ON THE LACK OF A UNIFORMLY CONSISTENT SEQUENCE OF ESTIMATORS OF A DENSITY FUNCTION IN CERTAIN CASES

By R. H. Farrell

Cornell University[1]

1. Introduction. We will let C_α be the set of distribution functions F on $R = (-\infty, \infty)$ with the following properties:
(1) F has a first derivative f_F defined and continuous at all points of R, such that $\sup_{x \varepsilon R} f_F(x) \leqq \alpha$.
(2) F has a second derivative defined and continuous at all points of R.

Let $\{\delta_n, n \geqq 1\}$ be a sequence of functions such that if $n \geqq 1$ then δ_N is a real valued Borel measurable function on R_N (Euclidean n-space.) Let $\{X_n, n \geqq 1\}$ be a sequence of independently and identically distributed random variables such that if F is the distribution function of X_1 then $F \varepsilon C_\alpha$. We let N be a stopping variable relative to $\{X_n, n \geqq 1\}$ and F, and consider the sequential estimator $\delta_N = \delta_N(X_1, \cdots, X_N)$ of $f_F(0)$.

Loss will be measured by square error, so that the risk function of $\delta = \{\delta_n, n \geqq 1, X_m, m \geqq 1, N\}$ is

$$R(F, \delta) = \sum_{n=1}^{\infty} \int \cdots \int_{\{N=n\}} (\delta_n(x_1, \cdots, x_n) - f_F(0))^2 \prod_{i=1}^{n} f_F(x_i)\, dx_i .$$

Theorem. *Suppose that* $\alpha \geqq 3$ *is a real number. If* $\sup_{F \varepsilon C_\alpha} E_F N < \infty$, *then* $\sup_{F \varepsilon C_\alpha} R(F, \delta) \geqq \frac{1}{16}$.

Since fixed sample size procedures satisfy the hypothesis of the theorem, we conclude that $\sup_{F \varepsilon C_\alpha} R(F, \delta) \geqq \frac{1}{16}$ for every choice of a fixed sample size procedure and choice of $\alpha \geqq 3$. It should be observed that this theorem does not deny the existence of a consistent sequence of estimators. For example, if $\{X_n, n \geqq 1\}$ are independently and identically distributed, if the distribution function of X_1 is F, and if F_n is the sample distribution function based on $X_1, \cdots, X_n$, $n \geqq 1$, then let $\beta = \frac{1}{3}$ and define $\delta_n = (F_n(n^{-\beta}) - F_n(-n^{-\beta}))/(2n^{-\beta})$. Then $\lim_{n\to\infty} E_F \delta_n = f_F(0)$ so that asymptotically the bias of δ_n goes to zero. Clearly the variance of δ_n goes to zero. The theorem says in effect that no *uniformly* consistent sequence of estimators exists relative to the class C_α.

In recent years there have appeared several papers discussing methods of estimation of the value of a density function. Papers which have come to our attention are Parzen [4] and Leadbetter [3]. In addition the author has had no difficulty in inventing several methods of estimation quite different from those of Parzen and Leadbetter (it is not our purpose to discuss these here.) All these methods have a common characteristic. If C_α^* is the set of F with continuous

Received 14 July 1966.

[1] Research supported in part by the Office of Naval Research under contract NONR 401(50).

second derivative bounded by α then $\lim_{n\to\infty} n^{2/3} \sup_{F \varepsilon C_\alpha^*} R(F, \delta_n) > 0$. The author believes $\lim_{n\to\infty} \inf_{\delta_n} \sup_{F\varepsilon C_\alpha^*} R(F, \delta_n) > 0$, he has not been able to prove this. The best result that he has been able to obtain is stated in the above theorem.

2. Proof of the theorem. First observe that if $F \varepsilon C_\alpha'$, the set of distributions with continuous first derivative bounded by α and sectionally smooth second derivative, then there exists a sequence $\{F_n, n \geqq 1\}$ of elements of C_α such that $\lim_{n\to\infty} f_{F_n}(x) = f_F(x)$ for almost all $x \varepsilon R$. It then follows at once by Fatou's lemma that $R(F, \delta) \leqq \lim_{n\to\infty} \inf R(F_n, \delta) \leqq \sup_{F' \varepsilon C_\alpha} R(F', \delta)$. Thus, since $C_\alpha' \supset C_\alpha$, it follows that $\sup_{F\varepsilon C_\alpha'} R(F, \delta) = \sup_{F\varepsilon C_\alpha} R(F, \delta)$. Therefore it suffices to prove the theorem relative to C_α'.

We prove our result by an argument similar to that used by Wald [5]. We give give an example of a two parameter family of distribution functions contained in C_α' for which the conclusion of the theorem holds, $\alpha \geqq 3$. We establish the desired result using the sequential Cramér-Rao inequality. See Lehmann [2] and Chow, Robbins and Teicher [1].

