Cramer-Rao Type Integral Inequalities for General Loss Functions

B. L. S. Prakasa Rao
Indian Statistical Institute, New Delhi
and
Australian National University, Canberra

July 1, 1999

Abstract

Cramer-Rao type integral inequalities where developed for loss functions $w(x)$ which are bounded below by functions of the type $g(x) = c|x|^l$, $l > 1$. As applications, we obtain lower bounds of Hajek-LeCam type for locally asymptotic minimax error for such loss functions.

Keywords and Phrases: Cramer-Rao type integral inequality; Bayes risk; Lower bound; Locally asymptotic minimax error; Hajek-LeCam lower bound.

1 Introduction

Cramer-Rao type integral inequalities for the Bayes risk of an estimator based on squared error loss function were derived earlier (cf. Prakasa Rao (1991, 1992, 1995)) following Borovkov and Sakhanenko (1980). Brown and Gajek (1990) and Brown and Low (1991) studied information bounds on the Bayes risk and minimax risk for squared error loss functions with applications to nonparametric regression. Ghosh(1993) discussed Cramer-Rao type integral inequalities for positive variances. Gill and Levit (1995) obtained bayesian Cramer-Rao bounds for some nonparametric inference problems as applications of Van Trees (1968) inequality. Babrovsky et al. (1987) extended and generalised Van Trees inequality to the multidimensional case. Prakasa Rao (1999a) studied Cramer-Rao type integral inequalities for parameters taking values in a Banach space with application to inference for stochastic processes. However all these results are based on a squared error loss function.

Our aim in this paper is to obtain lower bounds for Bayes risk for loss functions $w(x)$ which are bounded below by a function of the type $g(x) = c|x|^l$ for some $l \geq 2$ and a positive constant c. Our techniques generalize the methods used in Borovkov and Sakhanenko (1980) and the methods in the recent work on a generalization of Cramer-Rao inequality by Ibragimov (1999). We restrict the discussion to loss functions of the type $g(x) = |x|^l$. Results for any general loss function $w(x)$ bounded below by a function of the type $g(x) = c|x|^l$ are an immediate consequence of the results for a loss function of the type $g(x) = |x|^l$.

2 Lower Bound for Bayes Risk

Suppose $\mathbf{X}$ is a random element taking values in a measurable space $(\mathcal{X}, \mathcal{A})$ with probability measure P_θ, $\theta \in \Theta \subset \mathbb{R}$. Suppose that the family of probability measures of $\{P_\theta, \theta \in \Theta\}$ are dominated by a σ-finite measure ν on $\mathcal{A}$ with density $\frac{\partial P_\theta}{d\nu} = f(x, \theta)$. Let $\theta^\star = \theta^\star(\mathbf{X})$ be an estimator of θ based on $\mathbf{X}$. Let $q(\cdot)$ be a probability density function on $\mathbb{R}$. Let

$S_q = \{t : q(t) > 0\}$. Suppose that $S_q \subset \Theta^0$. Define, for $l > 1$,

$$\begin{aligned} R(\theta^\star) &= E|\theta^\star - \theta|^l \\ &= \iint_{\mathcal{X}\times\Theta} |\theta^\star(\mathbf{x}) - \theta|^l f(\mathbf{x},\theta) q(\theta) v(d\mathbf{x})\, d\theta. \end{aligned} \tag{2.1}$$

In the following, we denote the expectation with respect to the joint distribution of $(\mathbf{X}, \theta)$ by E and the expectation with respect to the conditional distribution of $\mathbf{X}$ given θ by E_θ. The conditional expectation given $\mathbf{X}$ will be denoted by $E(\cdot|\mathbf{X})$.

The problem of interest is to obtain a lower bound for $R(\theta^\star)$ independent of the estimator $\theta^\star$.

Theorem 1 *Assume that there exists a function $K(t, \mathbf{x})$ such that*

$$\int_{S_q} K(t, \mathbf{x})\, dt = 0 \tag{2.2}$$

for every $\mathbf{x} \in \mathcal{X}$. Define

$$G(\theta, \mathbf{x}) = \frac{K(\theta, \mathbf{x})}{f(\mathbf{x}, \theta) q(\theta)}, \qquad \theta \in S_q. \tag{2.3}$$

Suppose that $\{\mathbf{x} : f(\mathbf{x}, \theta) > 0\}$ does not depend on $\theta \in \Theta$ and

$$E|\theta G(\theta, \mathbf{X})| < \infty, \qquad 0 < E|G(\theta, \mathbf{X})|^\gamma < \infty \tag{2.4}$$

and

$$E|\theta^\star G(\theta, \mathbf{X})| < \infty \tag{2.5}$$

where $\theta^\star$ is an estimator of θ and $\gamma = \frac{l}{l-1}$. Then

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{|E(\theta G(\theta, \mathbf{X}))|^l}{\left\{E|G(\theta, \mathbf{X})|^\gamma\right\}^{l-1}}, \qquad \gamma = \frac{l}{l-1}. \tag{2.6}$$

