# A LOWER BOUND TO THE ASYMPTOTIC RISK OF AN ESTIMATOR FOR A FAMILY OF NON–REGULAR CASES

by

Tapas Samanta
Indian Statistical Institute and Purdue University

Technical Report # 91-32

Department of Statistics
Purdue University

June 1991

★

# A LOWER BOUND TO THE ASYMPTOTIC RISK OF AN ESTIMATOR FOR A FAMILY OF NON–REGULAR CASES

by

Tapas Samanta
Indian Statistical Institute and Purdue University

## ABSTRACT

Let $X_1, X_2, \ldots, X_n$ be i.i.d. observations with a common density $f(x, \theta)$ depending on a real parameter $\theta$. Our problem is to estimate $\theta$. In this paper, we consider a family of non–regular cases for which the support of the density is an interval depending on $\theta$ and obtain a lower bound for the asymptotic risk of an estimator using the notion of limits of experiments.

## 1. Introduction

Let $\{f_n(\cdot, \theta)\}$, $n \geq 1$, be a family of densities depending on a parameter $\theta$ taking values in $\Theta$, where $\Theta$ is an open subset of the real line $\mathbb{R}$. We consider the problem of estimation of $\theta$. Let $\{T_n\}$ be a sequence of estimators of $\theta$. We consider the asymptotic risk

$$\rho(\theta, \{T_n\}) = \lim_{A \to \infty} \liminf_{n \to \infty} \sup_{|\theta' - \theta| \leq A k_n^{-1}} L(k_n(T_n - \theta')) \tag{1.1}$$

as a measure of the asymptotic performance of the estimator $\{T_n\}$ at $\theta$, where $L$ is an appropriate loss function and $k_n(\uparrow \infty)$ is the normalizing factor (see, for example, Weiss and Wolfowitz (1974)) for the given family of distributions. The performance of $\{T_n\}$ at $\theta$ is evaluated not by the limiting risk at $\theta$ but by the limit of supremum of the risk at $\theta'$, over the set $|\theta' - \theta| \leq A k_n^{-1}$ where for any $A$, $A k_n^{-1}$ tends to zero. This is done to rule out superefficiency (see, Ghosh (1985)). For the regular cases a lower bound to the asymptotic risk (1.1) was obtained in Hajek (1972). The nonregular cases were studied by many authors including Weiss and Wolfowitz (1974), and Ibragimov and Hasminskii (1981). In Samanta (1989), a lower bound to the asymptotic risk was obtained for a family of non–regular cases using the notion of limiting experiment which is due to LeCam (1972). The examples given in Samanta (1989) includes the case where the observations are i.i.d. with density whose support is an interval $S(\theta) = [A_1(\theta),\ A_2(\theta)]$ which is monotone in $\theta$. In this paper we consider the case where the support $S(\theta)$ is not monotone in $\theta$. We first find in Section 2 the limit of the sequence of experiments $\{P^n_{\theta+\lambda/n}; \lambda \in \mathbb{R}\}$ where $P^n_{\theta'}$ is the distribution of the first $n$ observations when $\theta = \theta'$. In Section 3, we consider a wide class of loss functions and find the minimax risk in the limiting experiment which gives us a lower bound to the local asymptotic (minimax) risk by Hajek–LeCam asymptotic minimax theorem (see Section 3).

## 2. The Limiting Experiment

Let $X_1, X_2, \ldots, X_n$ be i.i.d. observations with a common distribution $P_\theta$ having density $f(x, \theta)$ with respect to Lebesgue measure where $\theta \in \Theta$, an open subset of $\mathbb{R}$. We assume that the support of the density is an interval $[A_1(\theta),\ A_2(\theta)]$ depending on $\theta$. We consider the sequence of experiments

$$E_n = \{P^n_{\theta_0+\lambda/n}; \lambda \in \Lambda\},\ n \geq 1$$

where $\Lambda$ is the whole real line $\mathbb{R}$ or some appropriate subset of $\mathbb{R}$, $P^n_\theta$ is the $n$–fold product of $P_\theta$ and $\theta_0$ is a fixed point in $\Theta$ which may be regarded as the true value of the parameter. We want to study the convergence of this sequence of experiments in the sense defined as follows (see Millar (1983)).

