Journal of the Korean
Statistical Society
Vol. 23, No. 1, 1994

A Sharp Cramer-Rao type Lower-Bound for Median-Unbiased Estimators

Beong-Soo So [1]

ABSTRACT

We derive a new Cramer-Rao type lower bound for the reciprocal of the density height of the median-unbiased estimators which improves most of the previous lower bounds and is attainable under much weaker conditions. We also identify useful necessary and sufficient condition for the attainability of the lower bound which is considerably weaker than those for the mean-unbiased estimators. It is shown that these lower bounds are attained not only for the family of continuous distributions with *monotone likelihood ratio* (MLR) property but also for the location and scale families with *strong unimodal* property.

KEYWORDS : Median-unbiased estimator, Cramer-Rao lower bound, Monotone likelihood ratio, Strongly unimodal family

1. INTRODUCTION

Let μ be a Lebesque measure on the Euclidean space R^n . Let $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$be the family of distributions on R^n which are absolutely continuous with respect to μ and depend upon a single parameter θ where the parameter space

[1] Dept. of Statistics, Ewha Womans University, Seodaemun, Seoul, 120-750, Korea

Θ is an open interval on the real line. We assume that for every $\theta \in \Theta$ $f(x;\theta)$ is a density function of P_θ with respect to μ where $x = (x_1, \cdots, x_n) \in R^n$. Let $X = (X_1, \cdots, X_n)$ be a random vector having a joint density function $f(x;\theta)$. Let $g(\theta)$ be a real valued differentiable function on Θ.

We define an estimator $\delta(X)$ of $g(\theta)$ is *median-unbiased* if

$$[1]\text{median}_\theta(\delta(X)) = g(\theta) \quad \textit{for all} \quad \theta \in \Theta. \tag{1.1}$$

For any estimator having absolutely continuous distribution condition (1.1) is equivalent to

$$P_\theta[\,\delta(X) \leq g(\theta)\,] = P[\,\delta(X) \geq g(\theta)\,] = 1/2 \quad \textit{for all} \quad \theta \in \Theta. \tag{1.2}$$

Let $\delta(X)$ be any median-unbiased estimator of $g(\theta)$ and let $f_\delta(y;\theta)$ be the density function of the random variable $\delta(X)$ with respect to the Lebesque measure in R.

In this framework several versions of the analogue of the Cramer-Rao lower bound for median-unbiased estimators were proposed in the literature. In a pioneering work in this direction, Alamo (1964) introduced the quantity $f_\delta(g(\theta);\theta)$ as a natural measure of *concentration* of the median-unbiased estimator $\delta(X)$ around the estimand $g(\theta)$ and proposed its reciprocal quantity $1/f_\delta(g(\theta);\theta)$ as a new measure of the *dispersion* of the estimator. Then he obtained the following lower bound for median-unbiased estimator:

$$[\,2f_\delta(g(\theta);\theta)\,]^{-1} \geq |g'(\theta)|/I_2(\theta)^{1/2} \tag{1.3}$$

where $I_2(\theta)$ is the usual Fisher information number

$$I_2(\theta) = E[\,(\partial/\partial\theta)\,\log f(X;\theta)\,]^2.$$

Here we note that the quantity $[2f_\delta(g(\theta);\theta)]^{-1}$ can be interpreted as a reasonable measure of the *dispersion* of the estimator $\delta(X)$ around the estimand $g(\theta)$ corresponding to the *concentration* measure $f_\delta(g(\theta);\theta)$. See Sung and et al (1990) for more detailed interpretation of quantity $[2f_\delta(g(\theta);\theta)]^{-1}$ as a measure of *diffusivity* of the estimator. Recently a sharper lower bound for the left hand side of the above inequality (1.3) was proposed by several authors including Stangenhaus (1977), Stangenhaus and David (1978) and Sung, Stangenhaus and David (1990) respectively under slightly different regularity conditions. Essentially they obtained

$$[2f_\delta(g(\theta);\,\theta)]^{-1} \geq |\,g'(\theta)\,|\,/\,I_1(\theta) \tag{1.4}$$

where $I_1(\theta)$ is an analogue of Fisher information

$$I_1(\theta) = E_\theta|\,(\partial/\partial\theta)\,\log\,f(X;\theta)\,|.$$

Sung, Stangenhaus and David (1990) also identified bound-achieving median-unbiased estimates for some special type of strongly unimodal location and scale families.

On the other hand Lehmann(1986) and Pfanzagl(1970) obtained strong optimality results for the best median-unbiased estimators with respect to arbitrary monotone loss function when the family of densities $f(x;\theta)$ has a *monotone likelihood ratio* (MLR) property .

