Generalized Efficient Estimates And Its Attainable Parametric Functions

Kakeshita, Shinichi
Kyushu University

https://doi.org/10.5109/13003

出版情報：統計数理研究. 10 (3/4), pp.5-15, 1962-10. Research Association of Statistical Sciences

GENERALIZED EFFICIENT ESTIMATES AND ITS ATTAINABLE PARAMETRIC FUNCTIONS

By

Shin'ichi KAKESHITA

(*Received* April 20th, 1962)

1. Introduction

In this paper the minimum variance unbiased estimates are discussed from the view-point of efficient estimates and its extensions (see Definition 3.1). The problem how to obtain the greatest lower bound for the variances of unbiased estimates was attacked by H. Cramér, C. R. Rao etc. And they obtained what is called as the Cramér-Rao inequality independently. As the base of this inequality, an estimate whose variance attains the lower bound given by Cramér-Rao inequality is called an efficient estimate, but unfortunately this lower bound does not always give the greatest lower bound. So restricting the class of distribution functions of the populations, A. Bhattacharyya [1] obtained the more exact inequality. Furthermore G. R. Seth [9] extended Bhattacharyya's results both to the sequential estimations and to the simultaneous estimations. The problem of finding the functions of parameters for each of which the lower bound given by Bhattacharyya's inequality is attainable when we take a suitable estimate, has not been solved. Y. Washio, H. Morimoto and N. Ikeda [10] solved the problem how to obtain the estimate of given parametric functions by making use of the theories of sufficient statistics and Laplace transforms and T. Kitagawa [3] discussed the problem from the standpoint of linear translatable operations. J. Neyman and E. Scott [7] obtained some results on the minimum variance unbiased estimates of transformed vatiables where the transformation is the second order entire function introduced by them and L. Schmetterer [8] extended their results. However the actual calculations of estimates in [10] are very difficult because of the difficulty of the inverse transforms of Laplace transforms. In this point, the calculation of the parametric functions and its estimates which attain the lower bound given by A. Bhattacharyya making use of the conditions of the attainment of the lower bound by A. Bhattacharyya [1] are easily obtained.

The main results in this paper are as follows,

(1) Characterization of the class of the parametric functions which have the N-th order efficient estimates and of the N-th order efficient estimate

(Theorem 3.2).

As the special case of (1), we can state,

(2) The class of parametric functions which have the (first order) efficient estimates is that of linear functions of expectation (Corollary of Theorem 3.2).

(3) Two examples of Theorem 3.2 are discussed in § 4.

(a) As to the parametric function h^2 of the parameter h of the Poisson distributions,

$$p_h(x=\nu)=\frac{e^{-h}h^{\nu}}{\nu!} \qquad (\nu=0,1,2,\cdots)$$

the (first order) efficient estimate does not exist, but the second order efficient estimate exists and its explicit form is given.

(b) For the parametric function θ^k of the parameter θ in normal distribution $N(\theta,1)$ for any assigned natural number k, the k-th order efficient estimate exists and its explicit form is given.

In the preparation of this paper, the author owed much to the discussion with Professor T. Kitagawa, Mr. T. Seguchi and other members of Kyushu Univ. I wish to thank them for their valuable advices.

§ 2. Notations and known results

In this section we shall introduce the notations to be used in this paper and the known results, some of which will be either elaborated or extended in this paper.

Definition 2.1. *If a statistic $T(x_1, x_2, \cdots, x_n)$ is an unbiased estimate of a parametric function $g(\theta)$, that is, the expectation of $T(x_1, x_2, \cdots, x_n)$ is $g(\theta)$, and we have, for any unbiased estimate $U(x_1, x_2, \cdots, x_n)$ and any θ in D,*

$$V_\theta[T(x_1, x_2, \cdots, x_n)] \leqq V_\theta[U(x_1, x_2, \cdots, x_n)] \tag{2.1}$$

then we call the statistic $T(x_1, x_2, \cdots, x_n)$ is the best estimate of $g(\theta)$.

