A Bayesian view of the Hammersley–Chapman–Robbins type inequality

M. AKAHIRA and N. OHYAUCHI*

Institute of Mathematics, University of Tsukuba, Ibaraki 305–8571, Japan

* E-mail: naopu@math.tsukuba.ac.jp

From the Bayesian viewpoint, the information inequality applicable to the non-regular case is discussed. It is shown to construct an estimator which minimizes locally the variance of any estimator satisfying weaker conditions than the unbiasedness condition, from which an information inequality is derived. The Hammersley-Chapman-Robbins inequality is also obtained as a special case of the inequality. An example is also given.

Keywords: Information inequality; Non-regular case; Unbiasedness condition; Uniform distribution

AMS Subject Classification: 62F10; 62F15; 62F30; 62F12

1 Introduction

There are various information inequalities in statistical estimation. For example, the Cramér-Rao inequality, the Bhattacharyya bound, etc. are well known as the fact that the variance of all unbiased estimators can not be smaller than the lower bound under suitable regularity conditions. On the other hand, in non-regular cases when the regularity conditions do not necessarily hold, the Hammersley-Chapman-Robbins inequality is known and plays a role in estimation [1–3]. Akahira and Takeuchi [4] also consider a one-directinal family of distributions with a parameter for which the support moves in the one direction (see also [5, 6]). And they show that the infimum of the bound for the variance of unbiased estimators is equal to zero at any specific point of the parameter (see also [7]). Further, in the monograph of Akahira and Takeuchi [7], the meanings and implications of regularity conditions are given as systematically as possible. A lower bound for the convex combination of the variances of all unbiased estimators at arbitrary two points of a parameter space is obtained by Vincze [8], using the Cramér-Rao inequality. Recently, the bound is also derived directly by Akahira and Ohyauchi [9] and

Ohyauchi [10], using the Lagrange method.

In this paper, we consider the information inequality from the Bayesian viewpoint. It is shown that an estimator minimizing locally the variance of any estimator with weaker conditions than the unbiasedness is constructed. And also the lower bound for the variance of estimators can be expressed by an information inequality. It is noted that the bound is global. If some special prior measure is chosen, the Hammersley-Chapman-Robbins inequality is shown to be represented as a special case of the inequality.

2 An information inequality

Let $\boldsymbol{X}$ be a real random vector with a joint probability density function (j.p.d.f.) $f_{\boldsymbol{X}}(\boldsymbol{x},\theta)$ with respect to (w.r.t.) a σ-finite measure μ, where $\theta \in \Theta$ and Θ is an open interval of $\boldsymbol{R}^1$. Let $g(\theta)$ be a real-valued function on Θ and $\mathcal{X}$ a sample space of $\boldsymbol{X}$. Suppose that an estimator $\hat{g} = \hat{g}(\boldsymbol{X})$ of $g(\theta)$ satisfies the condition

(A1) $$E_{\theta_0}(\hat{g}) = \int_{\mathcal{X}} \hat{g}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0) d\mu(\boldsymbol{x}) = g(\theta_0) =: g_0.$$

Let Π be a prior probability measure on Θ. Define

$$h_{\Pi}(\boldsymbol{x}) := \int_{\Theta} f_{\boldsymbol{X}}(\boldsymbol{x},\theta) d\Pi(\theta), \qquad \eta := \int_{\Theta} g(\theta) d\Pi(\theta).$$

We also have the following conditions:

(A2) $$\int_{-\infty}^{\infty} \hat{g}(\boldsymbol{x}) h_{\Pi}(\boldsymbol{x}) d\mu(\boldsymbol{x}) = \eta.$$

(A3) $$0 < J_{\Pi}(\theta_0) := E_{\theta_0}\left[\left\{\frac{h_{\Pi}(\boldsymbol{X})}{f_{\boldsymbol{X}}(\boldsymbol{X},\theta_0)}\right\}^2\right] - 1 < \infty.$$

Remark 1 If $\hat{g}(\boldsymbol{X})$ is unbiased for $g(\theta)$, then (A2) holds, since

$$\int_{\mathcal{X}} \hat{g}(\boldsymbol{x}) h_{\Pi}(\boldsymbol{x}) d\mu(\boldsymbol{x}) = \int_{\Theta} \left\{\int_{\mathcal{X}} \hat{g}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x},\theta) d\mu(\boldsymbol{x})\right\} d\Pi(\theta) = \int_{\Theta} g(\theta) d\Pi(\theta) = \eta.$$

