Matsuda, T.
Osaka J. Math.
21 (1984), 239–255

ASYMPTOTIC SUFFICIENCY I: REGULAR CASES

TADAYUKI MATSUDA

(Received October 4, 1982)

1. Introduction. The concept of asymptotic sufficiency of maximum likelihood (m.l.) estimator is due to Wald [16] and this work was succeeded by LeCam [4] and Pfanzagl [10]. Higher order asymptotic sufficiency has been subsequently studied by Ghosh and Subramanyam [3], Michel [7] and Suzuki [14], [15].

Let Θ be an open subset of the s-dimensional Euclidean space. Suppose that $x_1, \cdots, x_n$ are independent and identically distributed random variables with joint distribution $P_{n,\theta}$, $\theta \in \Theta$, which has a constant support and satisfies certain regularity conditions. For $\theta \in \Theta$ and $z_n = (x_1, \cdots, x_n)$ let $G_n^{(m)}(z_n, \theta)$ denote the m-th derivative relative to θ of the log-likelihood function. In Michel [7], it was shown that for $k \geqq 3$ a statistic $T_{n,k} = (T_n, G_n^{(2)}(z_n, T_n), \cdots, G_n^{(k)}(z_n, T_n))$, where $\{T_n\}$ is a sequence of asymptotic m.l. estimators of order $o(n^{-(k-2)/2})$ (see Definition in Section 3), is asymptotically sufficient up to order $o(n^{-(k-2)/2})$ in the following sense: For each $n \in N$, $T_{n,k}$ is sufficient for a family $\{Q_{n,\theta}; \theta \in \Theta\}$ of probability distributions and for every compact subset K of Θ

$$\sup_{\theta \in K} \|P_{n,\theta} - Q_{n,\theta}\| = o(n^{-(k-2)/2}),$$

where $\|\cdot\|$ means the total variation of a measure. Suzuki [14], [15] also showed that for $k \in N$ a statistic $(\hat{\theta}_n, G_n^{(1)}(z_n, \hat{\theta}_n), \cdots, G_n^{(k)}(z_n, \hat{\theta}_n))$, where $\hat{\theta}_n$ is a reasonable estimator including m.l. estimator, is asymptotically sufficient up to order $o(n^{-(k-1)/2})$ under a stronger moment condition than in Michel [7].

In this paper we give a refinement of their results on higher order asymptotic sufficiency. Our result includes that (1) $T_{n,k} = (T_n, G_n^{(2)}(z_n, T_n), \cdots, G_n^{(k)}(z_n, T_n))$ is asymptotically sufficient up to order $O(n^{-k/2})$ for any sequence $\{T_n\}$ of asymptotic m.l. estimators of order $O(n^{-k/2})$ and (2) a sequence of asymptotic m.l. estimators of order $O(n^{-r/2})$ with some $r \in (0, 1)$ is asymptotically sufficient up to order $O(n^{-r/2})$ under mild moment conditions for the first and the second derivatives of the log-likelihood function.

In the case $k=1$, Pfanzagl ([10], Theorem 1) proved that a sequence of estimators with properties analogous to those of asymptotic m.l. estimators of order $O(n^{-1/2})$ is asymptotically sufficient up to order $O(n^{-1/2})$, and showed in

[11] that this order of convergence cannot be improved in general. Thus our result is an extension of his and it seems to be impossible to improve the convergence order $O(n^{-k/2})$.

In Section 2 we present a result concerning probabilities of deviations for sums of independent and identically distributed random variables with a restricted moment. In Section 3 we investigate asymptotic sufficiency of $T_{n,k}$ constructed by asymptotic m.l. estimators T_n. In the final Section 4 we give conditions under which a sequence of m.l. estimators becomes the one of asymptotic m.l. estimators of order $O(n^{-r/2})$ with some $r>0$.

2. Probabilities of deviations. Let $Y_1, \cdots, Y_n$ be a sequence of random variables (r.v.'s) and put $S_m=\sum_{i=1}^m Y_i$, $1\leqq m\leqq n$. Using the elementary inequality

$$E|S_n|^r\leqq \sum_{i=1}^n E|Y_i|^r\,, \qquad r\leqq 1\,,$$

it follows from Markov's inequality that for $x>0$

$$P\{|S_n|\geqq x\}\leqq x^{-r}\sum_{i=1}^n E|Y_i|^r\,. \tag{2.1}$$

If the r.v.'s satisfy the relations

$$E(Y_{m+1}|S_m)=0 \quad \text{a.s.} \qquad 1\leqq m\leqq n-1\,, \tag{2.2}$$

then von Bahr and Esseen [1] showed that

$$E|S_n|^r\leqq 2\sum_{i=1}^n E|Y_i|^r\,, \qquad 1\leqq r\leqq 2\,. \tag{2.3}$$

The condition (2.2) is fulfilled if the r.v.'s are independent and have zero means. In this case, (2.3) together with Markov's inequality implies the following inequality

$$P\{|S_n|\geqq x\}\leqq 2x^{-r}\sum_{i=1}^n E|Y_i|^r\,, \qquad 1\leqq r\leqq 2\,, \tag{2.4}$$

for $x>0$.

The following theorem includes a uniform version of Corollary 2 in Nagaev [9].

Theorem 1. *Let $Y_1, \cdots, Y_n$ be a sequence of independent and identically distributed random variables with a common distribution P_θ, $\theta\in K$, where K is any set. Let $h(y, \theta)$ be a measurable function of y for any fixed $\theta\in K$ and put $S_{n,\theta}=\sum_{i=1}^n h(Y_i, \theta)$. If $E_\theta(h(Y_1, \theta))=0$ for all $\theta\in K$ and $\xi_r=\sup_{\theta\in K} E_\theta|h(Y_1, \theta)|^r<\infty$ for some $r>0$, then*

(2.5) $$\sup_{\theta\in K} P_\theta\{|S_{n,\theta}|\geqq x\} = O(nx^{-r}), \qquad 0<r\leqq 2,$$

for $x>0$, *and*

(2.6) $$\sup_{\theta\in K} P_\theta\{|S_{n,\theta}|>\xi_r^{1/r}x\} = O(nx^{-r}), \qquad r>2,$$

for $x\geqq\sqrt{8(r-2)n\log n}$.

Proof. (2.5) is an immediate consequence of (2.1) and (2.4).

