Statistica Sinica 4(1994), 679-692

ON FISHER'S BOUND FOR STABLE ESTIMATORS WITH EXTENSION TO THE CASE OF HILBERT PARAMETER SPACE

Zhongguo Zheng and Kai-Tai Fang

Peking University and Hong Kong Baptist College

Abstract: Fisher's bound for asymptotic variances has been shown to hold almost everywhere for several classes of estimators in LeCam (1953), Bahadur (1964) and Wong (1992). Wong's result applies to estimators with arbitrary asymptotic distributions, provided a certain continuity condition is satisfied. In this paper we improve on Wong's result by removing the continuity condition. We also generalize the result to the case where the parameter space is separable Hilbert.

Key words and phrases: Concentration probability, Fisher's bound of estimator, asymptotic efficiency.

1. Introduction

Fisher (1925) supposed that the maximum likelihood estimator (MLE) $\hat{\theta}_n$ of a real parameter θ is asymptotically optimal among a large class of estimators. The conjectured result can be formulated as following. Let $Y_1,\ldots,Y_n$ be a sample of size n from a distribution family $f_\theta(y)$ with the parameter $\theta\in\Theta\subset R^1$.

(1) The MLE $\hat{\theta}_n(y_1,\ldots,y_n)$ is asymptotically normal, i.e.

$$\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{\mathcal{L}}N\Big(0,\frac{1}{I(\theta)}\Big), \tag{1.1}$$

where $I(\theta)$ is the Fisher's information of the observation Y_1.

(2) If T_n is an asymptotically normal estimator, i.e., $\sqrt{n}(T_n-\theta)\xrightarrow{\mathcal{L}}N(0,v_\theta)$, then

$$v_\theta\geq I(\theta)^{-1}. \tag{1.2}$$

However, Hodges, J. L. (see LeCam (1953)) by a counter example pointed out that the conjecture is not always true. His example shows that "superefficiency" of the estimator can occur even in the smoothest model.

To state the result more precisely, we need to introduce some notation and conditions. For simplicity, assume, at first, that the parameter space is $\Theta=R^1$.

Let X_n be an observation taking values in the sample space $(\mathcal{X}_n, \mathcal{B}_{\mathcal{X}_n}, \nu_n)$, where ν_n is a σ-finite measure on the σ-field $\mathcal{B}_{\mathcal{X}_n}$. Denote by $P_\theta(B)$ the distribution of X_n under θ and $p_\theta(x_n)$ the corresponding density with respect to ν_n. According to Wong (1992) the following conditions on the model are needed.

Condition (M): For every n, $p_\theta(x_n)$ is a measurable function of (θ, x_n).

Condition (L): For all $h \in R^1$,

$$\frac{dP_{\theta+h/\sqrt{n}}}{dP_\theta}(x_n) = \exp\{hI(\theta)\Delta_{n,\theta} - \frac{1}{2}I(\theta)h^2 + R_n(\theta, h)\}, \tag{1.3}$$

where

$$\mathcal{L}_\theta\{\Delta_{n,\theta}\} \to N(0, I(\theta)^{-1}), \qquad R_n(\theta, h) = o_p(1)$$

and Fisher's information $I(\theta)$ is assumed to be a strictly positive and continuous function of θ (the notation "$\mathcal{L}_\theta\{\Delta_{n,\theta}\} \to N(0, I(\theta)^{-1})$" means that the distributions of random variables $\Delta_{n,\theta}$ converge to the distribution $N(0, I(\theta)^{-1})$.

Condition (L′): In Condition (L), the constant h is replaced by a series of real numbers h_n, which tends to a constant h.

LeCam (1953) first observed that, for asymptotically efficient estimators, superefficiency can occur only in sets of measure zero. Bahadur (1964) made the following extension.

Theorem 1.1. *Under conditions* (M) *and* (L), *if* $\mathcal{L}_\theta(\sqrt{n}(T_n - \theta)) \to N(0, v_\theta)$ *for all* $\theta \in \Theta$, *then*

$$\mu\{\theta : v_\theta < I(\theta)^{-1}\} = 0, \tag{1.4}$$

where μ *stands for the Lebesgue measure on the parameter space* Θ.

The result has been extended in several ways. First the restriction of the asymptotical normality of the estimator T_n is removed. A series of estimators T_n for the parameter θ is called stable, if for every $\theta \in \Theta$,

$$\mathcal{L}_\theta(\sqrt{n}(T_n - \theta)) \to F_\theta, \tag{1.5}$$

where F_θ is the asymptotic distribution of the estimator series. Since the asymptotic distribution in (1.5) is arbitrary, a natural way to consider the efficiency of stable series is to compare the concentration probabilities around the true value of the parameter. One may want to establish a result that the concentration probability $P_\theta\{\sqrt{n}\,|T_n - \theta| < \rho\}$ is asymptotically bounded above by $P\{|Z_\theta| < \rho\}$, where $Z_\theta \sim N(0, I(\theta)^{-1})$. Wong (1992) obtained the following extension.

Theorem 1.2. *Under conditions* (M) *and* (L), *if* T_n *is stable and* $r(\theta) = \lim P_\theta\{T_n < \theta\}$ *is continuous almost everywhere, then for almost all* $\theta \in \Theta$,

$$\lim P_\theta\{\sqrt{n}\,|T_n - \theta| < \rho\} \le P\{|Z_\theta| < \rho\} \tag{1.6}$$

for all $\rho \geq 0$.

