UNBIASED ESTIMATION OF LOCATION AND SCALE PARAMETERS[1]

By J. K. Ghosh[2] and Rajinder Singh

University of Illinois

1. Summary and introduction. It is well-known that there is a close connection between linear functionals on an appropriate Banach space and unbiased estimators. In Section 2 we prove some results concerning unbiased estimation of location and scale parameters. As application of these results we consider the case of Cauchy density with unknown location [scale] but known scale [location] parameter. We show that there exists no unbiased estimator for the location parameter, and none with finite variance for the scale parameters. If the Cauchy density involves both location and scale parameters, then it is shown that neither of these parameters has an unbiased estimator. Some information about other parametric functions is also given. The present results for the location parameter case were obtained previously by H. Pollard; we are grateful to Professor Kiefer for informing us of Pollard's work.

2. Estimation of location and scale parameters. Let X be a vector-valued random variable taking values in an Euclidean space E. Let μ be the Lebesgue measure on (E, B) where B is the Borel field. Let $P_\theta^X \equiv P_\theta$ define the distribution of X under $\theta \varepsilon \Theta$ and p_θ the density of P_θ with respect to μ. We shall always assume the family $\mathcal{P} = \{P_\theta : \theta \varepsilon \Theta\}$ is homogeneous, i.e., any two members have same null sets. Let $r_{\theta,\theta_0} = p_\theta / p_{\theta_0}$. Let $R_{\theta_0} = \{r_{\theta,\theta_0} : \theta \varepsilon \Theta\}$.

By $L^p(\mu)$, $1 \leqq p < \infty$, we denote as usual the Banach space of real-valued measurable functions whose pth power is integrable with respect to μ; we shall only need $L^1(\mu)$ and $L^2(\mu)$. We also need the space $L^\infty(\mu)$ of all essentially bounded measurable functions. The spaces $L^p(P_{\theta_0})$, $1 \leqq p \leqq \infty$, are similarly defined.

Let $R_{\theta_0} \subset L^2(P_{\theta_0})$. Then for any parametric function $\phi(\theta)$ there exists an unbiased estimator with finite variance under θ_0 if and only if

$$(1) \qquad |\textstyle\sum_1^m a_i \phi(\theta_i)| \leqq C \, \|\sum_1^m a_i r_{\theta_i,\theta_0}\|_{2,\theta_0}$$

for all m and all $a_1, \cdots, a_m$ where for $f \varepsilon L^2(P_{\theta_0})$, $\|f\|_{2,\theta_0}$ is its $L^2(P_{\theta_0})$-norm (for example see [1]). Necessity follows by an easy application of Schwarz inequality and it is this part that we shall require for applications later.

Suppose now that X is real-valued and $p_\theta(x) = p_0(x - \theta)$; θ being any real number. The following theorem relates unbiased estimators and completeness of the family $\mathcal{P} = \{p_\theta\}$. (We recall that a family $\mathcal{P} = \{p_\theta\}$ is complete iff $\int \phi p_\theta \, d\mu = 0$ for all θ implies $\phi = 0$ a.e. (μ).)

Received 10 February 1966; revised 1 July 1966.

[1] Partially supported by the National Science Foundation Grant GP-3814.

[2] On leave from the University of Calcutta.

THEOREM 2.1. *Let* $p_\theta(x) = p_0(x - \theta)$. *If* $E_0(X)$ *exists then there exists an unbiased estimator of* θ. *If* $E_0(X)$ *does not exist then existence of unbiased estimator for* θ *and completeness of the family* $\mathcal{P} = \{p_\theta\}$ *are incompatible.*

PROOF. Let $E_0(X) = b$. Then $X - b$ is an unbiased estimator of θ. To prove the second statement suppose $\mathcal{P} = \{p_\theta\}$ is complete and, if possible, let $T(X)$ be an unbiased estimator of θ. Then for each θ_0 completeness of $\mathcal{P}$ implies $T(x) = T(x + \theta_0) - \theta_0$ a.e. (μ). So $T(x) - x$ is almost invariant as defined in Lehmann (1959), p. 225, under translations and hence by Theorem 4, Lehmann (1959), p. 225, $T(x) = x + k$ a.e. (μ), where k is a constant. But this implies that $E_0(X) = E_0(T) - k$ exists contradicting our hypothesis. Hence the theorem is proved.

By exactly similar arguments we have the following theorem on scale parameters.

