

(1993e)

To appear in proceedings
of Penn State Conference on
Multivariate Analysis

An Information Inequality for the Bayes Risk under Truncated Squared Error Loss

Lawrence D. Brown*
Department of Mathematics
Cornell University
Ithaca, NY 14853

April 30, 1992

Abstract

A bound is given for the Bayes risk of an estimator under truncated squared error loss. The bound derives from an information inequality for the risk under this loss. It is then used to provide new proofs for some classical results of asymptotic theory.

Introduction

This paper develops a lower bound for the Bayes risk of an estimator under truncated squared error loss as defined in (1.1). The principle result is Theorem 2.1. This bound is then used in §3 to give new proofs for three classical theorems of asymptotic theory concerning the asymptotic Bayes property, local asymptotic minimaxity, and the set of superefficiency.

The development begins with an information inequality in §1 for the risk of an estimator under this loss. Another feature of the paper is Table 2.1 which provides a comparison of the bound with the actual Bayes risk in a case in which the latter can be conveniently computed.

This paper is a companion to Brown and Gajek (1990). (For convenience that paper will be referred to as BG.) Some of the relevant proofs appear in that paper; but otherwise the papers can be read independently. The results in the present paper were announced as part of my Wald Lecture in August,

*Research supported in part by NSF DMS 8506847 and NSF DMS 8809016. AMS-MOS Subject Classification: Primary 62F10, Secondary 62F15, 62C99, 60E15. Key Words and Phrases: Information inequality, Cramer-Rao inequality, truncated squared error efficiency, local asymptotic minimaxity, superefficiency.

1985 except that the conclusion of (what is now) Theorem 2.1 then read only $B_K(g) \geq C - D$ instead of the better statement, $B_K(g) \geq C^2(C + D)^{-1}$.

1 A Bound for the Risk

This section describes a lower bound for the risk of an estimator under truncated squared error loss. This bound has a peculiar and uncongenial nature. It is hard to conceive that standing alone it could be of any practical use. However it can be turned into a useful lower bound for the Bayes risk under this loss. This is accomplished in §2.

Only the bound for the univariate case will be described. Generalizations to the multivariate case undoubtedly exist but appear to be notationally awkward to state and prove.

Let X be an observable random variable with probability density $p_\theta, \theta \in \Theta$, relative to some Borel measure ν . Assume $\Theta \subset \mathbb{R}$ is an open interval. It is desired to estimate θ , and so the action space is $\bar{\Theta}$. The loss function is the truncated version of ordinary quadratic loss, defined by

$$L_K(\theta, a) = \min((a - \theta)^2, K^2) \quad (1.1)$$

for given $0 < K < \infty$. Denote the risk function for L_K by $R_K(\theta, \delta)$ and the corresponding expected risk and Bayes risk functions by $B_K(g, \delta)$ and $B_K(g) = \inf_{\delta} B_K(g, \delta)$.

This loss function is not convex; hence the non-randomized estimators are no longer necessarily a complete class. (See Brown, Cohen, and Strawderman (1976) for conditions under which they will be.) For notational simplicity only non-randomized estimators will be explicitly considered below. However, all the following results are valid also for randomized estimators.

Assume that $\{p_\theta\}$ satisfies the following two regularity conditions for every $\theta_0 \in \Theta$:

$$p_{\theta_0}(x) = 0 \Rightarrow p_\theta(x) = 0 \text{ a.e. } (\nu) \quad (1.2)$$

for all θ in a neighborhood of θ_0 , and

$$\begin{aligned} p_\theta/p_{\theta_0} \text{ is weakly differentiable in} \\ L_2(p_{\theta_0} d\nu) \text{ at } \theta = \theta_0 \text{ with weak derivative } q. \end{aligned} \quad (1.3)$$

(Generally, $q = \frac{\partial}{\partial \theta} \ln p_\theta(x)|_{\theta=\theta_0}$.) According to Fabian and Hannan (1977) these conditions imply the ordinary information inequality: If $\text{Var}_{\theta_0}(T) < \infty$ then

$$\text{Var}_{\theta_0}(T) \geq \frac{(e'(\theta_0))^2}{I(\theta_0)} \quad (1.4)$$

where $I(\theta_0) = E_{\theta_0}(q^2(X))$ denotes the ordinary Fisher information and $e(\theta) = E_{\theta}(T)$.