We use a symbol

$$f_\theta(x_1, x_2, \cdots, x_n) \tag{2.2}$$

as the joint probability density functions with respect to a certain σ-finite measure μ. We do not always assume the independence of random variables $x_1, x_2, \cdots, x_n$ in this section.

Definition 2.2. *The joint probability density function $f_\theta(x_1, x_2, \cdots, x_n)$ is said to satisfy the N-th order regularity condition, if the following five conditions are satisfied.*

(1) *The domain of θ, D, is an open interval on the real line, which may be infinite or semi-infinite interval.*

(2) *For any* $\theta \in D$ *and almost all* $(x_1, x_2, \cdots, x_n)$ *with respect to* μ-*measure*

$$\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n) \quad i=1, 2, \cdots, N \tag{2.3}$$

exist.

(3) $$\frac{d^i}{d\theta^i}\int_{R^n} f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$$=\int_{R^n} \frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n) \quad i=1, 2, \cdots, N \tag{2.4}$$

for all $\theta \in D$, *where* R^n *denotes the n dimensional Euclidean space.*

(4) $$J_{ij}=E_\theta\left[\frac{\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)\frac{\partial^j}{\partial\theta^j} f_\theta(x_1, x_2, \cdots, x_n)}{(f_\theta(x_1, x_2, \cdots, x_n))^2}\right]$$

exist for $i, j=1, 2, \cdots, N$ *and for all* $\theta \in D$ *and further*

$$|J_{ij}| \neq 0 \tag{2.5}$$

for all $\theta \in D$, *where* $|J_{ij}|$ *denotes the* $N\times N$ *determinante whose* $i-j$ *component is* J_{ij}.

(5) *If the expectations of* $T(x_1, x_2, \cdots, x_n)$ *exist for all* $\theta \in D$, *then*

$$\frac{d^i}{d\theta^i}\int_{R^n} T(x_1, x_2, \cdots, x_n)\, f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n) \tag{2.6}$$

$$=\int_{R^n} T(x_1, x_2, \cdots, x_n)\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$(i=1, 2, \cdots, n)$ *for all* $\theta \in D$.

Our N-th order regularity condition just given amounts to the assumptions given Bhattacharyya [1] under which he obtained the lower bound, although he did not given them explicitly.

Inverse matrix of (J_{ij}) is denoted by (J^{ij}), that is

$$\begin{pmatrix} J_{11} \cdots J_{1N} \\ J_{N1} \cdots J_{NN} \end{pmatrix}^{-1} = \begin{pmatrix} J^{11}, \cdots J^{1N} \\ J^{N1} \cdots J^{NN} \end{pmatrix} \tag{2.7}$$

The following theorem which gives a lower bound of variances of estimates is well known.

Theorem 2.1. (Cramér-Rao) *If* $T(x_1, x_2, \cdots, x_n)$ *is an unbiased estimate of* $g(\theta)$ *and the class of distributions satisfies the first order regularity condition, then*

$$V_\theta[T(x_1, x_2, \cdots, x_n)] \geq \frac{\left[\frac{d}{d\theta} g(\theta)\right]^2}{E_\theta\left[\frac{\frac{\partial}{\partial\theta} f_\theta(x_1, x_2, \cdots, x_n)}{f_\theta(x_1, x_2, \cdots, x_n)}\right]^2}. \tag{2.8}$$

The equality in (2.8) *holds if and only if the equality*

$$(2.9) \qquad T(x_1, x_2, \cdots, x_n) = g(\theta) + \frac{\frac{\partial}{\partial\theta} f_\theta(x_1, x_2, \cdots, x_n)}{f_\theta(x_1, x_2, \cdots, x_n)} J^{11}$$

holds in probability one, that is, the equality holds almost everywhere in R^n *with respect to measure* $f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$ *for all* θ. *(see* [2] *and* [4]).

