A Remark on the Efficient Estimation[1]

Hirokichi Kudō (工藤弘吉)

Department of Mathematics, Faculty of Science,
Ochanomizu University

In the classical theory of estimation, the efficiency of an unbiased estimate α^* depends on the parameter α chosen in the frequency function (fr. f.), so long as we define it as $\{nE(\alpha^*-\alpha)^2 \cdot E(\partial \log f(x,\alpha)/\partial\alpha)^2\}^{-1}$ in the one parameter case.[2] For example, in the case of the normal distribution with mean 0, and variance σ^2, there exists an efficient unbiased estimate for σ^2, but not for σ. This situation cannot be justified very well from the viewpoint where we are interested in the estimation of the distribution itself, but not of parameters. In this paper we shall give a condition to be able to choose a parameter admitting an efficient unbiased estimate, and further show the uniqueness of such a parameter under this condition. Throughout this paper the discussions will be restricted, for simplicity, to the case of one parameter and of continuous type.

1. Let A be a non-degenerate interval, and $X_1, X_2, \cdots, X_n$ a sample of size n from a population with a fr. f. $f(x;\alpha)$, which is derivable with respect to α for all α and x, and for which there exist two integrable functions $F_1(x)$, and $F_2(P)$ such that

$$\left|\frac{\partial f(x;\alpha)}{\partial\alpha}\right| < F_1(x), \qquad \text{and} \qquad \left|\frac{\partial L(P;\alpha)}{\partial\alpha}\right| < F_2(P)$$

over the whole space of the variable x and of $P=(x_1, x_2, \cdots, x_n)$ respectively, where $L(P;\alpha)=\Pi_{i=1}^n f(x_i;\alpha)$ is a likelihood function. And suppose that

$$\int_{-\infty}^{\infty}\left(\frac{\partial \log f(x;\alpha)}{\partial\alpha}\right)^2 f(x;\alpha)\,dx < \infty .$$

Further, suppose that a statistic $\alpha^*(P)$ of a sample P in the sample space R_n has a finite mean $\psi(\alpha)=E_\alpha(\alpha^*)=\int_{R_n}\alpha^*(P)\,L(P;\alpha)\,dP$, with its derivative $d\psi(\alpha)/d\alpha$, and also a finite $E_\alpha(\alpha^{*2})$.

Our principal tool will be the following two lemmas, which are modifications of a theorem in H. Cramér's book 1), page 479.

---

1 Contribution from Department of Mathematics, Faculty of Science, Ochanomizu University, No. 9

2 This definition of the efficiency will be found on page 482 in 1).

Lemma 1. (i) *For any function $\phi(\alpha)$ of a parameter α, the mean square deviation of a statistic $\alpha^*(P)$ from the function $\phi(\alpha)$ of the true value α satisfies the inequality*

$$(1)\qquad E_\alpha\big(\alpha^*(P)-\phi(\alpha)\big)^2 \geqq \frac{\left(\dfrac{d\psi(\alpha)}{d\alpha}\right)^2}{n\displaystyle\int_{-\infty}^{\infty}\left(\frac{\partial \log f(x;\alpha)}{\partial\alpha}\right)^2\cdot f(x;\alpha)\,dx}.$$

(ii) *The sign of equality holds here, for every α in A, if and only if*

$$(2)\qquad \frac{\partial \log L(P;\alpha)}{\partial\alpha}=k(\alpha)\cdot\big(\alpha^*(P)-\phi(\alpha)\big)$$

for all points P in $E_\alpha-N_\alpha$, where $E_\alpha=\{P; L(P;\alpha)>0\}$, and N_α is a set of measure 0 which depends on α.

The proof of this lemma can be carried out in the analogous way to the one of Lemma 1 of 1), page 475, and hence is omitted here.