Proof. Under the conditions stated above, it is easy to check that

$$E\big[\theta^\star G(\theta, \mathbf{X})\big] = 0$$

by an application of Fubini's Theorem. Hence

$$E\big[(\theta - \theta^\star) G(\theta, \mathbf{X})\big] = E\big[\theta G(\theta, \mathbf{X})\big] \tag{2.7}$$

which implies that

$$\begin{aligned} \big|E[\theta G(\theta, \mathbf{X})]\big| &= \Big|E\big[(\theta - \theta^\star) G(\theta, \mathbf{X})\big]\Big| \\ &= \left\{E|\theta - \theta^\star|^l\right\}^{1/l} \left\{E|G(\theta, \mathbf{X})|^\gamma\right\}^{1/\gamma} \end{aligned} \tag{2.8}$$

by Holder's inequality where $\frac{1}{l} + \frac{1}{\gamma} = 1$ or $\gamma = \frac{l}{l-1}$. Therefore

$$R(\theta^\star) = E|\theta - \theta^\star|^l \geq \frac{|E(\theta G(\theta, \mathbf{X}))|^l}{\left\{E|G(\theta, \mathbf{X})|^\gamma\right\}^{l-1}}, \qquad \gamma = \frac{l}{l-1}, \tag{2.9}$$

which proves that

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{|E(\theta G(\theta, \mathbf{X}))|^l}{\left\{E|G(\theta, \mathbf{X})|^\gamma\right\}^{l-1}}, \qquad \gamma = \frac{l}{l-1}. \tag{2.10}$$

■

Remark: In general for any function $g(\theta)$ and an estimator $g^\star$ of $g(\theta)$

$$E|g^\star - g(\theta)|^l \geq \frac{\Big(E\big[g(\theta)G(\theta,\mathbf{X})\big]\Big)^l}{\big\{E|G(\theta,\mathbf{X})|^\gamma\big\}^{l-1}}. \tag{2.11}$$

Special Cases

Let us not consider some special cases of the above theorem.

1. Let

$$K(t,bx) = \frac{d}{dt}\big\{f(\mathbf{x},t)h(t)\big\}, \tag{2.12}$$

where $h(\cdot)$ is an arbitrary function. Suppose $K(t,\mathbf{x})$ is well defined and

$$|f(\mathbf{x},t)h(t)| \to 0 \quad \text{as} \quad |t| \to \partial S_q, \qquad t \in \Theta$$

where ∂S_q denotes the boundary of S_q. Then

$$\int_{S_q} K(t,\mathbf{x})\,dt = 0. \tag{2.13}$$

Further more

$$G(t,\mathbf{x}) = \frac{h'(t)}{q(t)} + \frac{\partial \ \log f}{\partial t}\frac{h(t)}{q(t)}, \qquad t \in S_q.$$

Suppose that sufficient regularity conditions hold so that

$$E_\theta\Big[\frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\Big] = 0. \tag{2.14}$$

Then

$$E\big[\theta G(\theta,\mathbf{X})\big] = -E\Big[\frac{h(\theta)}{q(\theta)}\Big], \tag{2.15}$$

$$E|G(\theta,\mathbf{X})|^\gamma = E\Big|\frac{h'(\theta)}{q(\theta)} + \frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\frac{h(\theta)}{q(\theta)}\Big|^\gamma \tag{2.16}$$

and hence

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{|E\big[\frac{h(\theta)}{q(\theta)}\big]|^l}{\Big\{E\big|\frac{h'(\theta)}{q(\theta)} + \frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\frac{h(\theta)}{q(\theta)}\big|^\gamma\Big\}^{l-1}}, \qquad \gamma = \frac{l}{l-1}. \tag{2.17}$$

2. Let us consider further a special case of (**??**) with

$$h(\theta) \equiv \frac{q(\theta)}{I(\theta)} \tag{2.18}$$

where $I(\theta)$ is the Fisher information assuming that it exists and is positive. Then

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{|E\big[\frac{1}{I(\theta)}\big]|^l}{\Big\{E\big|\frac{(q(\theta)/I(\theta))'}{q(\theta)} + \frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\frac{1}{I(\theta)}\big|^\gamma\Big\}^{l-1}}, \qquad \gamma = \frac{l}{l-1}. \tag{2.19}$$

3. Let us consider another special case of (**??**) with

$$h(\theta) = q(\theta).$$

Then, it follows that,

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{1}{\left\{E\left|\frac{q'(\theta)}{q(\theta)} + \frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}}. \tag{2.20}$$

In general, for any differentiable function $g(\theta)$,

$$\inf_{g^\star} E|g^\star - g(\theta)|^l \geq \frac{\left(E[g'(\theta)]\right)^l}{\left\{E\left|\frac{q'(\theta)}{q(\theta)} + \frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}}. \tag{2.21}$$

