Ann. Inst. Statist. Math.
Vol. 55, No. 2, 309–317 (2003)

ON A FAMILY OF DISTRIBUTIONS ATTAINING THE BHATTACHARYYA BOUND

HIDEKAZU TANAKA AND MASAFUMI AKAHIRA

Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan,
e-mail: hide@math.tsukuba.ac.jp; akahira@math.tsukuba.ac.jp

(Received January 17, 2002; revised August 29, 2002)

Abstract. A family of distributions for which an unbiased estimator of a function $g(\theta)$ of a real parameter θ can attain the second order Bhattacharyya lower bound is derived. Indeed, we obtain a necessary and sufficient condition for the attainment of the second order Bhattacharyya bound for a family of mixtures of distributions which belong to the exponential family. Furthermore, we give an example which does not satisfy this condition, but where the Bhattacharyya bound is attainable for a non-exponential family of distributions.

Key words and phrases: Cramér-Rao bound, exponential family, normal mixture, Bessel differential equation.

1. Introduction

In the estimation of a function $g(\theta)$ of a real parameter θ, the Bhattacharyya bounds for the variance of an unbiased estimator are known as a generalization of the Cramér-Rao bound (see Zacks (1971), Lehmann and Casella (1998)). A necessary and sufficient condition for the attainment of the Cramér-Rao lower bound is that the density of the unbiased estimator $\hat{g}$ of $g(\theta)$ be of exponential type (see Wijsman (1973)). If an unbiased estimator $\hat{g}$ is a polynomial of degree k of the minimal sufficient statistic in an exponential family and if its variance does not attain the Bhattacharyya lower bounds up to the $(k-1)$-th order, then it attains the k-th order bound (see Fend (1959), Zacks (1971), and also Ishii (1976)). However, a family of distributions for which the Bhattacharyya bound can be attained seems to be still unknown. In this paper we consider a family of distributions for which the second order Bhattacharyya bound becomes sharp. A necessary and sufficient condition for the second order Bhattacharyya bound to be sharp is given for a family of linear combinations of distributions which belong to the exponential family. An example on the normal mixture is also given. Moreover, an example is presented where this condition is not satisfied, but the Bhattacharyya bound is attainable in the non-exponential family of distributions.

2. Bhattacharyya inequality

Suppose that $(\mathcal{X}, \mathcal{B})$ is a sample space and the family of probability distributions $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is dominated with respect to some σ-finite measure μ, where Θ is an open interval of $\boldsymbol{R}^1$. Denote by $f(x, \theta) = dP_\theta/d\mu$ $(\theta \in \Theta)$ a probability density function (p.d.f.). We consider an estimation problem of the U-estimable function $g(\theta)$, i.e. the function $g(\theta)$ for which its unbiased estimator with a finite variance exists, based on a sample X.

We consider the Bhattacharyya inequality under the following regularity conditions (A1) to (A4).

(A1) For almost all x $[\mu]$, there exist $(\partial^i/\partial\theta^i)f(x,\theta)$ for $i=1,\dots,k$.

(A2) For $i=1,\dots,k$, there exist $\mathcal{B}$-measurable functions $M_i(x)>0$ such that $|(\partial^i/\partial\theta^i)f(x,\theta)| \le M_i(x)$ for all $\theta\in\Theta$, and $\int M_i(x)d\mu(x)<\infty$.

(A3) $\int |((\partial^i/\partial\theta^i)f(x,\theta)(\partial^j/\partial\theta^j)f(x,\theta))/f(x,\theta)|d\mu(x)<\infty$, for $i,j=1,\dots,k$ and for all $\theta\in\Theta$.

(A4) For almost all x $[\mu]$ and for all $\theta\in\Theta$, $f(x,\theta)>0$.

