sigma)\, \lambda(d\sigma)}{\int\limits_{\Theta} A_n(\underline{x}, \sigma)\, \lambda(d\sigma)} - 1 \right| \geqq c_K\, n^{-1/2} \sim O_K(n^{-1})$$

and similar to the proof of Theorem 2 it can be shown that $s_K > 0$ may be chosen in such a way that

$$\left| \frac{\int\limits_{W_n^\vartheta(\underline{x}, s_K)} B_n^\vartheta(\underline{x}, \sigma)\, \lambda(d\sigma)}{\int\limits_{\Theta} B_n^\vartheta(\underline{x}, \sigma)\, \lambda(d\sigma)} - 1 \right| \geqq c_K\, n^{-1/2} \sim O_K(n^{-1}).$$

Therefore we need only show that

$$\int\limits_{W_n^\vartheta(\underline{x}, s_K)} \left| \frac{A_n(\underline{x}, \sigma)}{\int\limits_{W_n^\vartheta(\underline{x}, s_K)} A_n(\underline{x}, \sigma)\, \lambda(d\sigma)} - \frac{B_n^\vartheta(\underline{x}, \sigma)}{\int\limits_{W_n^\vartheta(\underline{x}, s_K)} B_n^\vartheta(\underline{x}, \sigma)\, \lambda(d\sigma)} \right| \lambda(d\sigma)$$
$$\geqq c_K (\log n)^{3/2}\, n^{-1/2} \sim o_K(n^{-1/2}).$$

It follows that (with simplified notation)

$$\int\limits_{W_n^\vartheta} \left| \frac{A_n}{\int\limits_{W_n^\vartheta} A_n\, d\lambda} - \frac{B_n^\vartheta}{\int\limits_{W_n^\vartheta} B_n^\vartheta\, d\lambda} \right| d\lambda$$
$$\leqq \int\limits_{W_n^\vartheta} \left[\frac{A_n}{(\int\limits_{W_n^\vartheta} A_n\, d\lambda)(\int\limits_{W_n^\vartheta} B_n^\vartheta\, d\lambda)} \left| \int\limits_{W_n^\vartheta} A_n\, d\lambda - \int\limits_{W_n^\vartheta} B_n^\vartheta\, d\lambda \right| + \frac{1}{\int\limits_{W_n^\vartheta} B_n^\vartheta\, d\lambda} |A_n - B_n^\vartheta| \right] d\lambda$$
$$\leqq 2 \frac{1}{\int\limits_{W_n^\vartheta} B_n^\vartheta\, d\lambda} \int\limits_{W_n^\vartheta} |A_n - B_n^\vartheta|\, d\lambda \leqq 2 \left(\sup_{W_n^\vartheta} \left| \frac{A_n}{B_n^\vartheta} - 1 \right| \right).$$

Let s_K and c_K be such that $\underline{x} \in M_n^{(3)}(s_K, c_K, \vartheta) \sim o_K(n^{-1/2})$. For $\underline{x} \notin M_n^{(3)}(s_K, c_K, \vartheta)$ we have

$$\sup_{\sigma \in W_n^\vartheta(\underline{x}, s_K)} \left| \log \frac{A_n(\underline{x}, \sigma)}{B_n^\vartheta(\underline{x}, \sigma)} \right|$$
$$= \sup_{\sigma \in W_n^\vartheta(\underline{x}, s_K)} \left| \frac{n}{2} \frac{(\sigma - \vartheta_n(\underline{x}))^2}{a(\vartheta)} \left[a(\vartheta) \frac{1}{n} \sum_{i=1}^n f''(x_i, \hat{\vartheta}_n(\underline{x}, \sigma)) - 1 \right] \right|$$
$$\leqq \frac{c_K}{2} a(\vartheta) (\log n)^{3/2}\, n^{-1/2}.$$

From $|e^{\xi}-1|\leqq c|\xi|$ for some $c\geqq 0$ and sufficiently small $|\xi|$ the assertion follows immediately. □

Proof of Theorem 4. Let $V_n^{\vartheta}(\underset{\sim}{x}, \Sigma)$ for any $\Sigma\in\mathscr{B}$ be defined as in Lemma 6. We have to show that

$$\sup_{\Sigma\in\mathscr{B}} |R_{n,\underset{\sim}{x}}(V_n^{\vartheta}(\underset{\sim}{x}, \Sigma)\cap W_n^{\vartheta}(\underset{\sim}{x}, s_K))-\sqrt{2\pi}\int_{\Sigma} \exp(-\tau^2/2)\,d\tau|$$
$$\geqq c_K(\log n)^{1/2}\, n^{-1/2}\sim o_K(n^{-1/2}).$$

Theorem 2 implies that we need only show

$$\sup_{\Sigma\in\mathscr{B}} \frac{\int\limits_{V_n^{\vartheta}(\underset{\sim}{x},\Sigma)\cap W_n^{\vartheta}(\underset{\sim}{x},s_K)} A_n(\underset{\sim}{x},\sigma)\,\lambda(d\sigma)}{\int\limits_{W_n^{\vartheta}(\underset{\sim}{x},s_K)} A_n(\underset{\sim}{x},\sigma)\,\lambda(d\sigma)}-\sqrt{2\pi}\int_{\Sigma}\exp(-\tau^2/2)\,d\tau$$
$$\geqq c_K(\log n)^{1/2}\, n^{-1/2}\sim o_K(n^{-1/2}).$$

