Ann. Inst. Statist. Math.
34 (1982), Part A, 19-37

UNBIASED ESTIMATORS IN THE SENSE OF LEHMANN AND THEIR DISCRIMINATION RATES

HISATAKA KUBOKI*

(Received Jan. 29, 1981; revised Aug. 10, 1981)

Summary

In this paper, we define the index of performance of unbiased estimators in the sense of Lehmann (L-unbiased), which evaluates the power for the estimators to discriminate any wrong values of a parametric function from a correct one. We shall call the index *discrimination rate* of the estimator. The larger discrimination rate the estimator has, the more desirable it is. An upper bound of discrimination rates is obtained, which is given by the *sensitivity* of the probability family under consideration. The discrimination rates of several L-unbiased estimators are investigated. Moreover we discuss the conditions under which the L-unbiased estimator is improved in the sense of discrimination rate by the L-unbiased estimator depending only on a sufficient statistic.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables with the common distribution P_θ, $\theta \in \Theta$, where Θ is an open interval in the real line. We write $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)$ and let $\delta(\boldsymbol{x})$ be an estimate of a given parametric function $\gamma(\theta)$. Let $W(\theta, \delta(\boldsymbol{x}))$ be a loss incurred by $\delta(\boldsymbol{x})$. We assume that the estimator $\delta(\boldsymbol{X})$ is unbiased in the sense of Lehmann [4], that is, it satisfies the inequality

$$(1.1) \qquad \mathrm{E}_\theta W(\theta, \delta(\boldsymbol{X})) \leqq \mathrm{E}_\theta W(\tau, \delta(\boldsymbol{X})) , \qquad \text{for } \theta, \tau \in \Theta .$$

We shall call such an estimator *L-unbiased* with respect to the loss function W, or simply L-unbiased. Since the inequality (1.1) implies that $\mathrm{E}_\theta W(\tau, \delta(\boldsymbol{X}))$, as a function of τ, takes the minimum value at θ, we have under suitable conditions,

* This research was supported in part by a Grant-in-Aid for Scientific Research of the Japanese Ministry of Education, Science and Culture.

$$E_\theta W'(\theta, \delta(\boldsymbol{X}))=0 , \tag{1.2}$$

where the symbol $'$ denote the differentiation with respect to θ. By differentiating (1.2) and using Schwarz's inequality, we obtain, under suitable conditions,

$$E_\theta W''(\theta, \delta(\boldsymbol{X}))\leqq \sqrt{E_\theta W'^2(\theta, \delta(\boldsymbol{X}))}\sqrt{nI(\theta)} ,$$

that is,

$$E_\theta W''(\theta, \delta(\boldsymbol{X}))/\sqrt{E_\theta W'^2(\theta, \delta(\boldsymbol{X}))}\leqq \sqrt{nI(\theta)} , \tag{1.3}$$

where $I(\theta)$ is the Fisher information. Note that (1.3) is reduced to the Cramér-Rao inequality when $W(\theta, \delta(\boldsymbol{x}))=\{\delta(\boldsymbol{x})-\gamma(\theta)\}^2$. We are interested in statistical implications of the inequality (1.3). We shall investigate this inequality in more general situations, and show that the left-hand side of (1.3) can be used for evaluating the performance of any L-unbiased estimator.

Usually, the performance of an estimator is evaluated by the degree of its nearness to the correct value of $\gamma(\cdot)$. When dealing with an L-unbiased estimator, however, the degree of its departure from the wrong values of $\gamma(\cdot)$ can be regarded as another measure for evaluating its performance. In fact, each L-unbiased estimator does not come, on the average, closer to any wrong value of $\gamma(\cdot)$ than to the correct one. A question naturally arises as to what measure should we use for evaluating the magnitude of their departure from the wrong values of $\gamma(\cdot)$.

Generalizing the condition (1.1), we shall consider those functions $c(\boldsymbol{x}, \theta)$ which satisfy the inequality

$$E_\theta c(\boldsymbol{X}, \theta)\leqq E_\theta c(\boldsymbol{X}, \tau) , \qquad \text{for } \theta, \tau \in \Theta . \tag{1.4}$$

If we take $c(\boldsymbol{x}, \theta)=W(\theta, \delta(\boldsymbol{x}))$, the condition (1.4) is reduced to (1.1). We shall say that the function $c(\boldsymbol{x}, \theta)$ with this property is *contrastive* or it has *contrastive power* (see Remark in this section). One of the natural ideas for evaluating the magnitude of contrastive power of $c(\boldsymbol{x}, \theta)$ is to observe the rate of change of $E_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}$ to small changes in τ at θ. But it must be normalized in a suitable scale, because any contrastive function multiplied by a positive constant is also a contrastive function. For a normalizing factor, we consider the standard deviation $\sqrt{\mathrm{Var}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}}$. Thus our measure for evaluating the magnitude of the contrastive power of $c(\boldsymbol{x}, \theta)$ is defined as follows.

DEFINITION 1.1. For a contrastive function $c(\boldsymbol{x}, \theta)$, $\theta \in \Theta$, we shall define its *discrimination rate* $D(\theta; c)$ at θ by

$$D(\theta\,;c)=2\lim_{\tau\to\theta}\Delta(\tau,\theta\,;c)/|\tau-\theta|\ ,$$

where

$$\Delta(\tau,\theta\,;c)=\mathrm{E}_\theta\{c(\boldsymbol{X},\tau)-c(\boldsymbol{X},\theta)\}/\sqrt{\mathrm{Var}_\theta\{c(\boldsymbol{X},\tau)-c(\boldsymbol{X},\theta)\}}\ .$$

We define $\Delta(\tau,\theta\,;c)=0$ whenever the numerator vanishes. In particular when $c(\boldsymbol{x},\theta)=W(\theta,\delta(\boldsymbol{x}))$, we shall write $D(\theta\,;\delta,W)$ instead of $D(\theta\,;c)$ and call it the discrimination rate of $\delta(\boldsymbol{X})$.

Remark. A function $\psi(x_1,\theta)$ with the property

$$\mathrm{E}_\theta\,\psi(X_1,\theta)<\mathrm{E}_\theta\,\psi(X_1,\tau)\ ,\qquad \theta,\tau\in\Theta$$

is called the *contrast function* by Pfanzagl [5]. It is used for constructing the *minimum contrast estimator* $\hat{\theta}_n(\boldsymbol{X})$ by solving the equation

$$\sum_{i=1}^n\psi(X_i,\hat{\theta}_n(\boldsymbol{X}))=\inf_{\theta\in\Theta}\sum_{i=1}^n\psi(X_i,\theta)\ .$$

It is well known that $\mathrm{E}_\theta\,\psi''(X_1,\theta)/\sqrt{\mathrm{E}_\theta\,\psi'^2(X_1,\theta)}$ is—under suitable regularity conditions—the reciprocal of the asymptotic variance of $\hat{\theta}_n(\boldsymbol{X})$ when $n\to\infty$ (Huber [1]).