The Appendix of BG contains several easier to verify conditions (labelled (A.7), (A.6), (A.5)) each of which implies (1.3).

Given an estimator δ let

$$\delta_{\theta,K}(x) = \begin{cases} \theta + K & \text{if } \theta + K < \delta(x) \\ \delta(x) & \text{if } \theta - K \leq \delta(x) \leq \theta + K \\ \theta - K & \text{if } \delta(x) < \theta - K. \end{cases} \quad (1.5)$$

Note that $\delta_{\theta,K}$ is not an estimator, since it depends on θ , but it is a measurable function. Then define

$$e_K(\theta) = E_{\theta}(\delta_{\theta,K}(X)). \quad (1.6)$$

Under the regularity conditions (1.2) and (1.3) on $\{p_{\theta}\}$ for every $\theta_0 \in \Theta$ the function $e_K(\theta)$ will be absolutely continuous on θ . In fact, $e'_K(\theta)$ will exist at all except at most a countable set of points in θ . (The points are those for which $p_{\theta}(|\delta(X) - \theta| = K) > 0$.) At those points both left and right hand derivatives $e'^{(\ell)}$ and $e'^{(r)}$ will exist, but will not be equal. To obtain an unambiguous statement in (1.9) and similar expressions which follow make the convention that $e'_K(\theta)$ is the value between $e'^{(\ell)}(\theta)$ and $e'^{(r)}(\theta)$ for which

$$|e'_K(\theta)| = \max(|e'^{(\ell)}(\theta)|, |e'^{(r)}(\theta)|). \quad (1.7)$$

(If $|e'^{(\ell)}(\theta)| = |e'^{(r)}(\theta)|$ either value can be chosen as $e'_K(\theta)$.)

Let $I(\theta)$ denote the Fisher information, assumed to be finite. For any $\alpha, 0 \leq \alpha < 1$, let

$$I_{K,\alpha}(\theta) = (I^{1/2}(\theta) + K^{-1}(1 - \alpha)^{-1/2})^2. \quad (1.8)$$

Note that $I_{K,\alpha}(\theta) > I(\theta)$. Here is the main result.

Theorem 1.1 *Make assumptions (1.2) and (1.3) on $\{p_{\theta}\}$. Then for any $\alpha, 0 \leq \alpha < 1$,*

$$R_K(\theta, \delta) \geq I_{K,\alpha}^{-1}(\theta)(e'_K(\theta))^2 + \alpha b_K^2(\theta) \quad (1.9)$$

where $b_K(\theta) = e_K(\theta) - \theta$.

Proof Under assumptions (1.2), (1.3) the right hand derivative $e'^{(r)}$ exists at θ_0 and is given by

$$e'^{(r)}(\theta_0) = \int \delta_{\theta_0,K}(x)q(x)p_{\theta_0}(x)\nu(dx) + p_{\theta_0}(\delta(X) \leq \theta_0 - K \text{ or } \delta(X) > \theta_0 + K).$$

There is a symmetric expression for $e_K^{(\ell)}(\theta_0)$. Hence $\max(|e_K^{(r)}(\theta_0)|, |e_K^{(\ell)}(\theta_0)|) = |e_K'(\theta_0)|$ (by 1.7) satisfies

$$\begin{aligned} |e_K'(\theta_0)| &\leq \left| \int \delta_{\theta_0, K}(x) q(x) p_{\theta_0}(x) \nu(dx) \right. \\ &\quad \left. + p_{\theta_0}(|\delta(X) - \theta_0| \geq K) \right| \\ &= \left| \int (\delta_{\theta_0, K}(x) - e_K(\theta_0)) q(x) p_{\theta_0}(x) \nu(dx) \right. \\ &\quad \left. + p_{\theta_0}(|\delta(X) - \theta_0| \geq K) \right|. \end{aligned} \quad (1.10)$$