THEOREM 2 *Under the conditions* (A1)–(A3), *there exists an estimator which minimizes the variance, i.e.*

$$\min_{\hat{g}:(A1),(A2)} V_{\theta_0}(\hat{g}) = V_{\theta_0}(\hat{g}_{\Pi}^*),$$

where

$$\hat{g}^*_\Pi(\boldsymbol{x}) := g_0 + \frac{g_0 - \eta}{J_\Pi(\theta_0)} \left\{ 1 - \frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} \right\} \tag{1}$$

if $f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) > 0$, *and* $\hat{g}^*_\Pi(\boldsymbol{x}) = 0$, *otherwise.*

Proof Since

$$V_{\theta_0}(\hat{g}) = \int_{\mathcal{X}} \{\hat{g}(\boldsymbol{x})\}^2 f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) - g_0^2, \tag{2}$$

it is enough to obtain the estimator $\hat{g}$ minimizing the first term of the right-hand side of (2). As in the Lagrange method, we start with the expression

$$\begin{aligned} F_{\hat{g}}(\lambda_1, \lambda_2) := &\int_{\mathcal{X}} \{\hat{g}(\boldsymbol{x})\}^2 f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) - \lambda_1 \left\{ \int_{\mathcal{X}} \hat{g}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) - g_0 \right\} \\ &- \lambda_2 \left\{ \int_{\mathcal{X}} \hat{g}(\boldsymbol{x}) h_\Pi(\boldsymbol{x}) d\mu(\boldsymbol{x}) - \eta \right\}. \end{aligned} \tag{3}$$

Then, we get the estimator $\hat{g}$ which minimizes $F_{\hat{g}}(\lambda_1, \lambda_2)$, since $\hat{g}$ satisfies the conditions (A1) and (A2). Since

$$\begin{aligned} F_{\hat{g}}(\lambda_1, \lambda_2) &= \int_{\mathcal{X}} \left[\{\hat{g}(\boldsymbol{x})\}^2 - \lambda_1 \hat{g}(\boldsymbol{x}) - \lambda_2 \frac{\hat{g}(\boldsymbol{x}) h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} \right] f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) + \lambda_1 g_0 + \lambda_2 \eta \\ &= \int_{\mathcal{X}} \left\{ \hat{g}(\boldsymbol{x}) - \frac{1}{2} \left(\lambda_1 + \lambda_2 \frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} \right) \right\}^2 f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) \\ &\quad - \frac{1}{4} \int_{\mathcal{X}} \left\{ \lambda_1 + \lambda_2 \frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} \right\}^2 f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) + \lambda_1 g_0 + \lambda_2 \eta, \end{aligned}$$

it follows that the estimator minimizing $F_{\hat{g}}(\lambda_1, \lambda_2)$ is of the form

$$\hat{g}_\Pi(\boldsymbol{x}) := \begin{cases} \frac{1}{2} \left\{ \lambda_1 + \lambda_2 \frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} \right\} & \text{if } f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

From (A1) we have

$$g_0 = \int_{\mathcal{X}} \hat{g}_\Pi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0) d\mu(\boldsymbol{x}) = \frac{\lambda_1}{2} + \frac{\lambda_2}{2} \int_{\mathcal{X}} h_\Pi(\boldsymbol{x}) d\mu(\boldsymbol{x}) = \frac{1}{2}(\lambda_1 + \lambda_2), \tag{5}$$

and from (A2)

$$\eta = \int_{\mathcal{X}} \hat{g}_\Pi(\boldsymbol{x}) h_\Pi(\boldsymbol{x}) d\mu(\boldsymbol{x}) = \frac{\lambda_1}{2} \int_{\mathcal{X}} h_\Pi(\boldsymbol{x}) d\mu(\boldsymbol{x}) + \frac{\lambda_2}{2} \int_{\mathcal{X}} \frac{\{h_\Pi(\boldsymbol{x})\}^2}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} d\mu(\boldsymbol{x}). \tag{6}$$

