AN INCONSISTENCY OF THE METHOD OF MAXIMUM LIKELIHOOD

By D. Basu

University of California, Berkeley

An example was given by Neyman and Scott [2] to show that there are situations where the method of maximum likelihood leads to inconsistent estimators. In their example considered, the observations were supposed to be drawn from an infinite sequence of distinct populations involving an infinite sequence of nuisance parameters.

An example is given here to demonstrate that even in simple situations where all the observations are independently and identically distributed and involve only one unknown parameter, the method of maximum likelihood may lead us astray. The example typifies the situations where the correct method of setting up a point estimate should begin with a test of hypothesis between two composite alternatives.

Let A be the set of all rational numbers in the closed interval (0, 1) and B any countable set of irrational numbers in the same interval. Let X be a random variable that takes the two values 0 and 1 with

$$P(X = 1) = \begin{cases} \theta & \text{if } \theta \varepsilon A, \\ 1 - \theta & \text{if } \theta \varepsilon B. \end{cases}$$

If r is the total number of 1's in a set of n random observations on X, then from the rationality of r/n it follows at once that the maximum likelihood estimator of θ is r/n. But r/n converges (in probability) to θ or $1 - \theta$ according as $\theta \varepsilon A$ or $\theta \varepsilon B$.

Now, since A and B are both countable sets, it follows [1] that there exists a consistent test for the composite hypothesis $\theta \varepsilon A$ against the composite alterna-

Received March 19, 1954.

A. DasGupta (ed.), *Selected Works of Debabrata Basu*, Selected Works in Probability and Statistics, DOI 10.1007/978-1-4419-5825-9_13,

tive $\theta \varepsilon B$. In other words, there exists a function φ_n of the first n observations such that φ_n takes only the two values 0 and 1 and such that

$$P(\varphi_n = 1 | \theta) \rightarrow \begin{cases} 0 & \text{if } \theta \varepsilon A, \\ 1 & \text{if } \theta \varepsilon B, \end{cases} \qquad n \rightarrow \infty.$$

It then follows readily that the estimator

$$t_n = (1 - \varphi_n)r/n + \varphi_n(1 - r/n)$$

is a consistent estimator of θ. Thus, though there exist consistent estimators for θ, the maximum likelihood estimator is not consistent.

For a simple construction for the function φ_n, let $\{\alpha_i\}$ and $\{\beta_i\}$, for $i = 1, 2, \cdots$, be two enumerations of the sets A and B, respectively, and let δ_k be the distance between the two sets $(\alpha_1, \cdots, \alpha_k)$ and $(1 - \beta_1, \cdots, 1 - \beta_k)$. Let $k(n)$ be the largest integer k such that $\delta_k > n^{-1/4}$. Note that $k(n)$ increases monotonically to infinity. Let I_{kn} and J_{kn} be the open intervals of length $n^{-1/4}$ centered around α_k and $1 - \beta_k$ respectively and let

$$R_n = \bigcup_{k \leq k(n)} I_{kn}, \qquad S_n = \bigcup_{k \leq k(n)} J_{kn}.$$

For every n, the sets R_n and S_n are clearly disjoint. Now consider a fixed k. For all n for which $k(n) \geq k$ we have

$$\begin{aligned} P(r/n \varepsilon S_n \mid \theta = \beta_k) &\geq P(r/n \varepsilon J_{kn} \mid \theta = \beta_k) \\ &= P(|r/n - [1 - \beta_k]| < -n^{-1/4} \mid \theta = \beta_k) \\ &\rightarrow 1 \text{ as } n \rightarrow \infty, \end{aligned}$$

because r/n is asymptotically normal with mean $1 - \beta_k$ and asymptotic s.d. $\sqrt{\beta_k(1 - \beta_k)/n}$. By the same argument we have

$$\begin{aligned} P(r/n \varepsilon S_n \mid \theta = \alpha_k) &\leq 1 - P(r/n \varepsilon R_n \mid \theta = \alpha_k) \\ &\rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

Now, if we define

$$\varphi_n = \begin{cases} 1 & \text{if } r/n \varepsilon S_n \\ 0 & \text{otherwise,} \end{cases}$$

then φ_n clearly has the required property.

Acknowledgement. I wish to thank Mr. J. L. Hodges, Jr., for his help in the writing of this note.

REFERENCES

[1] CHARLES KRAFT, "On the problem of consistent and uniformly consistent statistical procedures", Unpublished Ph.D. thesis submitted to the University of California (1954).

[2] J. NEYMAN AND E. L. SCOTT, "Consistent estimates based on partially consistent observations," *Econometrica*, Vol. 16 (1948), pp. 1–32.