A generalization of this Theorem is given by

Theorem 2.2. (A. Bhattacharyya [1]) *If* $T(x_1, x_2, \cdots, x_n)$ *is an unbiased estimate of* $g(\theta)$ *and the class of distributions satisfies the N-th order regularity condition, then*

$$(2.10) \qquad V_\theta[T(x_1, x_2, \cdots, x_n)] \geqq \sum_{ij=1}^{N} \frac{d^i}{d\theta^i} g(\theta) \frac{d^j}{d\theta^j} g(\theta) J^{ij}.$$

The equality in (2.10) *holds if and only if the equality*

$$(2.11) \qquad T(x_1, x_2, \cdots, x_n) = g(\theta) + \sum_{ij=1}^{N} \frac{\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n) \frac{d^j}{d\theta^j} g(\theta)}{f_\theta(x_1, x_2, \cdots, x_n)} J^{ij}$$

holds in probability one.

Definition 2.3. *When a parametric function* $g(\theta)$ *and an unbiased estimate* $T(x_1, x_2, \cdots, x_n)$ *of* $g(\theta)$ *which attains the lower bound given by the Cramér-Rao inequality, exist, the estimate* $T(x_1, x_2, \cdots, x_n)$ *is called an efficient estimate of* $g(\theta)$.

Theorem 2.3. (A. Bhattacharyya [1]) *If the efficient estimate exist, then the probability density function is the form of next type.*

$$(2.12) \qquad f_\theta(x_1, x_2, \cdots, x_n) = e^{A(\theta)T(x_1, x_2, \cdots, xn) + B(\theta) + X(x_1, x_2, \cdots, xn)},$$

where $A(\theta)$, $B(\theta)$ *are functions of* θ *and independent of* $x_1, x_2, \cdots, x_n$, *and* $X(x_1, x_2, \cdots, x_n)$ *is a function of* $x_1, x_2, \cdots, x_n$ *only.*

§ 3. *N*-th order efficient estimates

Definition 3.1. *If a parametric function* $g(\theta)$ *and the statistic* $T(x_1, x_2, \cdots, x_n)$ *achieve the N-th order lower bound given by Bhattacharyya's inequality* (2.10) *but not the k-th lower bound in* (2.10) *for at least one* θ_k *in D, for each k in* $1 \leqq k < N$, *then the statistic* $T(x_1, x_2, \cdots, x_n)$ *is called an N-th order efficient estimate of* $g(\theta)$.

Theorem 3.1. *If a statistic* $T(x_1, x_2, \cdots, x_n)$ *has the expectation, the class of distributions satisfies the N-th order regularity condition and further more*

$$(3.1)\qquad \sum_{i=1}^{N} a_i(\theta)\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n) = [T(x_1, x_2, \cdots, x_n) - g(\theta)]\, f_\theta(x_1, x_2, \cdots, x_n)$$

then $T(x_1, x_2, \cdots, x_n)$ *is an unique best estimate of* $g(\theta)$ *where the uniqueness means that if two best estimate exist, then they are equal in probability one.*

Proof. First let us prove the unbiasedness of $T(x_1, x_2, \cdots, x_n)$. Integrating the both side of (3.1) with respect to μ-measure on R^n, we have

$$(3.2)\qquad \int_{R^n} \sum_{i=1}^{N} a(\theta)\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$$= \int_{R^n} [T(x_1, x_2, \cdots, x_n) - g(\theta)] f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

According to the regularity condition, the differentiation with respect to θ and the integration with respect to μ measure can be exchanged, and hence the left hand side of (3.2) equals zero, consequently we have

$$(3.3)\qquad \int_{R^n} T(x_1, x_2, \cdots, x_n) f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n) = g(\theta)$$

The equation 3.3 means $T(x_1, x_2, \cdots, x_n)$ is an unbiased estimate of $g(\theta)$.