**Definition.** Let $E^n = \{(S^n, \mathcal{S}^n), Q^n_\lambda; \lambda \in \Lambda\},\ n \geq 1$, and $E = \{(S, \mathcal{S}), Q_\lambda;\ \lambda \in \Lambda\}$ be experiments with parameter set $\Lambda$. Then $E^n$ converges to $E$ if for every finite subset $\{\lambda_1, \ldots, \lambda_k\}$ of $\Lambda$,

$$\mathcal{L}\left\{\left(\frac{dQ^n_{\lambda_1}}{d\mu^n}, \ldots, \frac{dQ^n_{\lambda_k}}{d\mu^n}\right)\Bigg|\mu^n\right\} \Rightarrow \mathcal{L}\left\{\left(\frac{dQ_{\lambda_1}}{d\mu}, \ldots, \frac{dQ_{\lambda_k}}{d\mu}\right)\Bigg|\mu\right\}$$

where $\mu^n = \sum_{i=1}^{k} Q^n_{\lambda_i},\ \mu = \sum_{i=1}^{k} Q_{\lambda_i}$.

If one of the $A_i$'s is increasing (in $\theta$) and the other decreasing (one of them may be $\pm\infty$) so that the interval $[A_1(\theta),\ A_2(\theta)]$ is monotone in $\theta$, then the limit of the sequence of experiments $E_n$ with $\Lambda = \mathbb{R}^+$ or $\mathbb{R}^-$ is an exponential shift experiment under suitable assumptions stated in Samanta (1989). This follows from the general result proved in Samanta (1989, Sec. 2). In this paper we consider the case where either both the $A_i$'s are increasing or both decreasing.

To find the limit of $E_n$ we show that the Hellinger transform of the experiment $E_n$ converges to that of a certain experiment $E$ which is the limiting experiment.

**Definition.** The Hellinger transform of a statistical experiment $E = \{(\mathcal{X}, \mathcal{A}),\ P_\theta;\ \theta \in \Theta\}$ is defined on the set $S = \{z = (z_i)_{i \in I} \in [0,1]^I \colon \sum_{i \in I} z_i = 1,\ I$ is finite subset of $\Theta\}$ and for $I = \{\theta_1, \theta_2, \ldots, \theta_k\} \subset \Theta$ and $z = (z_{\theta_1}, \ldots, z_{\theta_k}) \in S$ it is given by

$$H_E(z) = \int \prod_{i=1}^{k} \left(\frac{dP_{\theta_i}}{d\mu}\right)^{z_{\theta_i}} d\mu$$

where $\mu$ is a $\sigma$–finite measure dominating $\{P_{\theta_i};\ i \in I\}$.

$H_E(z)$ is independent of the choice of $\mu$.

We shall use the following result to find the limiting experiment.

**Lemma.** A sequence of experiments $E_n$ converges to an experiment $E$ if and only if $H_{E_n}$ converges to $H_E$, pointwise on $S$.

We first find the limit of the Hellinger transform of $E_n = \{P^n_{\theta_0+\lambda/n};\ \lambda \in \mathbb{R}\}$.

We make the following assumptions.

(A1) $f(x,\theta)$ is strictly positive and jointly continuous in $(x,\theta)$ on the set $\{A_1(\theta) \leq x \leq A_2(\theta)\}$. We set $h(\theta) = f(A_1(\theta),\theta)$ and $g(\theta) = f(A_2(\theta),\theta)$.

(A2) $A_1(\theta)$ and $A_2(\theta)$ are continuously differentiable function of $\theta$.

(A3) The derivative $f'(x,\theta)$ of $f(x,\theta)$ with respect to $\theta$ exists and is continuous in $\theta$ on the set $\{A_1(\theta) < x < A_2(\theta)\}$.

We shall prove the following:

**Proposition 1.** For $\lambda_1 < \lambda_2 < \ldots < \lambda_k$, the Hellinger transform $H_n(z_1, z_2, \ldots, z_k)$ of $E_n = \{P^n_{\theta_0+\lambda/n};\ \lambda \in \mathbb{R}\}$ restricted to $\{\lambda_1, \ldots, \lambda_k\}$ converges to

$$H(z_1, z_2, \ldots, z_k) = \exp\left\{ c_2(\theta_0)\lambda_1 - c_1(\theta_0)\lambda_k - \sum_{i=1}^{k} \lambda_i z_i (c_2(\theta_0) - c_1(\theta_0)) \right\}$$

with $c_1(\theta_0) = A_1'(\theta_0)h(\theta_0)$ and $c_2(\theta_0) = A_2'(\theta_0)g(\theta_0)$.