For the proof of (2.6) we use the following inequality which is a slight modification of Theorem 1 in Nagaev [9]: For $x>0$ and $y>0$,

(2.7) $$P_\theta\{|S_{n,\theta}|>x\}<2nP_\theta\{|h(Y_1,\theta)|>y\}+2\left[\frac{n\xi_{r,\theta}d_r}{y^r}\right]^{x/y} \times\exp\left\{2n\left[\frac{r\log y-\log(n\xi_{r,\theta}d_r)}{y}\right]^2\xi_{r,\theta}^{2/r}+1\right\},$$

where $d_r=1+(r+1)^{r+2}\exp(-r)$ and $\xi_{r,\theta}=E_\theta|h(Y_1,\theta)|^r$. In order to show (2.7) it is enough to note that the relation (2.3) in [9] becomes

$$\left|\int_{-\infty}^{1/h}\exp\{h[h(Y_1,\theta)]\}dP_\theta-1\right|<2h^2\xi_{r,\theta}^{2/r}.$$

Setting $x=\xi_{r,\theta}^{1/r}n^{1/2}t$ and $y=x/2$ for $t\geqq\sqrt{8(r-2)\log n}$ in (2.7), then we obtain

(2.8) $$nP_\theta\{|h(Y_1,\theta)|>y\}\leqq n\xi_{r,\theta}y^{-r} = 2^rn^{(2-r)/2}t^{-r}$$

and

(2.9) $$\left[\frac{n\xi_{r,\theta}d_r}{y^r}\right]^{x/y}=2^{2r}d_r^2n^{2-r}t^{-2r}.$$

Let us assume that $n\geqq\exp\left\{\frac{1}{8(r-2)}\right\}$. $\Big($For $n<\exp\left\{\frac{1}{8(r-2)}\right\}$, (2.6) is trivially true.$\Big)$ Since $0\leqq t^{-1}\log t<1/2$ for $t\geqq 1$, we have

$$\begin{aligned}
&2n\left[\frac{r\log y-\log(n\xi_{r,\theta}d_r)}{y}\right]^2\xi_{r,\theta}^{2/r}\\
&=8t^{-2}[r\log t+\frac{r-2}{2}\log n-r\log 2-\log d_r]^2\\
&\leqq 8t^{-2}[r^2(\log t)^2+\frac{(r-2)^2}{4}(\log n)^2+r(r-2)\log n\log t+c_1]\\
&\leqq 2r^2+\frac{r-2}{4}\log n+r\log t+c_2,
\end{aligned}$$

where c_1 and c_2 denote positive constants depending only on r. From this fact and (2.9) it follows that the second term on the right side of (2.7) has an upper bound of the type $c_3 n^{3(2-r)/4} t^{-r}$. This, together with (2.8), implies (2.6).

REMARK. (1) In the case $r \geqq 3$, Michel [7] showed a result analogous to Theorem 1 (cf. also Lemma 1 in Pfanzagl [12]).

(2) Let $Y_1, \cdots, Y_n$ be a sequence of independent r.v.'s with zero means. It follows from an inequality due to Marcinkiewicz and Zygmund [5] that

$$E\left|\sum_{i=1}^{n} Y_i\right|^r \leqq c n^{(r-2)/2} \sum_{i=1}^{n} E|Y_i|^r, \qquad r \geqq 2, \tag{2.10}$$

where c is a positive constant depending only on r (see Chung [2], page 348). This leads to Lemma 2 in Pfanzagl [12] which requires a stronger moment condition than in Theorem 1 to evaluate probability of moderate deviations or large deviations.

3. Asymptotic sufficiency. Let Θ be an open subset of the s-dimensional Euclidean space $\boldsymbol{R}^s$ and for each $\theta \in \Theta$, let P_θ be a probability measure on a measurable space $(X, \mathcal{A})$. It is assumed that P_θ, $\theta \in \Theta$, is dominated by a σ-finite measure μ on $(X, \mathcal{A})$ and has a positive density $p(x, \theta)$. For each $n \in N = \{1, 2, \cdots\}$, let $(X^n, \mathcal{A}^n)$ be the Cartesian product of n copies of $(X, \mathcal{A})$ and $P_{n,\theta}$ be the product measure of n copies of P_θ. Furthermore, let μ_n denote the product measure of n copies of μ and write $p_n(z_n, \theta) = dP_{n,\theta}/d\mu_n$ for $\theta \in \Theta$ and $z_n = (x_1, \cdots, x_n) \in X^n$.

For a function $h(z, \cdot)$: $\boldsymbol{R}^s \to \boldsymbol{R}$ denote the m-th derivative relative to θ of $h(z, \theta)$ by

$$h^{(m)}(z, \theta) = \left(\frac{\partial^m}{\partial\theta_{i_1} \cdots \partial\theta_{i_m}} h(z, \theta);\ i_1, \cdots, i_m \in \{1, \cdots, s\}\right).$$

In particular, we write

$$h^{(1)}(z, \theta) = \left(\frac{\partial}{\partial\theta_1} h(z, \theta), \cdots, \frac{\partial}{\partial\theta_s} h(z, \theta)\right),$$

$$h^{(2)}(z, \theta) = \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} h(z, \theta)\right),$$

that is, $h^{(1)}$ means a row vector and $h^{(2)}$ a matrix. The Euclidean norm $\|\cdot\|$ of $h^{(m)}$ is defined by

$$\|h^{(m)}(z, \theta)\|^2 = \sum_{i_1, \cdots, i_m = 1}^{s} \left(\frac{\partial^m}{\partial\theta_{i_1} \cdots \partial\theta_{i_m}} h(z, \theta)\right)^2.$$

For any $\sigma = (\sigma_1, \cdots, \sigma_s) \in \boldsymbol{R}^s$ define

$$h^{(m)}(z, \theta)\sigma^m = \sum_{i_1, \cdots, i_m = 1}^{s} \frac{\partial^m}{\partial\theta_{i_1} \cdots \partial\theta_{i_m}} h(z, \theta) \prod_{p=1}^{m} \sigma_{i_p}.$$

Then, it is easy to see that

$$|h^{(m)}(z, \theta)\sigma^m| \leqq ||h^{(m)}(z, \theta)|| \, ||\sigma||^m .$$

Let $k \in N$ and $r>0$ be fixed. We shall impose the following Conditions A, B_r and $C_{k,r}$ on $p(x, \theta)$.

Condition A

(i) For each $x \in X$, $\theta \rightarrow p(x, \theta)$ admits continuous partial derivatives up to the order 2 on Θ.

Let $g(x, \theta)=\log p(x, \theta)$ and $g^{(m)}$ be the m-th derivative of g defined above. Moreover, for $\theta \in \Theta$ let $J(\theta)=E_\theta(-g^{(2)}(\cdot, \theta))$.

(ii) For every $\theta \in \Theta$

(a) $E_\theta(g^{(1)}(\cdot, \theta))=0$

(b) $J(\theta)$ is positive definite.

Condition B_r

For every compact $K \subset \Theta$

$$\sup_{\theta \in K} E_\theta(||g^{(1)}(\cdot, \theta)||^{r+2})<\infty .$$

Condition $C_{k,r}$

(i) For each $x \in X$, $\theta \rightarrow p(x, \theta)$ admits continuous partial derivatives up to the order $k+1$ on Θ.

(ii) For every $\theta \in \Theta$ there exist a neighborhood U_θ of θ and a measurable function $\lambda(x, \theta)$ such that

(a) for all $x \in X$, $\tau, \sigma \in U_\theta$, $||g^{(k+1)}(x, \tau)-g^{(k+1)}(x, \sigma)|| \leqq ||\tau-\sigma|| \lambda(x, \theta)$

(b) for every compact $K \subset \Theta$, $\sup_{\tau \in K} E_\tau(\lambda(\cdot, \theta)^{(r+2)/2})<\infty$

(c) $\sup_{\tau \in U_\theta} E_\tau(||g^{(k+1)}(\cdot, \tau)||^{\nu(r)})<\infty$,

where

$$\begin{aligned} \nu(r) &= \frac{2+r}{2-r}, \quad \text{if } 0<r<1 , \\ &= r+2 , \quad \text{if } r \geqq 1 . \end{aligned}$$

(iii) For every compact $K \subset \Theta$ there exist $\delta_K>0$ and $\eta_K>0$ such that $\theta \in K$ and $\tau \in \Theta$ with $||\theta-\tau||<\delta_K$ imply

$$||E_\theta(g^{(k+1)}(\cdot, \theta))-E_\tau(g^{(k+1)}(\cdot, \tau))|| \leqq \eta_K ||\theta-\tau|| .$$

REMARK. (1) Condition (iii) in $C_{k,r}$ follows from conditions (3)(a) and (3) (b) in Suzuki [15] (see also (3.4) in [14]).