Lemma 2. *Suppose that $\mathfrak{B}$ is a set of all values α in A where $\partial \log L(P;\alpha)/\partial\alpha$ is constant at almost all $P\in E_\alpha$. If (2) holds for almost all $P\in E_\alpha$, and if $\partial f(x;\alpha)/\partial\alpha$ is a continuous function of α whenever $f(x;\alpha)>0$, then $\mathfrak{A}=A-\mathfrak{B}$ is an open set and $k(\alpha)$ and $\phi(\alpha)$ must be continuous functions on $\mathfrak{A}$. Here we can choose N_α in (ii) of the above Lemma 1 independently of $\alpha\in\mathfrak{A}$.*

Proof. Denote the set $\{\alpha; L(P;\alpha)>0\}$ be B_P, and the set $\{P; L(P;\alpha)>0\}$ be E_α, then B_P is an open subset of A because of the continuity of $L(P;\alpha)$ with respect to α for all P. Moreover, by the assumption $\mathfrak{A}$ is the set of the values α, for which there exist two disjoint subsets $\varDelta_1$, $\varDelta_2$, with positive measures, of E_α, such that

$$(3)\qquad \inf_{P\in\varDelta_1}\frac{\partial \log L(P;\alpha)}{\partial\alpha}>\sup_{P\in\varDelta_2}\frac{\partial \log L(P;\alpha)}{\partial\alpha}.$$

This condition is also equivalent to that there exist two disjoint subsets D_1, D_2, with positive measures, of E_α and two real numbers $a<b$, such that

$$(4)\qquad \inf_{P\in D_1}\frac{\partial \log L(P;\alpha)}{\partial\alpha}>b>a>\sup_{P\in D_2}\frac{\partial \log L(P;\alpha)}{\partial\alpha}.$$

Therefore, because of the continuity of $\partial \log L(P;\alpha)/\partial\alpha$, we can see that $\mathfrak{A}$ is an open subset of A. In fact, let D_1 and D_2 be the sets with positive measures, for which (4) holds for an α_0 in $\mathfrak{A}$, and put

$$\varDelta_n=\left\{P;\left|\frac{\partial \log L(P;\alpha)}{\partial\alpha}-\frac{\partial \log L(P;\alpha_0)}{\partial\alpha}\right|<\frac{b-a}{2},\right.$$
$$\left.\textit{and } P\in E_\alpha \textit{ for all } \alpha \textit{ satisfying } |\alpha-\alpha_0|<\frac{1}{n}\right\}\cap D_1$$

for every positive integer n, then $\bigcup_{n=1}^{\infty} \Delta_n = D_1$, and hence there is an n for which Δ_n has a positive measure. Therefore we can see that, for all α in $U' = \left\{\alpha ; |\alpha - \alpha_0| < \frac{1}{n}\right\}$,

$$\inf_{P \in \Delta_n} \frac{\partial \log L(P ; \alpha)}{\partial \alpha} > \inf_{P \in D_1} \frac{\partial \log L(P ; \alpha_0)}{\partial \alpha} - \frac{b-a}{2} > \frac{a+b}{2}.$$

Similarly we can see that there is a neighborhood U'' of α_0, such that for every α in U''

$$\frac{a+b}{2} > \sup_{P \in \Delta'} \frac{\partial \log L(P ; \alpha)}{\partial \alpha},$$

where Δ' is a subset, with positive measure, of D_2. Therefore $U = U' \cap U''$ is contained in $\mathfrak{A}$.

Now let A' be a denumerable subset everywhere dense in A, and write $N = \bigcup_{\alpha_i \in A'} N_{\alpha_i}$, then the equation (2) holds for every α_i in A' and for all P in $E_\alpha - N$, and hence $k(\alpha)$ and $\phi(\alpha)$ are continuous on $A' \cap \mathfrak{A}$. Therefore the same results hold over the whole $\mathfrak{A}$ by the continuity of $\partial \log L(P ; \alpha)/\partial \alpha$ as a function of α. Thus the lemma is proved.

The following theorem is fundamental in this paper.

