The Annals of Statistics
1973, Vol. 1, No. 5, 944-947

MAXIMUM LIKELIHOOD ESTIMATION OF A TRANSLATION PARAMETER OF A TRUNCATED DISTRIBUTION

BY L. WEISS[1] AND J. WOLFOWITZ[2]

Cornell University and University of Illinois

$f(x)$ is a uniformly continuous density which equals zero for negative values of x, has a right-hand derivative equal to α at $x = 0$, where $0 < \alpha < \infty$, and satisfies certain regularity conditions. $X_1, \cdots, X_n$ are independent random variables with the common density $f(x - \theta)$, θ an unknown parameter. Let $\hat{\theta}_n$ denote the maximum likelihood estimator of θ, and define α_n by the equation $2\alpha_n^2 = \alpha n \log n$. It was shown by Woodroofe that the asymptotic distribution of $\alpha_n(\hat{\theta}_n - \theta)$ is standard normal. It is shown in the present paper that $\hat{\theta}_n$ is an asymptotically efficient estimator of θ.

Let f be a uniformly continuous density which vanishes on $(-\infty, 0]$ and is subject to regularity conditions to be described. Among these conditions is one which requires that $\alpha = \lim f'(x)$ exists as $x \to 0$ from the right, with $0 < \alpha < \infty$. Let $\Theta = (-\infty, \infty)$ be the parameter space of the unknown parameter θ. Let $X_1, \cdots, X_n$ be independent chance variables with the common density $f(x - \theta)$, at the point x of the real line. Let $\{\hat{\theta}_n\}$ be a consistent sequence of roots of the likelihood equation. It was proved by Woodroofe ([4]) under two different sets of regularity conditions (either of which we henceforth adopt) that, for any θ,

$$\lim_{n\to\infty} P_\theta\{\alpha_n(\hat{\theta}_n - \theta) < y\} = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{y} e^{-y^2/2}\, dy \tag{1}$$

for any y, where $2\alpha_n^2 = \alpha n \log n$.

It is obvious that we are here dealing with what is called a "non-regular" case (see, for example, [1]) since the normalizing factor is not $n^{\frac{1}{2}}$. Consequently the question remains open whether the maximum likelihood (m.l.) estimator $\hat{\theta}_n$ is asymptotically efficient. If a proof of this is not available then only faith in the eventual appearance of such a proof would justify the statistician's use of the m.l. estimator. In this note we prove the asymptotic efficiency of the m.l. estimator by proving that it is asymptotically equivalent to a maximum probability (m.p.) estimator ([2], [3]). The precise statement of efficiency is given in the theorem below. Our proof will be brief and will utilize some results of [4].

Let θ_0 now be any fixed point in Θ. We shall say that a sequence of functions $\{l_n(\cdot)\}$ converges in $H(h)$ (to a constant) if the following is true: Let $\{y_n, n = 1, 2, \cdots\}$ be any sequence of real numbers such that $|\alpha_n(y_n - \theta_0)| \leqq h$ for $n = 1, 2, \cdots$. Then the sequence of real numbers $\{l_n(y_n)\}$ converges (to the constant).

Received August 1972; revised January 1973.

[1] Research supported by NSF Grant GP 31430X.

[2] Research supported by the U. S. Air Force under Grant AF-AFOSR 70-1947, monitored by the Office of Scientific Research.

Let $R = (-r, r)$ be any interval centered at the origin. We now state and prove the following.

THEOREM. *Let $\{T_n\}$ be any competing sequence of estimators such that, for any $h > 0$, we have, for θ in $H(h)$,*

(2) $$\lim_{n\to\infty} [P_\theta\{\alpha_n(T_n - \theta) \text{ in } R\} - P_{\theta_0}\{\alpha_n(T_n - \theta_0) \text{ in } R\}] = 0 .$$

Then

(3) $$\limsup_{n\to\infty} P_{\theta_0}\{\alpha_n(T_n - \theta_0) \text{ in } R\} \leqq \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-r}^{r} e^{-y^2/2}\, dy .$$

This is the statement of efficiency for $\hat{\theta}_n$. If, as is usually required in the literature (but not necessary for us), $\alpha_n(T_n - \theta)$ is also asymptotically normally distributed (in P_θ-probability for every θ) with mean 0 and variance $\sigma_\theta^2(T)$, from (1) and the Theorem we immediately obtain that

(4) $$\sigma_\theta^2(T) \geqq 1 , \qquad \theta \text{ in } \Theta .$$

This is the classical statement of asymptotic efficiency.

PROOF OF THE THEOREM. An m.p. estimator Z_n can be obtained as follows. Let $\{\lambda_n\}$ be a sequence of positive numbers which approach zero. Then Z_n satisfies

(5) $$\alpha_n \int_{Z_n - r/\alpha_n}^{Z_n + r/\alpha_n} \prod_{i=1}^{n} f(X_i - \theta)\, d\theta \geqq \alpha_n(1 - \lambda_n) \sup_d \int_{d - r/\alpha_n}^{d + r/\alpha_n} \prod_{i=1}^{n} f(X_i - \theta)\, d\theta .$$

Define $t = \alpha_n(\theta - \hat{\theta}_n)$ and

(6) $$V_n(t) = [\prod_{i=1}^{n} f(X_i - \theta)][\prod_{i=1}^{n} f(X_i - \hat{\theta}_n)]^{-1} .$$

