THE ASYMPTOTIC EFFICIENCY OF A MAXIMUM LIKELIHOOD ESTIMATOR

H. E. DANIELS
UNIVERSITY OF BIRMINGHAM

1. Introduction and summary

The consistency of a maximum likelihood estimator has been established under very general conditions by Wald [6] and Wolfowitz [7]. Much more stringent conditions are needed for it to be asymptotically efficient, that is, consistent and asymptotically normal with variance equal to the Cramér-Rao lower bound. Typical conditions are given by Cramér [2], Gurland [3], Kulldorf [4], all of which restrict the behavior of at least the second derivative of the likelihood function. Authors such as, for example, Le Cam [5] and Bahadur [1] discuss large sample estimation in a more general context but still require regularity conditions on the second derivative of the likelihood for the maximum likelihood estimator to be asymptotically efficient.

However, cases are known which are not covered by these regularity conditions. The density function $f(x, \theta) = (1/2) \exp - |x - \theta|$ provides an example. The sample median is a maximum likelihood estimator of θ. It is known to be asymptotically normal with variance n^{-1}, which is the Cramér-Rao lower bound. But $\partial \log f/\partial\theta$ is discontinuous and $\partial^2 \log f/\partial\theta^2$ is zero for almost all x.

In the present paper weaker conditions for asymptotic efficiency are given which do not involve the second derivative of the likelihood. Two sets of sufficient conditions are stated. From the first, asymptotic efficiency can be proved directly without appeal to the Wald-Wolfowitz result but there is a convexity requirement which is frequently not satisfied. The second set of conditions dispenses with this requirement at the cost of some specialization elsewhere, but consistency has to be established by the Wald-Wolfowitz method. Finally a more general situation is considered where a modified maximum likelihood procedure is shown still to yield an asymptotically efficient estimator. The relation of this modified estimator to a class of smoothed estimators is indicated.

2. First set of sufficient conditions

We consider for simplicity a univariate distribution which has a probability density $f(x, \theta)$, where θ is a parameter which can take any value in an open interval Θ. With obvious changes the discussion will apply to discrete distributions also. Let $x_1, x_2, \cdots, x_n$ be a random sample S from such a distribution. Write $l(x, \theta) = \log f(x, \theta)$ and let $L(S, \theta) = \sum_{r=1}^{n} l(x_r, \theta)$ denote the log-likeli-