In view of these results it is clear that the lower bound (1.4) is not the best possible bound because lower bound in (1.4) is not attainable in the most of the monotone likelihood families except for the normal location family and this clearly contradicts the general optimality results of Lehmann(1986) and Pfanzagl(1970).

In this paper we will derive a sharper lower bound for the reciprocal quantity of the density height $[2f_\delta(g(\theta);\theta)]^{-1}$ of the median unbiased estimator $\delta(X)$ which improves the previous lower bound (1.4) and is attainable for the much wider families of density functions including arbitrary monotone likelihood ratio family and general strongly unimodal location family as important special cases. Under suitable regularity conditions, we will obtain the following lower bound :

$$[2f_\delta(g(\theta);\,\theta)]^{-1} \geq |\,g'(\theta)\,|\,/\,I_1^*(\theta) \tag{1.5}$$

where $I_1^*(\theta)$ is the *centered* analogue of the L_1-Fisher information

$$I_1^*(\theta) = E_\theta\,|\,(\partial/\partial\theta)\,\log\,f(X;\theta) - k\,|$$

and

$$k = \text{median}_\theta\,[\,(\partial/\partial\theta)\,\log\,f(X;\theta)\,].$$

Note that the lower bound (1.5) is always strictly greater than the previous bound (1.4) unless the constant k happens to be identically zero. We will also identify the exact necesssary and sufficient condition for the attainability of the lower bound (1.5) which is much weaker than that of the mean-unbiased estimators and also weaker than that of the previous result (1.4) for median-unbiased estimators. Then we show that these conditions are satisfied not only for the general family of density functions having *monotone likelihood ratio* property including the exponential class as an important special cases but also for the general location and scale families with *strongly unimodal* property including censored survival data as special cases.

2. LOWER BOUNDS

Our method of construction of the lower bound closely parallel the similar derivation of the analogue of the Chapman-Robbins inequality which is free of regularity conditions as is given by Sung et al. (1990) but differs fundamentally from all the previous derivations in using important centering argument which has never been considered before in this context. Let Δ be a real number such that both θ and $\theta + \Delta$ belong to Θ. By the definition of median-unbiased estimator, we have the following identities :

$$E_\theta\left[\, sgn(\delta(X) - g(\theta))\,\right] = 0. \tag{2.1}$$

$$E_{\theta+\Delta}\left[\, sgn(\delta(X) - g(\theta+\Delta))\,\right] = 0. \tag{2.2}$$

Subtracting (2.1) from (2.2) , we get

$$\begin{aligned}&\int [\, f(x;\theta+\Delta) - f(x;\theta)\,] sgn[\,\delta(X) - g(\theta+\Delta)\,]\, d\mu \\ &+ E_\theta \left\{ sgn[\,\delta(X) - g(\theta+\Delta)\,] - sgn[\,\delta(X) - g(\theta)\,] \right\} = 0.\end{aligned} \tag{2.3}$$

Multiplying (2.2) by $k\Delta$ and subtracting it from (2.3), we have the following identity :

$$\begin{aligned}&2\,[\, F_\delta(g(\theta+\Delta)\,;\,\theta) - F_\delta(g(\theta)\,;\,\theta)\,] \\ =\; &\int [\, f(x;\theta+\Delta) - f(x;\theta) - k\cdot\Delta f(x;\theta+\Delta)] \\ &sgn[\delta(X) - g(\theta+\Delta)] d\mu\end{aligned} \tag{2.4}$$

where $F_\delta(\cdot;\theta)$ is the distribution function of $\delta(X)$ under the distribution $f(x;\theta)$.

Following lemma summarizes above result in a form which will be more convenient for the derivation of lower bound.

Lemma 1. Let $g(\theta)$ be a real valued function on Θ. Let $\delta(X)$ be a median-unbiased estimator of $g(\theta)$ having absolutely continuous distribution function. Then for arbitrary constant k we have the following identity :

$$\begin{aligned} & 2\,[\,F_\delta(g(\theta+\Delta);\theta) - F_\delta(g(\theta);\theta)\,] \\ = & \int |\, f(x;\theta+\Delta) - f(x;\theta) - k\Delta f(x;\theta+\Delta)|s_1 \cdot s_2 d\mu \qquad (2.5) \end{aligned}$$

where $s_1 = sgn[f(x;\theta+\Delta) - f(x;\theta) - k\Delta f(x;\theta+\Delta)]$ and $s_2 = sgn[\delta(X) - g(\theta+\Delta)]$.

Proof. It follows Immediately from (2.4) .

