Statistics 16 (1985) 1, 89–95

Non-Existence of Unbiased Estimators of Ordered Parameters

ISHWARI D. DHARIYAL, DIVAKAR SHARMA, and K. KRISHNAMOORTHY

Indian Institute of Technology, Kanpur

Summary. For a certain general class of distributions and for the one-parameter exponential family of distributions, it is shown that ordered parameters do not admit unbiased estimators. It is also seen that these classes of distributions, for which non-existence of unbiased estimators of ordered parameters is exhibited, can be enlarged.

AMS 1980 *subject classification:* 62F10.

Key words: Ordered parameters, Unbiased estimators, Estimable functions, Exponential family of distributions.

1. Introduction

The problem of point estimation of ordered parameters has been studied extensively, among others, by ALAM (1967), BLUMENTHAL (1975, 1976), BLUMENTHAL and COHEN (1968a, b, c) and DUDEWICZ (1971a, b, 1972, 1973). Such a problem is of interest, for example, in the one-way Analysis of Variance setting with k treatments where the hypothesis of equal treatment effects has been rejected and the largest treatment effect is to be estimated. Different sampling schemes and a variety of estimators, for example, maximum likelihood, maximum probability, generalized BAYES, and PITMAN, especially in the case of ordered normal means, have been proposed by the above authors. In case there exists an unbiased estimator of an ordered parameter, a natural criterion is the variance of the estimator. However, in the absence of unbiased estimators a criterion is the bias in absolute value, as considered by BLUMENTHAL and COHEN (1968c). Briefly, the problem is as follows:—

Let $\pi_1, \ldots, \pi_k$ be $k(\geqq 2)$ populations with means $\mu_1, \ldots, \mu_k$ respectively and $\bar{X}_i$ be the mean of the sample from population π_i, $i=1, \ldots, k$. It is well-known that $\bar{X}_i$ is an unbiased estimator of μ_i. Let $\mu_{[1]} \leqq \ldots \leqq \mu_{[k]}$ and $\bar{X}_{[1]} \leqq \ldots \leqq \bar{X}_{[k]}$ be ordered population and sample means respectively. A natural estimator of $\mu_{[m]}$, $m=1, \ldots, k$, in this case is $\bar{X}_{[m]}$, which has some nice asymptotic properties, for example, asymptotic unbiasedness and strong consistency [see DUDEWICZ (1972)]. However, it may be noted that $\mathsf{E}\bar{X}_{[k]} \geqq \mu_{[k]}$, since $\bar{X}_{[k]}$ is a convex function of $(\bar{X}_1, \ldots, \bar{X}_k)$, with equality if and only if $\bar{X}_{[k]}$ is a linear function with probability 1. The latter condition is equivalent to the condition that there exists a k_0 with $\mathsf{P}(\bar{X}_{k_0} \geqq$

$\geqq \bar{X}_i$ for each $i) = 1$. Therefore, a natural question that arises is whether $\mu_{[m]}$ can be estimated unbiasedly.

In this paper, we establish that for certain families of distributions an unbiased estimator of an ordered parameter does not exist. BLUMENTHAL and COHEN (1968c) sketched a proof for the case of two normal population means. Here, we generalize and extend that result.

2. Non-existence of Unbiased estimators of ordered parameters in the case of one parameter exponential family of distributions

Let $X_{i1}, \ldots, X_{in_i}$, $i = 1, \ldots, k(\geqq 2)$ be a random sample of size n_i from a population whose distribution depends on a single real parameter θ_i. Let the space of the parameter $\theta = (\theta_1, \ldots, \theta_k)$ be a subset Ω of the k-dimensional EUCLIDean space and let $\mu_i = \mu_i(\theta_i)$, $i = 1, \ldots, k$, be non-constant estimable functions of the parameters. We assume that $\Omega \neq \{\theta : \theta_1 = \ldots = \theta_k\}$, otherwise the problem does not make sense. Let $\mu_{[1]} \leqq \ldots \leqq \mu_{[k]}$ denote $\mu_1, \ldots, \mu_k$ in their numerical order and let $T_i = T_i(X_{i1}, \ldots, X_{in_i})$ be a sufficient statistic for θ_i. In the following, for $\theta, \theta' \in \Omega$, $\mathsf{P}_\theta \ll \mathsf{P}_{\theta'}$ means, as usual, that the probability measure P_θ is absolutely continuous with respect to the probability measure $\mathsf{P}_{\theta'}$. For a given $m = 1, \ldots, k$ and for $i = 1, \ldots, k$, define