Apply Cauchy-Schwarz to the first term on the right of (1.10) and apply Chebyshev's inequality to the second term to find

$$|e_K'(\theta_0)| \leq I^{1/2}(\theta_0) \text{Var}_{\theta_0}^{1/2}(\delta_{\theta_0, K}(X)) + (1/K) E_{\theta_0}(|\delta_{\theta_0, K}(X) - \theta_0|).$$

Observe that $R_K(\theta_0, \delta) = \text{Var}_{\theta_0}(\delta_{\theta_0, K}(X)) + b_K^2(\theta_0)$ and also that $R_K(\theta_0, \delta) \geq E_{\theta_0}^2(|\delta_{\theta_0, K}(X) - \theta_0|)$ to get

$$|e_K'(\theta_0)| \leq (R_K(\theta_0, \delta) - b_K^2(\theta_0))^{1/2} I^{1/2}(\theta_0) + K^{-1} R_K^{1/2}(\theta_0, \delta). \quad (1.11)$$

Now, for any non-negative numbers $0 < \alpha < 1, 0 < b^2 < r, z \geq 0$

$$\frac{(r - b^2)^{1/2} z + K^{-1} r^{1/2}}{(r - \alpha b^2)^{1/2} z + \frac{(r - \alpha b^2)^{1/2}}{K(1-\alpha)^{1/2}}} \leq \max \left[\frac{(r - b^2)^{1/2}}{(r - \alpha b^2)^{1/2}}, \frac{r^{1/2}(1-\alpha)^{1/2}}{(r - \alpha b^2)^{1/2}} \right] < 1.$$

Hence

$$|e_K'(\theta_0)| \leq (R_K(\theta_0, \delta) - \alpha b_K^2(\theta_0))^{1/2} I^{1/2}(\theta_0) + K^{-1} (1 - \alpha)^{-1/2}. \quad (1.12)$$

Squaring both sides of (1.12) and rearranging terms yields (1.9). ■

Inequality (1.9) should be compared to the ordinary information inequality, which asserts

$$R(\theta, \delta) \geq I^{-1}(\theta) (e'(\theta))^2 + b^2(\theta). \quad (1.13)$$

Indeed, when $R(\theta, \delta) < \infty$ then $e_K'(\theta) \rightarrow e'(\theta)$ as $K \rightarrow \infty$ and (1.13) follows from (1.9) upon letting $K \rightarrow \infty$ then $\alpha \rightarrow 1$.

For the choice $\alpha = 0$ one obtains

$$R_K(\theta, \delta) \geq \frac{(e_K'(\theta_0))^2}{(I^{1/2} + K^{-1})^2}. \quad (1.14)$$

A variant of this inequality was obtained many years ago by H. Chernoff (private communication). Inequalities which like (1.14) do not involve a squared bias term will not suffice to yield Bayes risk bounds as in the next section.

2 Bayes Risk Lower Bound

Let g be an absolutely continuous prior density. Let $V_{K,\alpha}(\theta) = I_{K,\alpha}^{-1}(\theta)$: and assume that $V_{K,\alpha}$ is absolutely continuous on θ . Continue to assume (1.2) and (1.3). Let $sp(g) = \{\theta : g(\theta) > 0\}$ and let $csp(g)$ denote its closure.

Theorem 2.1 *If $csp(g)$ is a compact subset of θ and $0 < \alpha < 1$ then*

$$B_K(g) \geq \frac{C^2}{C+D} \geq C - D \quad (2.1)$$

where

$$C = \int V_{K,\alpha}(\theta)g(\theta)d\theta$$

$$D = \alpha^{-1} \int \left[\frac{d}{d\theta}(V_{K,\alpha}(\theta)g(\theta)) \right]^2 g^{-1}(\theta)d\theta.$$

Proof The proof begins with (1.9) in place of (1.13) and then runs exactly parallel to the proof of Corollary 2.3 in BG. ■

The condition that $csp(g)$ be a compact subset of θ is very much stronger than necessary. Here is a formal statement which relaxes that condition. To save space, this statement refers the reader to BG.