In this paper we will remove the condition of continuity of $r(\theta)$ under a slightly stronger condition (L′) on the model. We have

Theorem 1.3. *Under the conditions* (M) *and* (L′), *if* T_n *is stable, then for almost all* $\theta \in \Theta$, (1.6) *holds.*

A series of estimators T_n is called almost everywhere (a.e.) stable, if there exists a subset Θ_0 of Θ with Lebesgue measure zero such that for every $\theta \in \Theta\backslash\Theta_0$, (1.5) holds. For the a.e. stable estimators, we have

Theorem 1.4. *Under the conditions* (M) *and* (L′), *if* T_n *is a.e. stable, then for almost all* $\theta \in \Theta$, (1.6) *holds.*

Example 1.1. Let $X_1, \ldots, X_n \sim$ iid $N(\theta, 1), \theta \in (0,1)$. Let $A_n = \{\theta_{i_1}, \ldots, \theta_{i_{k_n}}\}$ be a series of finite subsets of parameters satisfying
(i) $A_i \subseteq A_{i+1} \subseteq \cdots$,
(ii) $A = \lim A_n$ is dense in (0,1),
(iii) there exists a series of constants $c_n \to \infty$ such that

$$\mu\{\bigcup_{n\geq n_0}\{\theta : \rho(\theta, A_n) \leq n^{-\frac{1}{4}} c_n\}\} \to 0 \text{ as } n_0 \to \infty,$$

where μ is the Lebesgue measure on $(0,1)$ and $\rho(\theta, A_n)$ stands for the distance between θ and the set A_n. Let $c_n = n^{\frac{1}{8}}$,

$$A = \{\theta_i, i \geq 1\} = \Big\{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}, \frac{1}{16}, \frac{3}{16}, \frac{5}{16}, \ldots\Big\}$$

and

$$D_i = \{\theta_1, \theta_2, \ldots, \theta_i\}, \quad i \geq 1.$$

Let

$$\begin{aligned} n_i &= \inf\{n : n^{-\frac{1}{8}}(i+1) \leq \frac{1}{2^i}\} \\ A_n &= D_i, \quad n_{i-1} < n \leq n_i, \quad (n_0 = 0). \end{aligned}$$

It is easy to show that for the c_n and A_n defined above, conditions (i), (ii) and (iii) are fulfilled. Therefore the Lebesgue measure of the set

$$N = \bigcap_{n_0 \geq 1} \bigcup_{n \geq n_0} \{\theta : \rho(\theta, A_n) \leq n^{-\frac{1}{4}} c_n\}$$

is equal to zero.

Let

$$T_n = \begin{cases} \bar{X}_n + \frac{1}{\sqrt{n}}, & \rho(\bar{X}_n, A_n) < n^{-\frac{1}{4}}, \\ \bar{X}_n, & \rho(\bar{X}_n, A_n) \geq n^{-\frac{1}{4}}. \end{cases}$$

When $\theta \in A$,

$$T_n = \Big(\bar{X}_n + \frac{1}{\sqrt{n}}\Big) - \frac{1}{\sqrt{n}} I_{\{\rho(\bar{X}_n, A_n) \geq n^{-\frac{1}{4}}\}}.$$

For n sufficiently large,

$$\begin{aligned} I_{\{\rho(\bar{X}_n, A_n) \geq n^{-\frac{1}{4}}\}} &\leq I_{\{|\bar{X}_n - \theta| \geq n^{-\frac{1}{4}}\}} \\ &= I_{\{\sqrt{n}|\bar{X}_n - \theta| \geq n^{\frac{1}{4}}\}} = o_p(1). \end{aligned}$$

Therefore

$$\begin{aligned} \sqrt{n}(T_n - \theta) &= \sqrt{n}(\bar{X}_n - \theta) + 1 - I_{\{\rho(\bar{X}_n, A_n) \geq n^{-\frac{1}{4}}\}} \\ &= \sqrt{n}(\bar{X}_n - \theta) + 1 + o_p(1) \to N(1,1), \quad \text{for } \theta \in A. \end{aligned} \tag{1.7}$$

Now suppose that $\theta \notin \{N \cup A\}$. Then

$$\sqrt{n}(T_n - \theta) = \sqrt{n}(\bar{X}_n - \theta) + I_{\{\rho(\bar{X}_n, A_n) < n^{-\frac{1}{4}}\}}. \tag{1.8}$$

Since $\theta \notin N$, i.e., $\theta \in \bigcap_{n \geq n_0} \{\tilde{\theta} : \rho(\tilde{\theta}, A_n) \geq n^{-\frac{1}{4}} c_n\}$ for large n_0, it is easy to show that

$$\{\rho(\bar{X}_n, A_n) < n^{-\frac{1}{4}}\} \subset \{|\bar{X}_n - \theta| > n^{-\frac{1}{4}}(c_n - 2)\}, \quad \text{for large } n,$$

which implies that $I_{\{\rho(\bar{X}_n, A_n) < n^{-\frac{1}{4}}\}} = o_p(1)$. Therefore

$$\sqrt{n}(T_n - \theta) \to N(0,1), \quad \text{for } \theta \in (0,1) \backslash \{N \cup A\}, \tag{1.9}$$

where the symbol $\backslash$ stands for the set operation subtraction. Equations (1.7) and (1.9) show that $r(\theta) = \lim\sup(P_\theta\{\sqrt{n}(T_n - \theta) \leq 0\}$ is not a.e. continuous. In fact every $\theta \in (0,1)$ is a discontinuous point of r. In this example, we can apply Theorem 1.4 to T_n, but we cannot use Wong's result on T_n.