$$\Omega_i = \{\theta \in \Omega : \mu_i = \mu_{[m]}\} .$$

Theorem 2.1. *If $(T_1, \ldots, T_k)$ is a complete statistic for θ belonging to some $\Omega_{i'}$ with $\Omega_{j'} - \Omega_{i'} \neq \emptyset$ for some j' and if $\mathsf{P}_\theta \ll \mathsf{P}_{\theta'}$ for all $\theta, \theta' \in \Omega$, then for any $m = 1, \ldots, k$, $\mu_{[m]}$ is not estimable.*

Proof. Since $\mu_{i'}$ is estimable, there exists a function $U(T_{i'})$ of $T_{i'}$ only which estimates $\mu_{i'}$ unbiasedly. Since $(T_1, \ldots, T_k)$ is a sufficient statistic for $\theta \in \Omega$, without loss of generality, suppose $g(T_1, \ldots, T_k)$ is an unbiased estimator of $\mu_{[m]}$. Then we must have

$$\mathsf{E}_\theta\,[g(T_1, \ldots, T_k) - U(T_{i'})] = 0 \quad \forall \theta \in \Omega_{i'}$$

which implies

$$\mathsf{P}_\theta\,[g(T_1, \ldots, T_k) = U(T_{i'})] = 1 \quad \forall \theta \in \Omega_{i'} .$$

But $\mathsf{P}_\theta \ll \mathsf{P}_{\theta'}$ for all $\theta\theta' \in \Omega$, hence

$$\mathsf{P}_\theta\,[g(T_1, \ldots, T_k) = U(T_{i'})] = 1 \quad \forall \theta \in \Omega ,$$

so that

$$\mathsf{E}_\theta\, g(T_1, \ldots, T_k) = \mu_{i'} \quad \forall \theta \in \Omega .$$

Corollary 2.1. *If $X_{i1}, \ldots, X_{in_i}$ is a random sample from a distribution with probability density*

$$f(x; \theta_i) = h(\theta_i)\, e^{-\theta_i T(x) + W(x)}, \quad i = 1, \ldots, k ,$$

that is, if we have one-parameter exponential distributions, and if some $\Omega_{i'}$ contains a k-dimensional rectangle, $\mu_{[m]}$ is not estimable.

The proof follows easily using Theorem 2.1 above, Lemma 8 on page 52 and Theorem 1 on page 132 of LEHMANN (1959).

Remark 2.1. In Theorem 2.1 above we have assumed that μ_i's are estimable. We justify this assumption as follows:

Let $X_i = (X_{i1}, \ldots, X_{in_i})$ be a random sample from a distribution with parameter θ_i, $i = 1, \ldots, k$, $g(X_i, \ldots, X_k)$ be an unbiased estimator of $\mu_{[m]}$ and $\{\theta : \theta_1 = \ldots = \theta_k\} \subset \subset \Omega$. Then

$$\mathsf{E}g(X_1, \ldots, X_k) = \mu_{[m]} \quad \forall\, \theta$$

implies

$$\mathsf{E}g(X_1, \ldots, X_k) = \mu_1 \quad \text{if} \quad \theta_1 = \ldots = \theta_k$$

so that $g(X_1, \ldots, X_k)$ is an unbiased estimator of μ_1 when $X_{ij} (i = 1, \ldots, k, j = 1, \ldots, n_i)$ are independently and identically distributed with parameter θ_1. Thus, if θ_1 admits a one-dimensional sufficient statistic $\mu_{[m]}$ is not estimable if μ_i is not ~~admissible.~~ estimable.