(2) It is easily seen that condition (ii) in $C_{k,r}$ and the following condition (iii)′ imply condition (iii) in $C_{k,r}$.

(iii)′ For every $\theta \in \Theta$ there exist a neighborhood U_θ of θ and a measurable function $\lambda^*(x, \theta)$ such that for all $x \in X$, $\tau \in U_\theta$

$$|p(x, \tau)/p(x, \theta)-1| \leqq \|\tau-\theta\|\lambda^*(x, \theta)$$

and for every compact $K \subset \Theta$

$$\sup_{\tau \in K} E_\tau(\lambda^*(\cdot, \theta)^{\nu(r)/(\nu(r)-1)}) < \infty .$$

The following definition is due to Michel [7].

DEFINITION. T_n, $n \in N$, is a sequence of *asymptotic maximum likelihood (m.l.) estimators of order* $O(n^{-r/2})$, $r>0$, if there exist positive constants π_1 and π_2 (depending on r) such that for every compact $K \subset \Theta$

(α_r) $\sup_{\theta \in K} P_{n,\theta}\{z_n \in X^n;\ n^{1/2}\|T_n(z_n)-\theta\| \geqq (\log n)^{\pi_1}\} = O(n^{-r/2})$

(β_r) $\sup_{\theta \in K} P_{n,\theta}\{z_n \in X^n;\ n^{r/2}\|\sum_{i=1}^n g^{(1)}(x_i, T_n(z_n))\| \geqq (\log n)^{\pi_2}\} = O(n^{-r/2})$.

Asymptotic m.l. estimators can be obtained from suitable initial estimators by applying a Newton-Raphson method (see Michel [6] and Pfanzagl [13]).

To simplify our notations we shall use n_K (depending on compact K) as a generic constant instead of the phrase "for all sufficiently large n". In the same manner we shall use c_K as a generic constant to denote factors occurring in the bounds which depend on compact K but not on $\theta \in K$ and $n \in N$.

Lemma 1. *Assume that Condition $C_{k,r}$ is fulfilled for some $k \in N$ and $r>0$. Let T_n, $n \in N$, be a sequence of estimators with the property (α_r). Then for every compact $K \subset \Theta$*

$$\sup_{\theta \in K} P_{n,\theta}\{\sup_{n^{1/2}\|T_n-\tau\| \leqq (\log n)^{\pi_1}} \|\sum_{i=1}^n [g^{(k+1)}(x_i, \tau)-E_\tau(g^{(k+1)}(\cdot, \tau))]\| \geqq \psi(n, r)\} = O(n^{-r/2}),$$

where

$$\psi(n, r) = n^{(2-r)/2}, \qquad \text{if } 0<r<1,$$
$$= n^{1/2}(\log n)^{\pi_1+1/2}, \qquad \text{if } r \geqq 1 .$$

Proof. Let $0<r<1$ and K be a compact subset of Θ. Condition (ii) implies that there exist $d_K>0$ and $\lambda_K(x)$ such that $\theta \in K$ and $\tau \in \Theta$ with $\|\theta-\tau\|<d_K$ imply $\|g^{(k+1)}(x, \theta)-g^{(k+1)}(x, \tau)\| \leqq \|\theta-\tau\|\lambda_K(x)$ for all $x \in X$, and such that $\sup_{\theta \in K} E_\theta(\lambda_K(\cdot)^{(r+2)/2}) < \infty$. Let

$$D_{n,\theta,K} = \{z_n \in X^n;\ |\sum_{i=1}^n [\lambda_K(x_i)-E_\theta(\lambda_K(\cdot))]| < n\} .$$

According to Theorem 1

$$\sup_{\theta \in K} P_{n,\theta}\{(D_{n,\theta,K})^c\} = O(n^{-r/2}) . \tag{3.1}$$

Furthermore, Theorem 1 together with condition (ii) (c) implies that

$$\sup_{\theta\in K} P_{n,\theta}\{(F_{n,\theta})^c\} = O(n^{-r/2}), \tag{3.2}$$

where

$$F_{n,\theta} = \{z_n\in X^n;\ \|\sum_{i=1}^n [g^{(k+1)}(x_i, \theta)-E_\theta(g^{(k+1)}(\cdot, \theta))]\|<1/2\, n^{(2-r)/2}\}.$$

Let $e_K>0$ be such that $\{\tau\in \boldsymbol{R}^s;\ \inf_{\theta\in K}\|\theta-\tau\|\leqq e_K\}\subset\Theta$. Choose n_K to satisfy $2n^{-1/2}(\log n)^{\pi_1}<\min\{d_K, e_K, \delta_K\}$ for all $n\geqq n_K$, where δ_K appears in condition (iii). Then, by conditions (ii) and (iii), for $n\geqq n_K$, $\theta\in K$, $\tau\in\boldsymbol{R}^s$ with $\|\theta-\tau\|\leqq 2n^{-1/2}(\log n)^{\pi_1}$ and $z_n\in D_{n,\theta,K}\cap F_{n,\theta}$

$$\begin{aligned}
&\|\sum_{i=1}^n [g^{(k+1)}(x_i, \tau)-E_\tau(g^{(k+1)}(\cdot, \tau))]\| \\
&\leqq\|\sum_{i=1}^n [g^{(k+1)}(x_i, \tau)-g^{(k+1)}(x_i, \theta)]\|+\|\sum_{i=1}^n [g^{(k+1)}(x_i, \theta)-E_\theta(g^{(k+1)}(\cdot, \theta))]\| \\
&\quad +n\|E_\theta(g^{(k+1)}(\cdot, \theta))-E_\tau(g^{(k+1)}(\cdot, \tau))\| \\
&\leqq n[1+\sup_{\theta\in K} E_\theta(\lambda_K(\cdot))+\eta_K]\|\theta-\tau\|+1/2\, n^{(2-r)/2} \\
&< n^{(2-r)/2}.
\end{aligned}$$

Taking account of (3.1) and (3.2), for every compact $K\subset\Theta$ we obtain

$$\begin{aligned}
&\sup_{\theta\in K} P_{n,\theta}\{\sup_{n^{1/2}\|\theta-\tau\|\leqq 2(\log n)^{\pi_1}} \|\sum_{i=1}^n [g^{(k+1)}(x_i, \tau)-E_\tau(g^{(k+1)}(\cdot, \tau))]\|\geqq n^{(2-r)/2}\} \\
&\quad = O(n^{-r/2}).
\end{aligned}$$

This together with the property (α_r) leads to the desired assertion.

In the case $r\geqq 1$, it is enough to show that there exists $c_K>0$ such that

$$\sup_{\theta\in K} P_{n,\theta}\{(F_{n,\theta,K})^c\} = o(n^{-r/2}), \tag{3.3}$$

where

$$F_{n,\theta,K} = \{z_n\in X^n;\ \|\sum_{i=1}^n [g^{(k+1)}(x_i, \theta)-E_\theta(g^{(k+1)}(\cdot, \theta))]\|\leqq c_K(n\log n)^{1/2}\}.$$

This follows from Theorem 1 and condition (ii) (c).