Let $\lambda(\theta)$ be a continuously differentiable function of θ. To establish bounds for estimating such a function, we have

Theorem 1.5. *Under conditions* (M) *and* (L′), *if T_n is a stable estimator of $\lambda(\theta)$, then for almost all $\theta \in \Theta$,*

$$\lim P_\theta\{\sqrt{n}\,|\,T_n - \lambda(\theta)| < \rho\} \leq P\{|\tilde{Z}_\theta| < \rho\} \tag{1.10}$$

holds for all $\rho \geq 0$, where

$$\tilde{Z}_\theta \sim N(0, \lambda'(\theta)^2 I(\theta)^{-1}).$$

Now suppose that the densities in the family $\{P_{\theta,n}\}$ are indexed by a possibly infinite dimensional parameter $\theta \in \Theta$, where Θ is an open set in a real Hilbert space $\mathcal{L}$, which is separable. Since there is no direct generalization of Lebesgue measure in Hilbert space, we consider Gaussian measure instead.

Definition 1.1. A probability measure Q, defined on a Hilbert space $\mathcal{L}$, is called a *Gaussian measure* if for every fixed $\theta \in \mathcal{L}$, the real valued random variable $\langle \theta, \eta \rangle$, defined on $\eta \in \mathcal{L}$, has a normal distribution, where the notation $\langle\ ,\ \rangle$ stands for the inner product in $\mathcal{L}$.

From the theory of Gaussian measure (see Kuo (1975)), we know that Gaussian measure is characterized by the mean value $m_Q \in \mathcal{L}$ and the covariance operator S_Q, which is a nonnegative, self-adjoint, compact operator in $\mathcal{L}$ with its eigenvalues λ_i satisfying $\sum_{i=1}^{\infty} \lambda_i < \infty$ and is given by

$$\langle S_Q x, y \rangle = \int_{\mathcal{L}} \langle x, z - m_Q \rangle \langle y, z - m_Q \rangle Q(dz).$$

So, we can use the notation $N(m, S)$ to denote a Gaussian measure in the Hilbert space $\mathcal{L}$, where $m \in \mathcal{L}$ is the mean, S is the operator mentioned above.

Without loss of generality, we assume that the space is given in the form of

$$\mathcal{L} = \{(c_1, c_2, \ldots), \sum_{i=1}^{\infty} c_i^2 < \infty\}$$

with $\langle \theta, \eta \rangle = \sum \theta_i \eta_i$, where $\theta = (\theta_1, \theta_2, \ldots)$, $\eta = (\eta_1, \eta_2, \ldots)$. As an example of a Gaussian measure, consider $N(0, S)$, where S is an operator in $\mathcal{L}$ with $S\theta = (\sigma_1^2 \theta_1, \sigma_2^2 \theta_2, \ldots)$, and $\sum \sigma_i^2 < \infty$. In this case, the coordinates θ_i of θ are independent of each other with each $\theta_i \sim N(0, \sigma_i^2)$.

Denote the set of unit vectors in $\mathcal{L}$ by $\mathcal{L}_0$. For the distribution family $\{P_\theta : \theta \in \Theta\}$ on $\mathcal{X}_n$, we require

Condition (L$''$). *For all $h_n \to h > 0$, $v_n \to v \in \mathcal{L}_0$ the following holds*

$$\frac{dP_{\theta + \frac{h_n}{\sqrt{n}} v_n}}{dP_\theta}(x_n) = \exp\{h_n I(\theta, v) \Delta_{n,\theta,v} - \frac{1}{2} h_n^2 I(\theta, v) + R_n(\theta, v_n, h_n)\}, \qquad (1.11)$$

where $\mathcal{L}_\theta\{\Delta_{n,\theta,v}\} \to N(0, I(\theta, v)^{-1})$, $R_n(\theta, v_n, h_n) = o_p(1)$, and $I(\theta, v)$ is a strictly positive and continuous function of θ and v.

From Condition (L$''$) we know that

$$I(\theta, -v) = I(\theta, v). \qquad (1.12)$$

Theorem 1.6. *Suppose that Condition* (M) *holds, that Condition* (L$''$) *holds for all $\theta \in \Theta$ and $v \in \mathcal{L}_0$ and that the measure Q is a Gaussian measure on the* *separable Hilbert space $\mathcal{L}$. Suppose that $\lambda : \theta \to R$ is continuously differentiable in the sense that, for all θ, the directional derivative*

$$\lambda'_\theta(v) = \frac{d\lambda(\theta + tv)}{dt}\Big|_{t=0}$$

is continuous in $v \in \mathcal{L}_0$ and that T_n is a stable estimators of $\lambda(\theta)$. Then for almost all $\theta \in \Theta$ with respect to Q and for all $\rho > 0$,

$$\lim P_\theta\{\sqrt{n}\,|\,T_n - \lambda(\theta)\,| < \rho\} \le P\{|\,Z\,| < \rho\}, \tag{1.13}$$

where

$$Z \sim N(0, \sigma^2),$$
$$\sigma^2 = \sup_{v \in \mathcal{L}_0} \frac{|\lambda'_\theta(v)|^2}{I(\theta, v)}.$$

From the expression for $\lambda'_\theta(v)$, we know that

$$\lambda'_\theta(-v) = -\lambda'_\theta(v). \tag{1.14}$$

2. Proofs

First, we need a lemma.