Subtracting both sides of (6) from (5), we have

$$\begin{aligned} g_0 - \eta &= -\frac{\lambda_2}{2}\left[\int_{\mathcal{X}}\left\{\frac{(h_\Pi(\boldsymbol{x}))^2}{(f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0))^2} - 1\right\} f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0)d\mu(\boldsymbol{x})\right] \\ &= -\frac{\lambda_2}{2}E_{\theta_0}\left[\left\{\frac{h_\Pi(\boldsymbol{X})}{f_{\boldsymbol{X}}(\boldsymbol{X},\theta_0)}\right\}^2 - 1\right] \\ &=: -\frac{\lambda_2}{2}J_\Pi(\theta_0) \quad \text{(say)}, \end{aligned}$$

which implies

$$\lambda_2 = -\frac{2(g_0-\eta)}{J_\Pi(\theta_0)} =: \lambda_2^* \quad \text{(say)}. \tag{7}$$

From equations (5) and (6), we obtain

$$\lambda_1 = 2g_0 - \lambda_2 = 2\left\{g_0 + \frac{g_0-\eta}{J_\Pi(\theta_0)}\right\} =: \lambda_1^* \quad \text{(say)}. \tag{8}$$

Letting λ_1^* in (8) and λ_2^* in (7) as λ_1 and λ_2 in (4), respectively, we have

$$\hat{g}_\Pi^*(\boldsymbol{x}) = \frac{1}{2}\left\{\lambda_1^* + \lambda_2^*\frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0)}\right\} = g_0 + \frac{g_0-\eta}{J_\Pi(\theta_0)}\left\{1 - \frac{h_\Pi(\boldsymbol{x})}{f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0)}\right\},$$

if $f_{\boldsymbol{X}}(\boldsymbol{x},\theta_0) > 0$. This completes the proof. □

From (1), it follows that the variance of $\hat{g}_\Pi^*$ is given by

$$V_{\theta_0}(\hat{g}_\Pi^*) = \frac{(g_0-\eta)^2}{\{J_\Pi(\theta_0)\}^2}E_{\theta_0}\left[\left\{1 - \frac{h_\Pi(\boldsymbol{X})}{f_{\boldsymbol{X}}(\boldsymbol{X},\theta_0)}\right\}^2\right] = \frac{(g_0-\eta)^2}{J_\Pi(\theta_0)}. \tag{9}$$

Then, we have the following.

COROLLARY 3 *For any estimator $\hat{g}$ satisfying the conditions* (A1) *and* (A2), *it holds that*

$$V_{\theta_0}(\hat{g}) \geq \frac{(g_0-\eta)^2}{J_\Pi(\theta_0)}. \tag{10}$$

Also,

$$\min_{\hat{g}:\text{unbiased}} V_{\theta_0}(\hat{g}) \geq \frac{(g_0-\eta)^2}{J_\Pi(\theta_0)}. \tag{11}$$

The proof of inequality (10) is straightforward from theorem 2 and equation (9) and that of (11) follows from the fact that an unbiased estimator of θ satisfies (A1) and (A2).

Remark 4 As is seen in theorem 2, the lower bound due to inequality (10) is sharp. But, the lower bound due to (11) is not generally sharp. Indeed, the totality of all the unbiased estimators of $g(\theta)$ is included in that of all the estimators satisfying (A1) and (A2). Since the difference between both sides of inequality (11) is larger than that of (10), the above arises.

3 The Hammersley–Chapman–Robbins type inequality

Let $\theta_1 \in \Theta$ with $\theta_1 \neq \theta_0$, and take Π_1 such that $\Pi_1(\{\theta_1\}) = 1$ as the prior measure Π. Then,

$$h_{\Pi_1}(\boldsymbol{x}) = \int_\Theta f_{\boldsymbol{X}}(\boldsymbol{x}, \theta) d\Pi_1(\theta) = f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_1),$$

and

$$\eta = \int_\Theta g(\theta) d\Pi_1(\theta) = g(\theta_1) =: g_1,$$

which implies

$$J_{\Pi_1}(\theta_0) = E_{\theta_0}\left[\left\{\frac{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_0)} - 1\right\}^2\right]. \tag{12}$$