Theorem I. *If $\partial \log L(P ; \alpha)/\partial \alpha$ is a continuous function of α whenever $L(P ; \alpha) > 0$, then it is a necessary and sufficient condition for existence of a statistic $\alpha^*(P)$, for which the sign of equality in* (1) *holds with a continuous function $\phi(\alpha)$ for all α in an open interval $I \subset \mathfrak{A}$, that for all $\alpha \in I$ the distribution of the universe has a form*

$$(5) \qquad \begin{aligned} f(x ; \alpha) &= \exp(p(\alpha) u(x) + q(\alpha) + r(x)) \quad \textit{in a set } \pi_\alpha, \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \textit{outside } \pi_\alpha, \end{aligned}$$

where π_α has a positive measure, $p(\alpha)$ and $q(\alpha)$ are continuously differentiable, and the relation

$$(6) \qquad q'(\alpha) = \phi(\alpha) p'(\alpha)$$

holds.

Proof. Necessity. From Lemma 1 and 2, indefinite integrals of the both sides of the equation (2) exist and are equal each other for all points $P \notin N$. Thus we have

$$(7) \qquad \log L(P ; \alpha) = np(\alpha) \alpha^*(P) + nq(\alpha) + r^*(P)$$

for all $P \notin N$, and $\alpha \in B_P$, where

$$(8) \qquad np(\alpha) = \int k(\alpha)\, d\alpha, \qquad nq(\alpha) = \int k(\alpha) \phi(\alpha)\, d\alpha,$$

and $r^*(P)$ is an integration constant. It follows that (6) holds.

Now let $R_{(i)}$ be the one-dimensional space of the i-th coordinate of the sample space R_n, and $R^{(i)}_{n-1}$ the $(n-1)$-dimensional space of the other rest ones of R_n, whose elements will be denoted by $Q=(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$. We shall take notations $\pi_\alpha=\{x;\ f(x;\alpha)>0\}$ and $\pi^{(i)}_\alpha=\underbrace{\pi_\alpha\times\cdots\times\pi_\alpha}_{n-1}$ in $R_{(i)}$, $R^{(i)}_{n-1}$ respectively, and since $L(P;\alpha)=\Pi^n_{i=1} f(x_i;\alpha)$, it is evident that $E_\alpha=\underbrace{\pi_\alpha\times\cdots\times\pi_\alpha}_{n}$. Moreover we shall denote $P=(x_1, \cdots, x_n)$ by (x_i, Q). By Fubini's theorem, the one-dimensional measure of the set $N_{i,Q}=\{x_i;\ (x_i, Q)\in N\}$ is zero for every $Q\in R^{(i)}_{n-1}$, except at most a set N_i of $(n-1)$-dimensional measure zero. Therefore

(9) $\log f(x_i;\alpha)=np(\alpha)\alpha^*(x_i;Q_i)+nq(\alpha)-\sum\limits_{j\neq i}\log f(x^{(i)}_j;\alpha)+r^*(x_i;Q_i)$

holds for all x_i in $\pi_\alpha-N_{i,Q_i}$, if $Q_i=(x^{(i)}_1, \cdots, x^{(i)}_{i-1}, x^{(i)}_{i+1}, \cdots, x^{(i)}_n)$ is fixed in $\pi^{(i)}_\alpha-N_i$. Thus, by noting that $\sum\limits_{i=1}^{n}\sum\limits_{j=1,j\neq i}^{n}\log f(x^{(i)}_j;\alpha)=\sum\limits_{i=1}^{n-1}\Big\{\sum\limits_{j=1,j\neq i}^{n}\log f(x^{(i)}_j;\alpha)+\log f(x^{(n)}_i;\alpha)\Big\}=\sum\limits_{i=1}^{n-1}\log L(P_i;\alpha)$, we have for all $P=(x_1, x_2, \cdots, x_n)\in E_a$,

$$(10)\qquad \begin{aligned}\log L(P;\alpha)&=\sum_{i=1}^{n}\log f(x_i;\alpha)\\ &=np(\alpha)\sum_{i=1}^{n}\alpha^*(x_i,Q_i)+n^2q(\alpha)-\sum_{i=1}^{n-1}\log L(P_i;\alpha)+\sum^{n} r^*(x_i,Q_i),\end{aligned}$$