Since the second factor of $V_n(t)$ does not depend on θ, we may rewrite (5) as

(7) $$\int_{\alpha_n(Z_n - \hat{\theta}_n) - r}^{\alpha_n(Z_n - \hat{\theta}_n) + r} V_n(t)\, dt \geqq (1 - \lambda_n) \sup_d \int_{\alpha_n(d - \hat{\theta}_n) - r}^{\alpha_n(d - \hat{\theta}_n) + r} V_n(t)\, dt .$$

Let $k(\cdot)$ be any positive function defined on the positive integers such that, as $n \to \infty$,

(8) $$k(n) \to \infty , \qquad \alpha_n^{-1} k(n) \to 0 .$$

It is a consequence of a remark made in [2] (proved in greater detail in [3]) that, in the present problem, Z_n remains asymptotically efficient if the supremum operation in the right member of (5) is performed with respect to d in the θ-interval

(9) $$\left[\Psi_n(X_1, \cdots, X_n) - \frac{k(n)}{2} \alpha_n^{-1}, \Psi_n(X_1, \cdots, X_n) + \frac{k(n)}{2} \alpha_n^{-1}\right]$$

where Ψ_n is any estimator of θ_0 such that

(10) $$|\Psi_n - \theta_0| = O_p(\alpha_n^{-1}) .$$

Of course, the length of the interval (9) approaches zero. As Ψ_n we choose $\hat{\theta}_n$, which, it follows from (1), satisfies (10). For $k(\cdot)$ we have a choice among many functions, and we choose

(11) $$k(n) = (\log n)^{\frac{1}{2}} , \quad \forall n .$$

It will be shown in the Appendix that Lemmas 3.4 and 3.5 of [4] hold with their constant k replaced by our $k(n)$; we assume this for the moment. It then follows from this version of these lemmas that

$$\sup_t \frac{|\log V_n(t) + \frac{1}{2}t^2|}{t^2} \tag{12}$$

converges to 0 in P_{θ_0}-probability as $n \to \infty$; the supremum in (12) is with respect to t in the t-interval

$$[\hat{\theta}_n - (\log n)^{\frac{1}{2}}, \hat{\theta}_n + (\log n)^{\frac{1}{2}}] . \tag{13}$$

(When $t = 0$ the expression being maximized in (12) becomes 0/0, and we define it as 0.) The t-interval (13) contains the t-interval into which the θ-interval (9) with the present choice of Ψ_n and $k(n)$ is transformed by the relation $t = \alpha_n(\theta - \hat{\theta}_n)$. From the above we conclude that, if we set $Z_n = \hat{\theta}_n$, there exists a sequence $\{\lambda_n\}$ such that (7) is satisfied with P_{θ_0}-probability approaching one. This proves that $\hat{\theta}_n$ is asymptotically equivalent to the m.p. estimator Z_n.

It remains only to prove that $\hat{\theta}_n$ satisfies the conditions (3.2) and (3.3) of [2] (or the conditions (3.5) and (3.6) of [3]). This follows immediately from (1) and the fact that θ is a translation parameter. This proves the Theorem, so $\hat{\theta}_n$ is asymptotically efficient.

As mentioned earlier, the $\{\hat{\theta}_n\}$ of [4] is a *consistent* sequence of roots of the likelihood equation. Since the statistician solves the likelihood equation for a particular value of n, how is he to recognize which roots belong to a consistent sequence and which do not? (As pointed out to one of us by the late Professor Abraham Wald, the same question can be raised about Cramér's proof of the consistency of the maximum likelihood estimator [1, Section 33.3].) The following procedure may be helpful; it is an application to the present case of the remark made in [2] and used by us earlier in this paper. Let $k_1(\cdot)$ be any function on the positive integers such that $k_1(n) \uparrow \infty$, $n^{-\frac{1}{2}}k_1(n) \to 0$. Then any consistent sequence will eventually lie in an interval of length $n^{-\frac{1}{2}}k_1(n)$ centered at $\min(X_1, \cdots, X_n)$.

APPENDIX

Define

$$M_n = \min(X_1, \cdots, X_n), \qquad N_n = \max(X_1, \cdots, X_n).$$

An examination of the proofs of Lemmas 3.4 and 3.5 of [4] shows that the proofs would remain unchanged if our $k(n)$ were substituted for their k, excepting only that we must now show for Lemma 3.4 that

$$P_{\theta_0}\left[M_n - \theta_0 \geqq \frac{\delta_n}{\varepsilon} = \frac{k(n)}{\varepsilon\alpha_n}\right] \to 1 \tag{14}$$

and for Lemma 3.5 that, in addition,

$$P_{\theta_0}\left[b - N_n \geqq \frac{\delta_n}{1-\beta}\right] \to 1 . \tag{15}$$

For large n the left member of (14) is greater than

$$\left(1 - \frac{\alpha \delta_n^2}{\varepsilon^2}\right)^n = \left(1 - \frac{2(\log n)^{\frac{1}{2}}}{\varepsilon^2 n \log n}\right)^n \to 1 .$$

This proves (14). As for (15), this statement is weaker than the conclusion of Lemma 2.2 of [4], so it certainly holds.

The above argument shows that $k(\cdot)$ could have been any positive function which satisfies (8), (14), and (15).

REFERENCES

[1] CRAMÉR, H. (1961). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[2] WEISS, L. and WOLFOWITZ, J. (1967). Maximum probability estimators. *Ann. Inst. Statist. Math.* **19** 193–206.

[3] WEISS, L. and WOLFOWITZ, J. *Asymptotic Methods in Statistics*, Chapter 3. In preparation.

[4] WOODROOFE, M. (1972). Maximum likelihood estimation of a translation parameter of a truncated distribution. *Ann. Math. Statist.* **43** 113–122.

DEPARTMENT OF OPERATIONS RESEARCH
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801