Lemma 2.1. *Suppose that $r(\theta)$ is a measurable function of θ, where θ is a real parameter. Then there exists a set N of θ with Lebesgue measure zero such that for every $\theta \notin N$ and every $h \in R^1$, there exists a series $h_n = h_n(\theta, h)$ satisfying the following conditions*

$$h_n \to h, \tag{2.1}$$
$$r\Big(\theta + \frac{h_n}{\sqrt{n}}\Big) \to r(\theta), \tag{2.2}$$
$$r\Big(\theta - \frac{h_n}{\sqrt{n}}\Big) \to r(\theta). \tag{2.3}$$

Proof. Without loss of generality, we may assume that the domain of r is R^1. According to the theory of real analysis (see Natanson (1950), p.297), the function $r(\theta)$ is almost everywhere approximately continuous, i.e., there exists a null set N such that for every $\theta \notin N$, there exists a set $M = M_\theta$ satisfying

$$\frac{\mu(M \cap (\theta - \delta, \theta + \delta))}{2\delta} \to 1, \quad \text{as} \quad \delta \to 0 \tag{2.4}$$

and

$$\lim_{\substack{\eta \to \theta \\ \eta \in M}} r(\eta) = r(\theta). \tag{2.5}$$

Let M_θ^s be the symmetric set of M_θ with respect to θ, i.e.,

$$M_\theta^s = \{\eta = \theta + \xi : \theta - \xi \in M_\theta\}.$$

It is easy to show from (2.4) that the set $M_\theta \cap M_\theta^s$ satisfies

$$\frac{\mu(M_\theta \cap M_\theta^s \cap (\theta - \delta, \theta + \delta))}{2\delta} \to 1, \quad \text{as} \quad \delta \to 0.$$

Therefore

$$\frac{\mu(M_\theta \cap M_\theta^s \cap (\theta - \frac{2h}{\sqrt{n}}, \theta + \frac{2h}{\sqrt{n}}))}{\frac{4h}{\sqrt{n}}} \to 1, \quad \text{as} \quad n \to \infty,$$

from which we know that there exists a series of positive numbers $\delta_n \to 0$ such that the intersection

$$M_\theta \cap M_\theta^s \cap \Big(\theta + \frac{h - \delta_n}{\sqrt{n}}, \theta + \frac{h + \delta_n}{\sqrt{n}}\Big)$$

is not empty for sufficiently large n. Choosing h_n such that $\theta + \frac{h_n}{\sqrt{n}} \in M_\theta \cap M_\theta^s \cap (\theta + \frac{h-\delta_n}{\sqrt{n}}, \theta + \frac{h+\delta_n}{\sqrt{n}})$, it follows that

$$h_n \to h, \qquad \theta + \frac{h_n}{\sqrt{n}} \in M_\theta, \quad \theta - \frac{h_n}{\sqrt{n}} \in M_\theta,$$

which, together with (2.5), shows that (2.1) – (2.3) hold.

Proof of Theorem 1.3. Let

$$r(\theta) = \lim_{n \to \infty} P_\theta\{T_n < \theta\}. \tag{2.6}$$

From Condition (M), it is easy to show that $r(\theta)$ is a measurable function of θ. By Lemma 2.1 we know that there exists a null set N such that for every $\theta \notin N$ and $\rho \in R^+ = (0, \infty)$, there exists a series $\rho_n = \rho_n(\theta, \rho) \to \rho$ for which

$$\lim_{n \to \infty} r\Big(\theta + \frac{\rho_n}{\sqrt{n}}\Big) = \lim_{n \to \infty} r\Big(\theta - \frac{\rho_n}{\sqrt{n}}\Big) = r(\theta). \tag{2.7}$$

To prove (1.6), it suffices to prove that there exists a countable set S which is dense in R^+ such that for every $\rho \in S$, every $\rho' \in (0, \rho)$ and every subseries $\{n_j\}$ of $\{n\}$, there exists a subseries $\{m_j\}$ of $\{n_j\}$ satisfying

$$\limsup_{m_j \to \infty} P_\theta\{\sqrt{m_j}\,|T_{m_j} - \theta| < \rho'\} \le P\{|Z_\theta| < \rho\}, \qquad \text{a.e.} \quad \theta \in \Theta. \tag{2.8}$$

The result can now be established by repeated use of the optimality of the Likelihood Ratio test. We omit the details because the argument is very similar to that in the proof of Proposition 4 in Wong (1992).

Proof of Theorem 1.4. The proof is the same as the proof of Theorem 1.3.