THEOREM 2.1. (Bhattacharyya (1946), Zacks (1971)) *Suppose that the conditions* (A1) *to* (A4) *hold. Assume that $g(\theta)$ is a U-estimable function which is k-times differentiable over Θ. Let $\boldsymbol{g}(\theta)={}^t(g^{(1)}(\theta),\dots,g^{(k)}(\theta))$, where $g^{(i)}(\theta)$ is the i-th order derivative of $g(\theta)$. Let $\hat{g}(X)$ be an unbiased estimator of $g(\theta)$ having a finite variance, and assume that, for $i=1,\dots,k$, there exists a function $N_i(x)$ such that $|\hat{g}(x)\cdot(\partial^i/\partial\theta^i)f(x,\theta)|\le N_i(x)$ for all $\theta\in\Theta$, and $\int N_i(x)d\mu(x)<\infty$. Furthermore, let $I(\theta)$ be a $k\times k$ non-negative definite matrix with elements*

$$I_{ij}(\theta)=E_\theta\left[\frac{\partial^i f(X,\theta)/\partial\theta^i}{f(X,\theta)}\cdot\frac{\partial^j f(X,\theta)/\partial\theta^j}{f(X,\theta)}\right],\qquad (i,j=1,\dots,k).$$

Then, if $I(\theta)$ is non-singular over Θ,

$$\mathrm{Var}_\theta(\hat{g}(X))\ge {}^t\boldsymbol{g}(\theta)I(\theta)^{-1}\boldsymbol{g}(\theta)=:B_k(\theta). \tag{2.1}$$

And the equality holds in (2.1) *if and only if*

$$\hat{g}(x)-g(\theta)=\sum_{i=1}^k a_i(\theta)\frac{\partial^i f(x,\theta)/\partial\theta^i}{f(x,\theta)}\qquad a.e.\ x\ [\mu] \tag{2.2}$$

for all $\theta\in\Theta$, where

$$\begin{pmatrix}a_1(\theta)\\ \vdots\\ a_k(\theta)\end{pmatrix}=I(\theta)^{-1}\begin{pmatrix}g^{(1)}(\theta)\\ \vdots\\ g^{(k)}(\theta)\end{pmatrix}.$$

For a proof, we refer the reader to Bhattacharyya (1946) and Zacks (1971). The lower bound $B_k(\theta)$ in (2.1) is called the k-th order Bhattacharyya (lower) bound. Note that $B_1(\theta)$ is consistent with the Cramér-Rao lower bound.

3. On the attainment of the Bhattacharyya bound

In the previous section, we stated the necessary and sufficient condition for an unbiased estimator to attain the Bhattacharyya lower bound, under some regularity conditions. The problem whether the Bhattacharyya bounds become sharp as $k\to\infty$

was investigated for some one-parameter cases by Blight and Rao (1974). Since (2.2) is a linear k-th order differential equation, its general solution is of the form

$$f(x,\theta)=\sum_{i=1}^{k}A_i(x)f_i(x,\theta), \tag{3.1}$$

where $f_i(x,\theta)$ $(i=1,\dots,k)$ are k linearly independent solutions of (2.2), and $A_i(\cdot)$ $(i=1,\dots,k)$ are arbitrary functions. Next, we consider the case when $k=2$ and $a_2(\theta)\neq 0$ for all $\theta\in\Theta$. Let $f_1(x,\theta)$ be a particular solution of (2.2), and define

$$f_2(x,\theta):=f_1(x,\theta)\int\frac{1}{f_1^2(x,\theta)}\exp\left\{-\int\frac{a_1(\theta)}{a_2(\theta)}d\theta\right\}d\theta.$$

Since

$$\frac{\frac{\partial}{\partial\theta}f_2(x,\theta)}{f_2(x,\theta)}=\frac{\frac{\partial}{\partial\theta}f_1(x,\theta)}{f_1(x,\theta)}+\frac{1}{f_1(x,\theta)f_2(x,\theta)}\exp\left\{-\int\frac{a_1(\theta)}{a_2(\theta)}d\theta\right\},$$

$$\frac{\frac{\partial^2}{\partial\theta^2}f_2(x,\theta)}{f_2(x,\theta)}=\frac{\frac{\partial^2}{\partial\theta^2}f_1(x,\theta)}{f_1(x,\theta)}-\frac{1}{f_1(x,\theta)f_2(x,\theta)}\frac{a_1(\theta)}{a_2(\theta)}\exp\left\{-\int\frac{a_1(\theta)}{a_2(\theta)}d\theta\right\},$$

it follows that

$$\begin{aligned}a_1(\theta)\frac{\frac{\partial}{\partial\theta}f_2(x,\theta)}{f_2(x,\theta)}+a_2(\theta)\frac{\frac{\partial^2}{\partial\theta^2}f_2(x,\theta)}{f_2(x,\theta)}&=a_1(\theta)\frac{\frac{\partial}{\partial\theta}f_1(x,\theta)}{f_1(x,\theta)}+a_2(\theta)\frac{\frac{\partial^2}{\partial\theta^2}f_1(x,\theta)}{f_1(x,\theta)}\\&=\hat{g}(x)-g(\theta).\end{aligned}$$

