Title	Higher order asymptotic sufficiency of posterior densities with respect to Kullback information
Author(s)	Kato, Takeshi
Citation	Osaka Journal of Mathematics. 1995, 32(2), p. 313-326
Version Type	VoR
URL	https://doi.org/10.18910/7547
rights	
Note	

Kato, T.
Osaka J. Math.
32 (1995), 313–326

HIGHER ORDER ASYMPTOTIC SUFFICIENCY OF POSTERIOR DENSITIES WITH RESPECT TO KULLBACK INFORMATION

TAKESHI KATO

(Received Sept. 8, 1993, Revised Feb. 28, 1994)

1. Introduction

Let $(X, \boldsymbol{B})$ be a measurable space and $\{P_\theta: \theta \in \Theta\}$, $\Theta \subset \boldsymbol{R}$, a family of probability measures on $(X, \boldsymbol{B})$. We suppose that there exists a σ-finite measure μ on $(X, \boldsymbol{B})$ which dominates P_θ for all $\theta \in \Theta$. Then we put $f(x, \theta) = (dP_\theta/d\mu)(x)$. Let $x_1, x_2, \cdots$ be independently and identically distributed in accurdance with $f(x, \theta_0)$ for some $\theta_0 \in \Theta$. We write $\underset{\sim}{x} = (x_1, x_2, \cdots)$ and $\underset{\sim}{x}_n = (x_1, \cdots, x_n)$. Let $(X^N, \boldsymbol{B}^N)$ be the countable Cartesian product of identical components $(X, \boldsymbol{B})$, and P_θ^N the independent product of a countable number of identical components P_θ. Then the definition of the higher order asmptotic sufficiency of posterior densities is as follows.

Definition. Let $k \geqq 1$ be a given integer, $\pi(\theta | \underset{\sim}{x}_n)$ a posterior density function of θ with respect to some prior density function $\rho(\cdot)$, and $\psi(\theta | T_n)$ a probability density function determined by some statistic $T = T_n(\underset{\sim}{x}_n)$. For every $\theta_0 \in \Theta$ and $\varepsilon > 0$ if

$$\lim_{n \to \infty} P_{\theta_0}^N \{\underset{\sim}{x} \in X^N: d(\pi(\theta | \underset{\sim}{x}_n), \psi(\theta | T_n)) \geqq \varepsilon n^{-(k-1)/2}\} = 0,$$

then T_n is called *the k-th order asymptotically sufficient statistic in Bayesianm sense under* $P_{\theta_0}^N$, where $d(\cdot,\cdot)$ is some criterion.

In this definition, it is not required that $\psi(\theta | T_n)$ should be formed out of the density function of T_n parameterized by θ and the prior density function ρ of θ by Bayes' theorem. As a matter of fact, it is too difficult to obtain the density function of T_n in general cases. Therefore, if there exists a density function approximating $\pi(\theta | \underset{\sim}{x}_n)$ and the observations $\underset{\sim}{x}_n = (x_1, \cdots, x_n)$ affect it through only a statistic $T_n = T_n(\underset{\sim}{x}_n)$, then we define T_n as the asymptotically sufficient statistic.

When $d(\cdot,\cdot)$ is total variation, it can be shown from the results in Ghosh et al [2] or Johnson [3] that $T_n^{(k)} = (\hat{\theta}_n, \bar{l}_{n2}(\hat{\theta}_n), \cdots, \bar{l}_{nk}(\hat{\theta}_n))$ $(k \geqq 2)$ is the k-th order asymptotically sufficient statistic in Bayesian sense, where $\hat{\theta}_n = \hat{\theta}_n(\underset{\sim}{x}_n)$ is the m.l.e. of θ and

$$\bar{I}_{nj}(\hat{\theta}_n)=n^{-1}\sum_{i=1}^{n}(\partial^j/\partial\theta^j)\log f(x_i,\theta)|_{\theta=\hat{\theta}_n}\ (j=2,\cdots,k).$$

And the idea in the case where $d(\cdot,\cdot)$ is Kullback information was suggested in Takeuchi [6] first, and the detailed argument has been done in Kato [4] in the second order case. According to them, the m.l.e. $\hat{\theta}_n$ is second order asymptotically sufficient with respect to Kullback information. This fact should be noted, because the second order can be attained by using $(\hat{\theta}_n, \bar{I}_{n2}(\hat{\theta}_n))$ with respect to total variation. An interesting problem arises from this difference, which is how high order the statistic $T_n^{(k)}$ attains when $d(\cdot,\cdot)$ is Kullback information. When it comes to this point, detailed description cannot be seen in [6]. Then the purpose of this paper is not only to solve it clearly but also to give the precise evaluation of the probability of the exceptional set.

2. Assumptions and the main theorem

Let Θ be an open interval in $\boldsymbol{R}=(-\infty, +\infty)$ and $\bar{\Theta}$the closure of Θ in $\bar{\boldsymbol{R}}=[-\infty, +\infty]$. In addition, let $k(\geqq 2)$ be a given integer and $r\geqq 1$ a real number, then we impose the following assumptions.

(A_1) There exist functions $\bar{f}(\cdot,\theta):X\to\bar{\boldsymbol{R}}$, $\theta\in\bar{\Theta}$, such that the following hold.
  (1) $\bar{f}(x, \theta)=f(x, \theta)$ for all $\theta\in\Theta$ and all $x\in X$.
  (2) $\bar{f}(x, \theta)$ is continous with respect to $\theta\in\bar{\Theta}$ for every $x\in X$.
  (3) For every $\theta\in\Theta$ and $t\in\bar{\Theta}$, provided $\theta\neq t$,

$$\int_X |\bar{f}(x, \theta)-\bar{f}(x, t)|d\mu(x)>0.$$

  (4) For every $\theta\in\Theta$ and $t\in\bar{\Theta}$, there exists a neighborhood V_t of t such that for all neighborhoods $V\subset V_t$ of t,

$$E_\theta[|\sup\{\log\bar{f}(x, \tau)\colon \tau\in V\}|^r]<+\infty.$$

(A_2) For every $\theta\in\Theta$, $E_\theta[|\log f(x, \theta)|^r]<+\infty$.
(A_3) For each is $x\in X, f(x, \theta)$ is $k+1$ times continuously differentiable with respect to θ in Θ.