Lemma 2. *Assume that Conditions A, B_r and $C_{1,r}$ are fulfilled for some $r>0$. Let T_n, $n\in N$, be a sequence of asymptotic m.l. estimators of order $O(n^{-r/2})$. Then for every compact $K\subset\Theta$*

$$\sup_{\theta\in K} P_{n,\theta}\{z_n\in X^n;\ \|T_n(z_n)-\theta-n^{-1}\sum_{i=1}^n g^{(1)}(x_i, \theta)J(\theta)^{-1}\|\geqq\omega(n, r)\}=O(n^{-r/2}),$$

where

$$\omega(n, r) = n^{-(r+1)/2}(\log n)^{2\pi_1}, \qquad \text{if } 0<r<1,$$
$$= n^{-1}(\log n)^{2\pi_1+1/2}, \qquad \text{if } r\geqq 1.$$

Proof. Let $0<r<1$ and K be a compact subset of Θ. Condition (ii) in $C_{1,r}$ implies that there exist $d_K>0$ and $\lambda_K(x)$ such that $x\in X$, $\theta\in K$ and $\tau\in\Theta$ with $||\theta-\tau||<d_K$ imply $||g^{(2)}(x, \theta)-g^{(2)}(x, \tau)||\leqq||\theta-\tau||\lambda_K(x)$ and such that $\sup_{\theta\in K} E_\theta(\lambda_K(\cdot)^{(r+2)/2})<\infty$. As in the proof of Lemma 1, we define

$$D_{n,\theta,K} = \{z_n\in X^n;\ |\sum_{i=1}^n [\lambda_K(x_i)-E_\theta(\lambda_K(\cdot))]|<n\},$$
$$F_{n,\theta} = \{z_n\in X^n;\ ||\sum_{i=1}^n [g^{(2)}(x_i, \theta)+J(\theta)||<1/2\, n^{(2-r)/2}\}.$$

It follows from Theorem 1, condition (ii) (a) in A and Condition B_r that there exists $c_K>0$ such that

$$\sup_{\theta\in K} P_{n,\theta}\{(H_{n,\theta,K})^c\} = o(n^{-r/2}), \tag{3.4}$$

where

$$H_{n,\theta,K} = \{z_n\in X^n;\ ||\sum_{i=1}^n g^{(1)}(x_i, \theta)||<c_K(n\log n)^{1/2}\}.$$

Let $U_{n,\theta}$ and $V_{n,r}$ be defined by

$$U_{n,\theta} = \{z_n\in X^n;\ n^{1/2}||T_n(z_n)-\theta||<(\log n)^{\pi_1}\},$$
$$V_{n,r} = \{z_n\in X^n;\ n^{r/2}||\sum_{i=1}^n g^{(1)}(x_i, T_n(z_n))||<(\log n)^{\pi_2}\}.$$

Choose $e_K>0$ such that $K^*=\{\tau\in \boldsymbol{R}^s;\ \inf_{\theta\in K}||\theta-\tau||\leqq e_K\}\subset\Theta$ and n_K such that $n^{-1/2}(\log n)^{\pi_1}<\min\{d_K, e_K, \delta_K\}$ for all $n\geqq n_K$, where δ_K is determined by condition (iii) in $C_{1,r}$. It is obvious that $n\geqq n_K$, $\theta\in K$ and $z_n\in U_{n,\theta}$ imply $T_n(z_n)\in K^*$. Since K^* is a compact subset of Θ, conditions (ii) (b) in A and (iii) in $C_{1,r}$ imply that

$$\rho_{K^*} = \sup_{\tau\in K^*}||J(\tau)^{-1}||<\infty.$$

Using the equality

$$\sum_{i=1}^n g^{(1)}(x_i, \theta) = \sum_{i=1}^n g^{(1)}(x_i, T_n)+(\theta-T_n)\sum_{i=1}^n \bar{g}^{(2)}(x_i, T_n, \theta)$$

with $\bar{g}^{(2)}(x, \theta, \sigma)=\int_0^1 g^{(2)}(x, (1-t)\theta+t\sigma)dt$, we obtain

$$T_n-\theta-n^{-1}\sum_{i=1}^n g^{(1)}(x_i, \theta)J(\theta)^{-1}+n^{-1}\sum_{i=1}^n g^{(1)}(x_i, T_n)J(T_n)^{-1}$$
$$= T_n-\theta-n^{-1}\sum_{i=1}^n g^{(1)}(x_i, \theta)[J(\theta)^{-1}-J(T_n)^{-1}]+(T_n-\theta)n^{-1}\sum_{i=1}^n \bar{g}^{(2)}(x_i, T_n, \theta)J(T_n)^{-1}$$

$$= (T_n-\theta)[J(T_n)-J(\theta)]J(T_n)^{-1}-n^{-1}\sum_{i=1}^{n} g^{(1)}(x_i, \theta)J(\theta)^{-1}[J(T_n)-J(\theta)]J(T_n)^{-1}$$

$$+(T_n-\theta)n^{-1}\sum_{i=1}^{n} [g^{(2)}(x_i, T_n, \theta)-g^{(2)}(x_i, \theta)]J(T_n)^{-1}$$

$$+(T_n-\theta)n^{-1}\sum_{i=1}^{n} [g^{(2)}(x_i, \theta)+J(\theta)]J(T_n)^{-1} .$$

Hence we have for $n \geqq n_K$, $\theta \in K$ and $z_n \in D_{n,\theta,K} \cap F_{n,\theta} \cap H_{n,\theta,K} \cap U_{n,\theta} \cap V_{n,r}$

$$||T_n-\theta-n^{-1}\sum_{i=1}^{n} g^{(1)}(x_i, \theta)J(\theta)^{-1}||$$

$$\leqq ||n^{-1}\sum_{i=1}^{n} g^{(1)}(x_i, T_n)J(T_n)^{-1}||+||T_n-\theta-n^{-1}\sum_{i=1}^{n} g^{(1)}(x_i, \theta)J(\theta)^{-1}$$

$$+n^{-1}\sum_{i=1}^{n} g^{(1)}(x_i, T_n)J(T_n)^{-1}||$$

$$\leqq \rho_{K^*}n^{-(r+2)/2}(\log n)^{\pi_2}+(c_K\eta_K\rho_{K^*}^2 n^{-1/2}(\log n)^{1/12}+1/2\ \rho_{K^*}n^{-r/2})||T_n-\theta||$$

$$+\rho_{K^*}(1+\eta_K+\sup_{\theta \in K} E_\theta(\lambda_K(\cdot)))||T_n-\theta||^2$$

$$\leqq c_K n^{-(r+1)/2}(\log n)^{\pi_1} .$$

This implies the desired result because of (3.1), (3.2), (3.4) and the properties (α_r), (β_r).

For the case $r \geqq 1$, the proof is also similar except that $F_{n,\theta}$ is replaced by $F_{n,\theta,K}$ in (3.3) with $k=1$.

