Journal of the Korean
Statistical Society
Vol. 26, No. 4, 1997

Chapman-Robbins-type and Bayesian Lower Bounds Based on Diffusivity for Median-unbiased Estimators †

Sung Nae Kyung[1]

Abstract

A more generalized version of the information inequality based on diffusivity which is a natural measure of dispersion for median-unbiased estimators developed by Sung *et al.* (1990) is presented. This non-Bayesian L_1 information inequality is free from regularity conditions and can be regarded as an analogue of the Chapman-Robbins inequality for mean-unbiased estimation. The approach given here, however, deals with a more generalized situation than that of the Chapman-Robbins inequality. We also develop a Bayesian version of the L_1 information inequality in median-unbiased estimation. This latter inequality is directly comparable to the Bayesian Cramér-Rao bound due to the van Trees inequality.

Key Words : Median-unbased estimation; Diffusivity; Van-Trees inequality; Information.

---

†The author wishes to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997 : Grant No. 1997-001-D00064.

[1]Statistics Department, Ewha Womans University, Seoul 120-750, Korea.

1. INTRODUCTION

There exist many criteria under which we determine a good estimator for a certain parametric function of interest. Whatever criterion we use, we need a measure of closeness of the estimator to the parametric function in order to assess the efficiency, or performance of the estimator. The usual practice of searching for a minimum-variance estimator within the class of mean-unbiased estimators is an example.

Because of the arbitrariness of selecting a resticted class of estimators, Lehmann (1951) considered a more general approach of finding an optimal estimator in the class of risk-unbiased estimators which include the usual class of L_p-unbiased estimators. Hence an optimal estimator in the class of L_p-unbiased estimators minimizes the Minkowsky metric.

The common practical choices of p would be either 2 or 1. For $p = 2$ the risk is restricted to the squared error loss and it is minimized under mean-unbiased estimators. For $p = 1$ the risk is restricted to the absolute error loss and it is minimized under median-unbiased or L_1-unbiased estimators.

Let X denote a random vector and $\tau(\theta)$ be a parametric function of interest.

Definition. $\delta(X)$ is called *median-unbiased* for $\tau(\theta)$ if

$$\text{median}_\theta \delta(X) = \tau(\theta) \quad \text{for all} \quad \theta \in \Theta$$

Let $\delta(X)$ be a median-unbiased estimator having a continuous density g_δ. Then, as was shown by Sung *et al.* (1990), under certain regularity conditions, the following information inequality holds:

$$\frac{1}{2g_\delta(\tau(\theta);\theta)} \geq \frac{|\tau'(\theta)|}{I_1(\theta)}, \tag{1.1}$$

where I_1 is the first absolute moment of the sample score:

$$I_1(\theta) = E_\theta \left| \frac{\partial \log f(x;\theta)}{\partial \theta} \right|.$$

The left-hand side term in (1.1) is called *diffusivity*, which is the reciprocal of twice the median-unbiased estimator's density height evaluated at its median point.

Refer to Sung (1993) and Sung (1990) for a general discussion of diffusivity and its multivariate form, respectively.

In this paper we generalize the L_1 information inequality (1.1) and present an analogue of the Chapman-Robbins (1951) inequality which is free from regularity conditions.

In addition, we derive a Bayesian L_1 information inequality which can be regarded as an L_1 version of the van Trees inequality. The van Trees inequality provides a Bayesian Cramér-Rao bound.

In this case, we define that $\delta(X)$ is median-unbiased for $\tau(\theta)$ if the median of $\delta(X)$ with respect to the posterior distribution coincides to $\tau(\theta)$. For more notations for Bayesian inferences, see Section 3.

Note that, in developing a Bayesian L_1 information inequalities, we will follow the notations used in Gill and Levit (1995).

2. GENERALIZED CHAPMAN-ROBBINS-TYPE INEQUALITY BASED ON DIFFUSIVITY

In this section we derive a generalized Chapman-Robbins-type information inequality based on diffusivity.