THEOREM 2.2. *Let* $p_\sigma(x) = \sigma^{-1}p_1(x/\sigma), \sigma > 0$. *If* $E_1(X)$ *exists then there exists an unbiased estimator of* σ. *If* $E_1(X)$ *does not exist, then existence of unbiased estimator of* σ *and completeness of the family* $\mathcal{P} = \{p_\sigma : \sigma > 0\}$ *are incompatible.*

If both location and scale parameters are unknown we have the following result:

THEOREM 2.3. *Let* $p_{\theta,\sigma}(x) = \sigma^{-1}p_{0,1}((x - \theta)/\sigma)$ *and let*

$$\mathcal{P} = \{p_{\theta,\sigma} : -\infty < \theta < \infty, \sigma > 0\}$$

be complete. If $E_{0,1}|X| < \infty$, *then* θ *has an unbiased estimator; if* $E_{0,1}|X| = \infty$ *then* θ *has no unbiased estimator. There is no unbiased estimator for* σ.

PROOF. The proof for the assertion about θ is similar to the proof of Theorem 2.1. That σ has no unbiased estimator can be seen by considering $T(x) - T(x + a)$ where $T(X)$ is, if possible, an unbiased estimator of σ. Using completeness and Theorem 4, Lehmann (1959), p. 225, we have $T(x) =$ const. a.e. (μ) contradicting $E_{\theta,\sigma}(T) = \sigma$.

While completeness seems to be a hard question to settle we do have the following which is a known formulation of a famous result of Wiener. We give the proof for the sake of completeness.

THEOREM 2.4. *Let* $p_\theta = p_0(x - \theta)$ *as in Theorem* 2.1. *Then the family* $\mathcal{P} = \{p_\theta : -\infty < \theta < \infty\}$ *is boundedly complete* (i.e. $\int \phi p_\theta \, d\mu = 0$ *for all* θ *and for bounded* ϕ *implies* $\phi = 0$ a.e. (μ)) *if and only if the Fourier transform* $\hat{p}_0(t)$ *of* p_0 *does not vanish for any* t.

PROOF. By the Hahn-Banach theorem and the representation theorem for the conjugate space of $L^1(\mu)$ as $L^\infty(\mu)$ it follows that $\mathcal{P}$ is boundedly complete if and only if the smallest closed linear space generated by $\mathcal{P}$ in $L^1(\mu)$ is $L^1(\mu)$ itself. The theorem now follows since Wiener's Tauberian theorem, Theorems 10 C, 10 E, [2], says the latter condition is equivalent to non-vanishing of $\hat{p}_0$ everywhere. (Wiener's theorem pertains to $L_c^1(\mu)$ the class of complex-valued integrable functions but it is trivially equivalent to the corresponding result for $L^1(\mu)$.)

If $\mathcal{P}$ is a family of probability densities of a nonnegative rv X involving only a scale parameter then by transforming to $\log X$ we get a family $\mathcal{P}'$ involving only an unknown location parameter. Hence a theorem similar to Theorem 2.4 can be proved for scale parameter also.

It is clear that given a real valued parametric function $\phi(\theta)$, finding an estimator ψ is equivalent to solving

$$p_0' * \psi = \phi \tag{2}$$

where $p_0'(-x) = p_0(x)$ and $*$ denotes convolution. If one could define Fourier transforms for ϕ and ψ, one would be led to solving

$$\hat{p}_0'\hat{\psi} = \hat{\phi}. \tag{3}$$

Unfortunately the technique of Fourier transforms is not available since ϕ usually does not lie in $L^1(\mu)$ or $L^2(\mu)$. Even if $\phi \in L^1(\mu)$ or $L^2(\mu)$ it is not necessarily true that an estimator, if it exists, would have similar integrability properties; e.g. a necessary condition is that $\hat{\phi} \to 0$ faster than $\hat{p}_0'$ as $t \to \pm\infty$. (That $\hat{p}_0' \to 0$ as $t \to \pm\infty$ is of course the Lebesgue-Riemann theorem.) However, the following result is not hard to show.