Corollary 2.1 *Assume $D < \infty$. Assume g satisfies (2.13a) or (2.13b) of BG with $V_{K,\alpha}(\theta)$ in place of $V(\theta)$, and with $g = h$. Then (2.1) is still valid.*

Proof See the proof of Corollary 2.6 of BG. ■

It is possible to also carry Theorems 2.7 and 2.9 of BG over to this situation and thereby obtain somewhat improved bounds.

Example 2.1: This example gives some idea of the precision of the bound in Theorem 2.1 since $B_K(g)$ can also be exactly computed. It will be seen that the bound only becomes reasonably precise as $K \rightarrow \infty$. This is basically to be expected because of the use of Chebyshev's inequality, which introduces appreciable imprecision in the proof of Theorem 2.1 except when K is large compared to the variance of X . We nevertheless found the rate of convergence surprisingly slow. This example employs conjugate priors and so should be quite favorable to the bound. In particular, the bound is exact when $K = \infty$. (In this regard see BG, Example 3.1.)

Consider the normal-mean location problem: so $X \sim N(\mu, 1)$. Let g be the normal density with mean 0, variance σ^2 . Then Theorem 2.1 and Corollary 2.1 yield

$$B_K(g) \geq [(1 + K^{-1}(1 - \alpha)^{-1/2})^2(1 + (\alpha\sigma^2)^{-1})]^{-1}. \quad (2.2)$$

The value of α which maximizes the right side of (2.2) can easily be determined numerically. Table 2.1 gives values of the resulting maximum, labelled BOUND. Finally, for comparison the table contains values of $B_K(g)$ computed numerically from the fact that

$$\begin{aligned}
 B_K(g) &= E(E((\theta - \gamma x)^2 \wedge K^2 | X = x)) \\
 &= \gamma - \frac{2\gamma}{\sqrt{2\pi}} \int_{K/\sqrt{\gamma}}^{\infty} (t^2 - K^2/\gamma) e^{-t^2/2} dt \quad (2.3)
 \end{aligned}$$

where $\gamma = \sigma^2/(1 + \sigma^2)$, since $(\theta | X = x) \sim N(\gamma x, \gamma)$. (Similar computations appear in Efron and Morris (1971).)

		σ^2					
		1	5	10	25	100	∞
$K = 1$	BOUND	.057	.127	.156	.187	.216	.250
	$B_K(g)$.371	.479	.497	.508	.514	.516
$K = 3$	BOUND	.161	.330	.391	.451	.506	.563
	$B_K(g)$.500	.832	.906	.958	.985	.995
$K = 10$	BOUND	.298	.560	.641	.713	.773	.826
	$B_K(g)$.500	.833	.909	.962	.990	1.000
$K = 50$	BOUND	.419	.570	.816	.881	.928	.961
	$B_K(g)$.500	.833	.909	.962	.990	1.000
$K = 100$	BOUND	.447	.770	.851	.913	.954	.980
	$B_K(g)$.500	.833	.909	.962	.990	1.000
$K = \infty$	BOUND	.500	.833	.909	.962	.990	1.000
	$B_K(g)$.500	.833	.909	.962	.990	1.000

Table 2.1: Values of the bound and of $B_K(g)$ in Example 2.1.

Some typical values of α maximizing the right side of (2.2) are .53, .77, .93 for $(K, \sigma^2) = (1, 1), (10, 1), (100, 1)$ resp.; .31, .47, .82 for $(1, 10), (10, 10), (100, 10)$ resp.; and .13, .26, .55 for $(1, 100), (10, 100), (100, 100)$. As $K \rightarrow \infty (\sigma^2 \rightarrow \infty, \text{ resp.})$ the best $\alpha \rightarrow 1 (0, \text{ resp.})$.