From inequalities (10) and (11), we have for any estimator $\hat{g}$ satisfying (A1) and (A2)

$$V_{\theta_0}(\hat{g}) \geq \frac{(g_0 - g_1)^2}{J_{\Pi_1}(\theta_0)} = \frac{(g_1 - g_0)^2}{E_{\theta_0}\left[\{(f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_1)/f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_0)) - 1\}^2\right]}, \tag{13}$$

which is the Hammersley-Chapman-Robbins (H-C-R)-type inequality. In relation to the above, an extension of inequality (13) from the non-Bayesian approach is also considered by Koike and Komatsu [11]. Now we represent $J_{\Pi_1}(\theta_0)$ as

$$J_n(\theta_0, \theta_1) := J_{\Pi_1}(\theta_0) = E_{\theta_0}\left[\left\{\frac{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_0)} - 1\right\}^2\right] = E_{\theta_0}\left[\left\{\frac{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_0)}\right\}^2\right] - 1. \tag{14}$$

We also define

$$I_n(\theta_0, \theta_1) = E_{\theta_0}\left[\left\{\frac{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_1)}{f_{\boldsymbol{X}}(\boldsymbol{X}, \theta_0)}\right\}^2\right] = \int_{\mathcal{X}} \frac{\{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_1)\}^2}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} d\mu(\boldsymbol{x}). \tag{15}$$

Suppose that $X_1, \cdots, X_n$ are independent and identically distributed (i.i.d.) real random variables with a p.d.f. $p(x, \theta)$ w.r.t. the Lebesgue measure. Then, the j.p.d.f. of $\boldsymbol{X} := (X_1, \cdots, X_n)$ is given by

$$f_{\boldsymbol{X}}(\boldsymbol{x}, \theta) := \prod_{i=1}^n p(x_i, \theta).$$

Since, by equations (14) and (15), $J_1(\theta_0, \theta_1) = I_1(\theta_0, \theta_1) - 1$, it follows that

$$\begin{aligned}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_1)\}^2}{f_{\boldsymbol{X}}(\boldsymbol{x}, \theta_0)} dx_1 \cdots dx_n &= \left[\int_{-\infty}^{\infty} \frac{\{p(x_1, \theta_1)\}^2}{p(x_1, \theta_0)} dx_1\right]^n \\ &= \{I_1(\theta_0, \theta_1)\}^n = \{1 + J_1(\theta_0, \theta_1)\}^n.\end{aligned}$$

From inequalities (11) and (13), it follows that for any unbiased estimator $\hat{\theta} = \hat{\theta}(\boldsymbol{X})$ of θ

$$\begin{aligned}V_{\theta_0}(\hat{\theta}) &\geq \frac{(\theta_1 - \theta_0)^2}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n (\{p(x_i, \theta_1)\}^2 / p(x_i, \theta_0)) dx_1 \cdots dx_n - 1} \\ &= \frac{(\theta_1 - \theta_0)^2}{\{1 + J_1(\theta_0, \theta_1)\}^n - 1}. \qquad (16)\end{aligned}$$

Here, note that the bigger the lower bound, the more the desire. Putting $\Delta := \theta_1 - \theta_0$, we have from inequality (16)

$$V_{\theta_0}(\hat{\theta}) \geq \sup_{\Delta:|\Delta|>0} \frac{\Delta^2}{\{1 + J_1(\theta_0, \theta_0 + \Delta)\}^n - 1} =: \sup_{\Delta:|\Delta|>0} B(\Delta) \qquad (17)$$

For a fixed Δ, we obtain for large n

$$\begin{aligned}\{1 + J_1(\theta_0, \theta_0 + \Delta)\}^n &= \left\{1 + \frac{1}{n} \cdot n J_1(\theta_0, \theta_0 + \Delta)\right\}^n \\ &\approx e^{nJ(\theta_0, \theta_0 + \Delta)},\end{aligned}$$

which implies

$$\begin{aligned}\sup_{\Delta:|\Delta|>0} B(\Delta) &= \sup_{\Delta:|\Delta|>0} \frac{\Delta^2}{\{1 + J_1(\theta_0, \theta_0 + \Delta)\}^n - 1} \\ &\approx \sup_{\Delta:|\Delta|>0} \frac{\Delta^2}{e^{nJ_1(\theta_0, \theta_0 + \Delta)} - 1}. \qquad (18)\end{aligned}$$

