Sankhyā : The Indian Journal of Statistics
1987, Volume 49, Series A, Pt. 1, pp. 37–42.

CONVERGENCE OF BHATTACHARYA BOUNDS-REVISITED

By J. K. GHOSH
Indian Statistical Institute
and
Y. S. SATHE
University of Bombay

SUMMARY. It is proved that for all estimable functions and all multiparameter exponential families, Bhattacharya bounds converge to the variance of the minimum variance unbiased estimate, provided the true value is an interior point of the parameter space. The special case of orthogonal Bhattacharya functions is briefly discussed and the variance is explicitly computed in a reliability problem.

1. Introduction

Suppose for an estimable $\tau(\theta)$ there exists a minimum variance unbiased estimate T. Let $B(m)$ be the Bhattacharya bound involving $(D^i_\theta f_\theta)/f_\theta$, $1 \leqslant i \leqslant m$. When is it true that $B(m) \rightarrow \text{var}_\theta(T)$? Blight and Rao (1974) show $\lim B(m) = \text{var}_\theta(T)$ if f_θ is normal, binomial, Poisson, negative binomial or gamma with scale parameter, (up to a linear transformation) and τ satisfies certain conditions. Apparently the same result is rediscovered by Khan (1984). See also Lehmann (1983, p. 130).

We prove the same result for all estimable τ and all multiparameter exponential families, provided the true values of the parameter is an interior point of the natural parameter space (Theorem 1 and Cor. 2). An offshoot is a result on the completeness of polynomials (Cor. 1). The argument is extremely simple, it only makes use of elementary ideas about Hilbert spaces.

A multiparameter example is worked out where τ is a function of interest in reliability theory and the Bhattacharya functions are orthogonal. The general question of orthogonality is studied briefly in the light of the results of Seth (1949) and Shanbhag (1972).

It is perhaps not inappropriate to point out that we still do not have a characterisation of densities for which the Bhattacharya bound $B(m)$ is attained.

AMS (1980) *subject classification* : Primary 62F10; Secondary 62N05.

Key words and phrases : Convergence; Bhattacharya bounds; Orthogonality; reliability.

2. Results and proof

Consider an exponential family. By reducing to the space of a minimal sufficient statistic, we may consider the family of densities to be of the form

$$f_{\theta}(x) = A(\theta)\exp\{ < \theta, x > \},\ \theta \in \Theta \subset R^k.$$

The dominating measure is some σ-finite μ. Here θ and x are $k \times 1$ real vectors. Fix θ_0 and assume

$$\theta_0 \text{ is an interior point of } \Theta. \qquad \dots \quad (1)$$

Let $D_i = D_{\theta_i}$ indicate differentiation with respect to the i-th co-ordinate of θ. Let

$$\phi_{j_1, \dots, j_k} = [(D_1^{j_1} \cdots D_k^{j_k} f_{\theta})/f_{\theta}]_{\theta = \theta_0}$$

be the functions used in computing Bhattacharya bounds. Let $\tau(\theta)$ be a real valued function which has an unbiased estimate T which has finite variance under θ_0. Note that T must be the unique minimum variance unbiased estimate. Let $B(m_1, \dots, m_k)$ be the Bhattacharya bound obtained by using $\phi_{j_1, \dots, j_k}$, $j_i \leqslant m_i$. Then the following is true.

Theorem 1 : $$var_{\theta_0}(T) = \lim_{m_i \to \infty,\ i = 1, \dots, k} B(m_1, \dots, m_k).$$