Secondly the minimum variance property will be proved. Let $U(x_1, x_2, \cdots, x_n)$ be any other unbiased estimate of $g(\theta)$, then

$$(3.4)\qquad \int_{R^n} [T(x_1, x_2, \cdots, x_n) - U(x_1, x_2, \cdots, x_n)] f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n) = 0.$$

According to (5) of the regularity condition and (3.1)

$$(3.5)\qquad \sum_{i=1}^{N} a_i(\theta)\frac{d^i}{d\theta^i}\int_{R^n} [T(x_1, x_2, \cdots, x_n) - U(x_1, x_2, \cdots, x_n)] f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$$= \int_{R^n} [T(x_1, x_2, \cdots, x_n) - U(x_1, x_2, \cdots, x_n)] \sum_{i=1}^{N} a_i(\theta)\frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$$= \int_{R^n} [T(x_1, x_2, \cdots, x_n) - U(x_1, x_2, \cdots, x_n)][T(x_1, x_2, \cdots, x_n) - g(\theta)]\, f_\theta(x_1, x_2, \cdots, x_n)\, d\mu(x_1, x_2, \cdots, x_n)$$

$$= 0.$$

Here using the relation of the variance and the covariance, each of which we can write V and *Cov* respectively,

$$(3.6)\qquad V[U(x_1, x_2, \cdots, x_n)]$$

$$=V[U(x_1, x_2, \cdots, x_n) - T(x_1, x_2, \cdots, x_n) + T(x_1, x_2, \cdots, x_n)]$$
$$=V[U(x_1, x_2, \cdots, x_n) - T(x_1, x_2, \cdots, x_n)] + 2Cov[U(x_1, x_2, \cdots, x_n) - T(x_1, x_2, \cdots, x_n), T(x_1, x_2, \cdots, x_n)] + V[T(x_1, x_2, \cdots, x_n)]$$

we obtain a relation

$$(3.7)\quad V[U(x_1, x_2, \cdots, x_n)] = V[U(x_1, x_2, \cdots, x_n) - T(x_1, x_2, \cdots, x_n)] + V[T(x_1, x_2, \cdots, x_n)] \geqq V[T(x_1, x_2, \cdots, x_n)].$$

The equation (3.7) means $T(x_1, x_2, \cdots, x_n)$ is a best estimate of $g(\theta)$.

Finally the proof of the uniqueness is evident because (3.7) comes an equality if and only if

$$(3.8)\qquad U(x_1, x_2, \cdots, x_n) = T(x_1, x_2, \cdots, x_n)$$

in probability one, which completes the proof.

Now the main theorem of this paper is given by

Theorem 3.2. *Let us assume that $x_1, x_2, \cdots, x_n$ are independent random variables, each having the same probability density function $e^{\theta x+\beta(\theta)+v(x)}$ and further more that $\frac{d}{dx}v(x)$, $\frac{d}{d\theta}\beta(\theta)$, each of which are differentiation of x and θ respectively, are well defined. If we put*

$$(3.9)\qquad g_0(\theta) = 1$$
$$g_{k+1}(\theta)e^{-n\beta(\theta)} = \frac{d}{d\theta} g_k(\theta)e^{-n\beta(\theta)} \quad (k=0, 1, 2, \cdots)$$

then the necessary and sufficient condition that the parametric function $g(\theta)$ have the N-th order efficient estimate is

$$(3.10)\qquad g(\theta) = \sum_{k=0}^{N} A_k g_k(\theta) \quad (A_N \neq 0)$$

and the N-th order efficient estimate of (3.10) *is given by*

$$(3.11)\qquad T(x_1, x_2, \cdots, x_n) = \sum_{k=0}^{N} A_k \left(\sum_{i=1}^{n} x_i\right)^k.$$

Proof. According to (2.11) if $T(x_1, x_2, \cdots, x_n)$ is an N-th order efficient estimate of $g(\theta)$ then

$$(3.12)\qquad [T(x_1, x_2, \cdots, x_n) - g(\theta)]f_\theta(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{N}\left[\sum_{j=1}^{N} \frac{d^j}{d\theta^j} g(\theta)\, J^{ij}\right] \frac{\partial^i}{\partial\theta^i} f_\theta(x_1, x_2, \cdots, x_n)$$

Considering $\sum_{j=1}^{N} \frac{d^j}{d\theta^j} g(\theta) J^{ij} = a_i(\theta)$, we obtain (3.1). This means that the class of all statistics satisfying (3.1) contains the class of N-th order efficient estimates of $g(\theta)$.