Some properties of contrastive functions and their discrimination rates are investigated in the following section, where we prove that, under suitable conditions, an upper bound of discrimination rates is given by $\sqrt{n}\,s(\theta)$, where $s(\theta)$ is defined by

$$s(\theta)=2\lim_{\tau\to\theta}\rho(P_\tau,P_\theta)/|\tau-\theta|\ ,$$

and where $\rho(P_\tau,P_\theta)$ is the Hellinger distance between P_τ and P_θ. $s(\theta)$ is the *sensitivity* of the family $\{P_\theta;\ \theta\in\Theta\}$ to small changes in τ at θ (Pitman [6]).

In Section 3, we treat the discrimination rates of L-unbiased estimators. We first establish the inequalities

$$D(\theta\,;\delta,W)\leqq s(\theta\,;\delta)\leqq\sqrt{n}\,s(\theta)\ ,$$

where $s(\theta\,;\delta)$ is the sensitivity of the family of distributions induced by $\delta(\boldsymbol{X})$. In particular this result implies the inequality (1.3). Since the larger discrimination rate $\delta(\boldsymbol{X})$ has, the more desirable it is, a reasonable index of its performance is $D(\theta\,;\delta,W)/\sqrt{n}\,s(\theta)$, the ratio of its discrimination rate to the maximum possible. We shall call it *efficiency* of $\delta(\boldsymbol{X})$. We apply this formulation to several loss functions and L-unbiased estimators.

In the last section, we consider the case where a sufficient statistic $\boldsymbol{T}$ for $\{P_\theta;\ \theta\in\Theta\}$ exists. For the quadratic loss function, each (L-)unbiased estimator can be improved by using the Rao-Blackwell theorem. So the following problem arises. Suppose that $\delta(\boldsymbol{X})$ is L-unbiased with respect to $W(\cdot, \cdot)$. What form of loss function W does permit the construction of the L-unbiased estimator which depends only on $\boldsymbol{T}$ and is better than $\delta(\boldsymbol{X})$ in the sense of discrimination rate? A result for this problem is given there. We also investigate the properties of the contrastive function $\bar{c}(\boldsymbol{t}, \theta)=\mathrm{E}\{c(\boldsymbol{X}, \theta)|\boldsymbol{T}=\boldsymbol{t}\}$, and especially the relation between $D(\theta; \bar{c})$ and $D(\theta; c)$.

2. An upper bound of discrimination rates

Let $f(x, \theta)$ be the density of P_θ with respect to a σ-finite measure μ. We shall first prove the following lemma, which is essential in our discussion.

LEMMA 2.1. *For any contrastive function $c(\boldsymbol{x}, \theta)$ and for all $\theta, \tau \in \Theta$, the following inequality holds*:

$$
\begin{aligned}
(2.1) \qquad & \mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\} \\
& \leqq [\mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/2+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\}^2/2]^{1/2} \\
& \quad \times\left[4n \int \{\sqrt{f(x, \tau)}-\sqrt{f(x, \theta)}\}^2 d\mu\right]^{1/2}.
\end{aligned}
$$

PROOF. We shall denote the joint density $f(x_1, \theta)f(x_2, \theta)\cdots f(x_n, \theta)$ by $f_n(\boldsymbol{x}, \theta)$. Using Schwarz's inequality, we find

$$
\begin{aligned}
& \mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\} \\
& =\int \{c(\boldsymbol{x}, \tau)-c(\boldsymbol{x}, \theta)\}\{\sqrt{f_n(\boldsymbol{x}, \tau)}+\sqrt{f_n(\boldsymbol{x}, \theta)}\} \\
& \quad \times\{\sqrt{f_n(\boldsymbol{x}, \theta)}-\sqrt{f_n(\boldsymbol{x}, \tau)}\}d\mu^n \\
& \leqq\left[\int \{c(\boldsymbol{x}, \tau)-c(\boldsymbol{x}, \theta)\}^2\{\sqrt{f_n(\boldsymbol{x}, \tau)}+\sqrt{f_n(\boldsymbol{x}, \theta)}\}^2 d\mu^n/4\right]^{1/2} \\
& \quad \times\left[4\int \{\sqrt{f_n(\boldsymbol{x}, \tau)}-\sqrt{f_n(\boldsymbol{x}, \theta)}\}^2 d\mu^n\right]^{1/2}.
\end{aligned}
$$

From the inequalities

$$
\{\sqrt{f_n(\boldsymbol{x}, \tau)}+\sqrt{f_n(\boldsymbol{x}, \theta)}\}^2 \leqq 2f_n(\boldsymbol{x}, \tau)+2f_n(\boldsymbol{x}, \theta)
$$

and

$$
\begin{aligned}
& \int \{\sqrt{f_n(\boldsymbol{x}, \tau)}-\sqrt{f_n(\boldsymbol{x}, \theta)}\}^2 d\mu^n=2\left[1-\left\{\int \sqrt{f(x, \tau)}\sqrt{f(x, \theta)}d\mu\right\}^n\right] \\
& \quad \leqq 2n\left[1-\int \sqrt{f(x, \tau)}\sqrt{f(x, \theta)}d\mu\right]=n\int \{\sqrt{f(x, \tau)}-\sqrt{f(x, \theta)}\}^2 d\mu ,
\end{aligned}
$$

the lemma follows.

Put

$$\lambda(\tau, \theta; c)=\frac{\mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\}}{[\mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/2+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\}^2/2]^{1/2}},$$

where we define $\lambda(\tau, \theta; c)=0$ whenever the numerator vanishes, and put

$$d(\theta; c)=\varliminf_{\tau\to\theta} \lambda(\tau, \theta; c)/|\tau-\theta| . \tag{2.2}$$

Moreover put

$$\rho(P_\tau, P_\theta)=\left[\int \{\sqrt{f(x, \tau)}-\sqrt{f(x, \theta)}\}^2 d\mu\right]^{1/2},$$

and

$$s(\theta)=2 \varliminf_{\tau\to\theta} \rho(P_\tau, P_\theta)/|\tau-\theta| .$$

$\rho(P_\tau, P_\theta)$ is the Hellinger distance between P_τ and P_θ, and $s(\theta)$ is the sensitivity of the family $\{P_\theta;\ \theta\in\Theta\}$ under consideration (Pitman [6]). From this lemma, for any contrastive function we have the inequality

$$d(\theta; c)\leqq\sqrt{n}\, s(\theta), \qquad \text{for all } \theta\in\Theta . \tag{2.3}$$

From now on, the following conditions are assumed.

(i) For almost all x, $f(x, \theta)$ has a θ derivative at each $\theta\in\Theta$, which will be denoted by $f'(x, \theta)$.

(ii) For all $\theta, \tau\in\Theta$, $\mathrm{E}_\theta c^2(\boldsymbol{X}, \tau)$ is finite.

(iii) For all $\boldsymbol{x}\in\boldsymbol{R}^n$, $c(\boldsymbol{x}, \theta)$ is twice continuously differentiable with respect to θ. The derivatives at each $\theta\in\Theta$ will be denoted by $c'(\boldsymbol{x}, \theta)$ and $c''(\boldsymbol{x}, \theta)$.

Then we shall establish the following theorem.

THEOREM 2.1. *If both* $\mathrm{E}_\theta c'^2(\boldsymbol{X}, \tau)$ *and* $\mathrm{E}_\theta c''(\boldsymbol{X}, \tau)$ *are continuous in* θ *and* τ, *then for all* $\theta\in\Theta$

$$D(\theta; c)\leqq\sqrt{n}\, s(\theta) . \tag{2.4}$$

For the proof of this theorem, we need the following lemma.