where $P_i=(x^{(n)}_i, Q_i)$. After $n-1$ points $Q_1, Q_2, \cdots, Q_{n-1}$ are arbitrarily chosen in $\pi^{(1)}_\alpha-N_1$, $\pi^{(2)}_\alpha-N_2, \cdots, \pi^{(n-1)}_\alpha-N_{n-1}$, respectively, we take a point Q_n in $(\pi_\alpha-N_{1,Q_1})\times\cdots\times(\pi_\alpha-N_{n-1,Q_{n-1}})-N_n$. By such a selection of $Q_1, Q_2, \cdots, Q_n$, P_i, $i=1, 2, \cdots, n-1$, are in $E_\alpha-N$, and hence from (10) and (7), we have

$$(11)\qquad \begin{aligned}&\log L(P;\alpha)\\ &=np(\alpha)\Big\{\sum_{i=1}^{n}\alpha^*(x_i,Q_i)-\sum_{i=1}^{n-1}\alpha^*(P_i)\Big\}+nq(\alpha)\\ &\qquad+\Big\{\sum_{i=1}^{n}r^*(x_i,Q_i)-\sum_{i=1}^{n-1}r^*(P_i)\Big\}.\end{aligned}$$

Therefore, for $x_1=x_2=\cdots=x_n=x\in\pi_\alpha-\bigcup^n_{i=1}N_{i,Q_i}$, we have

$$n\log f(x;\alpha)=np(\alpha)u(x)+nq(\alpha)+nr(x),$$

where $u(x)=\sum\limits_{i=1}^{n}\alpha^*(x,Q_i)-\sum\limits_{i=1}^{n-1}\alpha^*(P_i)$, and $nr(x)=\sum\limits_{i=1}^{n}r^*(x,Q_i)-\sum\limits_{i=1}^{n-1}r^*(P_i)$, that is, denoting $\pi_\alpha-\bigcup^n_{i=1}N_{i,Q_i}$ by π_α, we can obtain (5).

Sufficiency. From (5) we have for all $P\in\underbrace{\pi_\alpha\times\cdots\times\pi_\alpha}_{n}=E_\alpha$

$$(12)\qquad \log L(P;\alpha)=np(\alpha)\alpha^*(P)+nq(\alpha)-nr^*(P),$$

where $\alpha^*(P)=\Big(\sum\limits_{i=1}^{n}u(x_i)\Big)\Big/n$, and $r^*(P)=\Big(\sum\limits_{i=1}^{n}r(x_i)\Big)\Big/n$. And by the condition that $p(\alpha)$ and $q(\alpha)$ are differentiable, the equation (2) holds for $E_\alpha=\underbrace{\pi_\alpha\times\cdots\times\pi_\alpha}_{n}$.

In the case where the equation (2) holds for all E_α without any subset of exception the proof of the necessity can be rather easily carried out than the above. In fact, to obtain (5) we have only to put $x_1=x_2=\cdots=x_n=x$ in (7) and $u(x)=\alpha^*(x, x, \cdots, x)$.

The form (5) of fr. f. has been discussed by B. O. Koopman 2), and hence we call it Koopman's distribution. He has proved that (5) is the necessary and sufficient condition for existence of a sufficient statistic in the sample space of any size when E_α is independent of α.

Now we can state the following

Theorem II. *The value of $p(\alpha)$ in the form (5) is uniquely determined by the value of $q(\alpha)$. Denote this correspondence by $q(\alpha)=T(p(\alpha))$, then the moment generating function $\varphi(\theta;\alpha)$ of α^* is given by*

$$\varphi(\theta;\alpha)=\exp[-n\{T\big(\theta/n+p(\alpha)\big)-q(\alpha)\}]. \tag{13}$$

Proof. Integrating the both members of equation (5) and multiplying by $\exp(-q(\alpha))$, we have

$$\int_{\pi_\alpha}\exp\big(p(\alpha)\,u(x)\big)\exp\big(r(x)\big)\,dx=\exp\big(-q(\alpha)\big). \tag{14}$$

Hence the value of $q(\alpha)$ is uniquely determined by the one of $p(\alpha)$.

