Arab Gulf J. scient. Res. **2** (2), pp. 583-591 (1984)

On Bayesian Analogues to Bhattacharyya's Lower Bounds

M. Luisa Targhetta

Mathematical Institute, University of Cagliari, Italy

ABSTRACT. An inequality similar to the Bhattacharyya inequality is obtained in the presence of prior knowledge. Conditions under which equality holds are considered assuming a posterior density of the exponential type.

1. Introduction

Let $f(x|\theta)$ be a probability density function depending on an unknown parameter $\theta \in S$, where $S = (a,b)$ is an open interval, finite or not, on the real line.

Let $g(\theta)$ be a proper prior density for θ. Denote by $g(\theta|x)$ and $f(x)$ the posterior density function and the marginal density function of x, respectively.

Further, $E_{x|\theta}$, $E_{\theta|x}$, E_x and E_θ denote the expectations with respect to $f(x|\theta)$, $g(\theta|x)$, $f(x)$ and $g(\theta)$, respectively.

Let $\hat{\theta}(x)$ be an estimator of θ and let

$$B = \lim_{\theta \to b^-} g(\theta) E_{x|\theta}(\hat{\theta} - \theta) - \lim_{\theta \to a^+} g(\theta) E_{x|\theta}(\hat{\theta} - \theta). \tag{1}$$

In a Bayesian context, Ferreira (1981) has generalized the Cramér-Rao inequality as follows

$$E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}] \geq \frac{(1 + B)^2}{E_x\left[E_{\theta|x}\left\{\left(\frac{\partial}{\partial \theta} \log g(\theta|x)\right)^2\right\}\right]}, \tag{2}$$

where

$$I = E_x\left[E_{\theta|x}\left\{\left(\frac{\partial}{\partial\theta}\log g(\theta|x)\right)^2\right\}\right]$$

is defined as the amount of information about θ contained in the joint distribution of (x,θ).

In the present paper, we will find greater lower bounds than the right hand side of (2) for $E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}]$ in cases where $(1 + B)^2/I$ is not attainable, so extending the Bhattacharyya system of lower bounds (Kendall and Stuart 1961, Zacks 1971).

2. Bayesian Lower Bounds

In the following, we shall always require regularity conditions as differentiability and integrability; further, that the operation of differentiating under the integral sign and Fubini's theorem hold.

Consider

$$D_s = \hat{\theta} - \theta - \sum_{r=1}^{s} a_r \frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial\theta^r}, \tag{3}$$

where $a_1, \ldots, a_s$ are constants to be determined so that

$$E_\theta[E_{x|\theta}(D_s^2)] = \int g(\theta)d\theta \int \left\{\hat{\theta} - \theta - \sum_{r=1}^{s} a_r \frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial\theta^r}\right\}^2 f(x|\theta)dx \tag{4}$$

is a minimum.

Differentiating (4) with respect to a_j and equating the result to zero, we obtain

$$\int g(\theta)d\theta \int \left\{\hat{\theta} - \theta - \sum_{r=1}^{s} a_r \frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial\theta^r}\right\} \frac{1}{g(\theta|x)} \frac{\partial^j g(\theta|x)}{\partial\theta^j} f(x|\theta)dx = 0 \tag{5}$$

$$j = 1, \ldots, s.$$

Putting

$$\tau^{(j)} = \int g(\theta)d\theta \int (\hat{\theta} - \theta) \frac{1}{g(\theta|x)} \frac{\partial^j g(\theta|x)}{\partial\theta^j} f(x|\theta)dx \qquad j = 1, \ldots, s \tag{6}$$

$$K_{rj} = \int g(\theta)d\theta \int \frac{1}{g^2(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial\theta^r} \frac{\partial^j g(\theta|x)}{\partial\theta^j} f(x|\theta)dx \qquad r,j = 1, \ldots, s, \tag{7}$$

(5) can be written as

$$\tau^{(j)} = \sum_{r=1}^{s} a_r K_{rj} \qquad j = 1, \ldots, s \tag{8}$$

Let K be the matrix whose element is K_{rj}, r,j = 1, ..., s. If K is non-singular, we may invert the system (8) of linear equations to obtain

$$a_r = \sum_{j=1}^{s} \tau^{(j)} K^{-1}_{rj}, \qquad r = 1, \ldots, s \tag{9}$$

Substituting (9) in (4) we have

$$E_\theta[E_{x|\theta}(D_s^2)] = E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}] - \sum_{r=1}^{s} \sum_{j=1}^{s} \tau^{(j)} \tau^{(r)} K^{-1}_{rj}.$$

Since its left hand side is non-negative, we finally obtain the required inequality

$$E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}] \geqslant \sum_{r=1}^{s} \sum_{j=1}^{s} \tau^{(j)} \tau^{(r)} K^{-1}_{rj}. \tag{10}$$

Note that

$$\min_{(a_1, \ldots, a_{s+1})} E_\theta[E_{x|\theta}(D^2_{s+1})] \leqslant \min_{(a_1, \ldots, a_s)} E_\theta[E_{x|\theta}(D_s^2)] \qquad \forall s \geqslant 1,$$

hence it follows that the right hand side of (10) is a non decreasing function of s.