For simplicity, we use a notation

$$(\partial^j/\partial\theta^j)\log f(x, t)=(\partial^j/\partial\theta^j)\log f(x, \theta)|_{\theta=t}\quad (j=1,\cdots,k+1).$$

(A_4) For every $\theta\in\Theta$,

$$E_\theta[|(\partial^j/\partial\theta^j)\log f(x, \theta)|^r]<+\infty\quad (j=1,\cdots,k),$$

and there exist a neighborhood U_θ of θ and a measurable function $G(x, \theta)$

such that

$$|(\partial^{k+1}/\partial\theta^{k+1})\log f(x,\ t)|\leqq G(x,\ \theta)\qquad \text{for all } t\in U_\theta \text{ and all } x\in X,$$

and also $E_\theta[G(x,\ \theta)^r]<+\infty$.

(A_5) For every $\theta\in\Theta$, there exist a neighborhood U_θ of θ and a measurable function $H(x,\ \theta)$ such that

$$|(\partial^{k+1}/\partial\theta^{k+1})\log f(x,\ t)-(\partial^{k+1}/\partial\theta^{k+1})\log f(x,\ \tau)|\leqq|t-\tau|H(x,\ \theta)\qquad \text{for all } t,\ \tau\in U_\theta \text{ and all } x\in X,$$

and also $E_\theta[H(x,\ \theta)]<+\infty$.

(A_6) For every $\theta\in\Theta$,

$$I(\theta)=-E_\theta[(\partial^2/\partial\theta^2)\log f(x,\ \theta)]>0.$$

(A_7) The function

$$\lambda_{k+1}(\theta)=E_\theta[(\partial^{k+1}/\partial\theta^{k+1})\log f(x,\ \theta)]$$

is continuous in Θ.

Let $\rho(\cdot)$ be a prior density function of θ with respect to Lebesgue measure.

(A_8) $\int_{-\infty}^{+\infty}\theta^2\ \rho(\theta)d\theta<+\infty$.

(A_9) For every $\theta\in\Theta$, the function $g(x)=\int_{-\infty}^{+\infty}\rho(\tau)|\log f(x,\ \tau)|d\tau$ satisfies $E_\theta[g(x)^r]$ $<+\infty$.

When observations $\underset{\sim}{x}_n=(x_1,\cdots,x_n)$ are given, let a $\boldsymbol{B}^N$-measurable function $\hat{\theta}_n=\hat{\theta}_n(\underset{\sim}{x}_n)$ be the m.l.e. defined by

$$\prod_{i=1}^n \bar{f}(x_i,\ \hat{\theta}_n)=\sup\{\prod_{i=1}^n \bar{f}(x_i,\ \theta):\ \theta\in\bar{\Theta}\}.$$

For given $\hat{\theta}_n\in\bar{\Theta}$, we put

$$\text{(2.1)}\quad \pi(\xi|\underset{\sim}{x}_n)=\begin{cases}\rho(\hat{\theta}_n+\xi/\sqrt{n})\prod_{i=1}^n f(x_i,\ \hat{\theta}_n+\xi/\sqrt{n}) \\ \times[\int_{-\infty}^{+\infty}\rho(\hat{\theta}_n+\xi/\sqrt{n})\prod_{i=1}^n f(x_i,\ \hat{\theta}_n+\xi/\sqrt{n})d\xi]^{-1} & (\hat{\theta}_n\in\Theta) \\ 0 & (\hat{\theta}_n\in\bar{\Theta}\backslash\Theta),\end{cases}$$

where we assume that the range of ξ is restricted to the set $\{\xi\in\boldsymbol{R}|\hat{\theta}_n+\xi/\sqrt{n}\in\Theta\}$

whenever we write $f(x_i, \hat{\theta}_n + \xi/\sqrt{n})$ for $\hat{\theta}_n \in \Theta$. When $\hat{\theta}_n \in \Theta$, $\pi(\xi | \underset{\sim}{x}_n)$ is a centered and scaled posterior density of θ with respect to a prior density function ρ with a transformation $\xi = \sqrt{n}(\theta - \hat{\theta}_n)$. Formally we define $0/0 = 0$ and $0(-\infty) = 0$. On these assumptions, our final aim is to show the following theorem.

Theorem. *Let $\theta_0 \in \Theta$ such that $\rho(\theta_0) > 0$ and assumptions* (A_1) *through* (A_9) *be all satisfied for some $r \geqq 1$. Moreover if ρ is continuous in a neighborhood of θ_0, then for every $\varepsilon > 0$, there exists an integer $n_0 = n_0(\varepsilon, k) \geqq 1$ such that*

$$P_{\theta_0}^N\{\underset{\sim}{x} \in X^N \colon \int_{-\infty}^{+\infty} \pi(\xi | \underset{\sim}{x}_n) \log\{\pi(\xi | \underset{\sim}{x}_n)/\psi_k(\xi | T_n^{(k)})\} d\xi \geqq \varepsilon n^{-(k-1)}\}$$

$$< \varepsilon n^{1-r} \quad \textit{for all } n \geqq n_0,$$

where $T_n^{(k)} = (\hat{\theta}_n, \bar{l}_{n2}(\hat{\theta}_n), \cdots, \bar{l}_{nk}(\hat{\theta}_n))$ and $\psi_k(\xi | T_n^{(k)})$ has the form