For $l = 2$, the inequality reduces to

$$\inf_{g^\star} E|g^\star - g(\theta)|^2 \geq \frac{\left(E[g'(\theta)]\right)^2}{E\left[\frac{q'(\theta)}{q(\theta)}\right]^2 + E\left[\frac{\partial \ \log f(\mathbf{X},\theta)}{\partial\theta}\right]^2}. \tag{2.22}$$

4. Suppose that the components of $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ are independent and identically distributed with density $p(x,\theta)$. Then it follows from (**??**) that

$$R_n = \inf_{\theta^\star} R(\theta^\star) \geq \frac{1}{\left\{E\left|\frac{\partial \ \log q(\theta)}{\partial\theta} + \sum_{i=1}^n \frac{\partial \ \log p(X_i,\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}}. \tag{2.23}$$

3 Lower Bound for Minimax Risk

Let θ_0 be the true parameter and $\mathcal{D}_n = \{\theta : |\theta - \theta_0| \leq \epsilon_n\}$ be an ϵ_n-neighbourhood of θ_0. For any estimator $\theta^\star$ of θ_0

$$\begin{aligned}\sup_{\theta\in\mathcal{D}_n} E_\theta|\theta - \theta^\star|^l &\geq \int_{\mathcal{D}_n} E_\theta|\theta - \theta^\star|^l q(\theta)\, d\theta \\ &= R(\theta^\star) \\ &\geq \frac{1}{\left\{E\left|\frac{\partial \ \log q(\theta)}{\partial\theta} + \sum_{i=1}^n \frac{\partial \ \log p(X_i,\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}} \end{aligned} \tag{3.1}$$

from (**??**) and hence

$$\sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(\theta^\star - \theta)|^l \geq \frac{1}{\left\{E\left|\frac{1}{\sqrt{n}}\frac{\partial \ \log q(\theta)}{\partial\theta} + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial \ \log p(X_i,\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}} \tag{3.2}$$

Therefore

$$\begin{aligned}&\varliminf_{n\to\infty} \inf_{\theta^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(\theta^\star - \theta)|^l \\ &\qquad \geq \frac{1}{\varlimsup_{n\to\infty}\left\{E\left|\frac{1}{\sqrt{n}}\frac{\partial \ \log q(\theta)}{\partial\theta} + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial \ \log p(x_i,\theta)}{\partial\theta}\right|^\gamma\right\}^{l-1}} \end{aligned} \tag{3.3}$$

for any prior density $q(\cdot)$ on $\mathcal{D}_n$.

(i) Suppose $q(\cdot)$ is a prior density on $[\theta_0 - \epsilon_n, \theta_0 + \epsilon_n]$ such that

$$\frac{1}{n}E\left|\frac{\partial \ \log q(\theta)}{\partial\theta}\right|^\gamma \to 0 \quad \text{as} \quad n \to \infty.$$

We now consider a special case when γ is an even integer (say) $2k$, $k \geq 1$. Then $l = 2k/(2k-1) \leq 2$. In such a case (cf. Marcinkiewicz and Zygmund (1937) or Loéve (1977), p. 276), it follows that

$$E\Big|\frac{\partial \log q(\theta)}{\partial\theta} + \sum_{i=1}^{n} \frac{\partial \log p(X_i,\theta)}{\partial\theta}\Big|^{\gamma} \leq A_{\gamma}(n+1)^{(\gamma/2-1)}\Big\{\sum_{i=1}^{n} E\Big|\frac{\partial \log p(X_i,\theta)}{\partial\theta}\Big|^{\gamma} + E\Big|\frac{\partial \log q(\theta)}{\partial\theta}\Big|^{\gamma}\Big\} \tag{3.4}$$

where A_γ is a constant depending only on γ. Since X_i, $1 \leq i \leq n$ are i.i.d. random variables, it follows that

$$E\Big|\frac{\partial \log q(\theta)}{\partial\theta} + \sum_{i=1}^{n} \frac{\partial \log p(X_i,\theta)}{\partial\theta}\Big|^{\gamma} \leq A_{\gamma}(n+1)^{(\gamma/2-1)}\Big\{nE\Big|\frac{\partial \log p(X_1,\theta)}{\partial\theta}\Big|^{\gamma} + E\Big|\frac{\partial \log q(\theta)}{\partial\theta}\Big|^{\gamma}\Big\}. \tag{3.5}$$

Let

$$I_\gamma(\theta) = E_\theta\Big|\frac{\partial \log p(X_1,\theta)}{\partial\theta}\Big|^{\gamma}. \tag{3.6}$$

Suppose $I_\gamma(\theta)$ is continuous at $\theta = \theta_0$ and positive. Then, it follows that

$$E\Big|\frac{\partial \log q(\theta)}{\partial\theta} + \sum_{i=1}^{n} \frac{\partial \log p(X_i,\theta)}{\partial\theta}\Big|^{\gamma} \leq A_{\gamma}(n+1)^{(\gamma/2-1)}\Big\{nE(I_\gamma(\theta)) + E\Big|\frac{\partial \log q(\theta)}{\partial\theta}\Big|^{\gamma}\Big\}. \tag{3.7}$$