Proof of Theorem 1.5. To prove Theorem 1.5, it suffices to prove that there exists a countable set S, which is dense in R^+, such that for every $\rho \in S$, $\rho' \in R^+$, $\rho' < \rho$, and every sequence $n_j \to \infty$, there exists a subsequence m_j, satisfying

$$\limsup_{m_j \to \infty} P_\theta(\sqrt{m_j}\,|T_{m_j} - \lambda(\theta)| < \rho') \le P(|\tilde{Z}_\theta| < \rho) \quad \text{a.e.} \quad \theta \in \Theta, \tag{2.9}$$

where $\tilde{Z}_\theta \sim N(0, \lambda'(\theta)^2(1/I(\theta)))$. Let $r(\theta) = \lim_{\theta \to \infty} P_\theta(T_n < \lambda(\theta))$. From Lemma 2.1, we know there exists a sequence $\rho_n \to \rho$ such that

$$\lim r\Big(\theta + \frac{\rho_n}{\sqrt{n}}\Big) = \lim r\Big(\theta - \frac{\rho_n}{\sqrt{n}}\Big) = r(\theta), \qquad \text{a.e.}$$

Let $f_{n_j}(\theta) = P\{T_{n_j} < \lambda(\theta)\} - r(\theta)$. Then there exists a subsequence m_j such that

$$f\Big(\theta \pm \frac{\rho_{m_j}}{\sqrt{m_j}}\Big) \to 0, \qquad \text{a.e.}$$

Let N be the exceptional null set, i.e., for $\theta \notin N$,

$$f\Big(\theta \pm \frac{\rho_{m_j}}{\sqrt{m_j}}\Big) \to 0$$

as $m_j \to \infty$, or

$$\lim P_{\theta \pm \frac{\rho_{m_j}}{\sqrt{m_j}}} \Big\{T_{m_j} < \lambda(\theta \pm \frac{\rho_{m_j}}{\sqrt{m_j}})\Big\} = r(\theta). \tag{2.10}$$

When $\theta \in \{\lambda'(\theta) = 0\}$, (1.10) holds automatically. For $\theta \in \{\lambda'(\theta) \ne 0\} \cap N^c$, where N^c stands for the complement of N, we obtain, by a proof similar to treat for Proposition 4 in Wong (1992),

$$\lim\ \sup\ P_\theta\{\sqrt{m_j}\,|T_{m_j} - \lambda(\theta)| < \rho'\lambda'(\theta)\} \le P\{|Z_\theta| \le \rho\}$$

or

$$\lim\ \sup\ P_\theta\{\sqrt{m_j}\,|T_{m_j} - \lambda(\theta)| \le \rho'\} \le P\{|\tilde{Z}_\theta| < \rho\},$$

where $\tilde{Z}_\theta \sim N(0, (\lambda'(\theta))^2/I(\theta))$, which shows that (2.9) holds.

Let

$$\mathcal{L}_{(0,0)} = \{v : v = (c_1, \ldots, c_k, 0, 0, \ldots) \in \mathcal{L}_0, k = 1, 2, \ldots\}. \tag{2.11}$$

Lemma 2.2. *Suppose that $r(\theta)$ is a measurable function on $\Theta \subset \mathcal{L}$ and that $Q = N(m_Q, S)$ is a Gaussian measure on $\mathcal{L}$. Then there exists a set N with $Q(N) = 0$*

such that for every $\theta \notin N$, every $h \in R^+$ and every $v = (c_1, \ldots, c_k, 0, 0, \ldots) \in \mathcal{L}_{(0,0)}$ (k is an arbitrary positive integer), there exists a series of $h_n(\theta, h, v)$ and $v_n(\theta, h, v) = (c_1, \ldots, c_k, 0, 0, \ldots) \in \mathcal{L}_{(0,0)}$ such that

$$h_n \to h, \quad v_n \to v, \quad r\Big(\theta \pm \frac{h_n}{\sqrt{n}} v_n\Big) \to r(\theta). \tag{2.12}$$

Proof. This lemma is an R^k version of Lemma 2.1.

Proof of Theorem 1.6. Suppose that Q is a Gaussian measure on $\mathcal{L}$. We use the notation $Q = N(m_Q, S)$, where $m_Q \in \mathcal{L}$, S is a nonnegative, self-adjoint, compact operator with its eigenvalues λ_i satisfying $\sum \lambda_i < \infty$. Let e_i, $i = 1, 2, \ldots$ be the eigenvectors of S. Then S has the representation $Sx = \sum \lambda_i \langle x, e_i \rangle e_i$. Without loss of generality, we may assume that $\{e_i, i = 1, 2, \ldots\}$ is a base of $\mathcal{L}$, where $e_i = (0, \ldots, 0, \underset{i}{1}, 0, \ldots)$. To prove the theorem, we need only to prove the following conclusion: for almost all $\theta \in \Theta$

$$\lim P_\theta\{\sqrt{n}\,|T_n - \lambda(\theta)| < h\} \le P\{|Z_{(\theta,v)}| < h\}, \tag{2.13}$$

where $v \in \mathcal{L}_0$, $h > 0$, $Z_{(\theta,v)} \sim N(0, |\lambda'_\theta(v)|^2 / I(\theta, v))$.