To prove (13), we transform the likelihood function $L(P;\alpha)$ into

$$L(P;\alpha)=\exp[n\{p(\alpha)-p(\alpha_0)\}\alpha^*(P)+n\{q(\alpha)-q(\alpha_0)\}]\cdot L(P;\alpha_0). \tag{15}$$

If we put $\theta=n\big(p(\alpha)-p(\alpha_0)\big)$, and consequently $q(\alpha)-q(\alpha_0)=T\{\theta/n+p(\alpha_0)\}-T\{p(\alpha_0)\}$, then (13) follows from the equation $\int_{E\alpha}L(P;\alpha)\,dP=1$.

Theorem III. *If the sign of equality in (1) holds, and if $\partial f(x;\alpha)/\partial\alpha$ is continuous on α, then $\alpha^*(P)$ is an unbiased estimate of $\phi(\alpha)$ with variance $\phi'(\alpha)/p'(\alpha)$, whenever $p'(\alpha)\neq 0$. Moreover, if $\alpha^*(P)=const.$ almost everywhere on E_α, then the variance of $\alpha^*(P)$ is equal to zero.*

Proof. From $\int(\partial L(P;\alpha)/\partial\alpha)dP=0$,

$$0=\int_{-\infty}^{\infty}\frac{\partial\log L(P;\alpha)}{\partial\alpha}L(P;\alpha)\,dP=k(\alpha)\int_{-\infty}^{\infty}\big(\alpha^*(P)-\phi(\alpha)\big)L(P;\alpha)\,dP,$$

therefore, whenever $k(\alpha)\neq 0$, $\psi(\alpha)=E_\alpha\big(\alpha^*(P)\big)=\phi(\alpha)$.[3]
Moreover,

$$\phi'(\alpha)^2=\psi'(\alpha)^2=E_\alpha\big(\alpha^*(P)-\phi(\alpha)\big)^2\int_{E_\alpha}\big(\partial\log L(P;\alpha)/\partial\alpha\big)^2L(P;\alpha)\,dP$$

---

[3] I am indebted to Professor G. Maruyama for a remark leading to a simplification of my original proof of this.

$$=\{E_\alpha(\alpha^*(P)-\phi(\alpha))^2\}^2 k(\alpha)^2,$$

that is,

$$E_\alpha\left(\alpha^*(P)-\phi(\alpha)\right)^2=\frac{\phi'(\alpha)}{p'(\alpha)}.$$

If $\alpha^*(P)=const.$, then $\alpha^*(P)=\phi(\alpha)$ almost everywhere in E_α. Therefore the variance of $\alpha^*(P)$ is equal to zero.

In the rest of this section, we shall discuss how to choose a parameter admitting an efficient unbiased estimate. For this purpose we remember that

$$T'\left(p(\alpha)\right)=\frac{q'(\alpha)}{p'(\alpha)}=\phi(\alpha), \tag{16}$$

which follows from (13). Now suppose that $T'(\theta)$ has the unique inverse function $S(\beta)$, and put $V(\beta)=T\left(S(\beta)\right)$, then $p(\alpha)=S\left(\phi(\alpha)\right)$, $q(\alpha)=V\left(\phi(\alpha)\right)$, and therefore

$$f(x;\alpha)=\exp\{S(\beta)u(x)+V(\beta)+r(x)\},$$
$$L(P;\alpha)=\exp\left\{nS(\beta)\alpha^*(P)+nV(\beta)+\sum_{i=1}^n r(x_i)\right\},$$

where $\beta=\phi(\alpha)$. Denote these functions on the right hand sides of the last equations by $f_1(x;\beta)$ and $L_1(P;\beta)$ respectively, then we have

$$\frac{\partial\log L_1(P;\beta)}{\partial\beta}=\frac{k(\alpha)}{\dfrac{d\beta}{d\alpha}}\left(\alpha^*(P)-\beta\right). \tag{17}$$