$$\text{(2.2)} \qquad \psi_k(\xi | T_n^{(k)}) = \begin{cases} R_n(\xi) \Big/ \int_{-\infty}^{+\infty} R_n(\xi) d\xi & (\hat{\theta}_n \in \Theta) \\ 0 & (\hat{\theta}_n \in \bar{\Theta} \setminus \Theta), \end{cases}$$

$$R_n(\xi) = \begin{cases} \rho(\hat{\theta}_n + \xi/\sqrt{n}) \\ \qquad \times \exp\{\bar{l}_{n2}(\hat{\theta}_n)\xi^2/2 + \chi_{\delta n}(\xi)\lambda_3(\hat{\theta}_n)\xi^3/(6\sqrt{n})\} & (k=2) \\ \rho(\hat{\theta}_n + \xi/\sqrt{n}) \exp[\bar{l}_{n2}(\hat{\theta}_n)\xi^2/2 \\ \qquad + \chi_{\delta n}(\xi)\{\sum_{j=3}^k \bar{l}_{nj}(\hat{\theta}_n)\xi^j n^{-(j-2)/2}/j! \\ \qquad + \lambda_{k+1}(\hat{\theta}_n)\xi^{k+1} n^{-(k-1)/2}/(k+1)!\}] & (k \geqq 3), \end{cases}$$

$\chi_{\delta n}(\xi) = 1$ if $|\xi| < \delta\sqrt{n}$, $= 0$ if $|\xi| \geqq \delta\sqrt{n}$, and $\delta = \delta(\varepsilon, k) > 0$ is sufficiently small.

We know from this theorem that the statistic $T_n^{(k)}$ is $(2k-1)$-th order asymptotically sufficient with respect to Kullback information in Bayesian sense under $P_{\theta_0}^N$. And this assertion holds with probability one by the Borel Cantelli Lemma if the assumptions are satisfied for some $r > 2$.

REMARK 1. When $\lambda_{k+1}(\hat{\theta}_n)$ is omitted from $R_n(\xi)$, it can be shown that the bound $\varepsilon\, n^{-(k-1)}$ of the approximation to $\pi(\xi | \underset{\sim}{x}_n)$ by $\psi_k(\xi | T_n^{(k)})$ with respect to Kullback information should be replaced by $Mn^{-(k-1)}$ for some $M > 0$, and the probability of the exceptional set is εn^{1-r}, which is the same as the theorem. On the other hand, if we assume furthermore that $\lambda_{k+1}(\theta)$ is continuously differentiable in Θ, then it can be proved that $\psi_k(\xi | T_n^{(k)})$ approximates $\pi(\xi | \underset{\sim}{x}_n)$ up to order $O(n^{-k})$ with respect to Kullback information with probability approaching one. In this case, it is impossible to evaluate the probability of the exceptional set with order of n.

REMARK 2. The indicator function $\chi_{\delta n}$ used in the definition of ψ_k can be replaced by $\chi_{\delta n}(\xi)=1$ if $|\xi|<\delta\ n^{\alpha},=0$ if $|\xi|\geqq\delta n^{\alpha}$, $0<\alpha\leqq 1/2$. The smaller α is, the more simply the form of ψ_k becomes.

REMARK 3. It is a little bit difficult to make sure of the assumption (A_9). Let $\Theta_0=\{\theta:\rho(\theta)>0\}$ be an open interval in $\boldsymbol{R}$ and $\bar{\Theta}_0\subset\Theta$. Then (A_9) holds if for every $\theta\in\Theta$

$$E_\theta[[\sup\{|\log f(x,t)|:t\in\Theta_0\}]^r]<+\infty.$$

REMARK 4. Suppose that the density function $f(x,\theta)$ can be written in the form

$$f(x,\theta)=\exp\{\textstyle\sum_{j=1}^{L}\sum_{m=1}^{M_j}u_{jm}(x)h_{jm}(\theta)\}$$

and the following conditions are satisfied.

(i) Every $u_{jm}(x)$ is a measurable function on $(X, \boldsymbol{B})$.

(ii) For every $\theta\in\Theta$ and j, m, $E_\theta[|u_{jm}(x)|^r]<+\infty$.

(iii) For every j, m, $E_\theta[u_{jm}(x)]$ is continuous in Θ as a function of θ

(iv) For every j, m, $h_{jm}(\theta)$ is $k+2$ times continuously differentiable in Θ

(v) For every $\theta\in\Theta$ and j, m, there exists a neighborhood U_θ of θ such that

$$\sup\{|h_{jm}^{(k+2)}(t)|:t\in U_\theta\}<+\infty,$$

where $h_{jm}^{(l)}$ denotes the l-th derivative of h_{jm}

(vi) $\sum_{j=1}^{L}\sum_{m=1}^{M_j}E_\theta[u_{jm}(x)]h_{jm}^{(1)}(\theta)=0$ for all $\theta\in\Theta$.

(vii) $\sum_{j=1}^{L}\sum_{m=1}^{M_j}E_\theta[u_{jm}(x)]h_{jm}^{(2)}(\theta)<0$ for all $\theta\in\Theta$.

(viii) For every j, m, $\displaystyle\int_{-\infty}^{+\infty}\rho(\theta)|h_{jm}(\theta)|d\theta<+\infty$.

Then assumptions (A_2), (A_3), (A_4), (A_5), (A_6), (A_7) and (A_9) hold.

The following are examples of Remark 4.