Because the Wronskian of $f_1(x,\theta)$ and $f_2(x,\theta)$ is

$$W(f_1,f_2):=\begin{vmatrix}f_1(x,\theta) & f_2(x,\theta)\\ \frac{\partial}{\partial\theta}f_1(x,\theta) & \frac{\partial}{\partial\theta}f_2(x,\theta)\end{vmatrix}=\exp\left\{-\int\frac{a_1(\theta)}{a_2(\theta)}d\theta\right\}\neq 0,$$

we see that $f_1(x,\theta)$ and $f_2(x,\theta)$ are linearly independent, hence $\{f_1,f_2\}$ forms a fundamental system of solutions. By the theory of second order linear differential equation, it is well known that the general solution of (2.2) is given by

$$f(x,\theta)=A_1(x)f_1(x,\theta)+A_2(x)f_1(x,\theta)\int\frac{1}{f_1^2(x,\theta)}\exp\left\{-\int\frac{a_1(\theta)}{a_2(\theta)}d\theta\right\}d\theta.$$

Note that, even when $k=2$, for general functions $a_1(\theta)$ and $a_2(\theta)$, the solutions to (2.2) can not be expressed by means of elementary functions.

THEOREM 3.1. *Suppose that the conditions* (A1) *to* (A4) *hold. Assume that* $\mu(\{x\in\mathcal{X}|\hat{g}(x)=r\})=0$ *for all* $r\in\boldsymbol{R}^1$. *Let* $k=2$ *and* $a_2(\theta)\neq 0$ *for all* $\theta\in\Theta$. *Then the solution of* (2.2) *is expressed by a linear combination of distributions from the exponential family if and only if the following* (i) *and* (ii) *hold.*

(i) *There are a function* $t(x)$ *and constants* C_0, C_1 *and* C_2 *such that* C_2 *has the same sign as* $a_2(\theta)$, *and* $\hat{g}(x)$ *is of the form*

$$\hat{g}(x) = C_2 t^2(x) + C_1 t(x) + C_0.$$

(ii) *For* C_0, C_1 *and* C_2 *given in* (i), $g(\theta)$ *has the form*

$$g(\theta) = \frac{a_1^2(\theta)}{4a_2(\theta)} - \frac{a_1(\theta)a_2'(\theta)}{2a_2(\theta)} + \frac{a_1'(\theta)}{2} - \frac{a_2''(\theta)}{4} + \frac{3{a_2'}^2(\theta)}{16a_2(\theta)} + C_0 - \frac{C_1^2}{4C_2}.$$

PROOF. *Necessity.* Let

$$f(x,\theta) := \exp\{t(x)\psi_1(\theta) + \psi_2(\theta)\}$$

be a particular solution of (2.2). Then we have

$$\frac{\frac{\partial}{\partial\theta} f(x,\theta)}{f(x,\theta)} = t(x)\psi_1'(\theta) + \psi_2'(\theta),$$

$$\frac{\frac{\partial^2}{\partial\theta^2} f(x,\theta)}{f(x,\theta)} = t(x)\psi_1''(\theta) + \psi_2''(\theta) + (t(x)\psi_1'(\theta) + \psi_2'(\theta))^2.$$

By substituting into (2.2), we obtain

$$\begin{aligned} \hat{g}(x) - g(\theta) &= t^2(x)a_2(\theta){\psi_1'}^2(\theta) \\ &\quad + t(x)\{a_1(\theta)\psi_1'(\theta) + a_2(\theta)\psi_1''(\theta) + 2a_2(\theta)\psi_1'(\theta)\psi_2'(\theta)\} \\ &\quad + a_1(\theta)\psi_2'(\theta) + a_2(\theta)\psi_2''(\theta) + a_2(\theta){\psi_2'}^2(\theta). \end{aligned}$$