3 Asymptotic Bayes and Minimax Properties

The bound in §2 provides a very convenient means of proving certain well known classical asymptotic results. Again we treat here only the one dimensional case; multivariate extensions exist for all the following results. Versions of the following results were first formulated and proved by LeCam (1953). Lehmann (1983), Strasser (1985), or LeCam (1986) are good contemporary references.

Throughout this section let X_1, \dots, X_n be independent identically distributed variables satisfying the conditions of Theorem 2.1. Let $V(\theta) = I^{-1}(\theta)$ denote the variance bound for a single observation, X_1 .

Let g be any prior density having compact support, or, more generally, satisfying the conditions of Corollary 2.1. Then under ordinary squared error loss,

$$nB^{(n)}(g) \geq \int V(\theta)g(\theta)d\theta - n^{-1} \int \left[\frac{(V(\theta)g(\theta))'}{g(\theta)} \right]^2 g(\theta)d\theta \quad (3.1)$$

by BG, (Corollary 2.1) or Borovkov and Sakhanienko (1980). Here, and in similar expressions to follow, $B^{(n)}$ denotes the Bayes risk based on the sample of size n . Furthermore, under truncated squared error loss, (1.1) one has by Theorem 2.1

$$nB_{K/\sqrt{n}}^{(n)}(g) \geq \int V_{K,\alpha}(\theta)g(\theta)d\theta - (n\alpha)^{-1} \int \left[\frac{(V_{K,\alpha}(\theta)g(\theta))'}{g(\theta)} \right]^2 g(\theta)d\theta. \quad (3.2)$$

For what follows it is important only that for fixed $0 < \alpha < 1$

$$V_{K,\alpha}(\theta) \rightarrow V(\theta) \quad (3.3)$$

as $K \rightarrow \infty$, with a similar statement concerning the derivative of $V_{k,\alpha}$. To simplify the following proofs assume also that $V(\theta)$ is continuously differentiable on θ .

Here are some consequences of (3.1) and (3.2).

Consequence 3.1: Assume

$$\int \left[\frac{(V(\theta)g(\theta))'}{g(\theta)} \right]^2 g(\theta)d\theta < \infty.$$

Then the "limiting Bayes risk", $\liminf_{n \rightarrow \infty} nB(g)$, satisfies

$$\liminf_{n \rightarrow \infty} nB^{(n)}(g) \geq \int V(\theta)g(\theta)d\theta. \quad (3.4)$$

The "asymptotic Bayes risk", $\lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} nB_{K/\sqrt{n}}^{(n)}(g)$ also satisfies

$$\lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} nB_{K/\sqrt{n}}^{(n)}(g) \geq \int V(\theta)g(\theta)d\theta. \quad (3.4')$$

Proof (3.4) is immediate from (3.1) and (3.4') is immediate from (3.2) and (3.3). ■

Under suitable conditions on $\{p_\theta\}$ the maximum likelihood estimate, $\hat{\theta}$, has asymptotic risk $V(\theta)$. See, e.g., Lehmann (1983). Under slightly more stringent conditions it also has limiting risk $V(\theta)$. Thus, under suitable conditions the maximum likelihood estimator is asymptotically Bayes - i.e., satisfies

$$\lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} \int nR_{K/\sqrt{n}}^{(n)}(\theta, \hat{\theta})g(\theta)d\theta = \lim_{K \rightarrow \infty} \liminf_{n \rightarrow \infty} nR_{K/\sqrt{n}}^{(n)}(g).$$

Under slightly more stringent conditions it is also limiting Bayes - i.e., satisfies

$$\lim_{n \rightarrow \infty} \int nR^{(n)}(\theta, \hat{\theta})g(\theta)d\theta = \liminf_{n \rightarrow \infty} nB^{(n)}(g).$$

Hodges demonstrated the existence of a superefficient sequence of estimators. LeCam (1953) then proved (under suitable conditions) that the set of superefficiency has Lebesgue measure zero. For our purposes a sequence of estimators, $\{\delta_n\}$, is superefficient at $\theta \in \Theta$ if

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} nR_{K/\sqrt{n}}^{(n)}(\theta, \delta_n) < V(\theta). \quad (3.5)$$

Consequence 3.2: Let $\{\delta_n\}$ be a given sequence of estimators. Under the above assumptions the set of parameter points at which $\{\delta_n\}$ is superefficient has Lebesgue measure zero.