THEOREM 2.5. *Let* $p_\theta = p_0(x - \theta)$ *as in Theorem* 2.1. *and*

$$\mathcal{P} = \{p_\theta : -\infty < \theta < \infty) \subset L^2(\mu).$$

Suppose $|\hat{p}_0'| > 0$ a.e. (μ). *Then a function* $\phi(\theta)$ *has an unbiased estimator belonging to* $L^2(\mu)$ *if and only if* $\phi \in L^2(\mu)$ *and* $\hat{\phi}(\hat{p}_0')^{-1}$ *is square integrable with respect to* μ; *an estimator if it exists is unique.*

The theorem is a consequence of Plancherel's theorem and the fact that if $p_0 \in L^1(\mu)$, $\psi \in L^2(\mu)$ *then* $p_0' * \psi \in L^2(\mu)$ and $\hat{p}_0'\hat{\psi} = \phi$ where $\phi = p_0' * \psi$. It is of some interest that the proof of sufficiency in this case does provide a method of constructing or at least approximating the unbiased estimator in contrast to the preceding results of Section 2.

We conclude this section with another simple and interesting proposition, which may be compared with Theorem 2.1.

THEOREM 2.6. *Let* $p_\theta(x) = p_0(x - \theta)$. *If* $E_0(X^2) = \infty$ *then there does not exist an unbiased estimator of* θ *with bounded variance.*

PROOF. Suppose, if possible, there exists an unbiased estimator T with bounded variance, i.e., $E_\theta(T - \theta)^2 < k$ where $k < \infty$. By the Hunt-Stein theorem, Kiefer (1957), there exists c such that

$$E_\theta(X + c - \theta)^2 < k.$$

But this is impossible since $E_0(X^2) = \infty$.

We now discuss as application the completeness of the Cauchy densities and the estimation of the parameters therein.

Let $p_0(x) = \pi^{-1}(1 + x^2)^{-1}$ and $p_\theta(x) = p_0(x - \theta)$. Since $\hat{p}_0(t) = e^{-|t|}$ does not vanish anywhere, it follows from Theorem 2.4 that the family $\mathcal{P}_\theta = \{p_\theta\}$ is boundedly complete. By using Theorems 1 and 2 on Poisson transforms, page 123, Hoffman (1962), one can show that $\mathcal{P}_\theta$ is L^p-complete. In fact a stronger result is true. The family $\mathcal{P}_\theta$ is complete [6]. That the family $\mathcal{P}_{\theta,\sigma} = \{p_{\theta,\sigma} : p_{\theta,\sigma}(x) = \pi^{-1}\sigma/[\sigma^2 + (x - \theta)^2]\}$ of Cauchy densities involving both location and scale parameters is complete follows from this or can be proved by using the first theorem on page 123 of Hoffman (1962). The question of the completeness of the family $\mathcal{P}_\sigma = \{p_\theta : p_\theta(x) = \pi^{-1}\sigma/(\sigma^2 + x^2)\}$ is still open.

Since the family $\mathcal{P}_\theta$ is complete, it follows from Theorem 2.1 that there does not exist any unbiased estimator of θ based on a single observation. We shall give another proof of a slightly weaker form of this result. This proof has the advantage of being quite elementary and may be useful for sample size $\geqq 2$ when the method of completeness fails. For convenience we refer to the following results as propositions.

Proposition 1. *If $p_\theta(x) = \pi^{-1}[1 + (x - \theta)^2]^{-1}$, θ does not have an unbiased estimator with finite variance.*

Proof. Let $p_0(x) = \pi^{-1}(1 + x^2)^{-1}$ and $p_\theta(x) = p_0(x - \theta)$. Since $r_{\theta,0}$ is bounded for all θ, θ has an unbiased estimator with finite variance for all θ if and only if it has finite variance under $\theta = 0$. Also $R_0 \subset L^2(P_0)$. Hence we can use (1) to decide whether θ has an unbiased estimator with finite variance.

After squaring and simplification (1) becomes

$$\sum_1^m a_i^2\theta_i^2 + 2\sum\sum_{1\leqq i<j\leqq m} a_ia_j\theta_i\theta_j$$
$$\leqq C^2[2\sum\sum_{1\leqq i<j\leqq m} a_ia_j\{1 + 2\theta_i\theta_j[4 + (\theta_i - \theta_j)^2]^{-1}\} + \sum_{i=1}^m a_i^2(1 + \theta_i^2/2)].$$

If we let $\theta_j = m^\alpha + jm^\beta$, $j = 1, \cdots, m$, with $\alpha > \beta > 1$, it turns out that the left hand side of (1) is of higher order than the right hand side of (1) as $m \to \infty$. This shows that (1) does not hold and our proposition is proved.

It is worth noticing that with respect to the squared error loss the natural estimator X of θ is inadmissible compared with a trivial estimator $T_0(x) = \theta_0$ for all x. This indicates that for this problem the squared error loss is a very poor choice. Though we have not been able to find a loss function with respect to which X can be shown to be admissible, we believe that many such loss functions exist.