LEMMA 2.2. *Suppose that* $\mathrm{E}_\theta c'^2(\boldsymbol{X}, \theta)$ *is finite for all* $\theta\in\Theta$. *Then*

$$\lim_{\tau\to\theta} \mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/(\tau-\theta)^2=\mathrm{E}_\theta c'^2(\boldsymbol{X}, \theta)$$

implies

$$\lim_{\tau\to\theta} \mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}/(\tau-\theta)=\mathrm{E}_\theta c'(\boldsymbol{X}, \theta)=0 .$$

PROOF. Since the first equality is equivalent to

$$\lim_{\tau\to\theta} \mathrm{E}_\theta |\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}/(\tau-\theta)-c'(\boldsymbol{X}, \theta)|^2=0 \tag{2.5}$$

(Pitman [6], p. 99, Corollary), we have

$$\lim_{\tau\to\theta} \mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}/(\tau-\theta)=\mathrm{E}_\theta\, c'(\boldsymbol{X}, \theta) .$$

From the definition of contrastive function, $\mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}$, as a function of τ, takes the minimum value at θ. Hence we have

$$\lim_{\tau\to\theta} \mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}/(\tau-\theta)=0 ,$$

as was to be proved.

PROOF OF THEOREM 2.1. We have only to prove the following proposition: if for each $\theta\in\Theta$, both $\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \tau)$ and $\mathrm{E}_\theta\, c''(\boldsymbol{X}, \tau)$ are continuous in τ, then

$$d(\theta\,; c)=\lim_{\tau\to\theta} \frac{\mathrm{E}_\theta\, c''(\boldsymbol{X}, \theta)/2+\mathrm{E}_\tau\, c''(\boldsymbol{X}, \tau)/2}{[\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta)/2+\mathrm{E}_\tau\, c'^2(\boldsymbol{X}, \tau)/2]^{1/2}} \tag{2.6}$$

and

$$D(\theta\,; c)=\mathrm{E}_\theta\, c''(\boldsymbol{X}, \theta)/\sqrt{\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta)} . \tag{2.7}$$

In fact, if it is proved, we have from the assumptions of the theorem,

$$d(\theta\,; c)=D(\theta\,; c) ,$$

which together with (2.3) implies (2.4). To this end, notice that

$$\{c(\boldsymbol{x}, \tau)-c(\boldsymbol{x}, \theta)\}^2=\left\{\int_\theta^\tau c'(\boldsymbol{x}, \xi)d\xi\right\}^2$$

and

$$c(\boldsymbol{x}, \tau)-c(\boldsymbol{x}, \theta)=c'(\boldsymbol{x}, \theta)(\tau-\theta)+\int_\theta^\tau (\tau-\xi)c''(\boldsymbol{x}, \xi)d\xi .$$

Using Schwarz's inequality and Fubini's theorem, we have

$$\begin{aligned}\mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/(\tau-\theta)^2&=\mathrm{E}_\theta \left\{\int_\theta^\tau c'(\boldsymbol{X}, \xi)d\xi\right\}^2\Big/(\tau-\theta)^2\\ &\leqq \mathrm{E}_\theta \int_\theta^\tau c'^2(\boldsymbol{X}, \xi)d\xi/(\tau-\theta)\\ &=\int_\theta^\tau \mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \xi)d\xi/(\tau-\theta) .\end{aligned}$$

Since $\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \xi)$ is continuous in ξ, it follows that

$$\overline{\lim_{\tau\to\theta}}\, \mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/(\tau-\theta)^2\leqq \mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta) .$$

On the other hand, it follows from Fatou's lemma that

$$\varliminf_{\tau\to\theta} \mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/(\tau-\theta)^2 \geqq \mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta) .$$

Thus

$$\mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2 = \mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta)(\tau-\theta)^2 + o((\tau-\theta)^2) .$$

Therefore from Lemma 2.2, we have

$$\mathrm{E}_\theta\, c'(\boldsymbol{X}, \theta)=0 , \qquad \text{for all } \theta\in\Theta .$$

From this fact and Fubini's theorem, we obtain

$$\mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\} = \int_\theta^\tau (\tau-\xi)\, \mathrm{E}_\theta\, c''(\boldsymbol{X}, \xi)d\xi .$$

Since $\mathrm{E}_\theta\, c''(\boldsymbol{X}, \xi)$ is continuous in ξ, we have

$$\mathrm{E}_\theta \{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\} = \mathrm{E}_\theta\, c''(\boldsymbol{X}, \theta)(\tau-\theta)^2/2 + o((\tau-\theta)^2) .$$

From these facts and the definitions of $D(\theta; c)$ and $d(\theta; c)$, (2.6) and (2.7) are easily deduced. Thus the theorem is proved.

Remarks. (i) By Fatou's lemma, it is easy to see that in all cases

$$s^2(\theta) \geqq I(\theta) , \tag{2.8}$$

where $I(\theta)$ is the Fisher information, that is,

$$I(\theta) = \int f'^2(x, \theta)/f(x, \theta)d\mu .$$

Pitman [6] called the family of probability measures *smooth* at θ, if $s^2(\theta)$ is finite and equal to $I(\theta)$, and gave the following simple conditions for the smoothness. For all $\theta, \tau\in\Theta$,

(P-1) $\int f'(x, \theta)d\mu=0, \ = \frac{d}{d\theta}\int f(x, \theta)d\mu,$

(P-2) $\frac{d}{d\tau}\int \sqrt{f(x, \tau)}\sqrt{f(x, \theta)}\,d\mu = \int \frac{d}{d\tau}\sqrt{f(x, \tau)}\sqrt{f(x, \theta)}\,d\mu,$

(P-3) $I(\theta)$ is a continuous function of θ.

If in particular the supports of $f(x, \theta)$, $\theta\in\Theta$ are independent of θ, (P-3) is sufficient to ensure the smoothness. Indeed, by Schwarz's inequality and Fubini's theorem, we find

$$4\rho^2(P_\tau, P_\theta) = 4\int \left\{\int_\theta^\tau \sqrt{f(x, \xi)}'d\xi\right\}^2 d\mu \leqq (\tau-\theta)\int_\theta^\tau I(\xi)d\xi \tag{2.9}$$

(Ibragimov and Khas'minskii [2]), which implies

$$s^2(\theta) \leqq I(\theta) ,$$

and which in turn together with (2.8) implies

$$s^2(\theta)=I(\theta)\ .$$

An example in which $D(\theta\,;c)\leqq\sqrt{n}\,s(\theta)$ but $D(\theta\,;c)\geqq\sqrt{nI(\theta)}$ will be given in the next section.

(ii) Note that Jensen's inequality implies that the function

$$c_0(\boldsymbol{x},\theta)=-\log f_n(\boldsymbol{x},\theta)$$

is contrastive. It is easy to see that under the conditions of Theorem 2.1 and Pitman's conditions, the discrimination rate of c_0 attains the upper bound.