Then $\hat{g}(x)$ can be expressed as $\hat{g}(x) = C_2 t^2(x) + C_1 t(x) + C_0$, where

$$C_2 := a_2(\theta){\psi_1'}^2(\theta), \tag{3.2}$$

$$C_1 := a_1(\theta)\psi_1'(\theta) + a_2(\theta)\psi_1''(\theta) + 2a_2(\theta)\psi_1'(\theta)\psi_2'(\theta), \tag{3.3}$$

$$C_0 := a_1(\theta)\psi_2'(\theta) + a_2(\theta)\psi_2''(\theta) + a_2(\theta){\psi_2'}^2(\theta) + g(\theta). \tag{3.4}$$

Note that C_0, C_1 and C_2 are constants, since $\{t^2(x), t(x), 1\}$ is linearly independent. From (3.2) and (3.3) we deduce that

$$\psi_1'(\theta) = \pm\left(\frac{C_2}{a_2(\theta)}\right)^{1/2},$$

$$\psi_2'(\theta) = \pm\left(\frac{C_1}{2C_2}\right)\left(\frac{C_2}{a_2(\theta)}\right)^{1/2} - \frac{a_1(\theta)}{2a_2(\theta)} + \frac{a_2'(\theta)}{4a_2(\theta)},$$

consequently (3.4) implies (ii).

Sufficiency. Assume that (i) and (ii) hold. Let

$$\psi_{1\pm}(\theta) := \pm\int\left(\frac{C_2}{a_2(\theta)}\right)^{1/2} d\theta,$$

$$\psi_{2\pm}(\theta) := \pm\left(\frac{C_1}{2C_2}\right)\int\left(\frac{C_2}{a_2(\theta)}\right)^{1/2} d\theta - \frac{1}{2}\int\frac{a_1(\theta)}{a_2(\theta)}d\theta + \frac{1}{4}\int\frac{a_2'(\theta)}{a_2(\theta)}d\theta,$$

$$f_\pm(x,\theta) := \exp\{t(x)\psi_{1\pm}(\theta) + \psi_{2\pm}(\theta)\},$$

where + and − signs should be read consistently. By computation,

$$
\begin{aligned}
\frac{\frac{\partial}{\partial\theta} f_{\pm}(x,\theta)}{f_{\pm}(x,\theta)} &= \pm\left(\frac{C_2}{a_2(\theta)}\right)^{1/2} t(x) \pm \frac{C_1}{2C_2}\left(\frac{C_2}{a_2(\theta)}\right)^{1/2} - \frac{a_1(\theta)}{2a_2(\theta)} + \frac{a_2'(\theta)}{4a_2(\theta)}, \\
\frac{\frac{\partial^2}{\partial\theta^2} f_{\pm}(x,\theta)}{f_{\pm}(x,\theta)} &= \frac{C_2}{a_2(\theta)} t^2(x) + \left\{\frac{C_1}{a_2(\theta)} \mp \left(\frac{C_2}{a_2(\theta)}\right)^{1/2} \frac{a_1(\theta)}{a_2(\theta)}\right\} t(x) \\
&\quad + \frac{a_1(\theta)a_2'(\theta)}{4a_2^2(\theta)} + \frac{C_1^2}{4C_2 a_2(\theta)} - \frac{a_1'(\theta)}{2a_2(\theta)} + \frac{a_2''(\theta)}{4a_2(\theta)} \\
&\quad + \frac{a_1^2(\theta)}{4a_2(\theta)} - \frac{3{a_2'}^2(\theta)}{16a_2^2(\theta)} \mp \frac{C_1 a_1(\theta)}{2C_2 a_2(\theta)}\left(\frac{C_2}{a_2(\theta)}\right)^{1/2},
\end{aligned}
$$

which implies that $f_+(x,\theta)$ and $f_-(x,\theta)$ are solutions of (2.2). The Wronskian of $f_+(x,\theta)$ and $f_-(x,\theta)$ is

$$
\begin{aligned}
W(f_+, f_-) &= \begin{vmatrix} f_+(x,\theta) & f_-(x,\theta) \\ \frac{\partial}{\partial\theta} f_+(x,\theta) & \frac{\partial}{\partial\theta} f_-(x,\theta) \end{vmatrix} \\
&= -f_+(x,\theta) f_-(x,\theta) \left(\frac{C_2}{a_2(\theta)}\right)^{1/2} \left(2t(x) + \frac{C_1}{C_2}\right) \neq 0, \qquad \text{a.e. } x\ [\mu],
\end{aligned}
$$

since $C_2 \neq 0$ and $\mu(\{x \in \mathcal{X} | 2t(x) + C_1/C_2 = 0\}) = 0$. Hence $f_+(x,\theta)$ and $f_-(x,\theta)$ are fundamental solutions of (2.2), and the general solution $f(x,\theta)$ is given by (3.1). □