(3.13)
$$f_\theta(x_1, x_2, \cdots, x_n) = e^{\theta \sum_{i=1}^{n} x_i + n\beta(\theta) + \sum_{i=1}^{n} v(x_i)}$$
$$= b(\theta)\, e^{\theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} v(x_i)}$$

where $b(\theta) = e^{n\beta(\theta)}$

(3.14)
$$\frac{\partial^l}{\partial \theta^l} f_\theta(x_1, x_2, \cdots, x_n)$$
$$= e^{\theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} v(x_i)} \left[\sum_{k=0}^{l} {}_lC_k\, b^{(l-k)}(\theta) \left(\sum_{i=1}^{n} x_i \right)^k \right] \quad l = 1, 2, \cdots, N$$

(3.15)
$$\sum_{i=1}^{N} a_i(\theta) \frac{\partial^i}{\partial \theta^i} f_\theta(x_1, x_2, \cdots, x_n)$$
$$= e^{\theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} v(x_i)} \left[\sum_{k=1}^{N} \left(\sum_{i=k}^{N} a_i(\theta)\, {}_iC_k\, b^{(i-k)}(\theta) \right) \left(\sum_{i=1}^{n} x_i \right)^k + \sum_{i=1}^{N} a_i(\theta) b^{(i)}(\theta) \right]$$
$$= [T(x_1, x_2, \cdots, x_n) - g(\theta)]\, b(\theta)\, e^{\theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} v(x_i)}$$

Consequently, we obtain

(3.16)
$$T(x_1, \cdots, x_n) - g(\theta)$$
$$= \sum_{k=1}^{N} \left(\sum_{i=k}^{N} a_i(\theta)\, {}_iC_k \frac{b^{(i-k)}(\theta)}{b(\theta)} \right) \left(\sum_{i=1}^{n} x_i \right)^k + \sum_{i=1}^{N} a_i(\theta) \frac{b^{(i)}(\theta)}{b(\theta)}$$

and further

(3.17)
$$g(\theta) = - \sum_{i=1}^{N} a_i(\theta) \frac{b^{(i)}(\theta)}{b(\theta)} + A_0$$

(3.18)
$$T(x_1, \cdots, x_n) = \sum_{k=1}^{N} \left(\sum_{i=k}^{N} a_i(\theta)\, {}_iC_k \frac{b^{(i-k)}(\theta)}{b(\theta)} \right) \left(\sum_{i=1}^{n} x_i \right)^k + A_0$$

must be established. According to the independence of θ in the right hand side of (3.18), we have

(3.19)
$$\sum_{i=k}^{N} a_i(\theta)\, {}_iC_k \frac{b^{(i-k)}(\theta)}{b(\theta)} = A_k \qquad (k = 1, 2, \cdots, N)$$

(3.20)
$$T(x_1, \cdots, x_n) = \sum_{k=0}^{N} A_k \left(\sum_{i=1}^{n} x_i \right)^k .$$

Thus we obtain that the statistic satisfying (3.1) is written in the form (3.20).

Conversely, it will be proved that if $A_N \neq 0$ in (3.20) then the statistic defined (3.20) is the N-th order efficient estimate and if $A_N=A_{N-1}=\cdots=A_{M+1}=0$ and $A_M\neq 0$, then (3.20) gives an M-th order efficient estimate. According to [1] it is the sufficient condition for (2.10) comes an equality that

$$(3.21)\qquad T(x_1,\cdots,x_n)-g(\theta)=\sum_{i=1}^{N}\lambda_i(\theta)\frac{\frac{\partial^i}{\partial\theta^i}f_\theta(x_1,\cdots,x_n)}{f_\theta(x_1,\cdots,x_n)}$$

for some suitable $\lambda_i(\theta)$, $i=1,2,\cdots,N$.
So we calculate the (3.21) according to (3.14) and (3.20) thus

$$(3.22)\qquad \sum_{k=0}^{N}A_k\Big(\sum_{i=1}^{n}x_i\Big)^k-g(\theta)-\sum_{i=1}^{N}\lambda_i(\theta)\sum_{k=0}^{i}{}_iC_k\frac{b^{(i-k)}(\theta)}{b(\theta)}\Big(\sum_{i=1}^{n}x_i\Big)^k=0$$