For simplicity, we write

$$G_n^{(m)}(z_n, \theta) = \sum_{i=1}^{n} g^{(m)}(x_i, \theta) , \quad z_n = (x_1, \cdots, x_n) \in X^n , \quad \theta \in \Theta .$$

Now we can present a result on asymptotic sufficiency of the statistic

$$\begin{aligned} T_{n,k} &= T_n , & k&=1 , \\ &= (T_n, G_n^{(2)}(z_n, T_n), \cdots, G_n^{(k)}(z_n, T_n)) , & k&\geqq 2 , \end{aligned}$$

where T_n, $n \in N$, is a sequence of asymptotic m.l. estimators.

Theorem 2. *Assume that Conditions A, B_r, $C_{1,r}$ and $C_{k,r}$ hold for some $k \in N$ and $r>0$. Let T_n, $n \in N$, be a sequence of asymptotic m.l. estimators of order $O(n^{-r/2})$. Then there exists a sequence of families of probability measures $\{Q_{n,\theta}^k ;\ \theta \in \Theta\}$, $n \in N$, such that*

(a) *for each $n \in N$, $T_{n,k}$ is sufficient for $\{Q_{n,\theta}^k ;\ \theta \in \Theta\}$*

(b) *for every compact $K \subset \Theta$*

$$\sup_{\theta\in K}\|P_{n,\theta}-Q^k_{n,\theta}\| = O(n^{-r/2}), \qquad \text{if } r<k,$$
$$= O(n^{-k/2}), \qquad \text{if } r\geqq k.$$

Proof. Let

$$U_{n,\theta}=\{z_n\in X^n;\, n^{1/2}\|T_n-\theta\|<(\log n)^{\pi_1}\},$$
$$V_{n,r}=\{z_n\in X^n;\, n^{r/2}\|\sum_{i=1}^n g^{(1)}(x_i, T_n)\|<(\log n)^{\pi_2}\},$$
$$W_{n,r}=\{z_n\in X^n;\, \sup_{n^{1/2}\|T_n-\tau\|\leqq(\log n)^{\pi_1}}\|\sum_{i=1}^n [g^{(k+1)}(x_i,\tau)-E_\tau(g^{(k+1)}(\cdot,\tau))]\|<\psi(n,r)\},$$

where $\psi(n, r)$ is the same as in Lemma 1. We define

$$\bar{q}_{n,k}(z_n,\theta)=I_{U_{n,\theta}\cap V_{n,r}\cap W_{n,r}}(z_n)\exp\{G_n(z_n, T_n)+\sum_{m=2}^k\frac{1}{m!}G_n^{(m)}(z_n, T_n)(\theta-T_n)^m$$
$$+\frac{1}{(k+1)!}[E_\theta(G_n^{(k+1)}(\cdot,\theta))](\theta-T_n)^{k+1}\},$$

$$q_{n,k}(z_n,\theta)=v_n(\theta)\bar{q}_{n,k}(z_n,\theta),$$

where $v_n(\theta)=[\int_{X^n}\bar{q}_{n,k}(z_n,\theta)d\mu_n]^{-1}$. Here and hereafter $I_U(\cdot)$ means the indicator function of a set U. For $\theta\in\Theta$ and $n\in N$, we denote by $\bar{Q}^k_{n,\theta}$ and $Q^k_{n,\theta}$ the measures given by

$$\frac{d\bar{Q}^k_{n,\theta}}{d\mu_n}=\bar{q}_{n,k} \quad\text{and}\quad \frac{dQ^k_{n,\theta}}{d\mu_n}=q_{n,k}.$$

Then it follows from the factorization theorem that for each $n\in N$, $T_{n,k}$ is sufficient for $\{Q^k_{n,\theta};\theta\in\Theta\}$.

In order to prove the second assertion (b) we fix a compact subset K of Θ. Using the Taylor expansion

$$G_n(z_n,\theta)=G_n(z_n, T_n)+\sum_{m=1}^k\frac{1}{m!}G_n^{(m)}(z_n, T_n)(\theta-T_n)^m$$
$$+\frac{1}{(k+1)!}G_n^{(k+1)}(z_n, T_n^*)(\theta-T_n)^{k+1}$$

where $\max\{\|T_n^*-\theta\|, \|T_n^*-T_n\|\}\leqq\|T_n-\theta\|$, we have for $z_n\in U_{n,\theta}\cap V_{n,r}\cap W_{n,r}$

$$\left|\log\frac{\bar{q}_{n,k}(z_n,\theta)}{p_n(z_n,\theta)}\right| \tag{3.5}$$
$$\leqq\|G_n^{(1)}(z_n, T_n)\|\,\|T_n-\theta\|+\frac{1}{(k+1)!}\|G_n^{(k+1)}(z_n, T_n^*)$$
$$-E_\theta(G_n^{(k+1)}(\cdot,\theta))\|\,\|T_n-\theta\|^{k+1}.$$

Since

$$\begin{aligned}
&||G_n^{(k+1)}(z_n, T_n^*)-E_\theta(G_n^{(k+1)}(\cdot, \theta))|| \\
&\leqq ||G_n^{(k+1)}(z_n, T_n^*)-[E_\tau(G_n^{(k+1)}(\cdot, \tau))]_{\tau=T_n^*}|| \\
&\quad +||[E_\tau(G_n^{(k+1)}(\cdot, \tau))]_{\tau=T_n^*}-E_\theta(G_n^{(k+1)}(\cdot, \theta))||,
\end{aligned}$$

it follows from (3.5) and condition (iii) in $C_{k,r}$ that for $n\geqq n_K$, $\theta\in K$ and $z_n\in U_{n,\theta}\cap V_{n,r}\cap W_{n,r}$

$$\left|\log \frac{\bar{q}_{n,k}(z_n, \theta)}{p_n(z_n, \theta)}\right| \leqq n^{-(r+1)/2}(\log n)^{\pi_1+\pi_2}+n^{-(k+1)/2}(\log n)^{(k+1)\pi_1}\psi(n, r). \tag{3.6}$$

This implies that for $n\geqq n_K$, $\theta\in K$ and $z_n\in U_{n,\theta}\cap V_{n,r}\cap W_{n,r}$

$$\left|\log \frac{\bar{q}_{n,k}(z_n, \theta)}{p_n(z_n, \theta)}\right| \leqq \log 2. \tag{3.7}$$

Using the inequality $|1-\exp(x)|\leqq 2|x|$ for $|x|\leqq\log 2$, then from (3.7) we have for $n\geqq n_K$ and $\theta\in K$

$$\begin{aligned}
&||P_{n,\theta}-\bar{Q}_{n,\theta}^k|| \\
&\leqq \int_{U_{n,\theta}\cap V_{n,r}\cap W_{n,r}}\left|1-\frac{\bar{q}_{n,k}(z_n, \theta)}{p_n(z_n, \theta)}\right| dP_{n,\theta}+P_{n,\theta}\{(U_{n,\theta}\cap V_{n,r}\cap W_{n,r})^c\} \\
&\leqq 2E_\theta\left[\left|\log \frac{\bar{q}_{n,k}(\cdot, \theta)}{p_n(\cdot, \theta)}\right| I_{U_{n,\theta}\cap V_{n,r}\cap W_{n,r}}(\cdot)\right]+P_{n,\theta}\{(U_{n,\theta}\cap V_{n,r}\cap W_{n,r})^c\}.
\end{aligned} \tag{3.8}$$

By the properties (α_r), (β_r) and Lemma 1

$$\sup_{\theta\in K} P_{n,\theta}\{(U_{n,\theta}\cap V_{n,r}\cap W_{n,r})^c\} = O(n^{-r/2}). \tag{3.9}$$

Then it is obvious that the assertion (b) holds for the case $1\leqq r<k$ and for the case $k\geqq 2$ and $0<r<1$ because of (3.6), (3.8) and (3.9). It remains to prove the assertion (b) for the case $r\geqq k$ and for the case $k=1$ and $0<r<1$.

