Bhattacharyya Inequalities as an Extended version of the Rao-Cramer Inequalities

G. R. Mohtashami Borzadaran
Department of Mathematics and Statistics
Faculty of Sciences
University of Birjand
Birjand-IRAN

Abstract

In this paper, we consider characterizations based on the Bhattacharyya matrices. We obtain the structure of the r^{th} moment of the r.v. X about the origin for an NEF (natural exponential families) when the variance is a polynomial of the mean such that the mean is a linear function of the parameter of the family.

AMS subject classification : Primary 62E10, 62H10, Secondary 60E05.
Keyword: Exponential Families, Bhattacharyya Bounds, Rao-Cramer Inequality, Fisher Information, Diagonality of the Bhattacharyya matrices.

1 Introduction

A lower bound for the variance of an estimator is one of the fundamental things in the estimation theory because it gives us an idea about the accuracy of an estimator. Fisher showed that the asymptotic variance of a consistent estimator is bounded by the inverse of the Fisher information measure. The exact inequality giving the lower bound for the variance was established under some mild conditions by C. R. Rao (1945) and Cramer (1946). The Rao-Cramer inequality states that the variance of an unbiased estimator T of $\tau(\theta)$, satisfies

$$Var_\theta(T(\underline{X})) \geq \frac{(\tau^{(1)}(\theta))^2}{E_\theta\{-\frac{\partial^2 \ln f(\underline{X}|\theta)}{\partial\theta^2}\}}, \tag{1}$$

where $\tau^{(1)}(\theta)$ is the derivative of τ w.r.t. θ. The denominator of (**??**) is called the Fisher information. An important inequality to follow the Rao-Cramer inequality is that of A. Bhattacharyya (1946,1947,1948). The Bhattacharyya inequality achieves a greater lower bound for the variance of an unbiased estimator of a parametric function, and it becomes sharper and sharper as the order of the Bhattacharyya matrix increases. The closeness of the bound for the variance of an unbiased estimator, depends on the order of the Bhattacharyya matrix in such a way that increasing the order implies that the size of the Bhattacharyya matrix becomes bigger and the inverse of it requires a greater effort for calculation.

In this paper we obtain the structure of the r^{th} moment of the r.v. X about the origin for an NEF when the variance is a polynomial of the mean such that the mean is a linear function of the parameter of the family.

2 Main Results

The Bhattacharyya inequality involves the covariance matrix of the random vector

$$\frac{1}{f(\underline{X}|\theta)}[f^{(1)}(\underline{X}|\theta), f^{(2)}(\underline{X}|\theta), ..., f^{(n)}(\underline{X}|\theta)],$$

where $f^{(j)}(.|\theta)$ is the j^{th} derivative of the probability density function $f(.|\theta)$ w.r.t. the parameter θ. The covariance matrix of the above random vector is referred to as the $n \times n$ Bhattacharyya matrix. G. R. Mohtashami Borzadaran (1998, 2000) obtained characterizations related to the analogue of the Fisher information and Bhattacharyya matrices.

The Bhattacharyya inequality states that

$$V_\theta(T(\underline{X})) \geq \xi_\theta^t J^{-1} \xi_\theta, \tag{2}$$

where
(i) t is the notation for transpose and $\xi_\theta = (\tau^{(1)}(\theta), \tau^{(2)}(\theta), ..., \tau^{(k)}(\theta))^t$,
(ii) $\tau(\theta) = E_\theta(T(\underline{X}))$ and $\tau^{(j)} = \frac{\partial E_\theta(T(\underline{X}))}{\partial \theta^j}$ for $j = 1, 2, ..., k$,
(iii) J^{-1} is the inverse of the Bhattacharyya matrix where $J_{rs} = Cov\{\frac{f^{(r)}(X|\theta)}{f(X|\theta)}, \frac{f^{(s)}(X|\theta)}{f(X|\theta)}\}$ such that $E_\theta(\frac{f^{(r)}(X|\theta)}{f(X|\theta)}) = 0, \quad r, s = 1, 2, ..., n.$

Remark 2.1 If we substitute $k = 1$ in (**??**), then it indeed reduces to the Rao-Cramer inequality.