Hence

$$\begin{aligned}\overline{\lim_{n\to\infty}}\Big\{E\Big|\frac{1}{\sqrt{n}}\frac{\partial \log q(\theta)}{\partial\theta} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \frac{\partial \log p(X_i,\theta)}{\partial\theta}\Big|^{\gamma}\Big\}^{l-1} &\leq \overline{\lim_{n\to\infty}}\Big\{A_\gamma\frac{(n+1)^{\gamma/2-1}}{n^{\gamma/2}}\big[nE(I_\gamma(\theta)) + E\Big|\frac{\partial \log q(\theta)}{\partial\theta}\Big|^{\gamma}\big]\Big\}^{l-1} \\ &\leq A_\gamma\big\{I_\gamma(\theta)\big\}^{l-1}.\end{aligned} \tag{3.8}$$

Hence, for any integer $k \geq 1$, it follows from (**??**) that

$$\underline{\lim}_{n\to\infty} \inf_{\theta^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(\theta^\star - \theta)|^{\frac{2k}{2k-1}} \geq \frac{1}{A_\gamma\big\{I_\gamma(\theta_0)\big\}^{l-1}} \tag{3.9}$$

where $\gamma = \frac{l}{l-1}$ and $l = \frac{2k}{2k-1}$. If $k = 1$, then $l = 2$, $\gamma = 2$, A_γ can be chosen to be equal to one and we have the classical Hajek-Le Cam lower bound for quadratic loss function. In general, for any differentiable function $g(\theta)$,

$$\underline{\lim}_{n\to\infty} \inf_{g^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(g^\star - g(\theta))|^{\frac{2k}{2k-1}} \geq \frac{[g'(\theta_0)]^l}{A_\gamma\big\{I_\gamma(\theta_0)\big\}^{l-1}} \tag{3.10}$$

where $\gamma = 2k$, $l = \frac{2k}{2k-1}$ and k is an integer ≥ 1.

(ii) Suppose $l \geq 2$. Then a computable lower bound for the minimax risk can be obtained from (**??**) by following the relation

$$E|Z|^\gamma \leq \big(E|Z^2|\big)^{\gamma/2} \tag{3.11}$$

for any random variable Z with $E(Z^2) < \infty$ where $\gamma \leq 2$ or equivalently $l \geq 2$. Note that for any prior density $q(\cdot)$ on $\mathcal{D}_n$,

$$\varliminf_{n\to\infty} \inf_{\theta^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(\theta^\star - \theta)|^l$$

$$\geq \frac{1}{\varlimsup_{n\to\infty}\left\{E\left|\frac{1}{\sqrt{n}}\frac{\partial \log q(\theta)}{\partial\theta} + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial \log p(X_i,\theta)}{\partial\theta}\right|^\gamma\right\}^{(\gamma/2)(l-1)}} \tag{3.12}$$

$$\geq \frac{1}{\varlimsup_{n\to\infty}\left\{E\left[\frac{1}{\sqrt{n}}\frac{\partial \log q(\theta)}{\partial\theta}\right]^2\right\} + \varlimsup_{n\to\infty}\left(E\left[I_2(\theta)\right]\right)^{l/2}}. \tag{3.13}$$

If the prior density $q(\theta)$ is chosen to be

$$q(\theta) = \frac{1}{\epsilon_n}\cos^2\frac{\pi(\theta-\theta_0)}{2\epsilon_n}, \quad \theta\in\mathcal{D}_n = [\theta : |\theta-\theta_0| \leq \epsilon_n], \tag{3.14}$$

then

$$I_2(\theta) \equiv E_\theta\left[\frac{\partial \log q(\theta)}{\partial\theta}\right]^2 = \frac{\pi^2}{\epsilon_n^2}, \tag{3.15}$$

and let $\epsilon_n = n^{-\alpha}$ with $0 < \alpha < \frac{1}{2}$. Then it follows that, for $l \geq 2$,

$$\begin{aligned}\varlimsup_{n\to\infty} \inf_{\theta^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(\theta^\star - \theta)|^l &\geq \frac{1}{\varlimsup_{n\to\infty}\left(E(I_1(\theta))\right)^{l/2}} \\ &\geq \frac{1}{\varlimsup_{n\to\infty}\sup_{\theta\in\mathcal{D}_n}\left(I_2(\theta)\right)^{l/2}} = \frac{1}{\left(I_2(\theta_0)\right)^{l/2}},\end{aligned} \tag{3.16}$$

where $I_2(\theta)$ is the Fisher information in one observation defined above. This bound is not sharp in view of the approximation of the lower bound in (**??**) using (**??**). In general, it can be shown that, for $l \geq 2$,

$$\varliminf_{n\to\infty} \inf_{g^\star} \sup_{\theta\in\mathcal{D}_n} E_\theta|\sqrt{n}(g^\star - g(\theta))|^l \geq \frac{\left[g'(\theta_0)\right]^l}{\left(I_2(\theta_0)\right)^{l/2}} \tag{3.17}$$

whenever $g(\cdot)$ is differentiable following arguments similar to those in Prakasa Rao (1992).