EXAMPLE 1. $X=\boldsymbol{R}^2$. $\Theta=(-1,1)$. $x_i=(x_{i1}, x_{i2})$ is normally distributed with $E_\theta[x_{ij}]=0$, $E_\theta[x_{ij}^2]=1$ $(j=1,2)$ and $E_\theta[x_{i1}\ x_{i2}]=\theta$. Namely

$$f(x_i,\ \theta)=\{2\pi(1-\theta^2)^{1/2}\}^{-1}\exp\{-(x_{i1}^2-2\theta x_{i1}x_{i2}+x_{i2}^2)/(2(1-\theta^2))\}.$$

As an prior distribution, we can take a transformed Beta distribution of the form

$$\rho(\theta)=\{2^{1-\alpha-\beta}/B(\alpha,\beta)\}(1+\theta)^{\alpha-1}(1-\theta)^{\beta-1} \ (\alpha,\beta>1/2).$$

In particular, Remark 4 is often applicable to $(l,1)$ curved exponential families.

EXAMPLE 2. $X=\boldsymbol{R}^m$ $(m\geqq 1)$. $x_i=(x_{i1},\cdots,x_{im})$. $\Theta=(0,\infty)$.

$$f(x_i,\theta)=(2\pi)^{-m/2}\ \theta^{-m(m+1)/2}\ \exp\{-\textstyle\sum_{j=1}^{m}(x_{ij}-\theta^j)^2/(2\theta^{2j})\}.$$

This belongs to $(2m, 1)$ curved exponential families, and when $m=1$, this is a well known example of curved exponential families $N(\theta,\theta^2)$. In this case, we can take Gamma distribution, Gamma(α, β) $(\alpha>2m, \beta>0)$, where β is the scale parameter, or F-distribution, $F(\nu_1, \nu_2)$ $(\nu_1>2m, \nu_2>4)$, as a prior distribution.

These two examples also satisfy the rest of the assumptions (A_1), (A_8), which means that they are taken as concrete examples of the theorem.

3. Auxiliary results

The first thing, we quote the following Lemma 1 through 4 without their proofs. Because they are completely the same as or quite similar to their references. Roughly speaking, Lemma 4.2 in [1] is used to evaluate exceptional sets in each lemma.

Lemma 1 (cf. [5; Lemma 4]). *Let* (A_1) *and* (A_2) *be satisfied for some* $r\geqq 1$. *Then for every* $\theta_0\in\Theta$ *and* $\varepsilon_0>0$, *there exists an integer* $n_1=n_1(\varepsilon_0)\geqq 1$ *such that*

$$P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:|\hat{\theta}_n-\theta_0|\geqq\varepsilon_0\}>\varepsilon_0 n^{1-r} \textit{ for all } n\geqq n_1.$$

For simplicity, we write $\bar{l}_{nj}(t)=n^{-1}\sum_{j=1}^{n}(\partial^j/(\partial\theta^j)\log f(x_i,\theta)|_{\theta=t}$ for $j=1,\cdots,k+1$.

Lemma 2 (cf. [2; Lemma 4.6]). *Let* (A_3), (A_4) *hold for some* $r\geqq 1$, *and* (A_5) *be satisfied also for some* $r\geqq 1$ *when* $j=k+1$. *Then for every* $\theta_0\in\Theta$, $\varepsilon_0>0$ *and* $j=1,\cdots k+1$, *there exist* $\delta_1=\delta_1(\varepsilon_0)>0$ *and an integer* $n_1=n_1(\varepsilon_0)\geqq 1$ *such that for all* $\delta\in(0,\delta_1]$

$$P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\sup\{|\bar{l}_{nj}(t)-E_{\theta_0}[(\partial^j/(\partial\theta^j)\log f(x,\theta_0)]|: |t-\theta_0|<\delta\}\geqq\varepsilon_0\}<\varepsilon_0 n^{1-r} \textit{ for all } n\geqq n_1.$$

Lemma 3 (cf. [2; Lemma 4.7] and [3; Lemma 2.2]). *Let* (A_3), (A_4) *and* (A_6) *be satisfied for some* $r\geqq 1$. *Then for every* $\theta_0\in\Theta$, $\varepsilon_0>0$, *there exist* $\delta_1>0$ *and an integer* $n_1=n_1(\varepsilon_0)\geqq 1$ *such that for all* $\delta\in(0,\delta_1]$

$$P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\log\{\textstyle\prod_{i=1}^{n}[f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}$$

$$> -\xi^2 I(\theta_0)/4 \textit{ for } |\xi|<\delta\sqrt{n},\ |\hat{\theta}_n-\theta_0|<\delta/2\}<\varepsilon_0 n^{1-r} \textit{ for all } n \geqq n_1$$

Lemma 4 (cf. [2; Lemma 4.2] and [3; Lemma 2.3]). *Let* (A_1) *and* (A_2) *be satisfied for some* $r \geqq 1$. *Then for every* $\theta_0 \in \Theta$, $\delta_2>0$ *and* $\varepsilon_0>0$, *there exist a constant* $a>0$ *and an integer* $n_1=n_1\ (\varepsilon_0) \geqq 1$ *such that*

$$P_{\theta_0}^N\{\underset{\sim}{x} \in X^N : \log\{\textstyle\prod_{i=1}^n [f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}$$
$$> -na \textit{ for } |\xi| \geqq \delta_2\sqrt{n},\ |\hat{\theta}_n-\theta_0|<\delta_2/2\}<\varepsilon_0 n^{1-r} \textit{ for all } n \geqq n_1.$$

For simple description, when $\hat{\theta}_n \in \Theta$, we put

$$Q_n(\xi)=\rho(\hat{\theta}_n+\xi/\sqrt{n})\textstyle\prod_{i=1}^n [f(x_i,\ \hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\ \hat{\theta}_n)]. \tag{3.1}$$