Example 5 Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables with a p.d.f.

$$p(x, \theta) = \begin{cases} \dfrac{1}{\theta} & \text{for } 0 < x < \theta, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$. Let θ_0 be fixed in $\boldsymbol{R}_+ := (0, \infty)$ and $0 < \theta_1 < \theta_0$. Since

$$\begin{aligned}J_1(\theta_0, \theta_1) &= E_{\theta_0}\left[\left\{\frac{p(X_1, \theta_1)}{p(X_1, \theta_0)} - 1\right\}^2\right] \\ &= \int_0^{\theta_1} \theta_0 \left(\frac{1}{\theta_1} - \frac{1}{\theta_0}\right)^2 dx + \int_{\theta_1}^{\theta_0} \theta_0 \cdot \frac{1}{\theta_0^2} dx\end{aligned}$$

$$= \frac{\theta_0 - \theta_1}{\theta_1}, \tag{19}$$

it follows from inequality (16) that for any unbiased estimator $\hat{\theta}$ of θ

$$V_{\theta_0}(\hat{\theta}) \geq \left[\inf_{0<\theta_1<\theta_0} \frac{\{1 + J_1(\theta_0, \theta_1)\}^n - 1}{(\theta_1 - \theta_0)^2} \right]^{-1}. \tag{20}$$

Then,

$$\begin{aligned} \inf_{0<\theta_1<\theta_0} \frac{\{1 + J_1(\theta_0, \theta_1)\}^n - 1}{(\theta_1 - \theta_0)^2} &= \inf_{0<\theta_1<\theta_0} \frac{(\theta_0/\theta_1)^n - 1}{(\theta_1 - \theta_0)^2} = \frac{1}{\theta_0^2} \inf_{0<\theta_1<\theta_0} \frac{(\theta_0/\theta_1)^n - 1}{(1 - (\theta_1/\theta_0))^2} \\ &= \frac{1}{\theta_0^2} \inf_{\xi>1} \frac{\xi^2(\xi^n - 1)}{(\xi - 1)^2} = \frac{n^2}{\theta_0^2} \inf_{\xi>1} \frac{\xi^2(\xi^n - 1)}{n^2(\xi - 1)^2}. \end{aligned} \tag{21}$$

Letting $h(\xi) = \xi^2(\xi^n - 1)/\{n(\xi - 1)\}^2$ for $\xi > 1$, from inequality (20) and equation (21), we have for any unbiased estimator $\hat{\theta}$ of θ

$$V_{\theta_0}(\hat{\theta}) \geq \frac{\theta_0^2}{n^2} \left\{ \inf_{\xi>1} h(\xi) \right\}^{-1}. \tag{22}$$

Since

$$h'(\xi) = \frac{\xi}{n^2(\xi - 1)^3} \{n\xi^{n+1} - (n + 2)\xi^n + 2\}$$

for $\xi > 1$, the value of $\{\inf_{\xi>1} h(\xi)\}^{-1}$ for given n is obtained in table 1.

Table 1. The values of $\{\inf_{\xi>1} h(\xi)\}^{-1}$.

n	1	5	10	25	50	100	500	1000	∞
$\{\inf_{\xi>1} h(\xi)\}^{-1}$	0.25	0.4912	0.5586	0.6088	0.6276	0.6375	0.6456	0.6466	0.6476

Letting $\Delta = \theta_1 - \theta_0$, we have from equation (19)

$$J_1(\theta_0, \theta_0 + \Delta) = -\frac{\Delta}{\theta_0 + \Delta}.$$

Putting $t := n\Delta$, we obtain for large n

$$\frac{\Delta^2}{e^{nJ_1(\theta_0, \theta_0+\Delta)} - 1} = \frac{\Delta^2}{e^{-n\Delta/(\theta_0+\Delta)} - 1} \approx \frac{t^2}{n^2(e^{-t/\theta_0} - 1)} = \frac{\theta_0^2(t/\theta_0)^2}{n^2(e^{-t/\theta_0} - 1)}.$$