4 Lower Bound for the Posterior Moments

Ghosh (1993) obtained lower bounds for posterior variances applying the technique of Cramer-Rao inequality for posterior densities. We now obtain lower bounds for the posterior moments of even order. Suppose $h(\mathbf{X})$ is an arbitrary estimator of $\gamma(\theta)$. For any integer $k \geq 1$, consider

$$E\left[\left(h(\mathbf{X}) - \gamma(\theta)\right)^{2k}\Big|\mathbf{X}\right] \tag{4.1}$$

as a function of $h(\cdot)$. Suppose this is minimised by the function $C_k(\mathbf{X})$. Then, under suitable differentiable conditions, it follows that

$$E\left[\left(\gamma(\theta) - h(\mathbf{X})\right)^{2k}\Big|\mathbf{X}\right] \geq E\left[\left(\gamma(\theta) - C_k(\mathbf{X})\right)^{2k}\Big|\mathbf{X}\right] \tag{4.2}$$

and

$$E\left[\left(\gamma(\theta) - C_k(\mathbf{X})\right)^{2k-1}\Big|\mathbf{X}\right] = 0. \tag{4.3}$$

If $k=1$, then $C_k(\mathbf{X}) = E\big(\gamma(\theta)\big|\mathbf{X}\big)$ which is the posterior mean of $\gamma(\theta)$ given $\mathbf{X}$. Differentiating (**??**) under the integral sign with respect to θ, we have

$$E\big[\gamma'(\theta)(2k-1)(\gamma(\theta)-C_k(\mathbf{X}))^{2k-2}\big|\mathbf{X}\big] = E\big[(\gamma(\theta)-C_k(\mathbf{x}))^{2k-1}\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|\mathbf{X}\big] \tag{4.4}$$

where $\pi(\theta|\mathbf{X})$ is the posterior density of θ given $\mathbf{X}$. Applying Holder's inequality, we have

$$\begin{aligned}\Big|E\big[(\gamma(\theta)-C_k(\mathbf{x}))^{2k-1}\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|\mathbf{X}\big]\Big| \\ \le E^{\frac{2k-1}{2k}}\Big\{\big|\gamma(\theta)-C_k(\mathbf{X})\big|^{2k}\Big|\mathbf{X}\Big\}E^{\frac{1}{2k}}\Big\{\big|\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|^{2k}\Big|\mathbf{X}\Big\}\end{aligned} \tag{4.5}$$

and hence

$$E\Big\{\big|\gamma(\theta)-C_k(\mathbf{X})\big|^{2k}\Big|\mathbf{X}\Big\} \ge \frac{(2k-1)^{\frac{2k}{2k-1}}\big|E\big[\gamma'(\theta)(\gamma(\theta)-C_k(\mathbf{X}))^{2k-2}\big|\mathbf{X}\big]\big|^{\frac{2k}{2k-1}}}{E^{\frac{1}{2k-1}}\Big\{\big|\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|^{2k}\Big|\mathbf{X}\Big\}}. \tag{4.6}$$

Remark: In particular, for $k=1$, we have $C_1(\mathbf{x}) = E(\gamma(\theta)|\mathbf{X})$ and

$$E\Big\{\big|\gamma(\theta)-C_k(\mathbf{X})\big|^{2}\Big|\mathbf{X}\Big\} \ge \frac{\big(E\big[\gamma'(\theta)\big|\mathbf{X}\big]\big)^2}{E\Big\{\big|\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|^{2}\Big|\mathbf{X}\Big\}}. \tag{4.7}$$

Note that the previous inequality gives the lower bound for posterior variances derived by Ghosh (1993). If $h(\mathbf{X})$ is an arbitrary estimator of $\gamma(\theta)$, then

$$\begin{aligned}E\{|\gamma(\theta)-h(\mathbf{X})|^{2k}\big|\mathbf{X}\} \ge E\big[|\gamma(\theta)-h(\mathbf{X})|^{2k}\big|\mathbf{X}\big] - E\big[|\gamma(\theta)-C_k(\mathbf{X})|^{2k}\big|\mathbf{X}\big] \\ + \frac{(2k-1)^{\frac{2k}{2k-1}}E\big[\gamma'(\theta)(\gamma(\theta)-C_k(\mathbf{X}))^{2k-2}\big|\mathbf{X}\big]^{\frac{2k}{2k-1}}}{E^{\frac{1}{2k-1}}\Big\{\big|\frac{\partial\,\log\pi(\theta|\mathbf{X})}{\partial\theta}\big|^{2k}\Big|\mathbf{X}\Big\}}.\end{aligned} \tag{4.8}$$

Suppose $\mathcal{X} = \mathbb{R}$. An alternate way of obtaining another lower bound is as follows. Relation (**??**) can be written as

$$\int_{-\infty}^{\infty}\big[\gamma(\theta)-C_k(\mathbf{x})\big]^{2k-1}\pi(\theta|\mathbf{x})\,d\theta = 0. \tag{4.9}$$