3. Differentiability of estimable functions: applications to ordered parameters

Below we give a different method of proving the nonexistence of unbiased estimators of ordered parameters for a certain class of distributions. The results obtained in this section are extensions of the BLUMENTHAL and COHEN (1968c) result. In Theorem 3.1 below we establish that for populations with probability density functions of a certain form, any estimable function of the parameters is differentiable with respect to any one of the parameters. We shall need the following:

Lemma 3.1. *For $u \geqq 1$ and $a \geqq 1$*

$$u^{a-1} e^{-u^a} \leqq e^{-(u-1)^a}.$$

Proof. Since

$$\log u^{a-1} + (u-1)^a - u^a \leqq u^{a-1} + (u-1)^a - u^a ,$$

it suffices to show that

$$u^{a-1} + (u-1)^a - u^a \leqq 0 .$$

Since $u \geqq 1$, $a \geqq 1$, we have

$$\left(1 - \frac{1}{u}\right)^a \leqq 1 - \frac{1}{u}$$

which implies

$$u^{a-1} + (u-1)^a - u^a \leqq 0 .$$

Theorem 3.1. *Let $X_1, \ldots, X_k$ be $k(\geqq 2)$ independently distributed random variables with probability density functions $f(x-\theta_1), \ldots, f(x-\theta_k)$ respectively, where $f(\cdot)$ is of the form*

$$f(x) = Ce^{-|x|^a}, \quad a \geqq 1, \quad -\infty < x < \infty .$$

Let $\Phi(\theta_1, \ldots, \theta_k)$ be any estimable function of $(\theta_1, \ldots, \theta_k)$. Then $\Phi(\theta_1, \ldots, \theta_k)$ is differentiable with respect to θ_i, $i = 1, \ldots, k$.

Proof: Let $g(X_1, \ldots, X_k)$ be an unbiased estimator of $\Phi(\theta_1, \ldots, \theta_k)$. Without loss of generality we prove that $\Phi(\theta_1, \ldots, \theta_k)$ is differentiable with respect to θ_1.

We have

$$\underset{h\to 0}{\mathrm{Lt}}\, (\Phi(\theta_1 + h, \theta_2, \ldots, \theta_k) - \Phi(\theta_1, \ldots, \theta_k))/h$$

$$= \underset{h\to 0}{\mathrm{Lt}}\, C^k \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_k)\, e^{-\sum_{i=2}^{k} |x_i - \theta_i|^a} .$$

$$\times \frac{1}{h} \{e^{-|x_1-\theta_1-h|^a} - e^{-|x_1-\theta_1|^a}\}\, dx_1 \ldots dx_k .$$

Since

$$\underset{h\to 0}{\mathrm{Lt}}\, \frac{1}{h} \{e^{-|x_1-\theta_1-h|^a} - e^{-|x_1-\theta_1|^a}\} = \frac{de^{-|x_1-\theta_1|^a}}{d\theta_1}$$

exists for all $\theta_1 \neq x_1$, it suffices to prove that [see CRAMÉR (1946), pp. 67–68],

$$\left|\frac{de^{-|x_1-\theta_1|^a}}{d\theta_1}\right|_{\theta_1=\theta_1'} \leqq G(x_1, \theta_1), \quad \forall\, \theta_1 \neq x_1 \quad \text{and} \quad \forall\, \theta_1'$$

in some neighbourhood of θ_1, where $G(x_1, \theta_1)$ is such that

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} |g(x_1, \ldots, x_k)|\, G(x_1, \theta_1)\, e^{-\sum_{i=2}^{k} |x_i-\theta_i|^a}\, dx_1 \ldots dx_k < \infty .$$