Choose s_K such that $\underset{\sim}{x}\in M_n^{(3)}(s_K, c_K, \vartheta)\sim o_K(n^{-1/2})$. Choose n_K such that $n\geqq n_K$ implies

$$(s_K\, a(\vartheta)\log n)^{1/2}\, n^{-1/2}<d_K/2$$

for every $\vartheta\in K$ (where d_K is chosen according to condition (jjj)). Since $|\vartheta_n(\underset{\sim}{x})-\vartheta|\geqq d_K/2\sim O_K(n^{-1})$ we may restrict our attention to those $\underset{\sim}{x}\in X^n$ for which $n\geqq n_K$ and $\sigma\in W_n^{\vartheta}(\underset{\sim}{x}, s_K)$ imply $|\sigma-\vartheta|<d_K$. Then we obtain from condition (jjj) that for $n\geqq n_K$ and $\Sigma\in\mathscr{B}$

$$\Big|\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\lambda-p(\vartheta)\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\sigma\Big|\leqq p(\vartheta)\,c_K(\log n)^{1/2}\, n^{-1/2}\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\sigma$$

which implies

$$\frac{\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\lambda}{\int_{W_n^{\vartheta}} A_n\,d\lambda}=\frac{\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\sigma(1+\varepsilon_n')}{\int_{W_n^{\vartheta}} A_n\,d\sigma(1+\eta_n')}$$

where $\max\{|\varepsilon_n'|, |\eta_n'|\}\leqq c_K(\log n)^{1/2}\, n^{-1/2}$.

Define $\gamma_n=\sup_{\vartheta\in K} a(\vartheta)\,c_K(\log n)^{1/2}\, n^{-1/2}$. Then $\underset{\sim}{x}\notin M_n^{(3)}(s_K, c_K, \vartheta)$ implies

$$\exp\left[-\frac{n}{2}\frac{(\sigma-\vartheta_n(\underset{\sim}{x}))^2}{a(\vartheta)}(1+\gamma_n)\right]\leqq A_n(\underset{\sim}{x},\sigma)\leqq\exp\left[-\frac{n}{2}\frac{(\sigma-\vartheta_n(\underset{\sim}{x}))^2}{a(\vartheta)}(1-\gamma_n)\right].$$

We obtain with the aid of Lemma 6 that $n\geqq n_K$, $\Sigma\in\mathscr{B}$ and $\underset{\sim}{x}\notin M_n^{(3)}(s_K, c_K, \vartheta)$ imply

$$\frac{\int_{V_n^{\vartheta}\cap W_n^{\vartheta}} A_n\,d\lambda}{\int_{W_n^{\vartheta}} A_n\,d\lambda}=\frac{(1+\varepsilon_n')\left(\int_{\Sigma}\exp(-\tau^2/2)\,d\tau+\varepsilon_n\right)}{(1+\eta_n')(\sqrt{2\pi}+\eta_n)}$$

where $\max\{|\varepsilon_n|, |\eta_n|\}\leqq c_K|\gamma_n|$. Now easy computations finish the proof. □

References

1. Doob, J. L.: Applications of the theory of martingales. Colloque International Centre Nat. Rech. Sci., Paris, 22–28 (1949)
2. Johnson, R. A.: Asymptotic expansions associated with posterior distributions. Ann. Math. Statist. **41**, 851–864 (1970)
3. LeCam, L.: On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. Univ. Calif. Publ. Statist. **1**, 277–330 (1953)
4. LeCam, L.: Convergence of estimates under dimensionality restrictions. Ann. Statist. **1**, 38–53 (1973)
5. Michel, R., Pfanzagl, J.: The accuracy of the normal approximation for minimum contrast estimates. Z. Wahrscheinlichkeitstheorie verw. Gebiete **18**, 73–84 (1971)
6. Michel, R., Pfanzagl, J.: Asymptotic normality. Metrika **16**, 188–205 (1970)
7. Pfanzagl, J.: On the measurability and consistency of minimum contrast estimates. Metrika **14**, 249–272 (1969)
8. Pfanzagl, J.: The Berry-Esseen bound for minimum contrast estimates. Metrika **17**, 82–91 (1971)
9. Pfanzagl, J.: Asymptotic expansions related to minimum contrast estimators. Ann. Statist. **1**, 993–1026 (1973)
10. Schmetterer, L.: Einführung in die mathematische Statistik. Wien: Springer 1966
11. Schwartz, L.: On Bayes procedures. Z. Wahrscheinlichkeitstheorie verw. Gebiete **4**, 10–26 (1965)
12. Strasser, H.: On Bayes estimates. J. Multivariate Anal. **3**, 293–310 (1973)
13. Strasser, H.: A representation of consistent estimates relative to initial distributions. In: Proceedings of the Prague symposium on asymptotic statistics, p. 383–399 (1973)
14. van der Waerden, B.: Empirische Bestimmung von Wahrscheinlichkeiten und physiologische Konzentrationsauswertungen. Ber. Akad. Leipzig, math.-phys. **Kl.** 87, 353–364 (1935)

Received October 28, 1974; in revised form March 16, 1976