$$\sum_{k=0}^{N}A_k\Big(\sum_{i=1}^{n}x_i\Big)^k-g(\theta)-\sum_{k=1}^{N}\Big(\sum_{i=k}^{N}\lambda_i(\theta)\,{}_iC_k\frac{b^{(i-k)}(\theta)}{b(\theta)}\Big)\Big(\sum_{i=1}^{n}x_i\Big)^k-\sum_{i=1}^{N}\lambda_i(\theta)\frac{b^{(i)}(\theta)}{b(\theta)}$$

$$=0$$

Thus if

$$(3.23)\qquad A_k=\sum_{i=k}^{N}\lambda_i(\theta)\;{}_iC_k\frac{b^{(i-k)}(\theta)}{b(\theta)}\quad (k=1,2,\cdots,N)$$

then, (3.20) is the N-th order efficient estimate of $g(\theta)$ and (3.23) is easily verified $a_i(\theta)$ as $\lambda_i(\theta)$ $i=1,2,\cdots,N$.

When $A_n=A_{N-1}=\cdots=A_{M+1}=0$ and $A_M\neq 0$, we can similarly prove (3.20) is an M-th order efficient estimate.
$g_k(\theta)$ defined by (3.9) is also written

$$g_k(\theta)=\int_{R^n}\Big(\sum_{i=1}^{n}x_i\Big)^k e^{\theta\sum_{i=1}^{n}x_i+n\beta(\theta)+\sum_{i=1}^{n}v(x_i)}d\mu(x_1)\cdots d\mu(x_n)$$

as easily verified. Thus the only parametric functions of type

$$g(\theta)=\sum_{k=0}^{N}A_kg_k(\theta)\qquad (A_N\neq 0)$$

have the N-th order efficient estimate.

Corollary. *Under the assumptions to Theorem* 3.2 *the necessary and sufficient condition for the parametric function $g(\theta)$ have the (first order) efficient estimate, is that $g(\theta)$ is a linear function of the expectation of the distribution, that is*

$$g(\theta)=a\frac{d}{d\theta}\beta(\theta)+b$$

where a, b are any given real numbers.

§ 4. Applications

In this section we see two examples as applications of the result of this paper.

Example 1. (Poisson distribution)

$$p_h(x=\nu)=\frac{h^\nu e^{-h}}{\nu!} \qquad (\nu=0, 1, 2, \cdots) \tag{4.1}$$

Notice that if we consider $\theta=\log h$ then (4.1) are one of the exponential type of distributions that are defined in § 3. We consider the best estimate of h^2 in (4.1). We can not obtain the gratest lower bound by the Cramér-Rao inequality (2.8) according to the Corollary of Theorem 3.2. The lower bound of Cramér-Rao inequality is $\frac{4h^3}{n}$ [4].

$$\begin{aligned} g_0(\theta)&=1 \\ g_1(\theta)&=nh \\ g_2(\theta)&=n^2h^2+nh \end{aligned} \tag{4.2}$$

Therefore if we take a suitable linear combination of $g_0(\theta)$, $g_1(\theta)$ and $g_2(\theta)$, then h^2 is expressed by the form (3.10), that is

$$h^2=0g_0(\theta)-\frac{1}{n^2}g_1(\theta)+\frac{1}{n^2}g_2(\theta). \tag{4.3}$$

According to the Theorem 3.2, the best estimate of h^2 is given by the second order effinient estimate as follow

$$T(x_1, \cdots, x_n)=\frac{\left(\sum_{i=1}^{n}x_i\right)^2-\sum_{i=1}^{n}x_i}{n^2}. \tag{4.4}$$

In this case the second order Bhattacharyya's inequality is used, that is

$$V_\theta[T(x_1, \cdots, x_n)]=\frac{\left(\frac{d}{d\theta}g(\theta)\right)^2}{J_{11}}+\frac{\left(\frac{d}{d\theta}g(\theta)J_{12}-\frac{d^2}{d\theta^2}g(\theta)J_{11}\right)^2}{J_{11}\,(J_{11}\cdot J_{22}-J_{12}^2)} \tag{4.5}$$

see [1]. The variance of (4.4) and the right hand side of (4.5) are equal and these are

$$\frac{2h^2}{n^2}+\frac{4h^3}{n}. \tag{4.6}$$

Example 2. (Normal distribution $N(\theta, 1)$)

Assume that $x_1, \cdots, x_n$ are independent random variables each having the same probability density function $\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2}}$ with respect to Lebesgue measure.