In the case $r\geqq k$, we shall show that the first term on the right side of (3.8) has upper bound of order $O(n^{-k/2})$. Because of condition (ii) in $C_{k,r}$, choose $d_K>0$, $\lambda_K(x)$ and $D_{n,\theta,K}$ as in the proof of Lemma 1. Let

$$M_{n,\theta} = \{z_n\in X^n;\ ||T_n-\theta-n^{-1}\sum_{i=1}^n g^{(1)}(x_i, \theta)J(\theta)^{-1}||<n^{-1}(\log n)^{2\pi_1+1/2}\}.$$

According to Lemma 2

$$\sup_{\theta\in K} P_{n,\theta}\{(M_{n,\theta})^c\} = O(n^{-k/2}). \tag{3.10}$$

We must again estimate the second term on the right side of (3.5). Since for $\theta\in K$ and $z_n\in M_{n,\theta}$

$$||T_n-\theta||<\rho_K n^{-1}||\sum_{i=1}^n g^{(1)}(x_i, \theta)||+n^{-1}(\log n)^{2\pi_1+1/2}$$

with $\rho_K=\sup_{\theta\in K}||J(\theta)^{-1}||$, it follows from Minkowski's inequality that

$$[E_\theta(||T_n-\theta||^{k+2}I_{M_{n,\theta}}(\cdot))]^{1/(k+2)}\leqq\rho_K n^{-1}[E_\theta(||\sum_{i=1}^n g^{(1)}(\cdot,\theta)||^{k+2})]^{1/(k+2)} +n^{-1}(\log n)^{2\pi_1+1/2}.$$

(2.10) with Condition B_r implies that for $\theta\in K$

$$E_\theta(||\sum_{i=1}^n g^{(1)}(\cdot,\theta)||^{k+2})\leqq c_K n^{(k+2)/2},$$

which leads to

$$[E_\theta(||T_n-\theta||^{k+2}I_{M_{n,\theta}}(\cdot))]^{1/(k+2)}\leqq c_K n^{-1/2}. \tag{3.11}$$

Thus we have for $n\geqq n_K$ and $\theta\in K$

$$\begin{aligned}&E_\theta(||G_n^{(k+1)}(\cdot,T_n^*)-G_n^{(k+1)}(\cdot,\theta)||\,||T_n-\theta||^{k+1}I_{U_{n,\theta}\cap D_{n,\theta,K}\cap M_{n,\theta}}(\cdot))\\ &\leqq(1+\sup_{\theta\in K}E_\theta(\lambda_K(\cdot)))nE_\theta(||T_n-\theta||^{k+2}I_{M_{n,\theta}}(\cdot))\leqq c_K n^{-k/2}.\end{aligned} \tag{3.12}$$

By Hölder's inequality

$$\begin{aligned}&E_\theta(||G_n^{(k+1)}(\cdot,\theta)-E_\theta(G_n^{(k+1)}(\cdot,\theta))||\,||T_n-\theta||^{k+1}I_{M_{n,\theta}}(\cdot))\\ &\leqq[E_\theta(||G_n^{(k+1)}(\cdot,\theta)-E_\theta(G_n^{(k+1)}(\cdot,\theta))||^{k+2})]^{1/(k+2)}[E_\theta(||T_n-\theta||^{k+2}I_{M_{n,\theta}})]^{(k+1)/(k+2)},\end{aligned}$$

so that (2.10) with condition (ii) (c) in $C_{k,r}$ and (3.11) imply that for $\theta\in K$

$$E_\theta(||G_n^{(k+1)}(\cdot,\theta)-E_\theta(G_n^{(k+1)}(\cdot,\theta))||\,||T_n-\theta||^{k+1}I_{M_{n,\theta}}(\cdot))\leqq c_K n^{-k/2}. \tag{3.13}$$

Taking account of (3.7), we obtain for $n\geqq n_K$ and $\theta\in K$

$$\begin{aligned}&E_\theta\left(\left|\log\frac{\bar{q}_{n,k}(\cdot,\theta)}{p_n(\cdot,\theta)}\right|I_{U_{n,\theta}\cap V_{n,r}\cap W_{n,r}}(\cdot)\right)\\ &\leqq E_\theta\left(\left|\log\frac{\bar{q}_{n,k}(\cdot,\theta)}{p_n(\cdot,\theta)}\right|I_{U_{n,\theta}\cap V_{n,r}\cap W_{n,r}\cap D_{n,\theta,K}\cap M_{n,\theta}}(\cdot)\right)\\ &\quad+(\log 2)P_{n,\theta}\{(D_{n,\theta,K}\cap M_{n,\theta})^c\}.\end{aligned}$$

Thus, the first term on the right side of (3.8) has upper bound of order $O(n^{-k/2})$ because of (3.1), (3.5), (3.10), (3.12) and (3.13).

This, together with (3.8) and (3.9), implies that

$$\sup_{\theta\in K}||P_{n,\theta}-\bar{Q}_{n,\theta}^k||=O(n^{-k/2}).$$

Since

$$\begin{aligned}\sup_{\theta\in K}|1-v_n(\theta)^{-1}|&=\sup_{\theta\in K}|P_{n,\theta}\{X^n\}-\bar{Q}_{n,\theta}^k\{X^n\}|\\ &=O(n^{-k/2}),\end{aligned}$$

we have

$$\begin{aligned}\sup_{\theta\in K}\|P_{n,\theta}-Q_{n,\theta}^{k}\| &\leqq \sup_{\theta\in K}\|P_{n,\theta}-\bar{Q}_{n,\theta}^{k}\|+\sup_{\theta\in K}\|\bar{Q}_{n,\theta}^{k}-Q_{n,\theta}^{k}\| \\ &\leqq \sup_{\theta\in K}\|P_{n,\theta}-\bar{Q}_{n,\theta}^{k}\|+\sup_{\theta\in K}|1-v_n(\theta)^{-1}| \\ &= O(n^{-k/2}),\end{aligned}$$

which is the desired result.

In the case $k=1$ and $0<r<1$, $M_{n,\theta}$ is replaced by the following set $M_{n,\theta,r}$

$$M_{n,\theta,r}=\{z_n\in X^n;\ \|T_n-\theta-n^{-1}\sum_{i=1}^{n}g^{(1)}(x_i,\theta)J(\theta)^{-1}\|<n^{-(r+1)/2}(\log n)^{2\pi_1}\}.$$

Then, a similar argument shows that

$$\sup_{\theta\in K}\|P_{n,\theta}-Q_{n,\theta}^{1}\|=O(n^{-r/2}).$$

This completes the proof.

REMARK. (1) If $r\geqq k$, it is possible to choose $Q_{n,\theta}^{k}$ independent of r because $V_{n,r}$ and $W_{n,r}$ in the definition of $\bar{q}_{n,k}$ can be replaced by $V_{n,k}$ and $W_{n,k}$, respectively.