Remark 1. Introduction of the extra centering parameter k in (2.5) is the key difference between our approach and most of previous derivations which do not consider this possibility. This additional degree of freedom achieved by the introduction of the extra parameter k will be exploited crucially in the derivation of the sharper lower bound .

Now we are ready to present an analogue of Chapman-Robbins type inequality for the median-unbiased estimators.

Theorem 1. Let $g(\theta)$ be a real-valued differentiable function on $\Theta \subset R$. Let $\delta(X)$ be a median-unbiased estimator of $g(\theta)$ having a density function $f_\delta(\cdot;\theta)$ which is continuous at $g(\theta)$. Then we have :

$$\begin{aligned} [\,2f_\delta(g(\theta)\,;\,\theta)\,]^{-1} \geq \; & |g'(\theta)|/\inf_k \liminf_{\Delta\to 0} \int |\,\{\,f(x;\theta+\Delta) - f(x;\theta)\,\}/ \\ & \Delta - kf(x;\theta+\Delta)|\,d\mu. \qquad (2.6) \end{aligned}$$

Proof. Dividing (2.5) by Δ and taking limits as $\Delta \to 0$ for fixed k and noting $|s_1 \cdot s_2| \leq 1$, we get the result.

Remark 2. Actually we can replace the continuity of the density $f_\delta(\cdot;\theta)$ at the point $g(\theta)$ by the existence of the left and right limits of the $f_\delta(\cdot;\theta)$

at the point and can get the corresponding one-sided version of the inequality without difficulty.

Remark 3. The density of the uniform distribution

$$f(x\,;\,\theta) = I_{[\theta-1/2,\theta+1/2]}(x)$$

and the density of the double exponential distribution

$$f(x\,:\,\theta) = 1/2\exp(-|x-\theta|)$$

are two examples of the non-regular type of distributions for which we can apply the inequality (2.6).

In order to obtain an analogue of Cramer-Rao type inequality for the median-unbiased estimators we now introduce the following regularity conditions on the density $f(x;\theta)$:

A: (L^1 - differentiability) There exists a function $f'(x;\theta)$ such that $\int |f'(x;\theta)|d\mu < \infty$ and

$$\int |f(x;\theta+\Delta) - f(x;\theta) - f'(x;\theta)\Delta|d\mu = o(\Delta) \quad as \quad \Delta \to 0. \tag{2.7}$$

Remark 4. A simple sufficient condition for (2.7) is
$\mathbf{A_1}$: For every x , $\partial f(x;\theta)/\partial\theta$ exists and is continuous function of θ in Θ.
$\mathbf{A_2}$: For fixed θ_0, there exists a neighborhood $N(\theta_0)$ of θ_0 and non-negative function $G(x)$ such that

$$\sup_{\theta\in N} |(\partial/\partial\theta)f(x;\theta)| \le G(x) \quad for \;\; all \;\; x$$

and $\int G(x)d\mu < \infty$.

Under the regularity condition **A** , we have the following analogue of the Cramer-Rao type inequality for the median-unbiased estimators.

Theorem 2. Let $g(\theta)$ be a real-valued differentiable function on Θ. Let $\delta(X)$ be a median-unbiased of $g(\theta)$ having a density function $f_\delta(\cdot;\theta)$ which is continuous at $g(\theta)$. Then under the regularity condition **A** , we have :

$$[2f_\delta(g(\theta)\,;\,\theta)]^{-1} \geq |g'(\theta)| \,/\, I_1^*(\theta) \tag{2.8}$$

where $I_1^*(\theta) = \inf_k \int |f'(x;\theta) - kf(x;\theta)| d\mu$.

Moreover if the support of the $f(x;\theta)$ does not depend on θ , we can write $I_1^*(\theta)$ in the form

$$\begin{aligned} I_1^*(\theta) &= \inf_k E_\theta \,|\, f'(X;\theta)/f(X;\theta) - k| \\ &= E_\theta \,|\, f'(X;\theta)/f(X;\theta) - k(\theta)| \end{aligned} \tag{2.9}$$

where $k(\theta) = \text{median}_\theta[f'(X;\theta)/f(X;\theta)]$.

Proof. By the regularity condition **A** , the denominator of the right hand side of (2.6) has the limit which is the same as the denominator of the (2.8) for fixed k.

Remark 5. Most of previous lower bounds for the median-unbiased estimators ignored the possibility of improving the lower bound by centering the score function by its *median* instead of centering by *mean* which is identically zero. This is responsible for the limited applicability of most of previous lower bound in the non-symmetric family of density functions.