From this equation the factor $k(\alpha)\Big/\dfrac{d\beta}{d\alpha}$ of the right hand side is uniquely determined by β. We shall denote this factor by $k_1(\beta)$. Thus (17) becomes

$$\frac{\partial\log L_1(P;\beta)}{\partial\beta}=k_1(\beta)(\alpha^*-\beta).$$

This equation shows that $\alpha^*(P)$ is an efficient estimate of β, and that such β is unique. Thus we have the following

Theorem IV. *If $T'(\theta)$ has the unique inverse function in Koopman's distribution* (5), *there exists a parameter $T'\left(p(\alpha)\right)$ having an efficient unbiased estimate $\left(\sum_{i=1}^n u(x_i)\right)\Big/n$, and such a parameter is unique.*

Remark. If there exist two statistics $\alpha^*(P)$ and $\beta^*(P)$ for which holds the sign of equality in (1) with respect to some $\phi(\alpha)$ for any α in A, then we can easily see from (2) that the relation

$$\beta^*(P)=a\,\alpha^*(P)+b \qquad (a, b=\text{constant})$$

holds almost everywhere in R_n. And, further, if $\alpha^*(P)$ is an efficient unbiased estimate of α, then $a\alpha^*(P)+b$ is also the one of $a\alpha+b$. Therefore we *might* say that there exists *essentially* only one efficient unbiased estimate of distribution.

2. Examples. a) The variance σ^2 of the normal distribution with fr. f.

$$f(x;\sigma)=(2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2).$$

In this case we have $T'(\theta)=-(2\theta)^{-1}$, which has an inverse function, and $p(\sigma)=-(2\sigma^2)^{-1}$, and therefore $T'(p)=\sigma^2$ is a parameter, having an efficient unbiased estimate $\sum_{i=1}^{n} x_i^2\Big/n$.

b) Pearson's type III distribution. Writing

$$f(x;\alpha)=\big(\Gamma(\alpha)\big)^{-1}x^{(\alpha-1)}e^{-x},\quad x>0,\quad \alpha>1,$$

we have $p(\alpha)=\alpha-1$, and $T'\big(p(\alpha)\big)=\Gamma'(\alpha)/\Gamma(\alpha)$. This is the parameter, for which $\sum_{i=1}^{n}\log x_i\Big/n$ is an efficient unbiased estimate of size n.

c) Pearson's type III distribution. Writing

$$f(x;\alpha)=\big(\alpha/\Gamma(\lambda)\big)\,x^{\lambda-1}e^{-\alpha x},\quad x>0,\quad \lambda>1,\quad \alpha>0.$$

then $p(\alpha)=-\alpha$, and $T'(p)=-\lambda/\alpha$, which is the parameter having an efficient unbiased estimate $\sum_{i=1}^{n}x_i\Big/n$.

d) For Pearson's type IV distribution with fr. f.

$$f(x;\alpha)=e^{\pi/2}/G(\alpha)\cdot\exp(-\tan^{-1}x)\cdot(1+x^2)^{-\alpha/2+1},$$

where $G(\alpha)=\int_0^{\pi}e^t\sin^{\alpha}t\,dt$, $-\infty<t<\infty$, we have $p(\alpha)=-\alpha/2+1$, $q(\alpha)=-G(\alpha)$, and $u(x)=\log(1+x^2)$, therefore $T'(p)=-2G'(\alpha)$, the efficient unbiased estimate of which is $\sum_{i=1}^{n}\log(1+x_i^2)\Big/n$.

Many other distributions of this sort are found in Pearson's type, which can be transformed into Koopman's one.

Literature

1) Cramér, Harald. Mathematical Methods of Statistics, Princeton University Press, 1946

2) Koopman, B. O. On distributions admitting a sufficient statistic. Trans. Amer. Math. Soc. 39 (1936)