(2) In the case $k=1$, it follows from Theorem 2 that a sequence of asymptotic m.l. estimators of order $O(n^{-r/2})$ is asymptotically sufficient up to order $O(n^{-r/2})$ if $0<r<1$ and $O(n^{-1/2})$ if $r=1$. The latter result has been already shown by Pfanzagl [10] under similar circumstances to ours.

(3) Michel [7] showed that $T_{n,k}$, $k\geqq 3$, constructed by asymptotic m.l. estimators of order $o(n^{-(k-2)/2})$ is asymptotically sufficient up to order $o(n^{-(k-2)/2})$. According to Theorem 2, the convergence order concerning asymptotic sufficiency of $T_{n,k}$ can be improved up to $O(n^{-k/2})$ if $\{T_n\}$ is a sequence of asymptotic m.l. estimators with higher order than Michel's one.

(4) In [14], [15] Suzuki assumes the existence of moment generating function of $g^{(k+1)}(x,\theta)$ to evaluate probability of large deviations. Of course this condition is stronger than ours.

4. Properties of m.l. estimators. We shall investigate conditions under which a sequence of m.l. estimators has the properties (α_r) and (β_r) for some $r>0$.

Let $\bar{\Theta}$ denote the closure of Θ in $\bar{\boldsymbol{R}}^s=[-\infty,\infty]^s$. Assume that $g(\cdot,\theta)$: $X\to\boldsymbol{R}$, $\theta\in\Theta$, admits a measurable extension $g(\cdot,\theta)$: $X\to\bar{\boldsymbol{R}}$, $\theta\in\bar{\Theta}$.

Condition A^*

(i) $E_\theta(g(\cdot,\tau))<E_\theta(g(\cdot,\theta))$ for all $\theta\in\Theta$, $\tau\in\bar{\Theta}$, $\theta\neq\tau$.

(ii) For every $x\in X$, $\theta\to g(x,\theta)$ is continuous on $\bar{\Theta}$.

Condition B_r^*

(i) For every $\theta \in \Theta$ and every compact $K \subset \Theta$

$$\sup_{\tau \in K} E_\tau(|g(\cdot, \theta)|^{(r+2)/2}) < \infty .$$

(ii) For every $\theta \in \bar{\Theta}$ there exists a neighborhood U_θ of θ such that for every neighborhood U of θ, $U \subset U_\theta$, and every compact $K \subset \Theta$

$$\sup_{\tau \in K} E_\tau(|\sup_{\sigma \in U} g(\cdot, \sigma)|^{(r+2)/2}) < \infty .$$

(iii) For each $x \in X$, $\theta \to g(x, \theta)$ admits continuous partial derivatives up to the order 2 on Θ. For every $\theta \in \Theta$ there exist a neighborhood U_θ of θ and a measurable function $\lambda(x, \theta)$ such that

(a) for all $x \in X$, $\tau, \sigma \in U_\theta$, $||g^{(2)}(x, \tau) - g^{(2)}(x, \sigma)|| \leqq ||\tau - \sigma|| \lambda(x, \theta)$

(b) for every compact $K \subset \Theta$, $\sup_{\tau \in K} E_\tau(\lambda(\cdot, \theta)^{(r+2)/2}) < \infty$

(c) $\sup_{\tau \in U_\theta} E_\tau(||g^{(2)}(\cdot, \tau)||^{(r+2)/2}) < \infty$.

(iv) $\theta \to J(\theta)$ is continuous on Θ.

A *maximum likelihood estimator* for the sample size n is an estimator T_n for which $T_n \in \bar{\Theta}$ and

$$\sum_{i=1}^{n} g(x_i, T_n) = \sup_{\theta \in \bar{\Theta}} \sum_{i=1}^{n} g(x_i, \theta) .$$

Condition (ii) in A^* insures that m.l. estimators for the sample size n exist. Let $\hat{T}_n$, $n \in N$, be a sequence of m.l. estimators.

The following lemma can be obtained in a way analogous to the one used in the proof of Lemma 4 in Michel and Pfanzagl [8] except that Theorem 1 is used instead of Chebyshev's inequality.

Lemma 3. *Let Condition A^* and conditions* (i), (ii) *in B_r^* be satisfied for some $r>0$. Then for every $\varepsilon>0$ and every compact $K \subset \Theta$*

$$\sup_{\theta \in K} P_{n,\theta}\{z_n \in X^n;\ ||\hat{T}_n(z_n) - \theta|| \geqq \varepsilon\} = O(n^{-r/2}) .$$

The following proposition is an immediate consequence of Lemma 3.

Proposition 1. *Let Condition A^* and conditions* (i), (ii) *in B_r^* be satisfied for some $r>0$. Moreover, assume that for each $x \in X$, $\theta \to g(x, \theta)$ is continuously differentiable on Θ. Then for every compact $K \subset \Theta$*

$$\sup_{\theta \in K} P_{n,\theta}\{z_n \in X^n;\ ||\sum_{i=1}^{n} g^{(1)}(x_i, \hat{T}_n(z_n))|| > 0\} = O(n^{-r/2}) .$$

Lemma 4 (cf. Lemma 5 in Michel and Pfanzagl [8]). *Let Condition A^* and conditions* (i)–(iii) *in B_r^* be satisfied for some $r>0$. Then for every $\delta>0$*

and every compact $K\subset\Theta$ *there exists* $d>0$ *such that*

$$\sup_{\theta\in K} P_{n,\theta}\{z_n\in X^n;\ \sup_{\|\hat{T}_n-\tau\|\leqq d} \|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\tau)+J(\theta)]\|\geqq\delta\} = O(n^{-r/2})\,.$$

Proof. Let $\delta>0$ be given and K be a compact subset of Θ. By condition (iii) in B_r^* we may choose $d_K>0$, $\lambda_K(x)$ and $D_{n,\theta,K}$ as in Lemma 1 with $k=1$. We write

$$F_{n,\theta,\delta} = \{z_n\in X^n;\ \|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\theta)+J(\theta)]\|<\delta/2\}\,.$$

From condition (iii) (c) in B_r^* it follows that

$$\sup_{\theta\in K} P_{n,\theta}\{(F_{n,\theta,\delta})^c\} = O(n^{-r/2})\,.$$

Taking $2d=\min\{d_K,\ \delta/[2(1+\sup_{\theta\in K} E_\theta(\lambda_K(\cdot)))]\}$, we see that for $z_n\in D_{n,\theta,K}\cap F_{n,\theta,\delta}$, $\|\hat{T}_n-\theta\|<d$ and $\|\hat{T}_n-\tau\|\leqq d$

$$\|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\tau)+J(\theta)]\|\leqq\|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\tau)-g^{(2)}(x_i,\theta)]\|$$
$$+\|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\theta)+J(\theta)]\|<\delta\,.$$

This together with Lemma 3 implies the desired assertion.

Lemma 3 and Lemma 4 yield the following proposition.