3. Discrimination rates of L-unbiased estimators

The results of the preceding section make it possible to investigate the performance of L-unbiased estimators. Suppose that $\delta(\boldsymbol{X})$ is L-unbiased with respect to a given loss function $W(\theta,\cdot)$. Let $\{Q_\theta;\ \theta\in\Theta\}$ be the family of probability measures induced by $\delta(\boldsymbol{X})$ and let $g(\cdot,\theta)$, $\theta\in\Theta$ be the densities of Q_θ, $\theta\in\Theta$ with respect to a σ-field measure ν. We shall denote the sensitivity of $\{Q_\theta;\ \theta\in\Theta\}$ at θ by $s(\theta\,;\delta)$, that is,

$$\begin{aligned}s(\theta\,;\delta)&=2\varliminf_{\tau\to\theta}\rho(Q_\tau,Q_\theta)/|\tau-\theta|\\&=2\varliminf_{\tau\to\theta}\left[\int\{\sqrt{g(y,\tau)}-\sqrt{g(y,\theta)}\}^2d\nu\right]^{1/2}\Big/|\tau-\theta|\ .\end{aligned}$$

After the notation of $D(\theta\,;\delta,W)$, we shall write $d(\theta\,;\delta,W)$ instead of $d(\theta\,;c)$ when $c(\boldsymbol{x},\theta)=W(\theta,\delta(\boldsymbol{x}))$. Then we have the following theorem.

THEOREM 3.1. *If both* $\mathrm{E}_\theta W'^2(\tau,\delta(\boldsymbol{X}))$ *and* $\mathrm{E}_\theta W''(\tau,\delta(\boldsymbol{X}))$ *are continuous in* θ *and* τ, *then for all* $\theta\in\Theta$,

$$D(\theta\,;\delta,W)\leqq s(\theta\,;\delta)\leqq\sqrt{n}\,s(\theta)\ .$$

PROOF. A similar argument to the proof of Lemma 2.1 leads to

$$d(\theta\,;\delta,W)\leqq s(\theta\,;\delta)\ .$$

From the inequality

$$\rho(Q_\tau,Q_\theta)\leqq\rho(P_\tau^n,P_\theta^n)$$

(Pitman [6], p. 8), we have

$$d(\theta\,;\delta,W)\leqq s(\theta\,;\delta)\leqq\sqrt{n}\,s(\theta)\ .$$

The desired result follows from the proof of Theorem 2.1 and the assumptions of the theorem.

We investigate the attainability of the upper bound of the discrimination rate. Let $\{P_\theta;\ \theta \in \Theta\}$ be the exponential family with the natural parameter, i.e.,

$$f(x, \theta)=\exp\{\theta t(x)+\omega(\theta)+h(x)\} ,$$

and put

$$W_0(\theta, \delta(\boldsymbol{x}))=-\theta\delta(\boldsymbol{x})-\omega(\theta) .$$

Then the statistic

$$\delta_0(\boldsymbol{X})=\frac{1}{n}\sum_{i=1}^{n} t(X_i)$$

is L-unbiased with respect to W_0. Indeed, it follows from the equality

$$W_0(\theta, \delta_0(\boldsymbol{x}))=\frac{1}{n}\left\{-\log f_n(\boldsymbol{x}, \theta)+\sum_{i=1}^{n} h(x_i)\right\}$$

and Jensen's inequality. From the properties of the exponential family (Lehmann [4], p. 53), we can see that the conditions of Theorem 3.1 and Pitman's condition (P-3) are satisfied. Hence we have

$$D(\theta\,;\delta_0, W_0)=\sqrt{-n\omega''(\theta)}$$

and

$$\sqrt{n}\, s(\theta)=\sqrt{nI(\theta)}=\sqrt{-n\omega''(\theta)} .$$

Furthermore, notice that

$$\mathrm{E}_\theta\, \delta_0(\boldsymbol{X})=-\omega'(\theta) ,$$

which implies that $\delta_0(\boldsymbol{X})$ is L-unbiased with respect to

$$W_1(\theta, \delta)=\{\delta+\omega'(\theta)\}^2 .$$

Using the following Corollary 3.1, we find

$$D(\theta\,;\delta_0, W_1)=-\omega''(\theta)/\sqrt{\mathrm{Var}_\theta\, \delta_0(\boldsymbol{X})}=\sqrt{-n\omega''(\theta)} .$$

Thus both $D(\theta\,;\delta_0, W_0)$ and $D(\theta\,;\delta_0, W_1)$ attain the upper bound of the discrimination rate.

Since the larger discrimination rate $\delta(\boldsymbol{X})$ has, the more sensitive it is to the wrong values of $\gamma(\cdot)$ which come near to the correct one, we can use the discrimination rate as an index of the performance of $\delta(\boldsymbol{X})$.

DEFINITION 3.1. We shall define the *efficiency* of an L-unbiased estimator $\delta(\boldsymbol{X})$ with respect to W as

$$e(\theta;\delta,W)=D(\theta;\delta,W)/\sqrt{n}\,s(\theta)\ ,$$

the ratio of its discrimination rate to the maximum possible.

In the following, we study L-unbiased estimators with respect to the loss functions of special forms. To begin with, we treat the quadratic loss function

$$W_1(\theta,\delta)=\{\delta-\gamma(\theta)\}^2\ ,$$

where $\gamma(\theta)$ is differentiable. Then we have the following result.

COROLLARY 3.1. *Suppose that $\delta(\boldsymbol{X})$ is L-unbiased with respect to W_1, and that $\mathrm{E}_\theta\,\delta(\boldsymbol{X})$ takes one of the possible values of the function $\gamma(\cdot)$. If $\mathrm{E}_\theta\,\delta^2(\boldsymbol{X})$ is continuous in θ, then we have*

$$D(\theta;\delta,W_1)=|\gamma'(\theta)|/\sqrt{\mathrm{E}_\theta\,W_1(\theta,\delta(\boldsymbol{X}))}\leqq\sqrt{n}\,s(\theta)\ ,\qquad \textit{for all } \theta\in\Theta\ .$$

PROOF. From the assumptions of the corollary, we have

$$\mathrm{E}_\theta\,\delta(\boldsymbol{X})=\gamma(\theta)$$

(Lehmann [4], p. 12). Thus we find

$$\mathrm{E}_\theta\,\{W_1(\tau,\delta(\boldsymbol{X}))-W_1(\theta,\delta(\boldsymbol{X}))\}=\{\gamma(\tau)-\gamma(\theta)\}^2$$

and

$$\begin{aligned}&\mathrm{E}_\theta\,\{W_1(\tau,\delta(\boldsymbol{X}))-W_1(\theta,\delta(\boldsymbol{X}))\}^2\\&\quad=4\{\gamma(\tau)-\gamma(\theta)\}^2\,\mathrm{E}_\theta\,W_1(\theta,\delta(\boldsymbol{X}))+\{\gamma(\tau)-\gamma(\theta)\}^4\ .\end{aligned}$$

Using these equalities, we have

$$\lambda(\tau,\theta;\delta,W_1)=\frac{2|\gamma(\tau)-\gamma(\theta)|}{[2\,\mathrm{E}_\theta\,W_1(\theta,\delta(\boldsymbol{X}))+2\,\mathrm{E}_\tau\,W_1(\tau,\delta(\boldsymbol{X}))+\{\gamma(\tau)-\gamma(\theta)\}^2]^{1/2}}\ ,$$

hence

$$d(\theta;\delta,W_1)=\varliminf_{\tau\to\theta}\frac{2|\gamma'(\theta)|}{[2\,\mathrm{E}_\theta\,W_1(\theta,\delta(\boldsymbol{X}))+2\,\mathrm{E}_\tau\,W_1(\tau,\delta(\boldsymbol{X}))]^{1/2}}\ .$$

From the continuity of $\mathrm{E}_\theta\,\delta(\boldsymbol{X})$ and $\mathrm{E}_\theta\,\delta^2(\boldsymbol{X})$ in θ and a similar argument to the proof of Theorem 2.1, we have the desired result.