Lemma 5. *Let* (A_1), (A_2), (A_3), (A_4), (A_6) and (A_9) *be satisfied for some* $r \geqq 1$. *Then for every* $\theta_0 \in \Theta$ *and* $\varepsilon_0>0$, *there exist* $\delta_1>0$ *and an integer* $n_1=n_1(\varepsilon_0,k) \geqq 1$ *such that for all* $\delta \in (0,\delta_1]$

$$P_{\theta_0}^N\{\underset{\sim}{x} \in X^N : \left|\int_{|\xi| \geqq \delta\sqrt{n}} Q_n(\xi)\ \log\{\textstyle\prod_{i=1}^n [f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}d\xi\right| \geqq \varepsilon_0 n^{-(k-1)},$$
$$|\hat{\theta}_n-\theta_0|<\delta/2\}<\varepsilon_0 n^{1-r} \textit{ for all } n \geqq n_1.$$

Proof. Let $\varepsilon_0>0$ be given and $\delta_1>0$ sufficiently small. And we suppose that the following conditions hold.

$$|\hat{\theta}_n-\theta_0|<\delta/2,$$
$$\sup\{|\bar{l}_{n2}(\theta)+I(\theta_0)| : |\theta-\theta_0|<\delta/2\}<I(\theta_0)/2,$$
$$|n^{-1}\textstyle\sum_{i=1}^n |\log f(x_i,\theta_0)| - E_{\theta_0}[|\log f(x,\theta_0)|]|<1,$$
$$|n^{-1}\textstyle\sum_{i=1}^n g(x_i) - E_{\theta_0}[g(x)]|<1,$$
$$\log\{\textstyle\prod_{i=1}^n [f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\} \leqq -na \text{ for } |\xi| \geqq \delta\sqrt{n},$$

where $a>0$ appears in the statement of Lemma 4 and $\delta \in (0,\delta_1]$. Then we obtain

$$0 \geqq \int_{|\xi| \geqq \delta\sqrt{n}} Q_n(\xi)\log\{\textstyle\prod_{i=1}^n [f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}d\xi$$
$$\geqq \exp(-na)\ [\textstyle\sum_{i=1}^n \displaystyle\int_{|\xi| \geqq \delta\sqrt{n}} \rho(\hat{\theta}_n+\xi/\sqrt{n})\log f(x_i,\hat{\theta}_n+\xi/\sqrt{n})d\xi$$

$$+\{-\sum_{i=1}^{n}\log[f(x_i,\theta_0)+\sum_{i=1}^{n}\log\{f(x_i,\theta_0)/f(x_i,\hat{\theta}_n)\}\}$$

$$\times\int_{|\xi|\geqq\delta\sqrt{n}}\rho(\hat{\theta}_n+\xi/\sqrt{n})d\xi]$$

$$\geqq -n^{3/2}\exp(-na)\{n^{-1}\sum_{i=1}^{n}g(x_i)$$

$$+n^{-1}\sum_{i=1}^{n}|\log f(x_i,\theta_0)|+(\theta_0-\hat{\theta}_n)^2|\bar{l}_{n2}(\theta^*)|/2\}\quad(|\theta^*-\theta_0|\leqq|\hat{\theta}_n-\theta_0|)$$

$$> -n^{3/2}\exp(-na)\{E_{\theta_0}[g(x)]+1)+(E_{\theta_0}[|\log f(x,\theta_0)|]+1)+(\delta/2)^2(3I(\theta_0)/4)\}$$

$$> -\varepsilon_0 n^{-(k-1)}\text{ as }n\to+\infty.$$

Therefore it follows from Lemma 2, 4 and Lemma 4.2 in [1] that for every $\theta_0\in\Theta$ and $\varepsilon_0>0$, there exist $\delta_1>0$ and an integer $n_1=n_1(\varepsilon_0)\geqq 1$ such that

$$P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:|\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)$$

$$\times\log\{\prod_{i=1}^{n}[f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}d\xi|\geqq\varepsilon_0 n^{-(k-1)},|\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$\leqq P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\sup\{|\bar{l}_{n2}(\theta)+I(\theta_0)|:|\theta-\theta_0|<\delta_1/2\}\geqq I(\theta_0)/2\}$$

$$+P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:|n^{-1}\sum_{i=1}^{n}|\log f(x_i,\theta_0)|-E_{\theta_0}[|\log f(x_i,\theta_0)|]|\geqq 1\}$$

$$+P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:|n^{-1}\sum_{i=1}^{n}g(x_i)-E_{\theta_0}[g(x)]|\geqq 1\}$$

$$+P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\log\{\prod_{i=1}^{n}f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}$$

$$> -na\text{ for }|\xi|\geqq\delta\sqrt{n},\ |\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$<\varepsilon_0 n^{1-r}\text{ for all }n\geqq n_1.$$

Lemma 6. *Let $\theta_0\in\Theta$ such that $\rho(\theta_0)>0$ and* (A_3), (A_4) *and* (A_6) *be satisfied for some $r\geqq 1$. Moreover if ρ is continuous in a neighborhood of θ_0, then for every $\varepsilon_0>0$, there exists an integer $n_1=n_1(\varepsilon_0)\geqq 1$ such that*

$$P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\int_{-\infty}^{+\infty}Q_n(\xi)d\xi<\delta_3,\ |\hat{\theta}_n-\theta_0|<\delta_3\}<\varepsilon_0 n^{1-r}$$

for some $\delta_3>0$ and all $n\geqq n_1$, where $\delta_3>0$ is independent of ε_0.