Proof Let $S \subset \theta$ be the set of superefficiency and suppose the Lebesgue measure of S , $\lambda(S)$, is positive. Let $0 < \alpha < 1$. Then, since $V_{K,\alpha}(\theta) \rightarrow V(\theta)$ as $K \rightarrow \infty$, there must exist $N < \infty, 1 < K < \infty, 0 < \epsilon < 1$, and an $S' \subset S$ with $\lambda(S') > 0$ such that

$$nR_{K/\sqrt{n}}^{(n)}(\theta) < V_{K,\alpha}(\theta) - \epsilon, \theta \in S', \quad (3.6)$$

for all $n \geq N$. Since $\lambda(S') > 0$ there must exist a (possibly small) open interval $I = (a, b)$ such that

$$\frac{\lambda(I - S')}{\lambda(I)} < \frac{\epsilon}{9K^2}. \quad (3.7)$$

Let

$$g(\theta) = (12/(b-a)^3) \{[\min((\theta-a), (b-\theta))]^+\}^2. \quad (3.8)$$

Note that $\int \left[\frac{(V_{K,\alpha}(\theta)g(\theta))'}{g(\theta)} \right]^2 g(\theta) d\theta < \infty$. Hence (3.2) yields the existence of an N such that for $n > N$,

$$nB_{K/\sqrt{n}}^{(n)}(g) > \int V_{K,\alpha}(\theta)g(\theta)d\theta - \epsilon/3. \quad (3.9)$$

On the other hand $g(\theta) = 0$ for $\theta \notin I$ and $0 \leq g(\theta) < 3/(b-a)$. Note also that $nB_{K/\sqrt{n}}^{(n)}(g) < K^2$. Hence

$$\begin{aligned} & \int nR_{K/\sqrt{n}}^{(n)}(\theta, \delta_n)g(\theta)d\theta \\ & \leq \int_{S'} (V_{K,\alpha}(\theta) - \epsilon)g(\theta)d\theta + K^2 \int_{I-S'} g(\theta)d\theta \\ & \leq \int V_{K,\alpha}(\theta)g(\theta)d\theta - \epsilon \int_{S'} g(\theta)d\theta + K^2 \int_{I-S'} g(\theta)d\theta \\ & < \int V_{K,\alpha}(\theta)g(\theta)d\theta - \epsilon/3 \end{aligned} \quad (3.10)$$

since $\int_{S'} g(\theta)d\theta \geq \left[\frac{\lambda(S')}{b-a} \right]^3 \geq (1 - \epsilon/9K^2)^3 \geq (1 - 1/9)^3 > 2/3$, and $g(\theta) \leq 3/(b-a)$ so that $\int_{I-S'} g(\theta)d\theta \leq (3/(b-a))\lambda(I - S') < \epsilon/3K^2$.

But, (3.9) and (3.10) together claim that $\int nR_{K/\sqrt{n}}^{(n)}(\theta, \delta_n)g(\theta)d\theta < nB_{K/\sqrt{n}}^{(n)}(g)$, a contradiction. It follows that $\lambda(S) = 0$. ■

A third classic property about which (3.1) and (3.2) yield good bounds is local asymptotic minimaxity. For the purpose at hand define the local asymptotic minimax value, m_0 , at θ_0 , as

$$m_0 = \lim_{K \rightarrow \infty} \lim_{D \rightarrow \infty} \liminf_{n \rightarrow \infty} \inf_{\{\delta_n\}} \sup_{|\theta - \theta_0| < D/\sqrt{n}} R_{K/\sqrt{n}}^{(n)}(\theta, \delta_n). \quad (3.11)$$

Consequence 3.3: $m_0 = V(\theta_0)$.