Theorem 2.1. Let X be a random variable with density $f(x;\theta)$ on a σ-finite measure space $(\mathcal{X}, \mathcal{F}, \mu)$ with $\theta \in \Theta \subset \mathcal{R}$ open. Let $f(x;\theta)$ be absolutely continuous in θ for μ almost all x with Radon-Nikodym derivative $\tilde{f}(x;\theta)$ and let $I_1(\theta) = \int |\tilde{f}(x;\theta)| d\mu(x)$ be continuous in θ. Let $\tau : \Theta \to \mathcal{R}$ be differentiable with derivative τ'. If $\delta(X)$ is a median-unbiased estimator of $\tau(\theta)$ with distribution function $G_\delta(y;\theta)$, $y \in \mathcal{R}$, under θ, then

$$\limsup_{\epsilon \to 0} \frac{2}{\epsilon}\{G_\delta(\tau(\theta)+\epsilon;\theta) - G_\delta(\tau(\theta);\theta)\} \le \frac{I_1(\theta)}{|\tau'(\theta)|} \tag{2.1}$$

with equality *iff*

$$\tau'(\theta)\{\delta(x) - \tau(\theta)\}\tilde{f}(x;\theta)\} \ge 0 \qquad \text{for } \mu \text{ almost all } x. \tag{2.2}$$

Proof. The left hand side of the inequality (2.1) equals

$$\limsup_{\eta \to 0}\{\tau(\theta+\eta) - \tau(\theta)\}^{-1}\{2G_\delta(\tau(\theta+\eta);\theta) - 1\}$$
$$= \limsup_{\eta \to 0}\{\tau'(\theta)\}^{-1}\frac{1}{\eta}\{G_\delta(\tau(\theta+\eta);\theta) - [1 - G_\delta(\tau(\theta+\eta);\theta)]\}$$

$$
\begin{aligned}
&= \limsup_{\eta\to 0}\{\tau'(\theta)\}^{-1}\frac{1}{\eta}\left\{\int_{\delta(x)\le\tau(\theta+\eta)}[f(x;\theta)-f(x;\theta+\eta)]\mu(dx)\right.\\
&\qquad\left.+\int_{\delta(x)>\tau(\theta+\eta)}[f(x;\theta+\eta)-f(x;\theta)]\mu(dx)\right\}\\
&= \limsup_{\eta\to 0}\{\tau'(\theta)\}^{-1}\int\left\{1_{(\tau(\theta+\eta),\infty)}(\delta(x))-1_{(-\infty,\tau(\theta+\eta))}(\delta(x))\right\}\\
&\qquad\times\frac{1}{\eta}\int_{\theta}^{\theta+\eta}\tilde{f}(x;\xi)d\xi\mu(dx)\\
&\le \limsup_{\eta\to 0}\{|\tau'(\theta)|\}^{-1}\frac{1}{\eta}\int_{\theta}^{\theta+\eta}\int|\tilde{f}(x;\xi)|\mu(dx)d\xi\\
&= \frac{I_1(\theta)}{|\tau'(\theta)|}.
\end{aligned}
$$

Hence the inequality (2.1) is proved and one can verify that the equality (2.2) holds if and only if (2.2) is satisfied in view of the relations (14) and (15) given in Sung *et al.* (1990).

Theorem 2.2. If G_δ has a density g at $\tau(\theta)$, then (2.1) reduces to as

$$\frac{1}{2g(\tau(\theta);\theta)}\ge\frac{|\tau'(\theta)|}{I_1(\theta)}. \tag{2.3}$$

We remark that for the strongly unimodal densities of the form $f(x;\theta)\propto\exp h(x-\theta)$ for h strictly concave, one can verify that the MLE satisfies (2.2), as was shown in Sung *et al.* (1990).