If the family $\{P_\theta;\ \theta\in\Theta\}$ under consideration satisfies Pitman's conditions, then the Cramér-Rao inequality follows from Corollary 3.1. But see the following example.

Example 3.1. Let $X_1, X_2,\cdots, X_n$ be independent random variables with density function

$$f(x, \theta)=I\{x\geqq\theta\}\exp\{-(x-\theta)\}\ , \qquad -\infty<\theta<\infty\ .$$

We estimate θ with the loss $W_1(\theta, \delta)=(\delta-\theta)^2$. We can easily see that

$$\delta(\boldsymbol{X})=X_{(1)}-1/n$$

is L-unbiased with respect to W_1. From Corollary 3.1, we have

$$D(\theta\,;\delta, W_1)=n\ .$$

On the other hand, we have (Pitman [6], p. 12)

$$s^2(\theta)=\infty \qquad \text{and} \qquad I(\theta)=1\ .$$

Thus

$$D(\theta\,;\delta, W_1)\leqq\sqrt{n}\,s(\theta) \qquad \text{but} \qquad D(\theta\,;\delta, W_1)\geqq\sqrt{nI(\theta)}\ ,$$

and

$$e(\theta\,;\delta, W_1)=0\ .$$

From Lemma 2.1, the following generalized form of the Cramér-Rao inequality, which is due to Ibragimov and Khas'minskii [2], is easily obtained.

COROLLARY 3.2. *Let $\delta(\boldsymbol{X})$ be an estimator of θ. We assume that the density functions of the family have the supports which are independent of θ, and that $\gamma_n(\theta)=\mathrm{E}_\theta\,\delta(\boldsymbol{X})$ is continuous in θ. Then we have*

$$\begin{aligned}&[\mathrm{E}_\theta\{\delta(\boldsymbol{X})-\theta\}^2+\underline{\lim}_{\tau\to\theta}\mathrm{E}_\tau\{\delta(\boldsymbol{X})-\tau\}^2]/2\\ &\quad\geqq[\underline{\lim}_{\tau\to\theta}\{\gamma_n(\tau)-\gamma_n(\theta)\}^2/(\tau-\theta)^2]\Big/\Big[n\,\overline{\lim}_{\tau\to\theta}\int_\theta^\tau I(\xi)d\xi/(\tau-\theta)\Big]+b_n^2(\theta)\ ,\end{aligned}$$

where $b_n(\theta)=\gamma_n(\theta)-\theta$.

PROOF. It is easy to see that $c(\boldsymbol{x}, \theta)=\{\delta(\boldsymbol{x})-\gamma_n(\theta)\}^2$ is a contrastive function. From Lemma 2.1 and (2.9),

$$\begin{aligned}4\{\gamma_n(\tau)-\gamma_n(\theta)\}^2\leqq{}&[2\,\mathrm{E}_\theta\{\delta(\boldsymbol{X})-\theta\}^2+2\,\mathrm{E}_\tau\{\delta(\boldsymbol{X})-\tau\}^2-2b_n^2(\theta)-2b_n^2(\tau)\\ &+\{\gamma_n(\tau)-\gamma_n(\theta)\}^2]\Big[n(\tau-\theta)\int_\theta^\tau I(\xi)d\xi\Big]\ .\end{aligned}$$

Hence we have

$$\begin{aligned}&4\,\underline{\lim}_{\tau\to\theta}\{\gamma_n(\tau)-\gamma_n(\theta)\}^2/(\tau-\theta)^2\\ &\quad\leqq[2\,\mathrm{E}_\theta\{\delta(\boldsymbol{X})-\theta\}^2+2\,\underline{\lim}_{\tau\to\theta}\mathrm{E}_\tau\{\delta(\boldsymbol{X})-\tau\}^2-4b_n^2(\theta)]\\ &\quad\quad\times\Big[n\,\overline{\lim}_{\tau\to\theta}\int_\theta^\tau I(\xi)d\xi/(\tau-\theta)\Big]\ ,\end{aligned}$$

which implies the corollary.

Next we treat the quadratic loss function of the form

$$W_2(\theta, \delta)=(\delta-\theta)^2/\theta^2\,, \qquad \theta\in\Theta=(0, \infty)\,.$$

COROLLARY 3.3. *Suppose that $\delta(X)$ is L-unbiased with respect to W_2, and that $E_\theta\,\delta^i(X)$, $i=1, 2, 3, 4$ are continuous in θ. Then*

$$\begin{aligned} D(\theta\,;\,\delta, W_2)&=[1-E_\theta\,W_2(\theta, \delta(X))]/\sqrt{E_\theta\,\delta^2(X)W_2(\theta, \delta(X))} \\ &\leqq\sqrt{n}\,s(\theta)\,, \qquad \textit{for all } \theta\in\Theta\,. \end{aligned}$$

PROOF. We can easily see that the conditions of Theorem 3.1 are satisfied. Thus we have $E_\theta\,W_2'(\theta, \delta(X))=0$, that is,

$$E_\theta\,\delta^2(X)/\theta^2=E_\theta\,\delta(X)/\theta\,.$$

Using this equality, we find

$$E_\theta\,W_2''(\theta, \delta(X))=2\,E_\theta\,\delta(X)/\theta^3=2[1-E_\theta\,W_2(\theta, \delta(X))]/\theta^2$$

and

$$E_\theta\,W'^2(\theta, \delta(X))=4\,E_\theta\,\delta^2(X)\{\delta(X)-\theta\}^2/\theta^6=4\,E_\theta\,\delta^2(X)W_2(\theta, \delta(X))/\theta^4\,.$$

Thus the corollary follows from (2.7) and Theorem 3.1.

Example 3.2. Let $X_1, X_2, \cdots, X_n$ be a sample from the normal distribution $N(\mu, \sigma^2)$ with known μ. Then each of the estimators of σ^2,

$$\delta_1(X)=\frac{1}{n+2}\sum_{i=1}^n(X_i-\mu)^2\,, \qquad \delta_2(X)=\frac{1}{n+1}\sum_{i=1}^n(X_i-\bar{X})^2$$

and

$$\delta_3(X)=\frac{n}{3}(\bar{X}-\mu)^2$$

is L-unbiased with respect to the loss function

$$W_2(\sigma^2, \delta)=(\delta-\sigma^2)^2/(\sigma^2)^2\,.$$

The sensitivity of the family is given by

$$s(\sigma^2)=(\sqrt{2}\,\sigma^2)^{-1}\,.$$

We obtain

$$D(\sigma^2\,;\,\delta_1, W_2)=\sqrt{n}\,s(\sigma^2)\sqrt{1-4/(n+6)}\,; \qquad E_{\sigma^2}\,W_2(\sigma^2, \delta_1(X))=2/(n+2)\,,$$

$$D(\sigma^2\,;\,\delta_2, W_2)=\sqrt{n}\,s(\sigma^2)\sqrt{1-(5n+1)/(n^2+5n)}\,; \quad E_{\sigma^2}\,W_2(\sigma^2, \delta_2(X))=2/(n+1)$$

and

$D(\sigma^2; \delta_3, W_2)=\sqrt{n}\, s(\sigma^2)\sqrt{3/(7n)}$; $\qquad$ $E_{\sigma^2} W_2(\sigma^2, \delta_3(X))=2/3$.