Proof. Let $\delta_4>0$ be sufficiently small. Then we assume that

$$|\theta-\theta_0|<2\delta_4\text{ yields }|\rho(\theta)-\rho(\theta_0)|<\rho(\theta_0)/2.$$

If $|\hat{\theta}_n-\theta_0|<\delta_4$, then $\rho(\hat{\theta}_n+\xi/\sqrt{n})>\rho(\theta_0)/2$ for $|\xi|<\delta_4$ since $|(\hat{\theta}_n+\xi/\sqrt{n})-\theta_0|<|\hat{\theta}_n-\theta_0|+|\xi|<2\delta_4$. In addition, if

$$\sup\{|\bar{l}_{n2}(\theta)+I(\theta_0)|:|\theta-\theta_0|<2\delta_4\}<I(\theta_0)/2,$$

then when we put $\delta_3=\min\{\delta_4,\ \rho(\theta_0)\delta_4\exp\{-3I(\theta_0)\delta_4^2/4\}\}$, we obtain

$$\begin{aligned}
&\int_{-\infty}^{+\infty} Q_n(\xi)d\xi \\
&\geqq \int_{|\xi|<\delta_4} \rho(\hat{\theta}_n+\xi/\sqrt{n})\exp\{\bar{I}_{n2}(\theta^*)\xi^2/2\}d\xi \quad (|\theta^*-\hat{\theta}_n|\leqq|\xi|/\sqrt{n}) \\
&>2^{-1}\rho(\theta_0)\int_{|\xi|<\delta_4} \exp\{-3I(\theta_0)\xi^2/4\}d\xi \\
&>2^{-1}\rho(\theta_0)\exp\{-3I(\theta_0)\delta_4^2/4\}\int_{|\xi|<\delta_4} d\xi \\
&\geqq \delta_3.
\end{aligned}$$

Consequently, by Lemma 2, there exists an integer $n_1=n_1\ (\varepsilon_0)\geqq 1$ such that

$$\begin{aligned}
&P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\ \int_{-\infty}^{+\infty} Q_n(\xi)d\xi<\delta_3,\ |\hat{\theta}_n-\theta_0|<\delta_3\} \\
&\leqq P_{\theta_0}^N\{\underset{\sim}{x}\in X^N:\ \sup\{|\bar{I}_{n2}(\theta)+I(\theta_0)|:|\theta-\theta_0|<2\delta_4\}\geqq I(\theta_0)/2\} \\
&<\varepsilon_0 n^{1-r} \text{ for all } n\geqq n_1.
\end{aligned}$$

Now that we have finished proving lemmas as preparation, we will begin the proof of the main theorem.

4. Proof of the main theorem

When $\hat{\theta}_n\in\Theta$, by using (3.1), (2.1) can be rewritten as

$$\pi(\xi|\{\underset{\sim}{x}_n\})=Q_n(\xi)/\int_{-\infty}^{+\infty} Q_n(\xi)d\xi,$$

and we obtain

$$\begin{aligned}
(4.1)\qquad & 0\leqq\int_{-\infty}^{+\infty} \pi(\xi|\underset{\sim}{x}_n)\log\{\pi(\xi|\underset{\sim}{x}_n)/\psi_k(\xi|T_n^{(k)})\}d\xi \\
&=\int_{-\infty}^{+\infty} \pi(\xi|\underset{\sim}{x}^n)\log\{Q_n(\xi)/R_n(\xi)\}d\xi+\log\{\int_{-\infty}^{+\infty} R_n(\xi)d\xi/\int_{-\infty}^{+\infty} Q_n(\xi)d\xi\} \\
&\leqq\int_{-\infty}^{+\infty} Q_n(\xi)\log\{Q_n(\xi)/R_n(\xi)\}d\xi/\int_{-\infty}^{+\infty} Q_n(\xi)d\xi
\end{aligned}$$

$$
\begin{aligned}
&+\int_{-\infty}^{+\infty}\{R_n(\xi)-Q_n(\xi)\}d\xi\Big/\int_{-\infty}^{+\infty}Q_n(\xi)d\xi \\
&=[(\int_{|\xi|<\delta\sqrt{n}}+\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)\log\{Q_n(\xi)/R_n(\xi)\}d\xi \\
&\quad+\int_{|\xi|<\delta\sqrt{n}}Q_n(\xi)\{\exp\{\log\{R_n(\xi)/Q_n(\xi)\}\}-1\}d\xi \\
&\quad+\int_{|\xi|\geqq\delta\sqrt{n}}\{R_n(\xi)-Q_n(\xi)\}d\xi]\Big/\int_{-\infty}^{+\infty}Q_n(\xi)d\xi \\
&=[\int_{|\xi|<\delta\sqrt{n}}Q_n(\xi)\{\log\{Q_n(\xi)/R_n(\xi)\}+\log\{R_n(\xi)/Q_n(\xi)\}\}d\xi \\
&\quad+\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)\log\{Q_n(\xi)/R_n(\xi)\}d\xi \\
&\quad+2^{-1}\int_{|\xi|<\delta\sqrt{n}}Q_n(\xi)\{\log\{R_n(\xi)/Q_n(\xi)\}\}^2\exp\{\nu\log\{R_n(\xi)/Q_n(\xi)\}\}d\xi\quad(\nu\in(0,1)) \\
&\quad+\int_{|\xi|\geqq\delta\sqrt{n}}\{R_n(\xi)-Q_n(\xi)\}d\xi]\Big/\int_{-\infty}^{+\infty}Q_n(\xi)d\xi \\
&\leqq[|\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)\log\{\textstyle\prod_{i=1}^{n}[f(x_i,\hat{\theta}_n+\xi/\sqrt{n})/f(x_i,\hat{\theta}_n)]\}d\xi| \\
&\quad+2^{-1}|\bar{I}_{n2}(\hat{\theta}_n)|\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)\xi^2d\xi \\
&\quad+2^{-1}\int_{|\xi|<\delta\sqrt{n}}Q_n(\xi)\{\log\{R_n(\xi)/Q_n(\xi)\}\}^2\exp\{|\log\{R_n(\xi)/Q_n(\xi)\}|\}d\xi \\
&\quad+\int_{|\xi|\geqq\delta\sqrt{n}}\{R_n(\xi)+Q_n(\xi)\}d\xi]\Big/\int_{-\infty}^{+\infty}Q_n(\xi)d\xi,
\end{aligned}
$$

where $R_n(\xi)$ is defined in (2.2). We consider the evaluation of each term of the right hand side of (4.1). Let $\varepsilon>0$ be given. We suppose that there exists $\delta_3>0$ such that

$$\int_{-\infty}^{+\infty}Q_n(\xi)d\xi\geqq\delta_3 \text{ if } |\hat{\theta}_n-\theta_0|<\delta_3.$$