3. BAYESIAN INFORMATION BOUND

In this section we will derive a Bayesian information inequality for a dispersion measure called *Bayesian diffusivity.* However, before doing so, we review the van Trees inequality and introduce some notaions for Bayesian information inequalities.

Let $(\mathcal{X},\mathcal{F},P_\theta:\theta\in\Theta)$ be a family of distributions on some sample space $\mathcal{X}$, dominated by a measure μ. Let the parameter space Θ is a closed interval on the real line. Let $f(x|\theta)$ denote the density of P_θ with respect to μ. Let π be some probability distribution on Θ with a density $\lambda(\theta)$ with respect to Lebesgue measure.

We suppose that both λ and $f(x|\cdot)$ are absolutely continuous for μ almost surely. Assume further that λ converges to zero at the endpoints of the interval Θ. From now on a prime will denote a partial derivative with respect to θ.

Let $\delta = \delta(X)$ denote any estimator of θ, where $X \sim P_\theta$. We write E_θ for expectation with respect to θ. When $\underline{\theta}$ is drawn from the distribution π, and conditional on $\underline{\theta} = \theta$, X from P_θ, we write E for expectation with respect to the ensuing joint distribution of X and $\underline{\theta}$.

Define further

$$\mathcal{I}_2(\theta) = E_\theta \left[\frac{\partial \log f(X|\theta)}{\partial \theta} \right]^2$$

and

$$\mathcal{I}_2(\lambda) = E_\theta \left[\frac{\partial \log \lambda(\theta)}{\partial \theta} \right]^2,$$

where $\mathcal{I}_2(\theta)$ is the usual Fisher information for θ and $\mathcal{I}_2(\lambda)$ is the Fisher information for λ, respectively.

Then, as van Trees (1968) showed, under the condition that $\sqrt{\mathcal{I}_2(\theta)}$ is locally integrable in θ, we have the following van Trees inequality:

$$E[\delta(X) - \underline{\theta}]^2 \geq \frac{1}{E[\mathcal{I}_2(\underline{\theta})] + \mathcal{I}_2(\lambda)}. \tag{3.1}$$

The van Trees inequality provides a Bayesian Cramér-Rao bound for L_2 estimation.

Further studies concerning to the van Trees information inequality have been given in Brown and Gajek (1990), Bobrovsky *et al.* (1987) and Klaassen (1989).

Gill and Levit (1995) generalized the van Trees inequality and derived the following information inequality for estimating an absolutely continuous function τ of θ:

$$E[\delta(X) - \tau(\underline{\theta})]^2 \geq \frac{[E\{\tau'(\theta)\}]^2}{E[\mathcal{I}_2(\underline{\theta})] + \mathcal{I}_2(\lambda)}. \tag{3.2}$$

We now develop an analogue of the van Trees information inequality for median-unbiased estimators, where the measure of dispersion is the Bayesian diffusivity. The resulting inequality will be called the Bayesian L_1 information inequality.

Theorem 3.1. Let X be a random variable with density $f(x|\theta)$ on a σ-finite measure space $(\mathcal{X}, \mathcal{F}, \mu)$ with $\theta \in \Theta \subset \mathcal{R}$ open. Let π be some probability distribution on Θ with a density $\lambda(\theta)$ with respect to Lebesgue measure. Let $f(x|\cdot)$ be absolutely continuous in θ for μ almost surely with Radon-Nikodym derivative $\tilde{f}(x|\theta)$. Let λ be also absolutely continuous for μ almost surely. Define further

$$\mathcal{I}_1(\theta) = E_\theta \left| \frac{\partial \log f(X|\theta)}{\partial \theta} \right|$$

and

$$\mathcal{I}_1(\lambda) = E_\theta \left| \frac{\partial \log \lambda(\theta)}{\partial \theta} \right|.$$