Assuming that the differentiation under the integral sign with respect to $\mathbf{x}$ is possible, it follows that

$$\begin{aligned}-\int_{-\infty}^{\infty}C_k'(\mathbf{x})\big(\gamma(\theta)-C_k(\mathbf{x})\big)^{2k-2}(2k-1)\pi(\theta|\mathbf{x})\,d\theta \\ +\int_{-\infty}^{\infty}\big(\gamma(\theta)-C_k(x)\big)^{2k-1}\frac{\partial\pi(\theta|x)}{\partial x}\,d\theta = 0.\end{aligned} \tag{4.10}$$

Hence

$$\begin{aligned}E\Big[C_k'(X)\big(\gamma(\theta)-C_k(X)\big)^{2k-2}(2k-1)\Big|X\Big] \\ = E\Big[\big(\gamma(\theta)-C_k(X)\big)^{2k-1}\frac{1}{\pi(\theta|X)}\frac{\partial\pi(\theta|X)}{\partial X}\Big|X\Big].\end{aligned} \tag{4.11}$$

Applying Holder's inequality, we have

$$\begin{aligned}\Big|E\Big[\big(\gamma(\theta)-C_k(X)\big)^{2k-1}\frac{\partial\log\pi(\theta|X)}{\partial X}\Big|X\Big]\Big| \\ \le E^{\frac{2k-1}{2k}}\Big\{\big|\gamma(\theta)-C_k(X)\big|^{2k}\Big|X\Big\}E^{\frac{1}{2k}}\Big\{\big|\frac{\partial\log\pi(\theta|X)}{\partial X}\big|^{2k}\Big|X\Big\}.\end{aligned} \tag{4.12}$$

Hence

$$E\Big\{\big|\gamma(\theta)-C_k(X)\big|^{2k}\Big|X\Big\} \geq \frac{(2k-1)^{\frac{2k}{2k-1}}\big(C_k'(X)\big)^{\frac{2k}{2k-1}}E^{\frac{2k}{2k-1}}\Big\{\big|\gamma(\theta)-C_k(X)\big|^{2k-2}\Big|X\Big\}}{E^{\frac{1}{2k}}\Big\{\big|\frac{\partial\log\pi(\theta|X)}{\partial X}\big|^{2k}\Big|X\Big\}}$$

$$\geq \frac{(2k-1)^{\frac{2k}{2k-1}}\big(C_k'(X)\big)^{\frac{2k}{2k-1}}E^{\frac{2k}{2k-1}}\Big\{\big|\gamma(\theta)-C_{k-1}(X)\big|^{2k-2}\Big|X\Big\}}{E^{\frac{1}{2k-1}}\Big\{\big|\frac{\partial\log\pi(\theta|X)}{\partial X}\big|^{2k}\Big|X\Big\}}. \tag{4.13}$$

Repeating the process, we have

$$E\Big\{\big|\gamma(\theta)-C_k(X)\big|^{2k}\Big|X\Big\} \geq \prod_{l=0}^{k-1}\frac{\big\{|C_{k-l}'(X)|(2k-2l-1)\big\}^{\frac{2k(2k-2)\ldots(2k-2l)}{(2k-1)(2k-3)\cdots(2k-2l-1)}}}{E\Big(\big|\frac{\partial\ \log\pi(\theta|X)}{\partial X}\big|^{2k-2l}\Big|X\Big)^{\frac{2k}{2k-1}\cdots\frac{2k-2l+2}{2k-2l+1}\frac{1}{2k-2l-1}}}. \tag{4.14}$$

Remark: In particular, for $k=1$, we obtain that

$$E\big\{|\gamma(\theta)-C_1(X)|^2\big|X\big\} \geq \frac{|C_1'(X)|^2}{E\Big\{\big|\frac{\partial\ \log\pi(\theta|X)}{\partial X}\big|^2\Big|X\Big\}}. \tag{4.15}$$

Note that $C_1(X)=E[\gamma(\theta)|X]$ and hence, under suitable regularity conditions,

$$\begin{aligned}C_1'(x) &= \frac{d}{dx}\int_{-\infty}^{\infty}\gamma(\theta)\pi(\theta|x)\,d\theta\\ &= \int_{-\infty}^{\infty}\gamma(\theta)\frac{\partial\,\pi(\theta|x)}{\partial x}\,d\theta\\ &= \int_{-\infty}^{\infty}\gamma(\theta)\frac{\partial\ \log\pi(\theta|x)}{\partial x}\pi(\theta|x)\,d\theta\\ &= E\big[\gamma(\theta)\frac{\partial\ \log\pi(\theta|X)}{\partial X}\big|X\big]\end{aligned}$$

Hence

$$E\big\{|\gamma(\theta)-E(\gamma(\theta)|X)|^2\big|X\big\} \geq \frac{\Big(E\big[\gamma(\theta)\frac{\partial\ \log\pi(\theta|X)}{\partial X}\big|X\big]\Big)^2}{E\Big(\big|\frac{\partial\ \log\pi(\theta|X)}{\partial X}\big|^2\Big|X\Big)}. \tag{4.16}$$

The techniques used in this section are akin to those in Ibragimov (1999) generalizing Cramer-Rao inequality.