Remark 3.1. According to Theorems 2.1 and 3.1, the conditions (i) and (ii) of Theorem 3.1 are necessary and sufficient for the attainment of the second order Bhattacharyya bound for a family of mixtures of distributions belonging to the exponential family.

4. Examples

In this section we give an example on the normal mixture distribution and an example where the condition (ii) of Theorem 3.1 is not satisfied but the Bhattacharyya bound is attainable in the non-exponential family of distributions.

Example 4.1. (Normal mixture) Suppose that X is a random variable according to the normal mixture distribution with p.d.f.

$$f(x,\theta) = p\phi(x-\theta) + q\phi(x+\theta),$$

where $\phi(\cdot)$ is the p.d.f. of the standard normal distribution $N(0,1)$, $\theta \in \boldsymbol{R}^1$, $p + q = 1$ and $0 \leq p \leq 1$. Then we have

$$
\begin{aligned}
\frac{\partial}{\partial\theta} f(x,\theta) &= -\theta f(x,\theta) + xh(x,\theta), \\
\frac{\partial^2}{\partial\theta^2} f(x,\theta) &= (\theta^2 - 1) f(x,\theta) + x^2 f(x,\theta) - 2\theta x h(x,\theta),
\end{aligned}
$$

where

$$h(x,\theta) := p\phi(x-\theta) - q\phi(x+\theta),$$

and, consequently,

$$\left\{\frac{\frac{\partial}{\partial\theta}f(x,\theta)}{f(x,\theta)}\right\}^2 = \theta^2 - 2\theta x\frac{h(x,\theta)}{f(x,\theta)} + x^2 - 4pq\left\{\frac{x}{f(x,\theta)}\right\}^2\phi(x-\theta)\phi(x+\theta), \tag{4.1}$$

$$\begin{aligned}\frac{\frac{\partial}{\partial\theta}f(x,\theta)}{f(x,\theta)}\frac{\frac{\partial^2}{\partial\theta^2}f(x,\theta)}{f(x,\theta)} &= -\theta(\theta^2-1) - 3\theta x^2 + (3\theta^2-1)x\frac{h(x,\theta)}{f(x,\theta)} + x^3\frac{h(x,\theta)}{f(x,\theta)} \\ &\quad + 8pq\theta\left\{\frac{x}{f(x,\theta)}\right\}^2\phi(x-\theta)\phi(x+\theta),\end{aligned} \tag{4.2}$$

$$\begin{aligned}\left\{\frac{\frac{\partial^2}{\partial\theta^2}f(x,\theta)}{f(x,\theta)}\right\}^2 &= (\theta^2-1)^2 + 2(3\theta^2-1)x^2 + x^4 - 4\theta(\theta^2-1)x\frac{h(x,\theta)}{f(x,\theta)} \\ &\quad - 4\theta x^3\frac{h(x,\theta)}{f(x,\theta)} - 16pq\theta^2\left\{\frac{x}{f(x,\theta)}\right\}^2\phi(x-\theta)\phi(x+\theta).\end{aligned} \tag{4.3}$$

Let Y be a random variable according to the normal distribution $N(\theta,1)$ with mean θ and variance 1. Note that $-Y$ is distributed as $N(-\theta,1)$. Then we get

$$\begin{aligned}&E_\theta[X^2] = pE_\theta[Y^2] + qE_\theta[Y^2] = \theta^2+1,\\ &E_\theta[X^4] = pE_\theta[Y^4] + qE_\theta[Y^4] = \theta^4+6\theta^2+3,\\ &E_\theta\left[X\frac{h(X,\theta)}{f(X,\theta)}\right] = pE_\theta[Y] + qE_\theta[Y] = \theta,\\ &E_\theta\left[X^3\frac{h(X,\theta)}{f(X,\theta)}\right] = pE_\theta[Y^3] + qE_\theta[Y^3] = \theta^3+3\theta,\end{aligned}$$

therefore, by Theorem 2.1, (4.1), (4.2) and (4.3),

$$I(\theta) = \begin{pmatrix} 1-K(\theta) & 2\theta K(\theta) \\ 2\theta K(\theta) & 2(1-2\theta^2K(\theta)) \end{pmatrix},$$

where

$$K(\theta) := 4pqE_\theta\left[\left\{\frac{X}{f(X,\theta)}\right\}^2\phi(X-\theta)\phi(X+\theta)\right].$$