Since, as a function of v, $\lambda'_\theta(v)^2 / I(\theta, v)$ is continuous and the set $\mathcal{L}_{(0,0)} = \{v = (c_1, c_2, \ldots, c_k, 0, \ldots) \in \mathcal{L}_0,\ k = 1, \ldots\}$ is dense in $\mathcal{L}_0$, it is easy to show that

$$\sup_{\mathcal{L}_0} \frac{\lambda'_\theta(v)^2}{I(\theta, v)} = \sup_{\mathcal{L}_{(0,0)}} \frac{\lambda'_\theta(v)^2}{I(\theta, v)}. \tag{2.14}$$

From (2.14) we know that to prove (2.13), it is equivalent to prove that there exists a countable set S which is dense in R^+ such that for every $h \in S$, $h' \in (0, h)$, $v \in \mathcal{L}_{(0,0)}$ and every subseries $\{n_j\}$ of $\{n\}$, there exists a subseries $\{m_j\}$ of $\{n_j\}$ satisfying

$$\lim_{j\to\infty} P_\theta\{|T_{m_j} - \lambda(\theta)| < h'\} \le P\{|Z_{(\theta,v)}| < h\}, \quad \text{a.e. } Q. \tag{2.15}$$

Now suppose that $v = (c_1, \ldots, c_k, 0, 0, \ldots) = (v_k, 0) \in \mathcal{L}_{(0,0)}$. According to Lemma 2.2, for every $h \in R^+$, there exist series $h_{(n,k)}$ and $v_{(n,k)} \in R^k$ such that

$$h_{(n,k)} \to h, \quad v_{(n,k)} \to v_k \quad r\Big(\theta_k \pm \frac{h_{(n,k)}}{\sqrt{n}} v_{(n,k)}, \tilde{\theta}_k\Big) \to r(\theta), \quad \text{a.e. } Q, \tag{2.16}$$

where $\theta = (\theta_k, \tilde{\theta}_k)$ and $\theta_k \in R^k$. Let

$$f_{n_j}(\theta) = P_\theta\{T_{n_j} < \lambda(\theta)\} - r(\theta). \tag{2.17}$$

From the definition of $r(\theta)$ we know that

$$f_{n_j}(\theta) \to 0, \qquad \text{a.e. } Q.$$

Consider

$$Q\{A\} = \int dQ(\tilde{\theta}_k) \int_{A_{\tilde{\theta}_k}} dQ(\theta_k|\tilde{\theta}_k),$$

where $A = \{|f_{n_j}(\theta_k \pm \frac{h_{(n_j,k)}}{\sqrt{n_j}} v_{(n_j,k)}, \tilde{\theta}_k)| \geq \varepsilon\}$ and $A_{\tilde{\theta}_k} = \{\theta_k : (\theta_k, \tilde{\theta}_k) \in A\}$. Let $\theta = (\xi_1, \xi_2, \ldots)$ and $\theta_k = (\xi_1, \ldots, \xi_k)$. It is easy to show that under $Q, \xi_1, \xi_2, \ldots$ are independent and $\xi_i \sim N(m_{Q,i}, \lambda_i), i = 1, 2, \ldots$, where $m_{Q,i}$ is the ith coordinate of m_Q. Therefore

$$\begin{aligned} Q(A) &= \int dQ(\tilde{\theta}_k) \int_{A_{\tilde{\theta}_k}} \Big(\frac{1}{\sqrt{2\pi}}\Big)^k \prod_{i=1}^k \frac{1}{\sqrt{\lambda_i}} \exp\Big\{ -\frac{1}{2\lambda_i}(c_i - m_{Q,i})^2\Big\} dc_i \\ &= \int dQ(\tilde{\theta}_k) \int_{\{|f_{n_j}(\theta_k, \tilde{\theta}_k)| \geq \varepsilon\}} (\frac{1}{\sqrt{2\pi}})^k \\ &\quad \prod_{i=1}^k \frac{1}{\sqrt{\lambda_i}} \exp\{-\frac{1}{2\lambda_i}(c_i \mp \frac{h_{(n_j,k)}}{\sqrt{n_j}} v_{(n_j,k)}(i) - m_{Q,i})^2\} dc_i, \end{aligned}$$

where $v_{(n_j,k)}(i)$ is the ith coordinate of the vector $v_{(n_j,k)}$. Since $h_{(n_j,k)}$ and $v_{(n_j,k)}$ can be chosen so that they are uniformly bounded, there exists a pair of positive numbers b and B such that

$$\begin{aligned} Q(A) &\leq \int dQ(\tilde{\theta}_k) \int_{\{|f_{n_j}(\theta_k, \tilde{\theta}_k)| \geq \varepsilon\}} B \prod_{l=1}^k \exp\{bc_l^2\} \\ &\quad \Big(\frac{1}{\sqrt{2\pi}}\Big)^k \prod_{i=1}^k \frac{1}{\sqrt{\lambda_i}} \exp\Big\{ -\frac{1}{2\lambda_i}(c_i - m_{Q,i})^2\Big\} dc_i \\ &= \int_{\{|f_{n_j}(\theta)| \geq \varepsilon\}} B \prod_{l=1}^k \exp\{bc_l^2\} dQ(\theta). \end{aligned}$$

Since $f_{n_j} \to 0$, a.e. Q, and $\prod_{l=1}^k \exp\{bc_l^2\}$ is integrable with respect to Q, we find by the Lebesgue's dominated convergence theorem,

$$\int_{\{|f_{n_j}(\theta)| \geq \varepsilon\}} B \prod_{l=1}^k \exp\{bc_l^2\} dQ(\theta) \to 0,$$

therefore $Q\{A\} \to 0$ as $j \to \infty$ for every $\varepsilon > 0$, i.e., $f_{n_j}(\theta_k \pm \frac{h_{(n_j,k)}}{\sqrt{n_j}} v_{(n_j,k)}, \tilde{\theta}_k) \xrightarrow{Q} 0$. Therefore, there exists a subseries $\{m_j\}$ of $\{n_j\}$ such that as $m_j \to \infty$,

$$f_{n_j}\Big(\theta_k \pm \frac{h_{(n_j,k)}}{\sqrt{n_j}} v_{(n_j,k)}, \tilde{\theta}_k\Big) \to 0, \quad \text{a.e. } Q. \tag{2.18}$$