Proof : Let L_2 be the space of real valued functions ϕ which are square integrable with respect to $f_{\theta_0}\, d\mu$, with $\|\phi\|^2 = E_{\theta_0}(\phi^2)$. Let L and L_m be the linear subspaces spanned by $\phi_{j_1, \dots, j_k}$, $1 \leqslant j_i < \infty$ and $\phi_{j_1, \dots, j_k}$, $1 \leqslant j_i \leqslant m_i$. We denote by U the space of all ϕ in L_2 which are orthogonal to constants, i.e., $E_{\theta_0}(\phi) = 0$. We begin by proving that L is dense in U. To see this first note that if $\phi \in L_2$ then by the Cauchy-Schwartz inequality, $E_{\theta}(\phi)$ is well-defined and finite in a neighbourhood of θ_0. Moreover $E_{\theta}(\phi)$ is analytic (i.e., the restriction of an analytic function on a neighbourhood of θ_0 in C^k). Suppose now $\phi \in U$ and is orthogonal to L. Then $E_{\theta}(\phi)$ is an analytic function whose value and derivatives of all orders at θ_0 are zero. It follows that $E_{\theta}(\phi)$ is identically zero and hence, by a well-known result (vide Lehmann, 1959, p. 132) ϕ is identically zero. This implies L is dense in U.

Let T_m be the projection of $(T-\tau(\theta))$ on L_m. By the previous result $\|T-\tau(\theta)-T_m\| \to 0$. Hence,

$$\text{var}_{\theta_0}(T) = \|T-\tau(\theta)\|^2 = \|T_m\|^2+\|T-\tau(\theta)-T_m\|^2 = \lim \|T_m\|^2$$

which is what we have to prove since $B(m) = \|T_m\|^2$.

Corollary 1 : *The linear space of all polynomials is dense in* $L_2(f_{\theta_0})$.

Proof : The proof of the proposition shows that the linear space spanned by the constant function $\phi \equiv 1$ and $\phi_{j_1,\ldots,j_k}$ is dense in L_2. It is easily checked this linear space is identical with the linear space spanned by the polynomials.

It follows that a complete orthonormal basis for $L_2(f_{\theta_0})$ can be found from the set of all polynomials.

Let $\eta = E_\theta(X)$. Then in a neighbourhood of θ_0, $\theta \to \eta$ is a one-one analytic map (i.e., it has an extension to a k-dimensional complex domain that is analytic) and we may parametrise the density as g_η in terms of η. Let $\psi_{j_1,\ldots,j_k} = [(D^{j_1}_{\eta_1} \ldots D^{j_k}_{\eta_k} g_\eta)/g_\eta]_{\eta=\eta_0}$. Note that it is still true that if $\psi \in L_2(f_{\theta_0})$ then $E_\theta(\psi)$ is an analytic function of η. Hence the rest of the proof of Theorem 1 with η in place of θ remains valid and yields

Corollary 2 : *$var_{\theta_0}(T) = \lim C(m)$ where $C(m)$ is the Bhattacharya bound based on the function $\psi_{j_1,\ldots,j_k}$, and $m_i \to \infty$ $i = 1,\ldots,k$, $j_i \leqslant m_i$.*

As far as the computation of $B(m)$ or $C(m)$ is concerned, the five one-parameter families considered by Blight and Rao (1974) remain important. The Bhattacharya bounds are relatively easy to calculate if the Bhattacharya functions $\phi_{j_1,\ldots,j_k}$ or $\psi_{j_1,\ldots,j_k}$ are mutually orthogonal. For the one parameter exponential density,

$$(x-\eta) = K(\eta)\psi_1$$

at least in the interior of the parameter space. Seth (1949) had shown that orthogonality of ψ's is true if K is a quadratic in η. This last condition holds for the five families of [1], or slightly more generally, for random variable whose linear functions have this sort of distribution. Using Seth's calculations, it is not hard to show that orthogonality of ψ's for all θ in some open set implies K is a quadratic. This may be proved by noting that Seth's expressions (5.1.3) to (5.1.7) do not depend on his assumption (5.1.2) that K is a quadratic in η and expression (5.1.8) becomes

$$D_\eta\psi_3 = \psi_4 - \psi_1\psi_3 = -3Z_1\psi_3 - (3Z_2 + 3/K)\psi_2 - Z_3\psi_1 \quad \ldots \quad (2)$$

where $Z_i = K^{-1}D_\eta^i K$, as in Seth. Clearly (2) reduces to (5.1.8) when K is a quadratic. Differentiating (2) and making use of (5.1.3.) to (5.1.7) as well as (2), it can be shown that

$$D_\eta\psi_4 = \psi_5 - \psi_1\psi_4 = \text{a linear function of } \psi_i\text{'s } (i \geqslant 1). \quad \ldots \quad (3)$$