Let $\varepsilon_1=\varepsilon_1(\varepsilon, k)>0$ and $\delta=\delta(\varepsilon, k)>0$ be sufficiently small. Then we assume that $\delta/2>\delta_3$ and

$$|\theta-\theta_0|<3\delta/2 \quad \text{yields} \quad |\rho(\theta)-\rho(\theta_0)|<\rho(\theta_0),$$

$$|\theta-\theta_0|<\delta/2 \quad \text{yields} \quad |\lambda_{k+1}(\theta)-\lambda_{k+1}(\theta_0)|<\varepsilon_1.$$

In addition, we suppose that

$$|\hat{\theta}_n-\theta_0|<\delta/2,$$

$$\sup\{|\bar{l}_{n2}(\theta)+I(\theta_0)|:|\theta-\theta_0|<\delta/2\}>I(\theta_0)/2,$$

$$\sup\{|\bar{l}_{nk+1}(\theta)-\lambda_{k+1}(\theta_0)|:|\theta-\theta_0|<3\delta/2\}<\varepsilon_1,$$

$$\log\{\textstyle\prod_{i=1}^n[f(x_i, \hat{\theta}_n+\xi/\sqrt{n})/f(x_i, \hat{\theta}_n)]\}\leqq -I(\theta_0)\xi^2/4 \quad \text{for} |\xi|<\delta\sqrt{n},$$

$$\log\{\textstyle\prod_{i=1}^n[f(x_i, \hat{\theta}_n+\xi/\sqrt{n})/f(x_i, \hat{\theta}_n)]\}\leqq -na \quad \text{for} |\xi|\geqq\delta\sqrt{n},$$

$$\left|\int_{|\xi|\geqq\delta\sqrt{n}} Q_n(\xi)\log\{\textstyle\prod_{i=1}^n[f(x_i, \hat{\theta}_n+\xi/\sqrt{n})/f(x_i, \hat{\theta}_n)]\}d\xi\right|<\varepsilon_1 n^{-(k-1)},$$

where $a>0$ is that of Lemma 4. The probability of the complements of these events will be evaluated in the final step of the proof.

$$\text{(4.2)} \qquad 2^{-1}|\bar{l}_{n2}(\hat{\theta}_n)|\int_{|\xi|\geqq\delta\sqrt{n}} Q_n(\xi)\xi^2 d\xi$$

$$<(3/4)I(\theta_0)\exp(-na)\int_{-\infty}^{+\infty}\xi^2\rho(\hat{\theta}_n+\xi/\sqrt{n})d\xi$$

$$=(3/4)I(\theta_0)n^{3/2}\exp(-na)\int_{-\infty}^{+\infty}(y-\hat{\theta}_n)^2\rho(y)dy$$

$$<(3/2)I(\theta_0)n^{3/2}\exp(-na)\{\int_{-\infty}^{+\infty}y^2\rho(y)dy+(|\theta_0|+\delta/2)^2\}$$

$$<\varepsilon_1 n^{-(k-1)} \ (n\to+\infty),$$

where we have used $|\hat{\theta}_n|<|\theta_0|+\delta/2$. Next, we will show that

$$\text{(4.3)} \qquad 2^{-1}\int_{|\xi|<\delta\sqrt{n}} Q_n(\xi)\{\log\{R_n(\xi)/Q_n(\xi)\}\}^2 \times\exp\{|\log R_n(\xi)/Q_n(\xi)\}|\}d\xi<\varepsilon_1 n^{-(k-1)}.$$

In fact,

$$
\begin{aligned}
|\log\{R_n(\xi)/Q_n(\xi)\}| &= |\lambda_{k+1}(\hat{\theta}_n) - \bar{l}_{nk+1}(\theta^*)||\xi|^{k+1} n^{-(k-1)/2}/(k+1)! \\
&\quad (|\theta^* - \hat{\theta}_n| \leqq |\xi|/\sqrt{n}) \\
&\leqq \{|\lambda_{k+1}(\hat{\theta}_n) - \lambda_{k+1}(\theta_0)| + |\bar{l}_{nk+1}(\theta^*) - \lambda_{k+1}(\theta_0)|\}|\xi|^{k+1} n^{-(k-1)/2} \\
&< 2\varepsilon_1 |\xi|^{k+1} n^{-(k-1)/2}.
\end{aligned}
$$

Therefore the left hand side of (4.3) is dominated by

$$
\begin{aligned}
&4\rho(\theta_0)\varepsilon_1^2 n^{-(k-1)} \int_{|\xi|<\delta\sqrt{n}} \xi^{2(k+1)} \exp\{-I(\theta_0)\xi^2/4\} \\
&\qquad \times \exp\{2\varepsilon_1(\delta\sqrt{n})^{k-1} n^{-(k-1)/2}\xi^2\} d\xi \\
&\leqq 4\rho(\theta_0)\varepsilon_1^2 n^{-(k-1)} \int_{-\infty}^{+\infty} \xi^{2(k+1)} \exp\{-I(\theta_0)\xi^2/8\} d\xi \\
&< \varepsilon_1 n^{-(k-1)},
\end{aligned}
$$

which establishes (4.3).