From (2.16) and (2.18) we know that for almost all $\theta \in \Theta, (Q)$,

$$a = \lim_{j\to\infty} P_{(\theta_k + \frac{h_{(m_j,k)}}{\sqrt{m_j}} v_{(m_j,k)}, \tilde{\theta}_k)} \Big(T_{m_j} < \lambda(\theta_k + \frac{h_{(m_j,k)}}{\sqrt{m_j}} v_{(m_j,k)}, \tilde{\theta}_k) \Big)$$
$$= \lim_{j\to\infty} P_{\theta_k - \frac{h_{(m_j,k)}}{\sqrt{m_j}} v_{(m_j,k)}} \Big(T_{m_j} < \lambda(\theta_k - \frac{h_{(m_j,k)}}{\sqrt{m_j}} v_{(m_j,k)}, \tilde{\theta}_k) \Big) = b. \quad (2.19)$$

Without loss of generality, we assume that $\lambda'_\theta(v) = \lambda'_\theta(v_k, 0) > 0$.
Consider the test problem

$$H_0 : P_{(\theta + \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)}, 0))} \quad \text{vs} \quad H_A : P_\theta. \quad (2.20)$$

Let

$$B_{m_j} = \Big\{ T_{m_j} < \lambda(\theta + \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)}, 0)) \Big\}$$

be a rejection region of the test and let

$$A_{m_j} = \{ \Delta_{m_j,\theta,(v_k,0)} h_{(m_j,k)} I(\theta,(v_k,0)) + R_{m_j}$$
$$< I(\theta,(v_k,0)) h^2_{(m_j,k)} + \mathcal{Z}_{\alpha_j} \sqrt{h^2_{(m_j,k)} I(\theta,(v_k,0)} \},$$

where $\mathcal{Z}_{\alpha_j}$ is the α_j-quantile of the standard normal distribution. Let $\alpha_j \searrow a' > a$. Since the family $\{P_{\theta,m_j}\}$ satisfies Condition (L$''$), we have, using a standard technique,

$$\mathcal{L}_{\theta + \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)} (\Delta_{m_j,\theta,(v_k,0)} - h_{(m_j,k)}) \to N\Big(0, \frac{1}{I(\theta,(v_k,0))}\Big),$$

from which it follows that

$$P_{\theta + \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)} (A_{m_j}) \to a' > a.$$

Since A_{m_j} is the likelihood ratio rejection region,

$$\limsup P_\theta\{\sqrt{m_j}(T_{m_j} - \lambda(\theta)) \le \lambda'_\theta(v_k,0)h'\}$$
$$\le \limsup P_\theta \Big\{ T_{m_j} < \lambda(\theta + \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)) \Big\}$$
$$\le \lim P\{A_{m_j}\} = P\Big\{ Z_{(\theta,(v_k,0))} < \lambda'_\theta h + \frac{\mathcal{Z}_{a'}\lambda'_\theta}{\sqrt{I(\theta,(v_k,0))}} \Big\}, \quad (2.21)$$

where $h' < h$.

Now consider the test problem:

$$H_0 : P_\theta \quad \text{vs} \quad H_A : P_{\theta - \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)}. \tag{2.22}$$

Let

$$\tilde{B}_{m_j} = \Big\{T_{m_j} < \lambda(\theta - \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)}, 0))\Big\}$$

be the rejection region of (2.22). Let

$$\begin{aligned} \tilde{A}_{m_j} = \{&\Delta_{m_j,\theta,v}(-h_{(m_j,k)})I(\theta,(v_k,0)) + \tilde{R}_{m_j} \\ &> I(\theta,(v_k,0))h^2_{(m_j,k)} + \mathcal{Z}_{\alpha_j}\sqrt{I(\theta,(v_k,0))h^2_{(m_j,k)}}\}, \end{aligned}$$

where $\alpha_j \searrow (1-b') > 1-b$. It is easy to know that

$$\lim_{m_j\to\infty} P_{\theta - \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)}\{\tilde{A}_{m_j}\} = b' < b = \lim_{m_j\to\infty} P_{\theta - \frac{h_{(m_j,k)}}{\sqrt{m_j}}(v_{(m_j,k)},0)}\{\tilde{B}_{m_j}\}.$$

Since $\tilde{A}_{m_j}$ is the rejection region of the likelihood ratio test, we have

$$\begin{aligned} &\liminf P_\theta\{\sqrt{m_j}(T_{m_j} - \lambda(\theta)) \le -h'\lambda'_\theta((v_k,0))\} \ge \lim P_\theta\{\tilde{B}_{m_j}\} \\ \ge\ &\lim P_\theta\{\tilde{A}_{m_j}\} = P\Big\{Z_{(\theta,-(v_k,0))} < -\lambda'_\theta((v_k,0))h + \frac{\mathcal{Z}_{b'}\lambda'_\theta((v_k,0))}{\sqrt{I(\theta,(v_k,0))}}\Big\}. \end{aligned} \tag{2.23}$$

Comparing (2.21) and (2.23), we obtain (2.15).