Remark: It is easy to check that the bounds in (**??**) and (**??**) are sharp. This can be checked by considering the case when $X\simeq N(\theta,1)$ and $\theta\simeq N(0,1)$. The advantage of inequality (**??**) is that it gives an explicit lower bound for the posterior moments where as the bound in (**??**) gives a lower bound in terms of an expectation of a function involving a lower order moment.

5 Examples

Example 1. Let us suppose that $X\sim N(\theta,1)$ and $\theta\sim N(0,1)$. It is easy to check that $C_k(x)=\frac{1}{2}x$ for all integer $k\geq 1$ and hence an application of (**??**) gives the inequality

$$E\big\{|\theta-\frac{1}{2}X|^{2k}\big|X\big\} \geq (2k-1)^{\frac{2k}{2k-1}}\frac{\big|E\big[(\theta-\frac{1}{2}X)^{2k-2}\big|\mathbf{X}\big]\big|^{\frac{2k}{2k-1}}}{E^{\frac{1}{2k-1}}\big\{|\theta-\frac{1}{2}X|^{2k}\big|\mathbf{X}\big\}^{\frac{1}{2k-1}}} \tag{5.1}$$

and an application of (**??**) gives the inequality

$$E\big\{|\theta-\frac{1}{2}X|^{2k}\big|X\big\} \geq \prod_{l=0}^{k-1}\frac{\big\{\frac{1}{2}X(2k-2l-1)\big\}^{\frac{2k}{2k-1}\frac{2k-2}{2k-3}\cdots\frac{2k-2l}{2k-2l-1}}}{E\big(|\theta-\frac{1}{2}X|^{2k-2l}\big|X\big)^{\frac{2k}{2k-1}\cdots\frac{1}{2k-2l-1}}}. \tag{5.2}$$

However, for this problem, it is easy to write down the explicit expression for $E\{|\theta - \frac{1}{2}X|^{2k}|X\}$ since $\pi(\theta|x)$ is $N(\frac{1}{2}X, \frac{1}{2})$. Here $N(\mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2.

Example 2. Consider the stochastic differential equation

$$dX(t) = S(t,\theta)\, dt + dW(t), \qquad 0 \le t \le 1,$$

where $W = \{W(t), 0 \le t \le 1\}$ is the standard Weiner process. Let ν be the probability measure generated by the process W on $C[0,1]$. Let P_θ be the probability measure generated by the process $\mathbf{X} \equiv \{X(t), 0 \le t \le 1\}$ on $C[0,1]$ when θ is the parameter. Suppose θ has a prior density $q(\theta)$ on $\mathbb{R}$. Then the joint density of $(\mathbf{X}, \theta)$ is

$$\frac{dP_\theta}{d\nu} = q(\theta)$$

with respect to the measure $\nu \times \lambda$ on $C[0,1] \times \mathbb{R}$, where λ is the Lebesgue measure on $\mathbb{R}$ and the posterior density is given by

$$\pi(\theta|\mathbf{X}) = \frac{\frac{d\,P_\theta}{d\nu} q(\theta)}{\int_{\mathbb{R}} \frac{d\,P_\theta}{d\nu} q(\theta)\, d\theta}.$$

Note that

$$\frac{d\,P_\theta}{d\nu} = \exp\Big\{\int_0^1 S(t,\theta)\, dX(t) - \frac{1}{2}\int_0^1 S^2(t,\theta)\, dt\Big\} \tag{5.3}$$

$$= \exp\{Z(\theta) - \frac{1}{2}J(\theta)\} \tag{5.4}$$

where

$$Z(\theta) = \int_0^1 S(t,\theta)\, dX(t) \qquad \text{and} \qquad J(\theta) = \int_0^1 S^2(t,\theta)\, dt \tag{5.5}$$

(cf. Basawa and Prakasa Rao (1980), Prakasa Rao (1999b,c)). Suppose $q(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2}$, $-\infty < \theta < \infty$ and $S(\theta, t) = f(t)\theta$. Then

$$\pi(\theta|\mathbf{X}) = \frac{e^{-\frac{1}{2}\{\theta^2 + J(\theta) - 2Z(\theta)\}}}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}\{\theta^2 + J(\theta) - 2Z(\theta)\}}\, d\theta} \tag{5.6}$$

where

$$Z(\theta) = \theta \int_0^1 f(t)\, dX(t)$$

and

$$J(\theta) = \theta^2 \int_0^1 f^2(t)\, dt.$$

Hence

$$\pi(\theta|\mathbf{X}) = \frac{e^{-\frac{1}{2}\left\{\theta^2\left(1+\int_0^1 f^2(t)\, dt\right) - 2\theta\int_0^1 f(t)\, dX(t)\right\}}}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left\{\theta^2\left(1+\int_0^1 f^2(t)\, dt\right) - 2\theta\int_0^1 f(t)\, dX(t)\right\}}\, d\theta} \tag{5.7}$$

which shows that

$$\pi(\theta|\mathbf{X}) \quad \text{is} \quad N\Big(\frac{\int_0^1 f(t)\, dX(t)}{1+\int_0^1 f^2(t)\, dt}, \frac{1}{1+\int_0^1 f^2(t)\, dt}\Big).$$