$$
\begin{aligned}
(4.4) \qquad & \int_{|\xi|\geqq\delta\sqrt{n}} \{R_n(\xi) + Q_n(\xi)\} d\xi \\
&\leqq \int_{|\xi|\geqq\delta\sqrt{n}} \rho(\hat{\theta}_n + \xi/\sqrt{n})\{\exp\{-I(\theta_0)\xi^2/4\} + \exp(-na)\} d\xi \\
&\leqq \sqrt{n}\{\exp\{-I(\theta_0)\delta^2 n/4\} + \exp(-na)\} \\
&< \varepsilon_1 n^{-(k-1)} \quad (n \to +\infty).
\end{aligned}
$$

It follows from (4.1) through (4.4) that

$$
\begin{aligned}
&0 \leqq \int_{-\infty}^{+\infty} \pi(\xi|\underset{\sim}{x}_n) \log\{\pi(\xi|\underset{\sim}{x}_n)/\psi_k(\xi|T_n^{(k)})\} d\xi < 4\varepsilon_1 n^{-(k-1)}/\delta_3 \\
&< \varepsilon^{-(k-1)} \quad (n \to +\infty).
\end{aligned}
$$

Consequently, by Lemma 1, 2, 3, 4, 5 and 6, for every $\varepsilon > 0$, there exist $\delta = \delta(\varepsilon, k) > 0$ and an integer $n_0 = n_0(\varepsilon, k) \geqq 1$ such that

$$
P_{\theta_0}^N\{\underset{\sim}{x} \in X^N \colon \int_{-\infty}^{+\infty} \pi(\xi|\underset{\sim}{x}_n) \log\{\pi(\xi|\underset{\sim}{x}_n)/\psi_k(\xi|T_n^{(k)})\} d\xi \geqq \varepsilon n^{-(k-1)}\}
$$

$$\leqq P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ |\hat{\theta}_n-\theta_0|\geqq\delta/2\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\int_{-\infty}^{+\infty}\pi(\xi|\underset{\sim}{x}_n)\log\{\pi(\xi|\underset{\sim}{x}_n)/\psi_k(\xi|T^{(k)}_n)\}d\xi\geqq\varepsilon n^{-(k-1)},$$

$$|\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$\leqq P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ |\hat{\theta}_n-\theta_0|\geqq\delta/2\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ \sup\{|\bar{l}_{n2}(\theta)+I(\theta_0)|:|\theta-\theta_0|<\delta/2\}\geqq I(\theta_0)/2\}$$

$$\leqq P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ \sup\{|\bar{l}_{nk+1}(\theta)-\lambda_{k+1}(\theta_0)|:|\theta-\theta_0|<3\delta/2\}\geqq\varepsilon_1\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ \log\{\textstyle\prod_{i=1}^n[f(x_i,\ \hat{\theta}_n+\xi\sqrt{n})/f(x_i,\ \hat{\theta}_n)]\}$$

$$>-I(\theta_0)\xi^2/4\ \text{for}\ |\xi|<\delta\sqrt{n},\ |\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ \log\{\textstyle\prod_{i=1}^n[f(x_i,\ \hat{\theta}_n+\xi\sqrt{n})/f(x_i,\ \hat{\theta}_n)]\}$$

$$>-na\ \text{for}\ |\xi|\geqq\delta\sqrt{n},|\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ |\int_{|\xi|\geqq\delta\sqrt{n}}Q_n(\xi)$$

$$\times\log\{\textstyle\prod_{i=1}^n[f(x_i,\ \hat{\theta}_n+\xi\sqrt{n})/f(x_i,\ \hat{\theta}_n)]\}d\xi|\geqq\varepsilon_1 n^{-(k-1)},$$

$$|\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$+P^N_{\theta_0}\{\underset{\sim}{x}\in X^N:\ \int_{-\infty}^{+\infty}Q_n(\xi)d\xi<\delta_3,\ |\hat{\theta}_n-\theta_0|<\delta/2\}$$

$$<\varepsilon n^{1-r}\ \text{for all}\ n\geqq n_0.$$

This completes the proof of the theorem.

ACKNOWLEDGMENT. I'm grateful to Professor Takeru Suzuki for the careful examination of this study. He pointed out the improvement of the order of the approximation. I also would like to thank the referee for many valuble comments.

References

[1] D.M. Chibisov: *On the normal approximation for a certain class of statictics*, Proc. Sixth berkeley Symp. Math. Statist. Prob. **1**, Univ. of California Press, 1972, 153–174.
[2] J.K.Ghosh, B.K.Shinha and S.N.Joshi: *Expansions for posterior probability and integrated*

Bayes risk, Statistical Decision Theory and Related Topics III, Vol. 1. Academic Press, 1982, 403–456.

[3] R.A. Johnson: *Asymptotic expansions associated with posterior distributions*, Ann. Math. Statist. **41**, No. 3 (1970), 851–864.

[4] T. Kato: *Second order asymptotic sufficiency of posterior densities with respect to Kullback information.* (1994). (to appear in Statist. Decisions. **12**)

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

[6] K. Takeuchi: *Hi-Beyes no tachiba kara mita Bayes tôkeigaku* (*Bayesian statistics from a non-Bayesian point of view*). Bayes tôkeigaku to sono ôyô (Bayesian statistics and its applications), Chapter 5, Univ. of Tokyo Press, (1989), 119–136. (in Japanese)

Waseda University
Department of Mathematics, School of Science and Engineering,
Waseda University
Ohkubo, Shinjuku-ku, Tokyo 169, Japan