Now, for $\theta_1 \neq x_1$,

$$\left|\frac{de^{-|x_1-\theta_1|^a}}{d\theta_1}\right|_{\theta_1=\theta_1'} = a\,|x_1 - \theta_1'|^{a-1} e^{-|x_1-\theta_1'|^a}$$

$$\leqq a e^{-(|x_1-\theta_1'|-1)^a} \quad \text{if} \quad |x_1 - \theta_1'| \geqq 1 \quad \text{(from Lemma 3.1)}$$

$$\leqq a e^{-(|x_1-\theta_1|-2)^a} \quad \text{if} \quad |x_1 - \theta_1'| + 1 \geqq |x_1 - \theta_1| \geqq 2 .$$

Hence, for $|\theta_1 - \theta_1'| < 1$ and $|x_1 - \theta_1| \geqq 2$, we have

$$\left|\frac{de^{-|x_1-\theta_1|^a}}{d\theta_1}\right|_{\theta_1=\theta_1'} < a e^{-(|x_1-\theta_1|-2)^a} .$$

Also, it is easy to see that for $|\theta_1 - \theta_1'| < 1$ and $|x_1 - \theta_1| < 2$, we can choose a constant $M > 0$ such that

$$\left|\frac{de^{-|x_1-\theta_1|^a}}{d\theta_1}\right|_{\theta_1=\theta_1'} < M e^{-|x_1-\theta_1|^a} .$$

Thus, take

$$G(x_1, \theta_1) = \begin{cases} a e^{-(|x_1-\theta_1|-2)^a}, & |x_1-\theta_1| \geqq 2 \\ M e^{-|x_1-\theta_1|^a}, & |x_1-\theta_1| < 2 . \end{cases}$$

Then, clearly,

$$C^k \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |g(x_1, \dots, x_k)|\, G(x_1, \theta_1)\, e^{-\sum_{i=2}^{k} |x_i-\theta_i|^a} dx_1 \dots dx_k$$

$$= aC^k \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\theta_1-2} |g(x_1, \dots, x_k)|\, e^{-|x_1-\theta_1+2|^a - \sum_{i=2}^{k} |x_i-\theta_i|^a} dx_1 \dots dx_k$$

$$+ aC^k \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{\theta_1+2}^{\infty} |g(x_1, \dots, x_k)|\, e^{-|x_1-\theta_1-2|^a - \sum_{i=2}^{k} |x_i-\theta_i|^a} dx_1 \dots dx_k$$

$$+ MC^k \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{\theta_1-2}^{\theta_1+2} |g(x_1, \dots, x_k)|\, e^{-\sum_{i=1}^{k} |x_i-\theta_i|^a} dx_1 \dots dx_k$$

$$< a\mathsf{E}_{\theta_1+2,\theta_2,\dots,\theta_k} |g(X_1, \dots, X_k)|$$

$$+ a\mathsf{E}_{\theta_1-2,\theta_2,\dots,\theta_k} |g(X_1, \dots, X_k)| + M\mathsf{E}_{\theta_1,\dots,\theta_k} |g(X_1)| < \infty .$$

Corollary 3.1. *Let* $\theta_{[1]} \leqq \dots \leqq \theta_{[m]}$ *denote* $\theta_1, \dots, \theta_k$ *in numerical order. Under the assumptions of Theorem* 3.1, $\theta_{[m]}$ *is not estimable,* $m = 1, \dots, k$.

Remark 3.1. Proof of Theorem 3.1 suggests that the conclusions will hold for all distributions whose probability density functions $f(x; \theta)$ satisfy the following condition:

$$\left| \frac{df(x;\theta)}{d\theta} \right|_{\theta=\theta'} < \sum_{i=1}^{r} \alpha_i f(x; \theta^{(i)}) \ \forall \theta' |$$

in a neighbourhood of θ where the integer r, numbers α_i and parameter points $\theta^{(i)}$ may depend on the neighbourhood of θ but are independent of x. It is easily verified that the uniform distribution on $[0, \theta]$ and PARETO's distribution with probability density function

$$f(x; \theta) = \alpha\theta^\alpha / x^{\alpha+1}, \quad \alpha > 1, \quad x \geqq \theta > 0$$

satisfy the above condition.