Note that $K(\theta)$ is finite for all $\theta\in\Theta$. Let $g(\theta) := \theta^2$ be an estimand and $\hat{g}(X) := X^2-1$. Then Theorem 2.1 implies that the second order Bhattacharyya lower bound $B_2(\theta)$ is given by

$$B_2(\theta) = (2\theta, 2)\frac{1}{|I(\theta)|}\begin{pmatrix} 2(1-2\theta^2K(\theta)) & -2\theta K(\theta) \\ -2\theta K(\theta) & 1-K(\theta)\end{pmatrix}\begin{pmatrix}2\theta\\2\end{pmatrix} = 4\theta^2+2.$$

On the other hand, letting $a_1(\theta) = 2\theta$, $a_2(\theta) = 1$, $C_0 = \beta^2/\alpha^2 - 1$, $C_1 = -2\beta/\alpha^2$ and $C_2 = 1/\alpha^2$ for $\alpha \neq 0$, from Theorem 3.1 we see that the variance of the estimator $\hat{g}(X)$ is $\mathrm{Var}_\theta(\hat{g}(X)) = 4\theta^2 + 2$. Hence the variance of $\hat{g}(X)$ coincides with the second order Bhattacharyya lower bound $B_2(\theta)$. Therefore $\hat{g}(X)$ is a uniformly minimum variance unbiased (UMVU) estimator of $g(\theta)$. Note that $\hat{g}(X)$ does not attain the first order Bhattacharyya lower bound $B_1(\theta) = 4\theta^2/(1 - K(\theta))$, i.e., the Cramér-Rao lower bound.

Example 4.2. Let $J_x(\theta)$ be the x-th Bessel function of the first kind, that is,

$$J_x(\theta) = \left(\frac{\theta}{2}\right)^x \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+x+1)} \left(\frac{\theta}{2}\right)^{2n}$$

for $x = 0, 1, 2, \ldots$. Note that $J_x(\theta)$ is a particular solution of the Bessel differential equation

$$x^2 - \theta^2 = \theta \frac{\frac{\partial}{\partial\theta} w(x,\theta)}{w(x,\theta)} + \theta^2 \frac{\frac{\partial^2}{\partial\theta^2} w(x,\theta)}{w(x,\theta)}$$

(see, e.g. Abramowitz and Stegun (1965)). Let θ_0 be the smallest positive root of the equation $J_0(\theta) = 0$. For $0 < \theta < \theta_0$, we put

$$f(x,\theta) := \begin{cases} J_0(\theta) & \text{for } x = 0, \\ 2J_{2x}(\theta) & \text{for } x = 1, 2, \ldots. \end{cases}$$

Due to

$$2\sum_{x=1}^{\infty} J_{2x}(\theta) = 1 - J_0(\theta), \tag{4.4}$$

the function $f(x,\theta)$ is a probability mass function (p.m.f.). Then we have

$$I_{11}(\theta) = \frac{J_0'^2(\theta)}{J_0(\theta)} + 2L(\theta), \qquad I_{12}(\theta) = \frac{J_0'(\theta)J_0''(\theta)}{J_0(\theta)} + 2M(\theta),$$
$$I_{22}(\theta) = \frac{J_0''^2(\theta)}{J_0(\theta)} + 2N(\theta),$$

where

$$L(\theta) := \sum_{x=1}^{\infty} \frac{J_{2x}'^{\,2}(\theta)}{J_{2x}(\theta)}, \qquad M(\theta) := \sum_{x=1}^{\infty} \frac{J_{2x}'(\theta)J_{2x}''(\theta)}{J_{2x}(\theta)}, \qquad N(\theta) := \sum_{x=1}^{\infty} \frac{J_{2x}''^{\,2}(\theta)}{J_{2x}(\theta)}.$$