3. OPTIMALITY CONDITION

In this section we will identify the necessary and sufficient condition for the attainability of the lower bound (2.8) and then show that this bound is attained in two most important families of density functions which include the continuous *exponential family* and the *strongly unimodal* location family.

First suppose that the regularity condition **A** is satisfied by the family of density functions $\{f(x;\theta), \theta \in \Theta\}$. Then we define a median-unbaised estimator $\delta(X)$ of $g(\theta)$ to be *optimal* if

$$[2f_\delta(g(\theta)\,;\,\theta)]^{-1} = |g'(\theta)|/I_1^*(\theta) \quad \textit{for all } \theta.$$

As a first step for finding the optimal median-unbiasd estimator, we first characterize the necessary and sufficient condition for the attainability of the bound

(2.8).

Theorem 3. Let the conditions of the theorem 2 be satisfied by the family of density functions $\{f(x;\theta), \theta \in \Theta\}$. Then the following identity holds :

$$2f_\delta(g(\theta);\theta)|\,g'(\theta)\,| = \int |f'(x;\theta) - kfx;\theta)\,|s_1\, s_2\, s_3\, d\mu$$

where $s_1 = sgn[f'(x;\theta) - kf(x;\theta)]$, $s_2 = sgn[\delta(X) - g(\theta)]$ and $s_3 = sgn[g'(\theta)]$.

Moreover a median-unbiased estimator $\delta(X)$ is *optimal* if and only if for some k the identity

$$sgn\,[\,f'(x;\theta) - kf(x;\theta)\,] = sgn\,[\,\delta(X) - g(\theta)\,]\,sgn[\,g'(\theta)\,] \quad holds \quad a.e. \;\; \mu^* \qquad (3.1)$$

where $d\mu^*/d\mu = |f'(x;\theta) - kf(x;\theta)|$.

Proof. Applying L^1- differentiability to (2.5), we get immediately :

$$2\,f_\delta(g(\theta)\,;\,\theta)\,g'(\theta) = \int [\,f'(x;\theta) - kf(x;\theta]s_2\,d\mu. \qquad (3.2)$$

Then using the identity $x = |x|sgn(x)$, we can write (3.2) as

$$2\,f_\delta(g(\theta;\theta)\,|\,g'(\theta)| = \int |f'(x;\theta) - kf(x;\theta)\,|\,s_1\,s_2\,s_3\,d\mu = \int s_1\,s_2\,s_3\,d\mu^*.$$

This completes the proof because $s_1 s_2 s_3 \equiv 1$ holds if and only if $s_1 = s_2 s_3$ *a.e.* with respect to μ^*.

Remark 6. If $f(x;\theta)$ has the common support, then the optimality condition (3.1) is equivalent to the condition :

$$sgn[\,f'(x;\theta)/f(x;\theta) - k] = sgn[\delta(X) - g(\theta)]sgn[g'(\theta)] \quad holds \quad a.e. \;\mu^* \qquad (3.3)$$

and k is a median of the score function $f'(x;\theta)/f(x;\theta)$. If $f'(x;\theta)/f(x;\theta)$ has no atom at its median $k(\theta)$, then μ^* and μ are equivalent measures and *a.e.* μ^* can be replaced by *a.e.* μ in (3.3).

In general it is not trivial to find a family of density functions which satisfies the optimality condition (3.1) for some median-unbiased estimator. But we show that these optimality conditions are satisfied in the two important classes of density functions $f(x;\theta)$ which include the *exponential class* and the *strongly unimodal* location family respectively.

4. EXAMPLES

We first prove that optimality condition holds for the family of density functions with *monotone likelihood ratio* property.

Theorem 4. Let the family of density functions $\{f(x;\theta), \theta \in \Theta\}$ satisfy the monotone likelihood ratio property in $T(X)$ and let $T(X)$ have the density function having strictly increasing continuous score function. Assume also that the distribution function $F_\theta(t)$ of $T(X)$ is a continuous function of θ for fixed t . Let the regularity conditions $\mathbf{A_1}$ and $\mathbf{A_2}$ be satisfied by $f(x;\theta)$. Then there exists an optimal median-unbiased estimator of θ.

Proof. Monotone likelihood ratio property implies that the score function of X is the same as that of $T(X)$ which is assumed to be strictly increasing function of $T(X)$. By the same argument as in the Corollary 3 in p91 of Lehmann (1986), $m(\theta) = median_\theta[T(X)]$ is an strictly increasing function of θ. If we define $\delta(X) = m^{-1}(T(X))$, then $\delta(X)$ is an optimal median-unbiased estimator of θ which satisfies the optimality condition .