4. Concluding remarks

Remark 4.1. Whereas Theorem 2.1 covers a large class of distributions including the exponential family, certain important distributions like uniform on $[0, \theta]$, negative exponential with location parameter θ, double exponential with location parameter θ, etc. are left out. These other distributions are covered by Theorem 3.1 and the associated Remark 3.1.

Remark 4.2. When $X_i (i=1, ..., k)$ are independent and have probability density $c_i(\theta_i)\, h_i(x_i)\, e^{\theta_i T_i(x_i)}$ with respect to a sigma-finite measure ν, the joint density of $X_1, ..., X_k$ can be viewed as a density in the exponential family with parameter $(\theta_1, ..., \theta_k)$. Any estimable function $\Phi(\theta_1, ..., \theta_k)$ then is necessarily differentiable with respect θ_i (see, for example, LEHMANN (1959), Theorem 9, pp. 52–54, where the restriction "Let Φ be any bounded measurable function ..." can be replaced by "Let Φ be any integrable function . . ."). This provides an alternative proof for the results in Section 2. However, as the following remark shows, there are situations in which only one of the proofs can be applied.

Remark 4.3. When $k=2$, $\theta_{[2]}$ is not differentiable with respect to θ_1 at the point $\theta_1=\theta_2$ and this fact is used to prove non-existence of an unbiased estimator of $\theta_{[2]}$ if the distribution is such that any estimable function has to be differentiable. As such, in case parameter spaces for θ_1 and θ_2 do not intersect and the estimation of $\theta_{[2]}$ remains meaningful (it is not if $-\infty<\theta_1\leqq a$ and $a<\theta_2<\infty$), differentiability argument is not applicable. As an example, suppose X_1 and X_2 are independently $N(\theta_1, 1)$ and $N(\theta_2, 1)$ respectively with $-\infty<\theta_1<a$ or $b<\theta_1<\infty$ and $a<\theta_2<b$. Theorem 2.1 implies non-estimability of $\theta_{[2]}$, though $\theta_{[2]}$ is differentiable with respect to θ_1.

An example, where Theorem 2.1 cannot be applied but differentiability argument is available to us, is provided by X_1 and X_2 independently distributed as $N(\theta_1, 1)$ and $N(0, 1)$ respectively with $-\infty<\theta_1<\infty$.

Acknowledgement

The authors wish to thank the referees for some helpful comments.

References

ALAM, K. (1967). A two-sample estimate of the largest mean. *Ann. Inst. Statist. Math.* **19**, 271–283.

BLUMENTHAL, S. (1975). Sequential estimation of largest normal mean when variance is unknown. *Comm. Statist.* **4**, 655–669.

BLUMENTHAL, S. (1976). Sequential estimation of largest normal mean when variance is known. *Ann. Statist.* **4**, 1077–1087.

BLUMENTHAL, S. and COHEN, A. (1968a). Estimation of the larger translation parameter. *Ann. Math. Statist.* **39**, 502–516.

BLUMENTHAL, S. and COHEN, A. (1968b). Estimation of two ordered translation parameters. *Ann. Math. Statist.*, **39**, 517–530.

BLUMENTHAL, S. and COHEN, A. (1968c). Estimation of the larger of two normal means. *J. Amer. Statist. Assoc*, **63**, 861–876.

CRAMÈR, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton, N. J.

DUDEWICZ, E. J. (1971a). Maximum likelihood estimators for ranked means. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **19**, 29–42.

DUDEWICZ, E. J. (1971b). Maximum likelihood estimators for non-l-l functions. *Trabajos de Estadistica*, 65–70.

DUDEWICZ, E. J. (1972). Point estimation of ordered parameters: the general location parameter case. *Tamkang J. Math.*, **3**, 101–114.

DUDEWICZ, E. J. (1973). Maximum probability estimators for ranked means. *Ann. Inst. Statist. Math.*, **25**, 467–477.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses,* John Wiley and Sons, Inc., New York, N. Y.

Received August 1982; revised January 1984.

Indian Institute of Technology, Kanpur
Department of Mathematics
III Post Office
Kanpur - 208016 U.P.
India