Let $\tau : \Theta \to \mathcal{R}$ be differentiable with derivative τ'. Let $Y = \delta(X)$ be a median-unbiased estimator of $\tau(\theta)$ with distribution function $G_\delta(y|\theta)$, $y \in \mathcal{R}$, under θ. We suppose further that τ' and G' are either both nondecreasing or both nonincreasing. Then

$$\begin{aligned} &\limsup_{\epsilon \to 0} \frac{2}{\epsilon} \int \{G_\delta(\tau(\theta) + \epsilon|\theta) - G_\delta(\tau(\theta)|\theta)\}\pi(d\theta) \\ &\leq \{E[\mathcal{I}_1(\underline{\theta})] + \mathcal{I}_1(\lambda)\} E_\theta \left| \frac{1}{\tau'(\underline{\theta})} \right|. \end{aligned} \tag{3.3}$$

Proof. By taking $\epsilon = \tau(\theta+\eta) - \tau(\theta)$ and applying the covariance inequality, the left hand side of the inequality (3.3) becomes

$$\begin{aligned} &\limsup_{\eta \to 0} \int \left[\frac{\tau(\theta+\eta) - \tau(\theta)}{\eta} \right]^{-1} \frac{2}{\eta} \left[G_\delta(\tau(\theta+\eta)|\theta) - \frac{1}{2} \right] \pi(d\theta) \\ &\leq E_\theta \left| \frac{1}{\tau'(\underline{\theta})} \right| \limsup_{\eta \to 0} \int \frac{1}{\eta} [G_\delta(\tau(\theta+\eta)|\theta) - \{1 - G_\delta(\tau(\theta+\eta)|\theta)\}] \pi(d\theta) \\ &= E_\theta \left| \frac{1}{\tau'(\underline{\theta})} \right| \limsup_{\eta \to 0} \int \frac{1}{\eta} \int_{\delta(x) \leq \tau(\theta+\eta)} \{f^*(x,\theta) - f^*(x,\theta+\eta)\} \mu(dx) \\ &\quad + \int_{\delta(x) > \tau(\theta+\eta)} \{f^*(x,\theta+\eta) - f^*(x,\theta)\} \mu(dx) d\theta \\ &= E_\theta \left| \frac{1}{\tau'(\underline{\theta})} \right| \limsup_{\eta \to 0} \int \int \int \{1_{(\tau(\theta+\eta),\infty)}(\delta(x)) - 1_{(-\infty,\tau(\theta+\eta))}(\delta(x))\} \\ &\quad \times \frac{1}{\eta} \int_\theta^{\theta+\eta} \tilde{f}^*(x,\xi) d\xi \mu(dx) d\theta \end{aligned}$$

$$\le E_{\theta}\left|\frac{1}{\tau'(\underline{\theta})}\right| \limsup_{\eta\to 0} \int \frac{1}{\eta}\int_{\theta}^{\theta+\eta} |\tilde{f}^*(x,\xi)|\mu(dx)d\theta$$
$$= E_{\theta}\left|\frac{1}{\tau'(\underline{\theta})}\right| \int\int \left|\frac{\partial \log f(x|\theta)}{\partial\theta} + \frac{\partial \log\lambda(\theta)}{\partial\theta}\right| f(x|\theta)\mu(dx)\pi(d\theta)$$
$$\le E_{\theta}\left|\frac{1}{\tau'(\underline{\theta})}\right| \int\int \left\{\left|\frac{\partial \log f(x|\theta)}{\partial\theta}\right| + \left|\frac{\partial \log\lambda(\theta)}{\partial\theta}\right|\right\} f(x|\theta)\mu(dx)\pi(d\theta)$$
$$= E_{\theta}\left|\frac{1}{\tau'(\underline{\theta})}\right| \{E[\mathcal{I}_1(\underline{\theta})] + \mathcal{I}_1(\lambda)\}.$$

In the course of proof, f^* denotes the joint density of X and $\underline{\theta}$.

In Theorem 3.1, we call the reciprocal of the left hand side of (3.3) the *Bayesian diffusivity*, which is the expected diffusivity with respect to θ.