Thus when $n\geqq 2$,

$$E_{\sigma^2} W_2(\sigma^2, \delta_1(X))\leqq E_{\sigma^2} W_2(\sigma^2, \delta_2(X))\leqq E_{\sigma^2} W_2(\sigma^2, \delta_3(X))$$

and

$$D(\sigma^2; \delta_1, W_2)\geqq D(\sigma^2; \delta_2, W_2)\geqq D(\sigma^2; \delta_3, W_2) .$$

The efficiency of each estimator is tabulated below. The estimators $\delta_1(X)$ and $\delta_2(X)$ are asymptotically efficient in the sense of discrimination rate.

Table 3.1. The efficiencies of δ_1, δ_2, δ_3, δ_4 and δ_5

n	$e(\sigma^2; \delta_1, W_2)$	$e(\sigma^2; \delta_2, W_2)$	$e(\sigma^2; \delta_3, W_2)$	$e(\sigma^2; \delta_4, W_3)$	$e(\sigma^2; \delta_5, W_3)$
1	0.6546	0.0000	0.6546		
2	0.7071	0.4629	0.4629		
3	0.7453	0.5773	0.3779		
4	0.7745	0.6454	0.3273	0.0000	
5	0.7977	0.6928	0.2927	0.7071	0.0000
6	0.8164	0.7282	0.2672	0.8164	0.6454
7	0.8320	0.7559	0.2474	0.8660	0.7559
8	0.8451	0.7783	0.2314	0.8944	0.8100
9	0.8563	0.7968	0.2182	0.9128	0.8432
10	0.8660	0.8124	0.2070	0.9258	0.8660
⋮	⋮	⋮	⋮	⋮	⋮
∞	1.0000	1.0000	0.0000	1.0000	1.0000

Now we investigate the case where a loss function is not of a quadratic form. Consider the following loss function.

$$W_3(\theta, \delta)=\delta/\theta+\theta/\delta , \qquad \theta\in\Theta=(0, \infty) .$$

COROLLARY 3.4. *Suppose that $\delta(X)$ is L-unbiased with respect to W_3, and that $E_\theta\, \delta^i(X)$, $i=\pm 1, \pm 2$ are continuous in θ. Then*

$$D(\theta; \delta, W_3)=E_\theta W_3(\theta, \delta(X))/\theta\sqrt{E_\theta W_3^2(\theta, \delta(X))-4}$$
$$\leqq\sqrt{n}\, s(\theta) , \qquad \textit{for all } \theta\in\Theta .$$

PROOF. Since the proof is similar to that of Corollary 3.3, we omit it.

Example 3.2 (continued). Now the estimators of σ^2,

$$\delta_4(X)=\frac{1}{\sqrt{n(n-2)}}\sum_{i=1}^{n}(X_i-\mu)^2$$

and

$$\delta_5(X)=\frac{1}{\sqrt{(n-1)(n-3)}}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

are L-unbiased with respect to W_3. We obtain

$$D(\sigma^2;\,\delta_4,\,W_3)=\sqrt{n}\,s(\sigma^2)\sqrt{1-1/(n-3)}\,;$$
$$\mathrm{E}_{\sigma^2}\,W_3(\sigma^2,\,\delta_4(X))=2\sqrt{1+2/(n-2)}$$

and

$$D(\sigma^2;\,\delta_5,\,W_3)=\sqrt{n}\,s(\sigma^2)\sqrt{1-(2n-5)/(n^2-4n)}\,;$$
$$\mathrm{E}_{\sigma^2}\,W_3(\sigma^2,\,\delta_5(X))=2\sqrt{1+2/(n-3)}\,.$$

Thus when $n\geqq 5$,

$$\mathrm{E}_{\sigma^2}\,W_3(\sigma^2,\,\delta_4(X))\leqq \mathrm{E}_{\sigma^2}\,W_3(\sigma^2,\,\delta_5(X))$$

and

$$D(\sigma^2;\,\delta_4,\,W_3)\geqq D(\sigma^2;\,\delta_5,\,W_3)\,.$$

The efficiency of each estimator is tabulated in Table 3.1. Both $\delta_4(X)$ and $\delta_5(X)$ are asymptotically efficient in the sense of discrimination rate.

4. Sufficient statistics and discrimination rates

Suppose that a statistic $T=T(X)$ is sufficient for $\{P_\theta;\ \theta\in\Theta\}$. We shall write the conditional expectation $\mathrm{E}\{\cdot|T\}$ because T is sufficient.

Motivated by the Rao-Blackwell theorem, the following problem arises. Let $\delta(X)$ be L-unbiased with respect to $W(\cdot,\,\cdot)$. Then what form of loss function W does permit the construction of an L-unbiased estimator $\tilde{\delta}(T)$ which depends only on T and is better than $\delta(X)$ in the sense of discrimination rate, that is,

$$D(\theta;\,\delta,\,W)\leqq D(\theta;\,\tilde{\delta},\,W)\ ?$$

Klebanov [3] has treated the similar problem to ours in terms of the risk function. Following the *RB-condition* of [3], we define the *RBd-condition*.

DEFINITION 4.1. We shall say that a loss function $W(\cdot,\,\cdot)$ satisfies the *RBd-condition* if for any family of distributions $\{P_\theta;\ \theta\in\Theta\}$ having a sufficient statistic $\boldsymbol{T}$ for $\theta \in \Theta$ and any L-unbiased estimator $\delta(\boldsymbol{X})$ of the parametric function $\gamma(\theta)$, $\theta \in \Theta$, there exists an L-unbiased estimator $\tilde{\delta}$ of γ which depends only on $\boldsymbol{T}$ and satisfies the inequality

$$d(\theta;\delta, W) \leqq d(\theta;\tilde{\delta}, W) .$$

Remarks. (i) The RB-condition is defined by substituting the inequality $d(\theta;\delta, W) \leqq d(\theta;\tilde{\delta}, W)$ for the inequality $\mathrm{E}_\theta W(\theta, \delta(\boldsymbol{X})) \geqq \mathrm{E}_\theta W(\theta, \tilde{\delta}(\boldsymbol{T}))$.

(ii) Under some continuity conditions on δ and $\tilde{\delta}$, the inequality $d(\theta;\delta, W) \leqq d(\theta;\tilde{\delta}, W)$ is reduced to the inequality $D(\theta;\delta, W) \leqq D(\theta;\tilde{\delta}, W)$ (see the proof of Theorem 2.1).