**Corollary.** The sequence of experiments $E_n$ converges to the experiment $E = \{Q_\lambda;\ \lambda \in \mathbb{R}\}$ where $Q_\lambda$ is a probability distribution on $\mathbb{R}^2$ with density

$$q_\lambda(x,y) = \begin{cases} \exp\{(c_2(\theta_0) - c_1(\theta_0))\lambda\} \exp\{x - y\}, & \text{if } x < c_2(\theta_0)\lambda \\ & \text{and } y > c_1(\theta_0)\lambda, \\ 0, & \text{otherwise.} \end{cases}$$

The result stated in the corollary follows from the above lemma since the limit $H(z_1, z_2, \ldots, z_k)$ given in Proposition 1 is the Hellinger transform of the experiment $\{Q_\lambda;\ \lambda \in \Lambda\}$ restricted to $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$.

Now to prove the proposition it is enough to show that as $n \to \infty$

$$\begin{aligned} & n\left[\int \prod_{i=1}^{k} f^{z_i}(x, \theta_0 + \lambda_i/n)dx - 1\right] \\ \to & c_2(\theta_0)\lambda_1 - c_1(\theta_0)\lambda_k - \sum_{i=1}^{k} \lambda_i z_i (c_2(\theta_0) - c_1(\theta_0)) \end{aligned} \tag{2.1}$$

for any $(z_1, \dots, z_k)$ satisfying $0 \leq z_i \leq 1$ for all $i$ and $\sum_{i=1}^{k} z_i = 1$. We prove only for the case where both $A_1(\theta)$ and $A_2(\theta)$ are increasing in $\theta$ and $\lambda_i$'s are all $\geq 0$. The proof for the other cases are similar. Now, $n\left[\int \prod_{i=1}^{n} f^{z_i}(x, \theta_0 + \lambda/n)dx - 1\right]$

$$\begin{aligned} &= \int n\left[1_{\{A_1(\theta_0+\lambda_k/n) < x < A_2(\theta_0+\lambda_1/n)\}} \prod_{i=1}^{k} f^{z_i}(x, \theta_0 + \lambda_i/n) - f(x, \theta_0)\right] dx \\ &= -n \int_{A_1(\theta_0)}^{A_1(\theta_0+\lambda_k/n)} f(x, \theta_0)dx + n \int_{A_2(\theta_0)}^{A_2(\theta_0+\lambda_1/n)} \prod_{i=1}^{k} f^{z_i}(x, \theta_0 + \lambda_i/n)dx \\ &+ n \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} \left[\prod_{i=1}^{k} f^{z_i}(x, \theta_0 + \lambda_i/n) - f(x, \theta_0)\right] dx \\ &= I_{1n} + I_{2n} + I_{3n}, \text{ say.} \end{aligned} \tag{2.2}$$

By assumptions (A1) and (A2)

$$I_{1n} \to -c_1(\theta_0)\lambda_k \text{ and } I_{2n} \to c_2(\theta_0)\lambda_1.$$

It now remains to show that

$$I_{3n} \to \sum_{i=1}^{k} \lambda_i z_i (c_1(\theta_0) - c_2(\theta_0))$$

$$I_{3n} = \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} \left[\prod_{i=1}^{k} \left\{f(x, \theta_0) + \frac{\lambda_i}{n} f'(x, \theta_{in})\right\}^{z_i} - f(x, \theta_0)\right] dx$$

[where $\theta_{in}$ lies between $\theta_0$ and $\theta_0 + \lambda_i/n$]

$$= n \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} f(x, \theta_0) \left[\prod_{i=1}^{k} \left\{1 + \frac{\lambda_i}{n} \frac{f'(x, \theta_{in})}{f(x, \theta_0)}\right\}^{z_i} - 1\right] dx.$$