From Theorem 3.1, we have the following immediate consequences:

Theorem 3.2. If G_δ has a density g at $\tau(\theta)$, then (3.3) can be expressed as

$$\frac{1}{E_{\theta}2g(\tau(\underline{\theta})|\theta)} \ge \frac{1/[E_{\theta}|\{1/\tau'(\underline{\theta})\}|]}{E[\mathcal{I}_1(\underline{\theta})] + \mathcal{I}_1(\lambda)}. \tag{3.4}$$

Furthermore, if δ is a median-unbiased estimator of $\tau(\theta) = \theta$, then we have

$$\frac{1}{E_{\theta}2g(\tau(\underline{\theta})|\theta)} \ge \frac{1}{E[\mathcal{I}_1(\underline{\theta})] + \mathcal{I}_1(\lambda)}. \tag{3.5}$$

In view of Theorem 3.2, the Bayesian diffusivity has the form of $1/E_{\theta}[2g(\tau(\underline{\theta})\ |\theta)]$.

Note that (3.5) is a direct analogue of the van Trees inequality and (3.4) is an analogue of Gill-Levit-type inequality.

Example. Let $X|\theta \sim N(\theta, 1)$ and $\underline{\theta} \sim N(0,1)$. Then the posterior distribution of $\underline{\theta}$ is also normal such that

$$\underline{\theta}|x \sim N\left(\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}\right).$$

Take $\hat{\theta} = \delta(X) = X/2 + 1/2$. Then $E_{\theta}|\mathcal{I}_1(\underline{\theta})| = \mathcal{I}_1(\lambda) = \sqrt{2/\pi}$ and $E_{\theta}[2|\tau'(\underline{\theta}|\theta)|] = 2/\sqrt{\pi}$. Hence δ in this case does not achieve the bound.

It appears that finding an estimator which achieves the Bayesian bound is not easy. One may observe the same phenomenon as in Bayesian mean-unbiased estimation. For instance, see Gill and Levit (1995).

Finally we remark that exact computation of the Bayesian bound, whether it be for mean-unbiased or median-unbiased estimation, is very difficult except for a few trivial cases. In this regard it seems to be necessary to provide an efficient computer algorithm for proper numerical integration technique in order to obtain Bayesian bounds.

REFERENCES

(1) Bobrovsky, B. Z., Mayer-Wolf, E. and Zakai, M. (1987). Some classes of global Cramér-Rao bounds. *Annals of Statistics*, **15**, 1421-1438.

(2) Brown, L. D. and Gajek, L. (1990). Information inequalities for the Bayes risk. *Annals of Statistics*, **18**, 1578-1594.

(3) Chapman, D. G. and Robbins, H. (1951). Minimum variance estimation without regularity assumptions. *Annals of Mathematical Statistics*, **22**, 581-586.

(4) Gill, R. D. and Levit, B. Y. (1995). Applications of the van Trees inequality: a Bayesian Cramér-Rao bound. *Bernoulli*, **1**, 59-79.

(5) Klaassen, C. A. J. (1989). The asymptotic spread of estimators. *Journal of Statistical Planning and Inference*, **23**, 267-285.

(6) Lehmann, E. L. (1951). A general concept of unbiasedness. *Annals of Mathematical Statistics*, **22**, 587-592.

(7) Sung, N. K. (1990). A generalized Cramér-Rao analogue for median-unbiased estimators. *Journal of Multivariate Analysis*, **32**, 204-212.

(8) Sung, N. K. (1993). Improving L_1 information bound in the presence of a nuisance parameter for median-unbiased estimators. *Journal of Korean Statistical Society*, **22**, 1-12.

(9) Sung, N. K., Stangenhaus, G., and David, H. T. (1990). A Cramér-Rao analogue for median-unbiased estimators. *Trabajos de Estadística*, **5**, 83-94.

(10) van Trees, H. L. (1968). *Detection, estimation and modulation theory.* Part 1. Wiley, New York.