It is known that for a sample of size $n \geqq 5$ from the Cauchy density $p_\theta = \pi^{-1}[1 + (x - \theta)^2]^{-1}$, the sample median in an unbiased estimator of θ with finite variance, but we have not been able to settle the question for $1 < n < 5$. Incidentally the computation of the right hand side of (1) does not present any additional difficulties. It is easy to check that the right hand side of (1) becomes on squaring $\sum_i\sum_j a_ia_j\{1 + 2\theta_i\theta_j[4 + (\theta_i - \theta_j)^2]^{-1}\}^n$. For $n \geqq 7$ there exists an unbiased estimator which is admissible with respect to the squared error loss, Stein (1959).

Proposition 2. *If $p_\theta(x) = \pi^{-1}[1 + (x - \theta)^2]^{-1}$, then there does not exist an unbiased estimator for $\theta^{2+\delta}$ if $\delta > 0$.*

Proof. It is easily checked that $R_0 \subset L^\infty(\mu)$. It is obvious that if an unbiased estimator for $\phi(\theta) = \theta^{2+\delta}$ exists then

$$[\phi(\theta)]^2 \leqq C\|r_{\theta,\theta_0}\|^2_{\infty,\theta_0} \tag{4}$$

for some $C < \infty$ where $\|f\|_{\infty,\theta_0}$ is the essential supremum norm of $f \varepsilon L^{\infty}(P_{\theta_0})$. Now it is easily checked that $\|r_{\theta,0}\|_\infty = \sup_x \{(1 + x^2)[1 + (x - \theta)^2]^{-1}\} = O(\theta^2)$ as $\theta \to \infty$. Hence (4) fails to hold when $\theta \to \infty$ proving the proposition thereby.

Next we consider the case of unknown scale parameter (with known location parameter), i.e., we consider estimation of σ in $p_\sigma(x) = \pi^{-1}\sigma/(\sigma^2 + x^2)$.

PROPOSITION 3. *If* $p_\sigma(x) = \pi^{-1}\sigma/(\sigma^2 + x^2)$, *the scale parameter* σ *does not have an unbiased estimator with finite variance.*

PROOF. Since $r_{\sigma,1} = \sigma(1 + x^2)/(\sigma^2 + x^2)$ is bounded for each σ, σ has an unbiased estimator with finite variance if and only if it has finite variance under $\sigma = 1$. Since $R_1 \subset L^2(P_1)$, we shall use (1) to prove the proposition. Take $m = 1$. Then, after some simple calculations, we have

$$[\text{left hand side of } (1)/\text{right hand side of } (1)]^2 = 2\sigma^3/(1 + \sigma^2) \to \infty \quad \text{as} \quad \sigma \to \infty.$$

Consequently (1) does not hold and our proposition is proved.

Finally consider the case when both location and scale parameters are unknown, i.e., $p_{\theta,\sigma} = \pi^{-1}\sigma/(\sigma^2 + (x - \theta)^2)$. Then, as discussed above, the family $\mathcal{P}_{\theta,\sigma} = \{p_{\theta,\sigma}\}$ is complete. Also $E_{0,1}|X| = \infty$. Hency, by Theorem 2.3, we prove the following:

PROPOSITION 4. *If* $p_{\theta,\sigma}(x) = \pi^{-1}\sigma/(\sigma^2 + (x - \theta)^2)$, *then neither* θ *nor* σ *has an unbiased estimator.*

Acknowledgment. We thank Professors I. S. Luthar, M. Rajagopalan and P. K. Sen for some helpful discussions.

REFERENCES

[1] BARANKIN, E. W. (1949). Locally best unbiased estimates. *Ann. Math. Statist.* **20** 477–501.

[2] GOLDBERG, R. R. (1962). *Fourier Transforms.* Cambridge Tract 52.

[3] HOFFMAN, K. (1962). *Banach Spaces of Analytic Functions.* Prentice Hall, Englewood Cliffs.

[4] KIEFER, J. (1957). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28** 573,–601.

[5] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[6] POLLARD, H. (1955). The Poisson transform. *Trans. Amer. Math. Soc.* **78** 541–550.

[7] STEIN, C. (1950). Unbiased estimates with minimum variance, *Ann. Math. Statist.* **21** 406–415.

[8] STEIN, C. (1959). The admissibility of Pitmans' estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.