Proposition 2 (cf. Lemma 6 in [8] and Lemma 3 in Pfanzagl [12]). *Assume that Conditions* A, A^*, B_r *and* B_r^* *are fulfilled for some* $r>0$. *Then for every compact* $K\subset\Theta$ *there exists* $c_K>0$ *such that*

$$\sup_{\theta\in K} P_{n,\theta}\{z_n\in X^n;\ n^{1/2}\|\hat{T}_n(z_n)-\theta\|\geqq c_K(\log n)^{1/2}\} = O(n^{-r/2})\,.$$

Proof. Let K be a fixed compact subset of Θ. It follows from conditions (ii) (b) in A and (iv) in B_r^* that there exists $\delta_K>0$ such that $\theta\in K$ and matrix J with $\|J-J(\theta)\|<\delta_K$ imply that J is regular and $\|J^{-1}-J(\theta)^{-1}\|<1$. Let

$$W_{n,\theta}^* = \{z_n\in X^n;\ \sup_{\|\hat{T}_n-\tau\|\leqq d_K} \|n^{-1}\sum_{i=1}^{n}[g^{(2)}(x_i,\tau)+J(\theta)]\|<\delta_K\}\,,$$

where $d_K>0$ is chosen to satisfy that

$$\sup_{\theta\in K} P_{n,\theta}\{(W_{n,\theta}^*)^c\} = O(n^{-r/2})$$

because of Lemma 4. Choose $e_K>0$ such that $e_K\leqq d_K$ and $\{\tau\in \boldsymbol{R}^s;\ \inf_{\theta\in K}\|\theta-\tau\|\leqq e_K\}\subset\Theta$. Let

$$U_{n,\theta}^{*}=\{z_n\in X^n;\ ||\hat{T}_n-\theta||<e_K\}\ .$$

In view of Lemma 3 we have

$$\sup_{\theta\in K} P_{n,\theta}\{(U_{n,\theta}^{*})^c\}=O(n^{-r/2})\ .$$

Since for $\theta\in K$ and $z_n\in U_{n,\theta}^{*}$

$$\sum_{i=1}^{n} g^{(1)}(x_i,\theta)=(\theta-\hat{T}_n)\sum_{i=1}^{n} g^{(2)}(x_i,\hat{T}_n,\theta)\ ,$$

it follows that for $\theta\in K$ and $z_n\in U_{n,\theta}^{*}\cap W_{n,\theta}^{*}$

$$\begin{aligned}||n^{1/2}(\hat{T}_n-\theta)||&\leqq||n^{-1/2}\sum_{i=1}^{n} g^{(1)}(x_i,\theta)||\ ||(-n^{-1}\sum_{i=1}^{n} g^{(2)}(x_i,\hat{T}_n,\theta))^{-1}||\\&\leqq(1+\sup_{\theta\in K}||J(\theta)^{-1}||)n^{-1/2}||\sum_{i=1}^{n} g^{(1)}(x_i,\theta)||\ .\end{aligned}$$

In order to complete the proof it is enough to note that there exists $c_K>0$ such that

$$\sup_{\theta\in K} P_{n,\theta}\{z_n\in X^n;\ ||\sum_{i=1}^{n} g^{(1)}(x_i,\theta)||\geqq c_K(n\log n)^{1/2}\}=o(n^{-r/2})\ .$$

This follows from Theorem 1, condition (ii) (a) in A and Condition B_r.

REMARK. (1) Proposition 2 remains to hold for a sequence of minimum contrast estimators with obvious modification.

(2) If every $(r+2)/2$ in Condition B_r^* is replaced by a number greater than it, then Proposition 2 holds with $o(n^{-r/2})$ instead of $O(n^{-r/2})$.

(3) Proposition 2 improves Lemma 3 of Pfanzagl [12] in the following sense:

(a) This result still holds for $0<r<1$.

(b) In the case $r\geqq 1$, the moment conditions used in Proposition 2 are weaker than in [12] because of the use of Theorem 1 instead of Lemma 2 of [12] (see Remark (2) of Theorem 1).

From Theorem 2, Proposition 1 and Proposition 2, the following theorem is immediate.

Theorem 3. *Assume that Conditions A, A^*, B_r, (i), (ii) in B_r^*, $C_{1,r}$ and $C_{k,r}$ are fulfilled for some $k\in N$ and $r>0$. Then, $\hat{T}_{n,k}=(\hat{T}_n, G_n^{(2)}(z_n,\hat{T}_n),\cdots, G_n^{(k)}(z_n,\hat{T}_n))$ is asymptotically sufficient up to order $O(n^{-r/2})$ if $r<k$ and $O(n^{-k/2})$ if $r\geqq k$. Here $\hat{T}_{n,1}$ means $\hat{T}_n$.*

It is remarked that we need the $(2+r)$-th absolute moment of $g^{(1)}$ and the $(2+r)/(2-r)$-th absolute moment of $g^{(2)}$ in order to show that a sequence of m.l. estimators is asymptotically sufficient up to order $O(n^{-r/2})$ with $0<r\leqq 1$.

Acknowledgment. The author wishes to express his hearty thanks to Professor Takeru Suzuki for informing me of the literature [10].

References

[1] B. von Bahr and C.-G. Esseen: *Inequalities for the rth absolute moment of a sum of random variables*, $1\leqq r\leqq 2$, Ann. Math. Statist. **36** (1965), 299–303.

[2] K.L. Chung: *The strong law of large numbers,* Proceedings of Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press, 1951, 341–352.

[3] J.K. Ghosh and K. Subramanyam: *Second order efficiency of maximum likelihood estimators,* Sankhyā Ser. A **36** (1974), 325–358.

[4] L. LeCam: *On the asymptotic theory of estimation and testing hypotheses,* Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press, 1956, Vol. 1, 129–156.

[5] J. Marcinkiewicz and A. Zygmund: *Sur les fonctions indépendantes,* Fund. Math. **29** (1937), 60–90.

[6] R. Michel: *A multidimensional Newton-Raphson method and its applications to the existence of asymptotic F_n-estimators and their stochastic expansions,* J. Multivariate Anal. **7** (1977), 235–248.

[7] R. Michel: *Higher order asymptotic sufficiency,* Sankhyā Ser. A **40** (1978), 76–84.

[8] R. Michel and J. Pfanzagl: *The accuracy of the normal approximation for minimum contrast estimates,* Z. Wahrsch. Verw. Gebiete **18** (1971), 73–84.

[9] S.V. Nagaev: *Some limit theorems for large deviations,* Theor. Prob. Appl. **10** (1965), 214–235.

[10] J. Pfanzagl: *Further results on asymptotic normality* I, Metrika **18** (1972), 174–198.

[11] J. Pfanzagl: *Further results on asymptotic normality* II, Metrika **19** (1973), 89–97.

[12] J. Pfanzagl: *Asymptotic expansions related to minimum contrast estimators,* Ann. Statist. **1** (1973), 993–1026.

[13] J. Pfanzagl: *Asymptotically optimum estimation and test procedures,* Proceedings of the Prague Symposium on Asymptotic Statistics, 1974, Vol. 1, 201–272.

[14] T. Suzuki: *Asymptotic sufficiency up to higher orders and its applications to statistical tests and estimates,* Osaka J. Math. **15** (1978), 575–588.

[15] T. Suzuki: *A remark on asymptotic sufficiency up to higher orders in multi-dimensional parameter case,* Osaka J. Math. **17** (1980), 245–252.

[16] A. Wald: *Tests of statistical hypotheses concerning several parameters when the number of observations is large,* Trans. Amer. Math. Soc. **54** (1943), 426–482.

Faculty of Economics
Wakayama University
Nishitakamatsu
Wakayama 641, Japan