For s = 1 (10) becomes

$$E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}] \geqslant \frac{\{\tau^{(1)}\}^2}{K_{11}} ;$$

moreover, we have from (6) and (7)

$$\tau^{(1)} = \int f(x)dx \int_a^b (\hat{\theta} - \theta) \frac{\partial g(\theta|x)}{\partial \theta} d\theta = 1 + B$$

$$K_{11} = I,$$

hence, for s = 1, (10) reduces to (2). This motivates us to state that the lower bounds on the right hand side of (10) are an extension of the Bhattacharyya system to the case of Bayesian estimation.

The condition for the bound in (10) to be attained is

$$\hat{\theta} - \theta = \sum_{r=1}^{s} \sum_{j=1}^{s} \tau^{(j)} K_{rj}^{-1} \frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial \theta^r}, \tag{11}$$

up to sets of probability zero with respect to $f(x|\theta)g(\theta)$.

In the particular case s = 1, it follows from (11)

$$\hat{\theta} - \theta = \frac{1 + B}{I} \frac{\partial}{\partial \theta} \log g(\theta|x), \tag{12}$$

that is, the equality holds if and only if $g(\theta|x) = N(\hat{\theta},(1 + B)/I)$.

The inequality (10) can be obtained too by introducing the (s + 1)-dimensional random vector

$$\left\{\hat{\theta} - \theta, \frac{1}{g(\theta|x)} \frac{\partial g(\theta|x)}{\partial \theta}, \ldots, \frac{1}{g(\theta|x)} \frac{\partial^s g(\theta|x)}{\partial \theta^s}\right\}$$

whose marginal random variables are functions of (x,θ). The non-negative definite matrix Σ of the second order moments about the origin of the vector and the non-negative definite matrix K above introduced are such that

$$\det(\textstyle\sum) = \det(K)\left\{E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}] - \sum_{r=1}^{s} \sum_{j=1}^{s} \tau^{(r)}\tau^{(j)} K_{rj}^{-1}\right\},$$

where det(A) denotes the determinant of the matrix A. If K is non-singular, the inequality (10) follows again.

3. Posterior Distribution of the Exponential Type

Let

$$g(\theta|x) = h(x)e^{\psi_1(\theta)t(x) + \psi_2(\theta)} \tag{13}$$

where $\psi_1'(\theta) \neq 0$ and t(x) is the minimal sufficient statistic.

Suppose that $\hat{\theta}$ does not verify (12). Then there is no value of $s > 1$ for which the equality (11) holds.

Proof. According to the above results, the s-*th* bound, but not the (s − 1)-*th* bound, is attained if and only if there exist constants $a_1, \ldots, a_s$ such that

$$\hat{\theta} - \theta = \sum_{r=1}^{s} a_r \frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial \theta^r}. \tag{14}$$

Differentiating (13) r times with respect to θ, we can write

$$\frac{1}{g(\theta|x)} \frac{\partial^r g(\theta|x)}{\partial \theta^r} = \{t(x)\psi_1'(\theta) + \psi_2'(\theta)\}^r + \sum_{h=0}^{r-1} t^h U_{rh}(\theta), \tag{15}$$

where $U_{rh}(\theta)$ are linear combinations of products of $\psi_1(\theta)$ and $\psi_2(\theta)$'s derivatives.

Substituting in (14), we find that $\hat{\theta}$ is a polynomial of degree s in t(x). But we now prove that a polynomial of degree $s > 1$ never verifies (14) if $g(\theta|x)$ is given by (13). Let

$$\hat{\theta}(x) = \sum_{i=0}^{s} c_i t^i(x) \tag{16}$$

where c_i $i = 0, \ldots, s$ are arbitrary constants.

Making use of (15), equate the coefficient of the r-*th* power of t(x) in (14) to that of the corresponding power of t(x) in (16) for every $r = 0, \ldots, s$.

We get

$$a_s = \{\psi_1'(\theta)\}^{-s} c_s$$

$$a_{s-1} = \{\psi_1'(\theta)\}^{-(s-1)}\{c_{s-1} - a_s[s\{\psi_1'(\theta)\}^{s-1}\psi_2'(\theta) + m_s\psi_1''(\theta)\{\psi_1'(\theta)\}^{s-2}]\}$$

$$a_k = \{\psi_1'(\theta)\}^{-k}\left[c_k - \sum_{j=k+1}^{s} a_j\{\binom{j}{k}\{\psi_1'(\theta)\}^k\{\psi_2'(\theta)\}^{j-k} + U_{jk}(\theta)\}\right]$$

$$k = 1, \ldots, s-2$$

$$\theta = c_0 - \sum_{j=1}^{s} a_j[\{\psi_2'(\theta)\}^j + U_{j0}(\theta)], \tag{17}$$

where m_s is a positive integer for $s > 1$ and is equal to zero for $s = 1$.