Let

$$h(x) := \frac{x^2}{e^x - 1} \qquad \text{for} \quad x > 0.$$

Then $h(x)$ has the maximum value 0.6476 at $x \fallingdotseq 1.5936$, hence it follows from equation (18) that for large n

$$\sup_{\Delta:\Delta<0} B(\Delta) = \sup_{\Delta:\Delta<0} \frac{\Delta^2}{e^{nJ_1(\theta_0, \theta_0+\Delta)} - 1} \approx \frac{0.6476\theta_0^2}{n^2},$$

which coincides with the value obtained by Kiefer [3]. Now let $X_{(n)} := \max_{1\le i\le n} X_i$. Since $X_{(n)}$ is a complete sufficient statistic, the estimator

$$\hat{\theta}^* = \frac{n+1}{n} X_{(n)}$$

is uniformly minimum variance unbiased (UMVU) for θ. Then the variance of $\hat{\theta}^*$ is

$$V_{\theta_0}(\hat{\theta}^*) = \frac{1}{n(n+2)}\theta_0^2 = \frac{\theta_0^2}{n^2} + o\left(\frac{1}{n^2}\right). \tag{23}$$

From inequality (22) and equation (23), we have for large n

$$V_{\theta_0}(\hat{\theta}^*) = \frac{\theta_0^2}{n^2} + o\left(\frac{1}{n^2}\right) > 0.6476\frac{\theta_0^2}{n^2} + o\left(\frac{1}{n^2}\right),$$

hence the H-C-R lower bound can not be asymptotically attained by $\hat{\theta}^*$ up to the order $o(1/n^2)$.

4 Conclusion

The lower bound due to the H-C-R inequality is not generally attainable. But, as is seen in Section 2 the lower bound for the variance of estimators in some class, derived from the Bayesian viewpoint, is obtained and it is attained. As a special case of the information inequality giving the lower bound, we get the H-C-R inequality. In the example of the uniform distribution on an interval $(0, \theta)$, the H-C-R lower bound for the variance of all unbiased estimators is obtained for given n, and it is shown to be asymptotically unattainable from the comparison with the variance of the UMVU estimator. Finally, under suitable regularity conditions, the Cramér-Rao inequality can be derived from inequality (17).

Acknowledgements

The authors wish to thank the referees for their comments.

References

[1] Hammersley, J. M., 1950, On estimating restricted parameters. *Journal of the Royal Statistical Society, Series B*, **12**, 192–240.

[2] Chapman, D. G. and Robbins, H., 1951, Minimum variance estimation without regularity assumptions. *Annals of Mathematical Statistics*, **22**, 581–586.

[3] Kiefer, J., 1952, On minimum variance estimators. *Annals of Mathematical Statistics*, **23**, 627–629.

[4] Akahira, M. and Takeuchi, K., 1987, The lower bound for the variance of unbiased estimators for one-directional family of distributions. *Annals of the Institute of Statistical Mathematics*, **39**, Part A, 593–610.

[5] Akahira, M., 1991, Second order asymptotic sufficiency for a family of distributions with one-directionality. *Metron*, **49**, 133–143.

[6] Akahira, M., 1993, Asymptotics on the statistics for a family of non-regular distributions. In: Matusita *et al.* (Eds) *Statistical Science and Data Analysis* (Zeist: VSP), pp. 357–364.

[7] Akahira, M. and Takeuchi, K., 1995, Non-regular statistical estimation. *Lecture Notes in Statistics*, Vol. 107 (New York: Springer).

[8] Vincze, I., 1992, On nonparametric Cramér-Rao inequalities. In: P. K. Sen and I. A. Salama (Eds) *Order Statistics and Nonparametrics: Theory and Applications* (Amsterdam: North-Holland), pp. 439–454.

[9] Akahira, M. and Ohyauchi, N., 2003, An information inequality for the Bayes risk applicable to non-regular cases. Paper presented at the *Proceedings of the Symposium of Research Institute of Mathematical Science*, No. 1334, Kyoto University, 183–191.

[10] Ohyauchi, N., 2004, The Vincze inequality for the Bayes risk. *Journal of the Japan Statistical Society*, **34**, 65–74.

[11] Koike, K. and Komatsu, K., 2006, Comparison of some bounds in estimation theory. Unpublised.