It is easy to see that

$$C_k(\mathbf{X}) = \frac{\int_0^1 f(t)\,dX(t)}{1+\int_0^1 f^2(t)\,dt} \tag{5.8}$$

for all integers $k \geq 1$ and in fact

$$E\Big\{\big|\theta - \frac{\int_0^1 f(t)\,dX(t)}{1+\int_0^1 f^2(t)\,dt}\big|^{2k}\Big|\mathbf{X}\Big\} = \frac{1}{\big(1+\int_0^1 f^2(t)\,dt\big)^k}\alpha_k \tag{5.9}$$

where α_k is the $2k$-th moment of a standard normal distribution. An application of (**??**) proves that

$$E\Big\{\big|\theta - \frac{\int_0^1 f(t)\,dX(t)}{1+\int_0^1 f^2(t)\,dt}\big|^{2k}\Big|\mathbf{X}\Big\}$$
$$\geq \frac{(2k-1)^{\frac{2k}{2k-1}}\Big|E\Big\{\big[\theta - \frac{\int_0^1 f(t)\,dX(t)}{1+\int_0^1 f^2(t)\,dt}\big]\Big|\mathbf{X}\Big\}\Big|^{\frac{2k}{2k-1}}}{E\Big\{\Big|\Big\{\theta - \frac{\int_0^1 f(t)\,dX(t)}{1+\int_0^1 f^2(t)\,dt}\Big\}\Big/\Big\{\frac{1}{1+\int_0^1 f^2(t)\,dt}\Big|^{2k}\Big|\mathbf{X}\Big\}^{\frac{1}{2k-1}}}. \tag{5.10}$$

Repeating the process, one can obtain the lower bound. However, exact computation is possible in this example as indicated above.

Acknowledgement: The author wishes to thank Professor Peter Hall, Centre for Mathematics and its Applications, Australian National University, for his invitation to visit during which time this work was completed.

References

[1] Basawa, I. V. and Prakasa Rao, B. L. S. (1980) *Statistical Inference for Stochastic Processes*, Academic Press, New York.

[2] Bobrovsky, B. Z., Mayer-Wolf, E., and Zakai, M. (1987) Some classes of global Cramer-Rao bounds, *Ann. Statistics* **15**, 1421–1438.

[3] Borovkov, A. A. and Sakhanenko, A. I. (1980) On estimates for the average quadratic risk, *Probab. Math. Statist.* **1**, 185–195 (In Russian).

[4] Brown, L. D. and Gajek, L. (1990) Information inequalities for the Bayes risk, *Ann. Statistics* **18**, 1578–1594.

[5] Ghosh, M. (1993) Cramer-Rao bound for posterior variances, *Stat. Prob. Letters* **17**, 173–178.

[6] Gill, R. D. and Levit, B. Y. (1995) Application of the Van Trees inequality; a Bayesian Cramer-Rao bound, *Bernoulli* **1**, 59–79.

[7] Ibragimov, I. A. (1999) A generalization of the Cramer-Rao inequality, *J. Math. Sci.* **93**, 349–353.

[8] Loéve, M. (1977) *Probability Theory*, Vol I, Springer-Verlag, New York.

[9] Marcinkiewicz, M. and Zygmund, A. (1937) Sur les fonctions independantes, *Fund. Math.* **29**, 60–90.

[10] Prakasa Rao, B. L. S. (1991) On Cramer-Rao type integral inequalities, *Calcutta Stat. Assoc. Bull.* **40**, 183–205.

[11] Prakasa Rao, B. L. S. (1992) Cramer-Rao type integral inequalities for estimators of functions of multidimensional parameter, *Sankhya Ser. A.* **54**, 53–73.

[12] Prakasa Rao, B. L. S. (1995) Remarks on Cramer-Rao type integral inequality for randomly censored data, *IMS Lecture Notes No. 27, Analysis of censored data* (Ed. H. L. Koul and J. V. Deshpande), pp. 163–176.

[13] Prakasa Rao, B. L. S. (1999a) Cramer-Rao type integral inequality in Banach-spaces, *Proc. 3rd Triennial Calcutta Symposium on Prob. and Stat.* (to appear).

[14] Prakasa Rao, B. L. S. (1999b) *Statistical Inference for Diffusion Type Processes*, Arnold Publishers, London.

[15] Prakasa Rao, B. L. S. (1999c) *Semimartingales and Their Statistical Inference*, CRC Press, Boca Raton, Florida.

[16] Van Trees, H. (1968) *Detection, Estimation and Modulation Theory*, Vol. I, Wiley, New York.