We restrict our consideration to the case where the loss function is of the form

$$W(\theta, \delta) = \phi(\delta - \gamma(\theta)) , \tag{4.1}$$

where ϕ is a twice continuously differentiable non-negative strictly convex function such that $\phi(0)=0$, and γ is a continuously differentiable function. Then we have the following result, which is similar to that of Klebanov with respect to the RB-condition.

THEOREM 4.1. *In order that a loss function of the form* (4.1) *satisfies the RBd-condition, it is necessary and sufficient that either*

$$\phi(x) = A[\exp(\alpha x) - \alpha x - 1] \tag{4.2}$$

or

$$\phi(x) = a x^2 , \tag{4.3}$$

where A, a and α are constants.

PROOF. Since the proof of necessity is carried out similarly to that of Klebanov [3], we proceed to the proof of sufficiency. If $\gamma(\theta) \equiv c$ (constant) for all $\theta \in \Theta$, then every estimator of γ is L-unbiased and its discrimination rate is zero. So we have only to consider the case where $\gamma(\theta) \not\equiv c$ for all $\theta \in \Theta$. First we investigate the loss function given by (4.2), that is,

$$W(\theta, \delta) = A[\exp\{\alpha(\delta - \gamma(\theta))\} - \alpha(\delta - \gamma(\theta)) - 1] .$$

Put $w(\gamma(\theta), \delta) = W(\theta, \delta)$ and denote the derivative of $w(\gamma(\theta), \delta)$ with respect to γ by $w'(\gamma(\theta), \delta)$. Let $\delta(\boldsymbol{X})$ be L-unbiased. Since $\mathrm{E}_\theta w'^2(\zeta, \delta(\boldsymbol{X}))$ is continuous in ζ, it follows from Lemma 2.2 that

$$\mathrm{E}_\theta w'(\gamma(\theta), \delta(\boldsymbol{X})) = A\alpha[-\exp\{-\alpha\gamma(\theta)\}\, \mathrm{E}_\theta \exp\{\alpha\delta(\boldsymbol{X})\} + 1] = 0 ,$$

that is,

$$\mathrm{E}_\theta \exp\{\alpha\delta(\boldsymbol{X})\} = \exp\{\alpha\gamma(\theta)\}\ , \qquad \text{for all } \theta\in\Theta\,.$$

Using this equality, we can easily verify that

$$\mathrm{E}_\theta\{W(\tau, \delta(\boldsymbol{X}))-W(\theta, \delta(\boldsymbol{X}))\} = A\alpha^2\{\gamma(\tau)-\gamma(\theta)\}^2\{1+o(1)\}/2$$

and

$$\begin{aligned}&\mathrm{E}_\theta\{W(\tau, \delta(\boldsymbol{X}))-W(\theta, \delta(\boldsymbol{X}))\}^2\\ &\quad = A^2\alpha^2\{\gamma(\tau)-\gamma(\theta)\}^2[\exp\{-2\alpha\gamma(\theta)\}\,\mathrm{E}_\theta\exp\{2\alpha\delta(\boldsymbol{X})\}-1+o(1)]\,.\end{aligned}$$

From the definition of $d(\theta\,;\delta, W)$,

$$d(\theta\,;\delta, W)=\lim_{\tau\to\theta}\frac{|\alpha\gamma'(\theta)|\exp\{\alpha\gamma(\theta)\}}{[\mathrm{E}_\theta\exp\{2\alpha\delta(\boldsymbol{X})\}/2+\mathrm{E}_\tau\exp\{2\alpha\delta(\boldsymbol{X})\}/2-\exp\{2\alpha\gamma(\theta)\}]^{1/2}}\,.$$

Set

$$\tilde{\delta}(\boldsymbol{T})=\frac{1}{\alpha}\log \mathrm{E}\{\exp(\alpha\delta(\boldsymbol{X}))|\boldsymbol{T}\}\,.$$

From the equality

$$\mathrm{E}_\theta\exp\{\alpha\tilde{\delta}(\boldsymbol{T})\} = \mathrm{E}_\theta\,\mathrm{E}\{\exp(\alpha\delta(\boldsymbol{X}))|\boldsymbol{T}\} = \exp\{\alpha\gamma(\theta)\}$$

and the inequality

$$\exp(\alpha x)-\alpha x\geqq 1\ ,$$

we can easily prove that $\tilde{\delta}(\boldsymbol{T})$ is L-unbiased. An argument similar to the above yields

$$d(\theta\,;\tilde{\delta}, W)=\lim_{\tau\to\theta}\frac{|\alpha\gamma'(\theta)|\exp\{\alpha\gamma(\theta)\}}{[\mathrm{E}_\theta\exp\{2\alpha\tilde{\delta}(\boldsymbol{T})\}/2+\mathrm{E}_\tau\exp\{2\alpha\tilde{\delta}(\boldsymbol{T})\}/2-\exp\{2\alpha\gamma(\theta)\}]^{1/2}}\,.$$

From Jensen's inequality,

$$\begin{aligned}\mathrm{E}_\theta\exp\{2\alpha\tilde{\delta}(\boldsymbol{T})\} &= \mathrm{E}_\theta\{\mathrm{E}\{\exp(\alpha\delta(\boldsymbol{X}))|\boldsymbol{T}\}\}^2\\ &\leqq \mathrm{E}_\theta\,\mathrm{E}\{\{\exp(\alpha\delta(\boldsymbol{X}))\}^2|\boldsymbol{T}\} = \mathrm{E}_\theta\exp\{2\alpha\delta(\boldsymbol{X})\}\,.\end{aligned}$$

Hence

$$d(\theta\,;\delta, W)\leqq d(\theta\,;\tilde{\delta}, W)\,.$$

Thus the loss function given by (4.2) satisfies the RBd-condition. We now proceed to the proof for the loss function of the form (4.3), that is,

$$W(\theta, \delta)=a\{\delta-\gamma(\theta)\}^2\,.$$

Let $\delta(\boldsymbol{X})$ be L-unbiased. Using an argument similar to the previous case, we have

$$\mathrm{E}_\theta\, \delta(\boldsymbol{X})=\gamma(\theta) .$$

Put

$$\tilde{\delta}(\boldsymbol{T})=\mathrm{E}\,\{\delta(\boldsymbol{X})|\boldsymbol{T}\} .$$

Obviously it is L-unbiased. By an argument similar to the proof of Corollary 3.1, we can prove that

$$d(\theta\,;\, \delta,\, W)=\varliminf_{\tau\to\theta} 2|\gamma'(\theta)|/[2\,\mathrm{E}_\theta\, \delta^2(\boldsymbol{X})+2\,\mathrm{E}_\tau\, \delta^2(\boldsymbol{X})-4\gamma^2(\theta)]^{1/2}$$

and

$$d(\theta\,;\, \tilde{\delta},\, W)=\varliminf_{\tau\to\theta} 2|\gamma'(\theta)|/[2\,\mathrm{E}_\theta\, \tilde{\delta}^2(\boldsymbol{T})+2\,\mathrm{E}_\tau\, \tilde{\delta}^2(\boldsymbol{T})-4\gamma^2(\theta)]^{1/2} .$$

Therefore using Jensen's inequality, we have

$$d(\theta\,;\, \delta,\, W)\leqq d(\theta\,;\, \tilde{\delta},\, W) .$$

Thus we complete the proof.