To construct the density functions, let $\gamma > 0$ be a real number and $C(\gamma) > 0$ be the constant defined by $1 = 2C(\gamma)(e^{-\gamma} + \gamma e^{-\gamma} + \gamma/2)$. Let $f(\cdot, \gamma)$ be the density function satisfying $\lim_{t\to-\infty} f(t, \gamma) = 0$ and having partial derivative

$$\begin{aligned} f_1(t, \gamma) &= C(\gamma)e^{-|t|} && \text{if } t > \gamma; \\ f_1(t, \gamma) &= C(\gamma)\gamma^{-1} && \text{if } 0 < t < \gamma; \\ f_1(t, \gamma) &= -f_1(-t, \gamma) && \text{for all real } t. \end{aligned}$$

Integration then gives $f(t, \gamma) = C(\gamma)e^{-|t|}$ if $|t| \geqq \gamma$, and $f(t, \gamma) = C(\gamma)(e^{-\gamma} + 1 - |t|/\gamma)$ if $0 \leqq |t| \leqq \gamma$. It is easily checked that the choice of $C(\gamma)$ given above makes this a density function.

We will apply the Cramér-Rao inequality to the two parameter family $g(t, \theta, \gamma) = f(t + \theta, \gamma)$, $-\infty < t, \theta < \infty$ and $\gamma > 0$. If $n \geqq 1$ let δ_n be a real valued Borel measurable function on R_n (Euclidean n-space) such that

$$\int \cdots \int (\delta_n(x_1, \cdots, x_n))^2 \prod_{i=1}^n f_F(x_i)\, dx_i < \infty \qquad \text{for all } F \varepsilon C_\alpha .$$

Since if $\theta_1 \neq \theta_2$ then $\sup_{-\infty<t<\infty} f(t + \theta_1, \gamma)/f(t + \theta_2, \gamma) < \infty$, the finiteness of

$$\int \cdots \int |\delta_n(x_1, \cdots, x_n)| \prod_{i=1}^n f(x_i + \theta, \gamma)\, dx_i \qquad \text{implies}$$

$$\int \cdots \int |\delta_n(x_1, \cdots, x_n)| \prod_{i=1}^n f(x_i + \theta_i, \gamma)\, dx_i < \infty$$

$$\text{for all } (\theta_1, \cdots, \theta_n) \varepsilon R_n .$$

Let $a(\theta_1, \cdots, \theta_n) = \int \cdots \int \delta_n(x_1, \cdots, x_n) \prod_{i=1}^n f(x_i + \theta_i, \gamma)\, dx_i$ and let a_i be the partial derivative of a on the ith variable. Then $d/d\theta a(\theta, \cdots, \theta) = \sum_{i=1}^n a_i(\theta, \cdots, \theta)$. Therefore to verify the differentiability conditions of the Cramér-Rao inequality it is sufficient to verify for $j = 1, \cdots, n$ that

$$* \quad a_j(\theta_1, \cdots, \theta_n) = \int \cdots \int \delta_n(x_1, \cdots, x_n)(f_1(x_j + \theta_j, \gamma)/f(x_j + \theta_j, \gamma)) \cdot \prod_{i=1}^n f(x_i + \theta_i, \gamma)\, dx_i$$

From the definition of f it is easy to verify that there exists a constant $C_1(\gamma)$ such that for all real x_1, and x_2 such that $|x_1 - x_2| < 1$, it follows that $|f(x_1, \gamma) - f(x_2, \gamma)| \leqq C_1(\gamma)|x_1 - x_2|f(x_1, \gamma)$. Therefore using the definition of a derivative and the bounded convergence theorem, formula (*) follows.

From this discussion we may conclude that

$$0 = \int \cdots \int \left(\sum_{j=1}^{n} (f_1(x_j + \theta_j, \gamma)/f(x_j + \theta_j, \gamma)) \right) \prod_{i=1}^{n} f(x_i + \theta_i, \gamma)\, dx_i$$

and that

$$(d/d\theta)a(\theta, \cdots, \theta) = \int \cdots \int \delta_n(x_1, \cdots, x_n) \sum_{j=1}^{n} (f_1(x_j + \theta, \gamma)/f(x_j + \theta, \gamma))) \cdot \prod_{i=1}^{n} f(x_i + \theta, \gamma)\, dx_i .$$

It follows that the differentiability conditions for application of the Cramér-Rao inequality are satisfied and that the fixed sample size inequality may be applied.