Proof Let $\theta_0 = 0$, with no loss of generality, and

$$g_{n,D}(\theta) = 12(\sqrt{n}/2D)^3 [(D/\sqrt{n} - |\theta|)^+]^2, \quad (3.12)$$

similarly to (3.8). Then

$$\begin{aligned} \inf_{\{\delta_n\}} \sup_{|\theta| < D/\sqrt{n}} R_{K/\sqrt{n}}^{(n)}(\theta, \delta_n) &\geq B_{K/\sqrt{n}}^{(n)}(g_n, D) \\ &\geq \int V_{K,\alpha}(\theta) g_{n,D}(\theta) d\theta - (n\alpha)^{-1} \int \left[\frac{(V_{K,\alpha}(\theta) g_{n,D}(\theta))'}{g(\theta)} \right]^2 g(\theta) d\theta \end{aligned}$$

by (3.2). Now, for some $C < \infty$, $V_{K,\alpha}(\theta) < C$ and $V'_{K,\alpha}(\theta) < C$ for $|\theta| < D/\sqrt{n}$, as a consequence of the assumptions on $I(\theta)$ and the definition of $V_{K,\alpha}$. Hence

$$\begin{aligned} (n\alpha)^{-1} \int \left[\frac{(V_{K,\alpha}(\theta) g_{n,D}(\theta))'}{g(\theta)} \right]^2 g(\theta) d\theta \\ \leq (n\alpha)^{-1} \left[C^2 + C^2 \int \frac{(g'_{n,D}(\theta))^2}{g_{n,D}(\theta)} d\theta \right] \\ = (n\alpha)^{-1} C^2 [1 + 12n/D^2]. \end{aligned}$$

Consequently,

$$\begin{aligned} \lim_{D \rightarrow \infty} \liminf_{n \rightarrow \infty} B_{K/\sqrt{n}}^{(n)}(g_n, D) &\geq \liminf_{n \rightarrow \infty} \int V_{K,\alpha}(\theta) g_{n,D}(\theta) d\theta \\ &= V_{K,\alpha}(0). \end{aligned}$$

It follows that $m_0 = V(0)$ since $\lim_{K \rightarrow \infty} V_{K,\alpha}(\theta) = V(\theta)$. ■

Under suitable conditions on $\{p_\theta\}$ the maximum likelihood estimator, $\hat{\theta}$, satisfies

$$V(\theta_0) = \lim_{K \rightarrow \infty} \lim_{D \rightarrow \infty} \lim_{n \rightarrow \infty} \sup_{|\theta - \theta_0| < D/\sqrt{n}} R_{K/\sqrt{n}}^{(n)}(\theta, \hat{\theta})$$

and hence is locally asymptotically minimax. See Hajek (1972) and, again, Lehmann (1983) for a good general account of the theory plus many further references.

References

- Brown, L.D., Cohen, A., and Strawderman, W.E. (1976). A complete class theorem for monotone likelihood ratio with applications. *Ann. Statist.* 4, 712-722.
- Brown, L.D. and Gajek L. (1990). Information inequalities for the Bayes risk (referred here to as BG). *Ann. Statist.* 18, 1578-1594.
- Borovkov, A.A. and Sakhanienko, A.U. (1980). On estimates of the expected quadratic risk. *Prob. and Math. Statist.* 1, 185-195 (in Russian).

- Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and empirical Bayes Estimators - Part I: The Bayes case. *J. Amer. Statist. Assoc.* 66, 807-815.
- Fabian, V. and Hannan, J. (1977). On the Cramer Rao inequality. *Ann. Statist.* 5, 197-205.
- Hajek, J. (1972). Local asymptotic minimaxity and admissibility in estimation. *Proc. Sixth Berk. Symp. Math. Statist. and Prob.* 1, 175-194.
- LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. of Calif. Publ. in Statist.* 1, 277-330.
- LeCam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer.
- Lehmann, E.L. (1983). *Theory of Point Estimation*. Wiley and Sons.
- Strasser, H. (1985). *Mathematical Theory of Statistics-Statistical Experiments and Asymptotic Decision Theory*. W. de Gruyter.