Now we examine the property of the discrimination rate of a contrastive function when a sufficient statistic exists. Notice that if the function $c(\boldsymbol{X}, \theta)$ is contrastive, so is the function $\bar{c}(\boldsymbol{T}, \theta)=\mathrm{E}\,\{c(\boldsymbol{X}, \theta)|\boldsymbol{T}\}$. Then we prove the following theorem.

THEOREM 4.2. *Suppose that both* $\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \tau)$ *and* $\mathrm{E}_\theta\, c''(\boldsymbol{X}, \tau)$ *are continuous in* θ *and* τ. *If* $\mathrm{E}_\theta\{\mathrm{E}\,\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}\}^2$ *is continuous in* θ, *then for all* $\theta \in \Theta$,

$$D(\theta\,;\, c)\leqq D(\theta\,;\, \bar{c})\leqq \sqrt{n}\, s(\theta) .$$

Before proceeding to the proof, we establish the following lemma.

LEMMA 4.1. *Suppose that* $\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta)$ *is finite for all* $\theta \in \Theta$. *If*

$$\lim_{\tau\to\theta} \mathrm{E}_\theta\,\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/(\tau-\theta)^2=\mathrm{E}_\theta\, c'^2(\boldsymbol{X}, \theta) ,$$

then

$$\lim_{\tau\to\theta} \mathrm{E}_\theta\,\{\bar{c}(\boldsymbol{T}, \tau)-\bar{c}(\boldsymbol{T}, \theta)\}^2/(\tau-\theta)^2=\mathrm{E}_\theta\,\{\mathrm{E}\,\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}\}^2$$

and

$$\lim_{\tau\to\theta} \mathrm{E}_\theta\,\{\bar{c}(\boldsymbol{T}, \tau)-\bar{c}(\boldsymbol{T}, \theta)\}/(\tau-\theta)=\mathrm{E}_\theta\,\mathrm{E}\,\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}=0 .$$

PROOF. Using Jensen's inequality, we find

$$\begin{aligned}&\mathrm{E}_\theta|\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}/(\tau-\theta)-c'(\boldsymbol{X}, \theta)|^2\\&\quad\geqq \mathrm{E}_\theta|\{\bar{c}(\boldsymbol{T}, \tau)-\bar{c}(\boldsymbol{T}, \theta)\}/(\tau-\theta)-\mathrm{E}\,\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}|^2 .\end{aligned}$$

Thus from (2.5),

$$\lim_{\tau\to\theta} \mathrm{E}_\theta |\{\tilde{c}(\boldsymbol{T}, \tau)-\tilde{c}(\boldsymbol{T}, \theta)\}/(\tau-\theta)-\mathrm{E}\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}|^2=0 .$$

Following the proof of Lemma 2.2, we can easily see the lemma.

PROOF OF THEOREM 4.2. From Lemma 2.1 and Jensen's inequality, the following inequalities hold:

$$\begin{aligned}
&\mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\} \\
&\quad=\mathrm{E}_\theta\{\tilde{c}(\boldsymbol{T}, \tau)-\tilde{c}(\boldsymbol{T}, \theta)\}+\mathrm{E}_\tau\{\tilde{c}(\boldsymbol{T}, \theta)-\tilde{c}(\boldsymbol{T}, \tau)\} \\
&\quad\leqq[\mathrm{E}_\theta\{\tilde{c}(\boldsymbol{T}, \tau)-\tilde{c}(\boldsymbol{T}, \theta)\}^2/2+\mathrm{E}_\tau\{\tilde{c}(\boldsymbol{T}, \theta)-\tilde{c}(\boldsymbol{T}, \tau)\}^2/2]^{1/2}[4n\rho^2(P_\tau, P_\theta)]^{1/2} \\
&\quad\leqq[\mathrm{E}_\theta\{c(\boldsymbol{X}, \tau)-c(\boldsymbol{X}, \theta)\}^2/2+\mathrm{E}_\tau\{c(\boldsymbol{X}, \theta)-c(\boldsymbol{X}, \tau)\}^2/2]^{1/2}[4n\rho^2(P_\tau, P_\theta)]^{1/2} .
\end{aligned}$$

Thus from (2.2), we have

$$d(\theta; c)\leqq d(\theta; \tilde{c})\leqq\sqrt{n}\, s(\theta) .$$

Since it has already proved that $d(\theta; c)=D(\theta; c)$ (see the proof of Theorem 2.1), we have only to prove that $d(\theta; \tilde{c})=D(\theta; \tilde{c})$. This is easy to see. Indeed, from Lemma 4.1 and the proof of Theorem 2.1, we obtain at once that

$$d(\theta; \tilde{c})=\lim_{\tau\to\theta}\frac{\mathrm{E}_\theta c''(\boldsymbol{X}, \theta)/2+\mathrm{E}_\tau c''(\boldsymbol{X}, \tau)/2}{[\mathrm{E}_\theta\{\mathrm{E}\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}\}^2/2+\mathrm{E}_\tau\{\mathrm{E}\{c'(\boldsymbol{X}, \tau)|\boldsymbol{T}\}\}^2/2]^{1/2}}$$

and

$$D(\theta; \tilde{c})=\mathrm{E}_\theta c''(\boldsymbol{X}, \theta)/\sqrt{\mathrm{E}_\theta\{\mathrm{E}\{c'(\boldsymbol{X}, \theta)|\boldsymbol{T}\}\}^2} ,$$

and therefore from assumptions, we have

$$d(\theta; \tilde{c})=D(\theta; \tilde{c}) ,$$

as was to be proved.

Remark. Let $\delta(\boldsymbol{X})$ be L-unbiased with respect to the loss function $W(\theta, \delta)=\{\delta-\gamma(\theta)\}^2$. Applying Theorem 4.2 to the contrastive functions $c(\boldsymbol{X}, \theta)=\{\delta(\boldsymbol{X})-\gamma(\theta)\}^2$ and $\tilde{c}(\boldsymbol{T}, \theta)=\mathrm{E}\{\{\delta(\boldsymbol{X})-\gamma(\theta)\}^2|\boldsymbol{T}\}$, we can prove that

$$D(\theta; \delta, W)\leqq D(\theta; \tilde{\delta}, W) ,$$

where $\tilde{\delta}(\boldsymbol{T})=\mathrm{E}\{\delta(\boldsymbol{X})|\boldsymbol{T}\}$.

Acknowledgement

The author would like to thank Professor R. Shimizu for helpful suggestions and observations and a critical reading of the manuscript.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Huber, P. J. (1967). The behavior of maximum likelihood estimator under nonstandard conditions, *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, **1**, 221-233.

[2] Ibragimov, I. A. and Khas'minskii, R. Z. (1973). Information-theoretic inequalities and superefficient estimates, *Problems of Information Transmission*, **9**, 216-227.

[3] Klebanov, L. B. (1974). Unbiased estimators and sufficient statistics, *Theory Prob. Appl.*, **19**, 379-383.

[4] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, John Wiley & Sons, New York.

[5] Pfanzagl, J. (1969). On the measurability and consistency of minimum contrast estimates, *Metrika*, **14**, 249-272.

[6] Pitman, E. J. G. (1979). *Some Basic Theory for Statistical Inference*, Chapman and Hall, London.