It appears evident from (17) that $a_1, \ldots, a_s$ may be constants only if $\psi'_1(\theta)$ and $\psi_2'(\theta)$ are independent on θ, in which case the last relation in (17) is not verified.

Hence, if $g(\theta|x)$ is of the exponential type, there exists no estimator $\hat{\theta}(x)$ such that $E_\theta[E_{x|\theta}\{(\hat{\theta} - \theta)^2\}]$ attains the s-*th* lower bound (10); but not the (s − 1)-*th* bound. (Q.E.D.)

For $s = 1$ it follows from (17)

$$a_1 = \{\psi_1'(\theta)\}^{-1}c_1$$

$$\theta = c_0 - a_1\psi_2'(\theta),$$

hence the equality (12) is possible only if $\psi_1'(\theta) = \text{constant} = k$, $\psi_2'(\theta) = \frac{k}{c_1}(c_0 - \theta)$.

Finally, we remark that in a non-Bayesian context, Fend's theorem (Fend 1959, Zacks 1971) states that the variance of an estimator which is a polynomial of degree s in the minimal sufficient statistic in an exponential family attains the s-*th* Bhattacharyya lower bound, but not the (s − 1)-*th* bound; and conversely.

4. A Decisional Bayesian Inequality

In a decisional Bayesian context, let $L(\theta,\hat{\theta})$: $S \times D \to R^+$ be the loss function which associates a real value ≥ 0 to every $(\theta \in S, \hat{\theta} \in D)$.

Suppose that regularity conditions on $L(\theta,\hat{\theta})$ and $\sqrt{L(\theta,\hat{\theta})}$ are verified.

Consider

$$E_x\left[E_{\theta|x}\left\{\frac{\partial\sqrt{L(\theta,\hat{\theta})}}{\partial\theta}\right\}\right] = \int f(x)dx \int_a^b \frac{\partial\sqrt{L(\theta,\hat{\theta})}}{\partial\theta} g(\theta|x)d\theta. \tag{18}$$

Following Ferreira (1981), that is integrating the right hand side of (18) by parts, then applying to the result the Cauchy-Schwarz inequality, we get

$$W \geqslant \frac{\left\{-E_x\left[E_{\theta|x}\left\{\frac{\partial\sqrt{L(\theta,\hat{\theta})}}{\partial\theta}\right\}\right] + B(\sqrt{L})\right\}^2}{I}, \tag{19}$$

where

$$W = E_x[E_{\theta|x}\{L(\theta,\hat{\theta})\}]$$

denotes the prior risk and

$$B(\sqrt{L}) = \lim_{\theta\to b^-} E_{x|\theta}\{\sqrt{L}\}g(\theta) - \lim_{\theta\to a^+} E_{x|\theta}\{\sqrt{L}\}g(\theta).$$

Convex loss functions of the following form

$$L(\theta,\hat{\theta}) = |\theta - \hat{\theta}|^m, \quad m > 0$$

are often of interest in statistical decision problems.

Assuming

$$\sqrt{L}(\theta,\hat{\theta}) = |\theta - \hat{\theta}|^{m/2}\text{sgn}(\hat{\theta} - \theta),$$

where

$$\text{sgn}\,y = \begin{cases} +1 & \text{if } y > 0 \\ -1 & \text{if } y < 0 \end{cases},$$

W attains the right hand side of (19) if and only if there exists a constant K such that

$$|\theta - \hat{\theta}|^{m/2}\text{sgn}(\hat{\theta} - \theta) = K\frac{\partial}{\partial\theta}\log g(\theta|x),$$

from which it follows

$$g(\theta|x)\ \alpha\ e^{-\frac{2}{K(m+2)}|\theta-\hat{\theta}|^{\frac{1}{2}m+1}}$$

In particular for $m = 2$ we find again $g(\theta|x) = N(\hat{\theta},K)$.

Finally, the inequality (19) coincides with the Ferreira inequality (Ferreira 1982) concerning estimating functions, assuming there $g^2(x,\theta) = L(\theta,\hat{\theta})$.

References

Fend, A.V. (1959) On the attainment of Cramér-Rao and Bhattacharyya bounds for the variances of an estimate, *Ann. Math. Statist.* **30**: 381-388.

Ferreira, P.E. (1981) Extending Fisher's measure of information, *Biometrika* **68**: 695-698.

Ferreira, P.E. (1982) Estimating equations in the presence of prior knowledge, *Biometrika* **69**: 667-669.

Kendall, M.G. and **Stuart, A.** (1961) *The advanced Theory of Statistics*, vol. 2, Griffin, London.

Zacks, S. (1971) *The Theory of Statistical Inference*, Wiley, New York.

(Received 14/02/1983;
in revised form 21/03/1984)

مُناظِرات بايزيَّة لحدود بهاتاتشاريه الصغرى

م. لويزا تارجِتَّا
معهد الرياضيات - جامعة كاجلياري - إيطاليا

تم إيجاد لامتساوية شبيهة بلامتساوية بهاتاتشاريه باستخدام معلومات سابقة.

وقد أخذت فى الاعتبار الظروف التى تحقق المساواة بافتراض كثافة احتمالية لاحقة من النوع الأسى.