Statistic $U=\sum_{i=1}^{n}x_i$ obeys the distribution $N(n\theta, n)$

$$E_\theta[U^k]=n^{k/2}k!\sum_{i=0}^{[k/2]}\frac{(\sqrt{n}\theta)^{k-2i}}{2^i i!(k-2i)!} \quad (k=0,1,2,\cdots). \tag{4.7}$$

Thus for $k=1,2,\cdots$, θ^k is expressed as a linear combination of $E_\theta[U^j]$ $j=0, 1,\cdots,k$, that is

$$\theta^k=\frac{1}{n^{k/2}}k!\sum_{i=0}^{[k/2]}\frac{(-1)\frac{E_\theta[U^{k-2i}]}{(\sqrt{n})^{k-2i}}}{2^i\ i!(k-2i)!} \quad (k=1,2,\cdots). \tag{4.8}$$

According to the Theorem 3.2

$$\frac{1}{n^{k/2}}k!\sum_{i=0}^{[k/2]}\frac{(-1)^i\left(\frac{\sum_{\nu=1}^{n}x_\nu}{\sqrt{n}}\right)^{k-2i}}{2^i\ i!(k-2i)!} \tag{4.9}$$

$$=\left(\frac{1}{\sqrt{n}}\right)^k H_k\left(\frac{\sum_{\nu=1}^{n}x_\nu}{\sqrt{n}}\right) \quad (k=1,2,\cdots)$$

is the k-th order efficient estimate of θ^k, where $H_k(x)$ is Hermite polynomial, that is

$$H_k(x)=(-1)^k e^{x^2/2}\frac{d^k}{dx^k}\ e^{-x^2/2}$$

$$=k!\sum_{i=1}^{[k/2]}\frac{(-1)^i x^{k-2i}}{2^i i!\,(k-2i)!} \quad (k=0,1,2,\cdots)$$

and the variance of (4.9) is given

$$\sum_{i=1}^{k}[{}_kC_i\sqrt{(k-1)!}\ \theta^i(\sqrt{n})^{i-k}]^2-\theta^{2k}.$$

References

[1] A. BHATTACHARYYA: *On some analogues of the amount of information and their use in statistical estimation,* Sankya vol. **8** (1946). 1–14.
[2] H. CRAMÉR: *Mathematical methods of statistics,* Princeton Univ. press (1946).
[3] T. KITAGAWA: *The operational calculas and the estimations of parameter admitting sufficient statistics,* Bull. Math. Stat. Vol. **6** (1956). 95–108.
[4] E. L. LEHMANN: *Notes on the theory of estimation,* Univ. of Cal, (1950).
[5] S. MORIGUCHI: *Tôkei Kaiseki,* Iwanami Shoten (1957).
[6] H. MORIMOTO: *Jûsoku tôkeiryo ni tsuite,* Lecture on Symposium of Japan Math. Soc. (1958).

[7] J. Neyman and E. J. *Scott.: Correction for bias introduced by a transformation of variables,* Ann. Math. Stat. Vol. **31** (1960). 643-655.
[8] L. Schmetterer: *On a problem of J. Neyman and E. Scott,* Ann. Math. Stat. Vol. **31** (1960). 656-661.
[9] G. R. Seth: *On the variance of estimates,* Ann. Math. Stat. Vol. **20** (1949). 1-27.
[10] Y. Washio, H. Morimoto and N. Ikeda: *Unbiased estimation based on sufficient statistics,* Bull. Math. Stat. Vol. **6** (1956). 69-93.

Kyushu University