The next two steps are to show

$$\begin{aligned} & n \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} f(x, \theta_0) \left[\prod_{i=1}^{k} \left\{1 + \frac{\lambda_i}{n} \frac{f'(x, \theta_{in})}{f(x, \theta_0)}\right\}^{z_i} - \prod_{i=1}^{k} \left\{1 + \frac{\lambda_i}{n} \frac{f'(x, \theta_{in})}{f(x, \theta_0)} z_i\right\}\right] dx \\ & \to 0 \text{ as } n \to \infty \end{aligned} \tag{2.3}$$

and

$$
\begin{aligned}
&n\int_{A_1(\theta_0+\lambda_{k/n})}^{A_2(\theta_0)} f(x,\theta_0)\left[\prod_{i=1}^{k}\left\{1+\frac{\lambda_i}{n}\frac{f'(x,\theta_{in})}{f(x,\theta_0)}z_i\right\}-1\right]dx \\
&\simeq n\int_{A_1(\theta_0+\lambda_{k/n})}^{A_2(\theta_0)} f(x,\theta_0)\left[\sum_{i=1}^{k}\frac{\lambda_i}{n}\frac{f'(x,\theta_{in})}{f(x,\theta_0)}z_i\right]dx.
\end{aligned}
\tag{2.4}
$$

By "$a_n \simeq b_n$" we mean $a_n - b_n \to 0$. (2.3) and (2.4) are valid if we assume, for example, the following

(A4) There exists a neighbourhood $N(\theta_0)$ of $\theta_0$ such that for all $\theta \in N(\theta_0)$,

$$\left|\frac{f'(x,\theta)}{f(x,\theta_0)}\right| \leq M(\theta_0)$$

for some constant $M(\theta_0)$ and for all $x \in [A_1(\theta_0),\ A_2(\theta_0)]$ for which $f'(x,\theta)$ exists.

To prove (2.3) we use the expansion

$$(1+y)^z = 1 + zy + \frac{z(z-1)}{2}y^2(1+\xi)^{z-2}$$

where $\xi$ lies between 0 and $y$.

From (2.3) and (2.4) we have by (A3) and (A4)

$$
\begin{aligned}
I_{3n} &\simeq \sum_{i=1}^{k}\lambda_i z_i \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} f'(x,\theta_{in})dx \\
&\simeq \sum_{i=1}^{k}\lambda_i z_i \int_{A_1(\theta_0+\lambda_k/n)}^{A_2(\theta_0)} f'(x,\theta_0)dx \\
&\simeq \sum_{i=1}^{k}\lambda_i z_i \int_{A_1(\theta_0)}^{A_2(\theta_0)} f'(x,\theta_0)dx.
\end{aligned}
\tag{2.5}
$$

Now

$$
\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\left[\int_{A_1(\theta)}^{A_2(\theta)} f(x,\theta)dx\right] \\
&= f(A_2(\theta),\theta)A_2'(\theta) - f(A_1(\theta),\theta)A_1'(\theta) + \int_{A_1(\theta)}^{A_2(\theta)} f'(x,\theta)dx.
\end{aligned}
$$

Therefore, from (2.5) we have

$$I_{3n} \to \sum_{i=1}^{k}\lambda_i z_i(c_1(\theta_0) - c_2(\theta_0))$$

and thus the proposition is proved under assumptions (A1)–(A4).

## 3. A Lower Bound to the Asymptotic Risk

We now find the minimax risk of the limiting experiment obtained in Section 2 to find a lower bound to the asymptotic risk $\rho(\theta, \{T_n\})$ of an estimator $\{T_n\}$. We use the following general result which is known as the Hajek–LeCam asymptotic minimax theorem. Suppose we have experiments $E^n = \{(S^n, \mathcal{S}^n),\ Q_\lambda^n;\ \lambda \in \Lambda\}$, $n \geq 1$, and $E = \{(S, \mathcal{S}),\ Q_\lambda;\ \lambda \in \Lambda\}$. Let $D$ be a fixed decision space and $L$ a loss function on $\Lambda \times D$ which is lower semicontinuous on $D$ for each fixed $\lambda \in \Lambda$. Let $\rho_n(b, \lambda)$ and $\rho(b, \lambda)$ be the risk functions of a procedure $b$ in the decision theoretic structures $(E^n, D, L)$ and $(E, D, L)$ respectively. Then we have

**Theorem** (Hajek–LeCam asymptotic minimax theorem). If $E^n$ converges to $E$, then

$$\liminf_{n\to\infty} \inf_b \sup_\lambda \rho_n(b, \lambda) \geq \inf_b \sup_\lambda \rho(b, \lambda)$$

where the infimum in either side is over all "generalized" procedures (see Millar (1983, Ch. II)) for the corresponding experiment.