Let X be a non-degenerate r.v. distributed according to distribution with density :

$$f(x|\theta) = \frac{exp\{xg(\theta)\}}{\beta(g(\theta))}\psi(x), \tag{3}$$

where $\theta \in \Theta$, Θ is an open interval, and g is thrice differentiable. Shanbhag (1972) proved that a 3×3 Bhattacharyya Matrix is diagonal if and only if

$$\begin{cases} E_\theta(X) = c_{11} + c_{21}\theta \\ E_\theta(X^2) = c_{12} + c_{22}\theta + c_{23}\theta^2. \end{cases} \tag{4}$$

(**??**) implies that $V_\theta(X) = c_{13} + c_{23}\theta + c_{33}\theta^2$ and $g'(\theta) = \frac{c_{21}}{c_{13}+c_{23}\theta+c_{33}\theta^2}$, where c_{ij} $i, j = 1, 2, 3$, are constants. Shanbhag characterized the normal, gamma, binomial, negative binomial, and Poisson distributions as in the case of Laha and Lukacs (1960) via (**??**). In Shanbhag (1979), the result is extended by arriving at a characterization of a larger class of distributions including Meixner (hyperbolic) distribution. The six distributions involved here are referred to by Morris (1982,1983) as NEF-QVF. Lai (1982) defined the class of distributions consisting of these six distributions as Meixner class. Thus, it follows that Shanbhag (1972,1979) proved that a 3×3 Bhattacharyya matrix is diagonal if and only if the distribution is a member of the Meixner class.

Blight & Rao (1974) considered Bhattacharyya bounds for an unbiased estimator of a parametric function for an exponential family and illustrated via exponential and negative binomial distributions, that the Bhattacharyya bound actually converges to the variance of the best unbiased estimator for members of the Meixner class.

Fosam (1993) used essentially the same idea to characterize the Letac-Mora (1990)class. More recently, Fosam & Shanbhag (1996) extended the Laha-Luckas characterization result, to have a charaterization based on a cubic regression property, subsuming, the Letac-Mora characterization of the NEF.

Theorem 2.2 *If for the exponential family (??), $V_\theta(X)$ is the k^{th} degree polynomial in $E_\theta(X)$, then, under the assumption that $E_\theta(X)$ is linear in θ, $E_\theta(X^r)$ for $r \geq 1$ is a polynomial in θ of degree $1 + (r-1)(k-1)$.*

Proof: On noting that $E_\theta(X)$ is a linear function of θ, and $g'(\theta) = \frac{1}{b_0^{(k)} + b_1^{(k)}\theta + b_2^{(k)}\theta^2 + \ldots + b_k^{(k)}\theta^k}$, we have

$$\begin{aligned} E_\theta(X^2) &= \frac{\beta^{(2)}(g(\theta))}{\beta(g(\theta))} \\ &\equiv P_k(\theta), \end{aligned}$$

where $P_n(\theta) = c_0^{(n)} + c_1^{(n)}\theta + c_2^{(n)}\theta^2 + \ldots + c_n^{(n)}\theta^n$ is a representation of a polynomial of θ of order n such that $c_i^{(n)}, i = 1, 2, \ldots, n$ are real constants. Further

$$\begin{aligned} \frac{d}{d\theta} E_\theta(X^2) &= g'(\theta)\{\frac{\beta^{(3)}(g(\theta))}{\beta(g(\theta))} - \frac{\beta^{(2)}(g(\theta))}{\beta(g(\theta))} \frac{\beta^{(1)}(g(\theta))}{\beta(g(\theta))}\} \\ &\equiv P_{k-1}(\theta), \end{aligned}$$

and

$$E_\theta(X^3) = \frac{\beta^{(3)}(g(\theta))}{\beta(g(\theta))} \equiv P_{2k-1}(\theta).$$

If we assume now that for a fixed integer $r > 1$

$$E_\theta(X^{r-1}) = \frac{\beta^{(r-1)}(g(\theta))}{\beta(g(\theta))} \equiv P_{(r-2)(k-1)+1}(\theta),$$

then

$$\begin{aligned} \frac{d}{d\theta} E_\theta(X^{r-1}) &= g'(\theta)\{\frac{\beta^{(r)}(g(\theta))}{\beta(g(\theta))} - \frac{\beta^{(r-1)}(g(\theta))}{\beta(g(\theta))} \frac{\beta^{(1)}(g(\theta))}{\beta(g(\theta))}\} \\ &\equiv P_{(r-2)(k-1)}(\theta), \end{aligned}$$

yielding

$$\begin{aligned} E_\theta(X^r) &= \frac{\beta^{(r)}(g(\theta))}{\beta(g(\theta))} \\ &= P_k(\theta) P_{(r-2)(k-1)}(\theta) \equiv P_{(r-1)(k-1)+1}(\theta). \end{aligned}$$

Hence, by induction, we have the assertion.