A proof of (3) is obtained as follows. The first equality in (3) follows from (5.1.3). Also by differentiating (2), $D_\eta\psi_4$ is seen to be $D_\eta\psi_1\psi_3$ plus the derivative of the right hand side of (2). Each term is now reduced to a linear combination of ψ_i's and a constant. We illustrate this by considering only $D_\eta\psi_1\psi_3$,

$$D_\eta\psi_1\psi_3 = \psi_1 D_\eta\psi_3 + \psi_3 D_\eta\psi_1. \quad \ldots \quad (4)$$

By (2), the first term above is a linear combination. $\psi_3\psi_1$, $\psi_2\psi_1$ and ψ_1^2 of which the first two are linear combination of ψ_i's by (2) and (5.1.7) and the last is a linear combination of ψ_i's and a constant by (5.1.6). The second term in (4) can be handled in a similar way. Finally collecting all the constant terms appearing in such reductions one notes they cancel each other. This proves (3). It follows from (2) and (3) that

$$\begin{aligned} \psi_1^2\psi_3 &= -Z_3\psi_1^2 + \text{a linear combination of } \psi_i\text{'s } (i \geqslant 1) \\ &\quad + (\text{a linear combination of } \psi_1\psi_3 \text{ and } \psi_1\psi_2) \\ &= -Z_3\psi_1^2 + \text{a linear combination of } \psi_i\text{'s } (i \geqslant 1) \end{aligned}$$

by another application of (2) and (5.1.7). Hence $E_\eta(\psi_1^2\psi_3) = -Z_3 E_\eta(\psi_1^2)$ which is zero iff $Z_3 = 0$, i.e., iff K is a quadratic in η.

The fact that orthogonality of ψ's implies K is a quadratic in η has also been noted by Shanbhag (1972) with somewhat different computations. The main result of Shanbhag (1972) is that (under some conditions) orthogonality holds for a one parameter exponential iff it is one of the five listed in [1] upto a linear transformation. It also follows from his calculations that the only Bhattacharya functions which can be orthogonal are the ψ's obtained by differentiating the density with respect to η.

Combining the above facts it follows that (under mild conditions) the following are equivalent : (1) Seth's condition holds, (2) ψ_j's are orthogonal for an exponential family and (3) the exponential family is one of the five listed in [1] up to a linear transformation. A multiparameter version of this result would be interesting. Of course if the components of $\boldsymbol{X}$ are independent with each of the marginal distributions belonging to one of the above five families, up to a linear transformation, then (taking $k=2$, say) ψ_{j_1,j_2}'s are orthogonal and $\psi_{j_1,j_2} = \psi_{j_1,0} \cdot \psi_{0,j_2}$. Then[1]

$$\text{var}_{\theta}(T) = \Sigma\, \Sigma(\tau_{j_1,j_2}/A_{j_1}\, B_{j_2})^2$$

where $\tau_{j_1,j_2} = D_{\theta_1}^{j_1} D_{\theta_2}^{j_2} \tau$, $A_{j_1}^2 = E_{\theta}(\psi_{j_1,0})^2$, $B_{j_2}^2 = E_{\theta}(\psi_{0,j_2})^2$ and the sum is over $j_1 \geqslant 0$, $j_2 \geqslant 0$, $j_1+j_2 \geqslant 1$. The following example of this kind illustrates the computation of $\text{var}_{\theta}(T)$ in a reliability problem.