We consider a loss function $L(\cdot)$ satisfying the following

(i) $L(0) = 0,\ L(x) \geq 0$ for all $x$

(ii) $L(x) = L(-x)$

(iii) $\{x\colon L(x) \leq c\}$ is closed and convex for all $c > 0$.

Using the above theorem and proceeding as in the proof of Theorem VII 2.6 of Millar (1983), we can now prove that

$$\begin{aligned} &\lim_{A\to\infty} \liminf_{n\to\infty} \inf_{T_n} \sup_{|\theta-\theta_0|\leq An^{-1}} E_\theta L[n(T_n - \theta)] \\ &\geq \inf_\delta \sup_{\lambda\in\mathbf{R}} \rho(\delta, \lambda) \end{aligned} \tag{3.1}$$

where the infimum in the left hand side is over all estimators $T_n$ of $\theta$ and the infimum in the right hand side is over all randomized procedures for the limiting experiment $E$ obtained in Section 2.

We now compute the minimax risk in the limiting experiment $E$. Let $(X, Y)$ have joint density $q_\lambda(x, y)$ given in Section 2. Let us consider the transformation

$$U = \frac{c_1 X - c_2 Y}{c_1 + c_2}$$
$$V = \frac{X + Y}{c_1 + c_2}.$$

(We write $c_i$ in place of $c_i(\theta_0)$). Then $U$ and $V$ have joint density of the form $g(u, v - \lambda)$ where

$$g(u,v) = \begin{cases} (c_1 + c_2)\exp\{2u + (c_2 - c_1)v\}, & \text{if } u < 0 \text{ and } \frac{u}{c_1} < v < -\frac{u}{c_2}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

The problem of estimation of $\lambda$ is invariant under the group of transformation $(u, v) \to (u, v + c)$, $c \in \mathbb{R}$. Any equivariant estimator is of the form $V + \phi(U)$. By the result of Kiefer (1957), the best equivariant estimator is minimax and therefore the minimax risk in experiment $E$ is

$$\inf_\phi \int\int L(v + \phi(u))g(u,v)dudv$$
$$= \int\int L(v + \phi_0(u))g(u,v)dudv, \text{ say.}$$

We assume that a minimizing $\phi_0$ exists. If, for example, $L$ is convex and the integral is finite for some $\phi$, then such a $\phi_0$ exists. If we have squared error loss, $\phi_0(u)$ is the conditional expectation of $V$ given $u$. If $c_1(\theta) \equiv c_2(\theta)$, $\phi_0(u) \equiv 0$ for any loss function $L$ satisfying (i)–(iii) above.

Now, from (3.1) we have

$$\lim_{A\to\infty} \liminf_{n\to\infty} \inf_{T_n} \sup_{|\theta-\theta_0|\le An^{-1}} E_\theta L[n(T_n - \theta)]$$
$$\ge EL(V + \phi_0(U))$$

where $U$ and $V$ have joint density $g(u, v)$ as given in (3.2) and $\phi_0$ is as defined above. In particular, for any sequence of estimators $\{T_n\}$, we have

$$\rho(\theta_0, \{T_n\}) \ge EL(V + \phi_0(U)).$$

## References


Ghosh, J.K. (1985). Efficiency of Estimates–Part I. *Sankhyā Ser. A*, **Pt. 3**, 310–325.

Ibragimov, I.A. and Hasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

Hajek, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proc. Sixth Berkeley Symp.*, **Vol. 1**, 175–194 (L.M. LeCam, J. Neyman and E.J. Scott, eds.). University of California Press, Berkeley and Los Angeles.

Kiefer, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Statist.* **28**, 573–601.

LeCam, L.M. (1972). Limits of experiments. *Proc. Sixth Berkeley Symp.* **Vol. 1**, 245–261 (L.M. LeCam, J. Neyman and E.L. Scott, eds.). University of California Press, Berkeley and Los Angeles.

Millar, P.W. (1983). The minimax principle in asymptotic statistical theory. Ecole d'Ete de Probabilites de Saint–Flour XI (P.L. Hennequin, ed.), *Lecture Notes in Mathematics.* Springer–Verlag, Berlin.

Samanta, T. (1989). Local asymptotic minimax estimation in non–regular cases. *Sankhyā, Ser. A*, **Vol. 51**.

Weiss, L. and Wolfowitz, J. (1974). Maximum probability estimators and related topics. *Lecture Notes in Mathematics*, Springer–Verlag, Berlin.