Remark 7. The most important family of densities $f(x;\theta)$ which has the monotone likelihood ratio property is the exponential class of continuous densities of the form :

$$f(x\,;\,\theta) = \exp[\,x\theta - b(\theta)]\,h(x)$$

for some functions $b(\cdot)$ and $h(x)$. As a specific example which demonstrates the difference between our result and previous bound (1.4), we consider the problem of estimating the mean θ of the exponential distribution $f(x;\theta) = \theta^{-1}exp(-x/\theta)$, $x > 0$, $\theta > 0$ from the random sample $X_1, \cdots, X_n$ of size n. Then we can easily show that the estimator $\delta(X) = \sum_{i=1}^{n} X_i/C_n$ is an optimal median-unbiased estimator of θ which attains the bound (1.5) but not the previous bound (1.4) which is strictly less than the best possible bound (1.5).

Here C_n denotes the median of the gamma distribution with shape parameter n and scale parameter 1. Following table provides typical Monte Carlo simulation results for the comparison of several information bounds for some values of n with $\theta = 1$.

n	C_n	$2f_\delta(\theta;\theta)$	$I_1^*(\theta)$	$I_1(\theta)$	$\sqrt{I_2(\theta)}$
1	.693	.69	.69	.75	1.00
2	1.68	1.05	1.05	1.09	1.41
3	2.67	1.32	1.32	1.34	1.73
10	9.67	2.49	2.49	2.52	3.16

Note that the median-unbiased estimator $\delta(X)$ does not attain the information bound $I_1(\theta)$ but attains the new information bound $I_1^*(\theta)$.

Next we consider the location family of density functions $f(x - \theta)$ and provide sufficient conditions for the existence of optimal median-unbiased estimator of θ.

Theorem 5. Let the random vector $X = (X_1, \cdots, X_n)$ have density function of the form $f(x - \bar{\theta})$ where $\bar{\theta} = (\theta, \cdots, \theta) \in R^n$ and $log f(x - \bar{\theta})$ is *strictly concave* function of $\theta \in R$. Assume also that the regularity conditions $\mathbf{A_1}$ and $\mathbf{A_2}$ are satisfied. Then there exists an optimal median-unbiased estimator $\delta(X)$ of θ.

Proof. Let k be the median of the score function which is independent of θ. Let $\delta(X)$ be the unique solution of the equation

$$\partial/\partial\theta\,[\log\, f(x - \bar{\theta})\,] = k.$$

Then the monotonicity of the score function implies that $\delta(X) > \theta$ if and only if $\partial/\partial\theta[log f(x - \bar{\theta})] > k$. Therefore $\delta(X)$ is an optimal median-unbiased estimator of θ by (3.4).

Remark 8. Let $X = (X_1, \cdots, X_n)$ be a random sample from a *symmetric strongly unimodal* density $f(x - \theta), \theta \in R$. Then optimal median-unbiased estimator $\delta(X)$ of θ is the unique *maximum likelihood estimator* (MLE) as is first noted by Stangenhaus and David (1978). But we note that our results provide complete answers to the *asymmetric* distrbutions as well as the symmetric distributions as long as they are strongly unimodal.

Remark 9. As an important special case of above result, we can mention the application of above result to the problem of analysis of the *type II censored* data with strong unimodal density $f(x-\theta)$.

Remark 10. By the useful *invariance* property of the optimal median-unbiased estimator, above result can be easily applied to the *scale* family of densities of the form :

$$f(x\,;\,\theta)=f(x_1/\theta,\cdots,x_n/\theta)\,/\,\theta^n, \qquad \theta>0$$

if it can be transformed to the location family with strongly unimodal density function by the *log* transformation.

REFERENCES

(1) Bejar Alamo, J. (1964). Estimacion en Mediana. *Trabajos de Estadistica* , 15. 93-102.

(2) Lehmann, E.L. (1986) . *Testing Statistical Hypotheses* , 2nd ed., Wiley , New York.

(3) Pfanzagl. J. (1970). Median Unbiased Estimates for m.l.r. Families. *Metrika* , 15, 30-39.

(4) Stangenhaus, G. (1977). *Optimal Estimation under Generalized Unbiasedness* , Unpublished Ph.D. dissertation, Department of Statitics, Iowa State University, Ames, Iowa.

(5) Stangenhaus, G. and David, H.T. (1978). Estimacao Nao Tendenciosa Mediana. *Proceedings of the 3-rd Sinape San Paulo Brasil* ,99-102.

(6) Sung, N.K. , Stangenhaus, G. and David, H.T. (1990). A Cramer-Rao Analogue for Median-Unbiased Estimators. *Trabajos de esadistica* , Vol. 5, 89-94.