Example : Let U, V indicate the stimulus and tolerance for a system. The system survives if $U < V$. Suppose U, V are independent, exponential,

$$f_{\theta}(u, v) = \frac{1}{\theta_1\theta_2} \exp\left\{ \frac{-u}{\theta_1} - \frac{v}{\theta_2} \right\}.$$

Let $\tau(\boldsymbol{\theta}) = P_{\theta}\{U < V\}$. Then $\tau = \theta_2(\theta_1+\theta_2)^{-1}$. We want to estimate τ given n i.i.d. copies of U and m i.i.d. copies of V. Clearly we may work with the sufficient statistics $X_1 = \sum_1^n U_i$, $X_2 = \sum_1^m V_i$. Then X_1, X_2 are independent with marginal gamma distributions. If T is the minimum variance unbiased estimate of τ then

$$\tau_{j_1,j_2} = (-1)^{j_1+j_2} \frac{(j_1+j_2)\,!}{(\theta_1+\theta_2)^{j_1+j_2+1}} \, \frac{(j_1\theta_2 - j_2\theta_1)}{(j_1+j_2)}$$

$$A_{j_1}^2 = \frac{(n+j_1-1)!\, j_1!}{(n-1)!\, \theta_1^{2j_1}}, \quad B_{j_2}^2 = \frac{(m+j_2-1)!\, j_2!}{(m-1)!\, \theta_2^{2j_2}}.$$

Thus, though $T = P(U_1 < V_1 | \boldsymbol{X})$ is hard to write down explicitly, there is an explicit expression for its variance. In the special case $n = m = 1$, $\text{var}_{\theta}(T) = \tau(1-\tau)$, so that one gets the curious identity

$$\frac{\theta_1\theta_2}{(\theta_1+\theta_2)^2} = \sum_{\substack{j_1 \geqslant 0,\, j_2 \geqslant 0 \\ j_1+j_2 \geqslant 1}} \binom{j_1+j_2}{j_1}^2 \frac{(j_1\theta_2 - j_2\theta_1)^2\, \theta_1^{2j_1}\, \theta_2^{2j_2}}{((j_1+j_2))^2 (\theta_1+\theta_2)^{2(j_1+j_2+1)}}$$

Another curious identity of the same type may be obtained if we take $\tau = e^{-u/\theta_1} = P_{\theta_1}\{u < U_1\}$ and $n = 1$. Then

$$e^{-u/\theta_1}(1 - e^{-u/\theta_1}) = \sum_{i \geqslant 1} \frac{(D_{\theta_1}^i \tau)^2 \theta_1^{2i}}{(\tau\,!)^2}$$

[1] After the paper was prepared, the authors came to know that this special results and the following example have appeared in Bartoszowicz (1980). Zastosowania Mathematicae, Vol. 16, pp. 601–60.

Professor S. Bagchi has shown us a direct proof of this using Parseval's result for Fourier transforms.

Incidentally if one takes the normal with a location parameter or the exponential with a scale parameter the ψ's (for $\theta_0 = 0$ and $\theta_0 = 1$) are the Hermite and Laguerre polynomials (up to a constant). Corollary 1 implies that these classical polynomials form a basis for $L_2(f_{\theta_0})$. The present proof seems to be different from the classical one, videCo urant and Hilbert (1953, 94-96).

Acknowledgement. The authors would like to thank the co-editor and the referee for improvement in presentation.

References

Blight, B. J. N. and Rao, P. V. (1974): The convergence of Bhattacharya bounds. *Biometrika*, 61, 137-142.

Courant, R. and Hilbert, D. (1953): *Methods of Mathematical Physics*, vol. 1, Interscience, New York.

Khan, R. A. (1984): On UMVU estimators and Bhattacharya bounds in exponential distributions. *JSPI*, 9, 199-206.

Lehmann, E. L. (1959): *Testing Statistical Hypotheses*, Wiley, New York.

——— (1983): *Theory of Point Estimation*, Wiley, New York.

Seth, G. R. (1949): On the variance of estimates. *Ann. Math. Statist.*, 20, 1-27.

Shanbhag, D. N. (1972): Some characterisations based on the Bhattacharya matrix. *J. Appl. Prob.*, 9, 580-587.

Paper received : April, 1986.

Revised : May, 1986.