Corollary 2.3 *For the Letac-Mora class of distributions, under the assumption that the mean is a linear function of θ, for each positive integer r, $E_\theta(X^r)$ is a polynomial of degree $2r - 1$ in θ. For the Seth-Shanbhag-Morris class of distributions, under the assumption for each positive integer r, $E_\theta(X^r)$ is a polynomial of degree r in θ.*

Proof: The corollary follows from the theorem on noting that for two classes in question the assumption of the theorem is met with $k = 3$ and 2 respectively.

Remark 2.4 Via the Shanbhag's method, we have

$$\begin{aligned} J_{ir} &= E_\theta\{\frac{f^{(r)}}{f}\sum_{l=0}^{i} d_l^{(i)}(\theta)X^l\} \\ &= \sum_{l=0}^{i} d_l^{(i)}(\theta)\frac{d^r E_\theta(X^l)}{d\theta^r}, \end{aligned}$$

where $d_l^{(i)}(\theta)$ is the coefficient of X^l in $\frac{f^{(i)}}{f}$. This yields, in view of Theorem **??** and (symmetry), under the assumption that $E_\theta(X)$ is linear in θ, the assertion that $J_{ir} = J_{ri} = 0$ for $r > (i-1)(k-1)+1$; this latter result holds for the Letac-Mora family of distributions with $k = 3$ and the Shanbhag-Morris family with $k = 2$.

3 Conclusion

We have shown that for an exponential family of the form (**??**), if $V_\theta(T(\underline{X}))$ is the k^{th} degree polynomial in $E_\theta(X)$, then $E_\theta(X^r)$ is a polynomial in θ of degree $1 + (r-1)(k-1)$. The Bhattacharyya matrices relative to the Letac-Mora and Shanbhag-Morris classes are as in the conclusion of the above results with $k = 3$ and $k = 2$ respectively.

Aknowledgment
I am grateful to Prof. D. N. Shanbhag for his evaluable suggestions.

References

[1] Bhattacharyya, A. (1946). On some analogues of the amount of information and their use in statistical estimation. *Sankhya, Ser. A.*, **8**, 1-14.

[2] Bhattacharyya, A. (1947). On some analogues of the amount of information and their use in statistical estimation. *Sankhya, Ser. A.*, **8**, 201-218.

[3] Bhattacharyya, A. (1948). On some analogues of the amount of information and their use in statistical estimation (concluded). *Sankhya, Ser. A*, **8**, 315-328.

[4] Blight, B. J. N. & Rao, P. V. (1974). The convergence of the Bhattacharyya bounds. *Biometrika*, **61**, 137-142.

[5] Cramer, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton.

[6] Fosam, E. B. (1993). *Characterizations and Structural Aspects of Probability Distributions.* Ph.D. Thesis, Sheffield University.

[7] Fosam, E. B., Shanbhag, D. N. (1996). An extended Laha-Lukacs characterization result based on a regression property. *Special Issue on characterizations of JSPI, ed. D.N. Shanbhag (to appear).*

[8] Laha, R. G. & Lukacs, E. (1960). On a problem connected with quadratic regression. *Biometrika*, **47, 3**, 335-343.

[9] Lai, C. D. (1982). Meixner classes and Meixner hypergeometric distributions. *Austral. J. Statist.*, **24**, 221-233.

[10] Letac, G., & Mora, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statist.*, **Vol. 18, No. 1**, 1-37.

[11] Mohtashami Borzadaran, G. R. (1998). Analogues of the Fisher information. *Submitted to Metrika.*

[12] Mohtashami Borzadaran, G. R. (2000). Results related to Bhattacharyya Matrices.. *Accepted for publication in Sankhya Ser. A.*

[13] Morris, C. N. (1982). Natural real exponential families with quadratic variance functions. *Ann. Statist.*, **10**, 65-80.

[14] Morris, C. N. (1983). Natural real exponential families with quadratic variance functions : statistical theory. *Ann. Statist.*, **Vol. 11, No. 2**, 515-529.

[15] Rao, C. R. (1945). Information and accuraccy attainable in estimation of statistical parameters. *Bull. Cal. Math. Soc.*, **37(3)**, 81-89.

[16] Shanbhag, D. N. (1972). Some characterizations based on the Bhattacharyya Matrix. *J. Appl. Prob.*, **9**, 580-587.

[17] Shanbhag, D. N. (1979). Diagonality of the Bhattacharyya Matrix as a characterization. *Theory. Prob. Application*, **24**, 430-433.