For brevity we let

$$(C_2(\gamma))^{-1} = \int ((f_1(x, \gamma))^2/f(x, \gamma))\, dx = 2C(\gamma)(e^{-\gamma} + \gamma^{-1} \log (1 + e^{\gamma})).$$

Therefore $\lim_{\gamma\to 0+} \gamma(C_2(\gamma))^{-1} = \log 2$. In the sequential case we need to know that

$$\sum_{n=1}^{\infty} \int \cdots \int_{\{N=n\}} \delta_n(x_1, \cdots, x_n) \prod_{i=1}^{n} f(x_i + \theta, \gamma)\, dx_i$$

can be differentiated term by term. The classical theorem of advanced calculus says that this may be done provided

$$\sum_{n=1}^{\infty} \int \cdots \int_{\{N=n\}} |\delta_n(x_1, \cdots, x_n)|\, \left|\sum_{j=1}^{n} f_1(x_j + \theta, \gamma)/f(x_j + \theta, \gamma)\right| \cdot \prod_{i=1}^{n} f(x_i + \theta, \gamma)\, dx_i \leqq (\text{var } \delta_N)^{\frac{1}{2}}((EN)C_2(\gamma)^{-1})^{\frac{1}{2}} < \infty$$

The finiteness of the last expression follows from results given in Chow, Robbins and Teicher [1]. Therefore the differentiability conditions for application of the sequential Cramér-Rao inequality hold.

Let $F \varepsilon C_\alpha$ and suppose δ_N is an estimator of $f_F(0)$. In the sequel we let $b(\theta, \gamma)$ the bias of δ_N when F has as density function $f(\cdot + \theta, \gamma)$, so that the value to be estimated is $f(\theta, \gamma)$. As we have seen above, if $\theta \neq 0, \pm\gamma$, then $\partial/\partial\theta b(\theta, \gamma)$ exists.

If $R(\theta, \gamma, \delta)$ is the risk function of δ_N then from the Cramér-Rao inequality we obtain, using the assumption $\sup_{F\varepsilon C_\alpha} E_F N \leqq \beta$,

$$R(\theta, \gamma, \delta) \geqq (b(\theta, \gamma))^2 + \beta^{-1}C_2(\gamma)(f_1(\theta, \gamma) + \partial/\partial\theta b(\theta, \gamma))^2,$$

this being valid so long as the derivatives in question exist. Let $0 < \delta < \gamma/2$. Then $\sup_{|\theta-\gamma/2|<\delta} R(\theta, \gamma, \delta)$ is greater than or equal the larger of $\sup_{|\theta-\gamma/2|<\delta} \cdot (b(\theta, \gamma))^2$ and $\sup_{|\theta-\gamma/2|<\delta} \beta^{-1}C_2(\gamma)(f_1(\theta, \gamma) + \partial/\partial\theta b(\theta, \gamma))^2$. If $\partial/\partial\theta b(\theta, \gamma) \leqq -\epsilon f_1(\theta, \gamma) = \epsilon C(\gamma)/\gamma$ for some θ satisfying $|\theta - \gamma/2| < \delta$ then $\sup_{|\theta-\gamma/2|<\delta} R(\theta, \gamma, \delta) \geqq \beta^{-1}C_2(\gamma)(1 - \epsilon)^2C(\gamma)^2/\gamma^2$. In the other case it must be that $\partial/\partial\theta b(\theta, \gamma) \geqq \epsilon C(\gamma)/\gamma$ for all θ in $(\gamma/2 - \delta, \gamma/2 + \delta)$. Thus in this interval $b(\cdot, \gamma)$ is a strictly increasing function and $b(\gamma/2 + \delta, \gamma) - b(\gamma/2 - \delta, \gamma) > 2\epsilon C(\gamma)\delta/\gamma$. Therefore there exists a θ in $(\gamma/2 - \delta, \gamma/2 + \delta)$ such that $(b(\theta, \gamma))^2 > \epsilon^2 C(\gamma)^2\delta^2/\gamma^2$. Therefore we find that

$$\sup_{|\theta-\gamma/2|<\delta} R(\theta, \gamma, \delta) \geqq \min(\epsilon^2 C(\gamma)^2\delta^2/\gamma^2, \beta^{-1}C_2(\gamma)C(\gamma)^2(1 - \epsilon)^2/\gamma^2).$$

If we give δ its largest possible value, $\gamma/2$, then the first term of the minimum becomes $\epsilon^2 C(\gamma)^2/4$ which tends to $\epsilon^2/16$ as $\gamma \to 0+$. Also

$$\lim_{\gamma\to 0+} \beta^{-1}C_2(\gamma)C(\gamma)^2(1 - \epsilon)^2/\gamma^2 = \infty.$$

Since $0 < \epsilon < 1$ but ϵ is otherwise arbitrary we find

$$\lim_{\gamma\to 0+} \sup \sup_{|\theta-\gamma/2|<\gamma/2} R(\theta, \gamma, \delta) \geqq \tfrac{1}{16},$$

which implies the conclusion of the theorem.

REFERENCES

[1] Chow, Y. S., Robbins, Herbert and Teicher, Henry (1965). Moments of randomly stopped sums. *Ann. Math. Statist.* **36** 789–799.

[2] Lehmann, E. L. (1950). *Notes on the theory of estimation.* Associated Students Store, Berkeley, California.

[3] Leadbetter, M. R. and Watson, G. S. (1963). On the estimation of a probability density, I. *Ann. Math. Statist.* **34** 480–491.

[4] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

[5] Wald, A. (1951). Asymptotic minimax solutions of sequential point estimation problems. *Proc. Second Berkeley Symp. on Math. Statist. Prob.* Univ. of California Press.