Next, we represent $M(\theta)$ and $N(\theta)$ in terms of $L(\theta)$. Since

$$4x^2 - \theta^2 = \theta \frac{J_{2x}'(\theta)}{J_{2x}(\theta)} + \theta^2 \frac{J_{2x}''(\theta)}{J_{2x}(\theta)}$$

and

$$8\sum_{x=1}^{\infty} x^2 J_{2x}(\theta) = \theta^2, \tag{4.5}$$

we can conclude that

$$M(\theta) = \sum_{x=1}^{\infty} \left(4\frac{x^2}{\theta^2} - 1 - \frac{J'_{2x}(\theta)}{\theta J_{2x}(\theta)} \right) J'_{2x}(\theta) = \frac{1}{\theta} - \frac{J_1(\theta)}{2} - \frac{L(\theta)}{\theta}.$$

Taking into account

$$32 \sum_{x=1}^{\infty} x^4 J_{2x}(\theta) = \theta^4 + 4\theta^2, \tag{4.6}$$

we obtain, in a similar way as above,

$$N(\theta) = -\frac{J_0(\theta)}{2} + \frac{J_1(\theta)}{\theta} + \frac{L(\theta)}{\theta^2},$$

which leads us to

$$I(\theta) = \begin{pmatrix} \dfrac{Q(\theta)}{J_0(\theta)} & \dfrac{2J_0(\theta) - Q(\theta)}{\theta J_0(\theta)} \\ \dfrac{2J_0(\theta) - Q(\theta)}{\theta J_0(\theta)} & \dfrac{Q(\theta)}{\theta^2 J_0(\theta)} \end{pmatrix},$$

where

$$Q(\theta) := J_1^2(\theta) + 2J_0(\theta)L(\theta).$$

Note that (4.4), (4.5) and (4.6) are quoted from Abramowitz and Stegun ((1965), formulas 9.1.46 and 9.1.87). We consider the estimating problem of $g(\theta) = \theta^2$. Then, it follows from (2.1) that the second order Bhattacharyya lower bound $B_2(\theta)$ is $4\theta^2$. Since the estimator $\hat{g}(X) = 4X^2$ is unbiased for $g(\theta)$ and $\mathrm{Var}_\theta(\hat{g}(X)) = 4\theta^2$, the variance of $\hat{g}(X)$ coincides with the second order Bhattacharyya lower bound $B_2(\theta)$, hence $\hat{g}(X)$ is the UMVU estimator of $g(\theta)$. We remark that in this case, the p.m.f. $f(x,\theta)$ is not expressed as a linear combination of distributions belonging to the exponential family, because $a_1(\theta) = \theta$, $a_2(\theta) = \theta^2$ and the condition (ii) of Theorem 3.1 is not satisfied, but the estimator $\hat{g}(X)$ attains the second order Bhattacharyya lower bound $B_2(\theta)$.

Remark 4.1. The term of the Bessel function distribution seems to be conventionally used as distributions of $\sigma_1^2 X_1 \pm \sigma_2^2 X_2$, where X_1 and X_2 are mutually independent random variables, each distributed as χ^2 with same degree of freedom (McKay (1932), Bhattacharyya (1942), and Johnson *et al.* (1994)). The distribution in this example is different from the Bessel function distribution.

Acknowledgements

The authors wish to thank the referees for the comments.

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York.

Bhattacharyya, A. (1946). On some analogues of the information and their use in statistical estimation, *Sankhyā*, **8**, 1–14.

Bhattacharyya, B. C. (1942). The use of McKay's Bessel function curves for graduating frequency distributions, *Sankhyā*, **6**, 175–182.

Blight, J. N. and Rao, P. V. (1974). The convergence of Bhattacharyya bounds, *Biometrika*, **61**, 137–142.

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

Ishii, G. (1976). On the attainment of Bhattacharyya bound, *Chinese J. Math.*, **4**, 37–45.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., Wiley, New York.

Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, 2nd ed., Springer, New York.

McKay, A. T. (1932). A Bessel function distribution, *Biometrika*, **24**, 39–44.

Wijsman, R. A. (1973). On the attainment of the Cramér-Rao lower bound, *Ann. Statist.*, **1**, 538–542.

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