3. An Example

Let

$$Y = ax + \phi(u) + \varepsilon, \tag{3.1}$$

where $x \in [0,1]$, $\phi(\cdot) \in \mathcal{L}_2(0,1)$ and the error distribution is standard normal. Let $(X,U) \sim \mu \times \nu$, where μ and ν are known distributions on $[0,1]$. Without loss of generality, we suppose that $\text{var}(X) = \int x^2 d\mu - (\int x d\mu)^2 > 0$. Suppose that X, U and ε are independent. It is easy to see that (3.1) is a semiparametric model with parameter $\theta = (a,\phi)$. Consider the parameter θ as a member of a Hilbert space $\mathcal{L}$ with inner product

$$\langle\theta_1,\theta_2\rangle = a_1a_2 + \int_0^1 \phi_1\phi_2 du. \tag{3.2}$$

Let $d\mu \times d\nu \times dy$ be a σ-finite measure on $[0,1]^2 \times R^1$. The density of (X,U,Y) with respect to $d\mu \times d\nu \times dy$ is

$$\frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{1}{2}(y - ax - \phi(u))^2\Big\}, \tag{3.3}$$

which, as a function of (x,u,y,θ), is measurable. Let (X_i,U_i,Y_i), $i=1,\ldots,n$, be a sample of size n from the model (3.1). It is easy to see that the joint density of (X_i,U_i,Y_i), $i=1,\ldots,n$, satisfies Condition (M).

Let $v=(a_0,\phi_0)$ be a vector with $\|v\|^2=a_0^2+\int_0^1\phi_0^2(u)du=1$. Let $v_n=(a_n,\phi_n)$ be a series of vectors in $\mathcal{L}$. Suppose that $v_n\to v=(a_0,\phi_0)$ in $\mathcal{L}$. Let h_n be a series of real numbers with limit h. The likelihood ratio of the sample $((x_1,u_1,y_1),\ldots,(x_n,u_n,y_n))$ has the form

$$\begin{aligned}\frac{dP_{\theta+\frac{h_n}{\sqrt{n}}v_n}}{dP_\theta} &= \prod_{i=1}^n \exp\Big\{-\frac{1}{2}[y_i-ax_i-\phi(u_i)-\frac{h_n}{\sqrt{n}}(a_nx_i+\phi_n(u_i))]^2\Big\}\\ &\quad \Big/\prod_{i=1}^n \exp\Big\{-\frac{1}{2}[y_i-ax_i-\phi(u_i)]^2\Big\}\\ &= \exp\Big\{h_nI(\theta,v)\Delta_{n,\theta,v}-\frac{1}{2}h_n^2I(\theta,v)+o_p(1)\Big\},\end{aligned}\tag{3.4}$$

where

$$I(\theta,v)=E(a_0X+\phi_0(U))^2$$

and

$$\Delta_{n,\theta,v}=\sum_{i=1}^n\frac{1}{\sqrt{n}}(Y_i-a_0X_i-\phi_0(U_i))(a_nX_i+\phi_n(U_i)/I(\theta,v))\xrightarrow{\mathcal{L}}N\Big(0,\frac{1}{I(\theta,v)}\Big),$$

which shows that the family $\{P_\theta,\theta\in\Theta\}$ satisfies Condition (L$''$). Let $\lambda(\theta)=a$. It is easy to show that $\lambda_\theta'(v)=a_0$, where $v=(a_0,\phi_0)\in\mathcal{L}$. So,

$$\sup_{v\in\mathcal{L}}\frac{\lambda_\theta'(v)^2}{I(\theta,v)}=\frac{1}{\mathrm{var}(X)}.$$

Now Suppose that $Q\sim N(m_Q,S)$ is a Gaussian measure in the Hilbert space $\mathcal{L}=\{\theta:\theta=(a,\phi),a\in R,\phi\in\mathcal{L}_2(0,1)\}$. By using Theorem 1.5, we conclude that for every stable estimator T_n of $\lambda(\theta)=a$, there exists a null set N in $\mathcal{L}$, such that for every $\theta=(a,\phi)\notin N$ and every $\rho>0$,

$$\lim P_\theta\{\sqrt{n}\,|T_n-a|<\rho\}\le P\{|Z|<\rho\},$$

where $Z\sim N(0,\sigma^2)$ and $\sigma^2=(\mathrm{var}(X))^{-1}$.

Acknowledgement

Research of Z. Zheng was supported partially by Hong Kong Baptist College, the NSFC and the Doctoral Program Foundation of China. Research of K. T. Fang was supported by the Hong Kong UPGC Grant. The authors are grateful to the referees and editors for their useful comments which led to improvement of this paper.

References

Bahadur, R. R. (1964). On Fisher's bound for asymptotic variances. *Ann. Math. Statist.* **35**, 1545-1552.

Fisher, R. A. (1925). Theorey of statistical estimation. *Proc. Cambridge Philos. Soc.* **22**, 700-725,

Kuo, H. H. (1975). *Gaussian Measure in Banach Space.* Lecture Notes in Math. **463**, Springer Verlag, New York.

LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **1**, 175-194.

Natanson, E. P. (1950). *Theory of Real Function*, Vol. **1** (In Chinese, Translated from Russian), 297, The Higher Education Press.

Wong, W. H. (1992). On asymptotic efficiency in estimation theory. *Statistica Sinica* **2**, 47-68.

Department of Probability and Statistics, Peking University, Beijing 100871.
Department of Mathematics, Hong Kong Baptist Colledge, Kowloon, Hong Kong.

(Received June 1992; accepted January 1994)