SUFFICIENCY AND APPROXIMATE SUFFICIENCY[1]

By L. Le Cam

University of California, Berkeley

1. Introduction. The present paper is essentially an investigation of the relations between various definitions of sufficiency and approximate sufficiency. Although some of the arguments described here may be applicable to sequential experimentation, the paper is concerned only with the case where the observations are taken in one single step.

Our main theorem is a version of the Blackwell-Sherman-Stein theorem on the comparison of experiments ([4], p. 328) in which a possibility of "sufficiency within ϵ" has been introduced.

The study of definitions of sufficiency is marred by technical difficulties of a measure theoretic nature, which may be judged rather irrelevant for ordinary statistical purposes. To avoid these difficulties we have been led to generalize the usual description of what is meant by an experiment, ignoring σ-additivity and other regularity conditions. The bulk of the paper is intended to show that such a generalization is very convenient in many respects. Furthermore, there is no essential difficulty in returning to the usual system after the main results have been proved.

The introduction of the term "sufficient statistic" is usually attributed to R. A. Fisher who gave several definitions of the concept (see [10], p. 316 and [11], 713). For the present purposes, Fisher's most relevant statement seems to be the requirement ". . . that the statistic chosen should summarize the whole of the relevant information supplied by the sample." Such a requirement may be made precise in various ways, some of which are loosely described below. Accurate definitions will be found in the main body of the present paper. The following three interpretations are the most common.

(A) *The classical, or operational definition of sufficiency* claims that a statistic S is sufficient if given the value of S one can proceed to a post-experimental randomization reproducing variables which have the same distributions as the originally observable variables.

(B) *The Bayesian interpretation.* A statistic S is sufficient if for every a priori distribution of the parameter the a posteriori distributions of the parameter given S is the same as if the entire result of the experiment was given.

(C) *The economic interpretation.* A statistic S is sufficient if for every decision problem and every decision procedure made available by the experiment there

Received 8 March 1961; revised 30 June 1964.

[1] This paper was started while the author was a fellow of the Alfred P. Sloan Foundation and terminated with the partial support of the U. S. Army Research Office (Durham), grant DA-ARO(D)-31-124-G83. It is a revision of a Special Invited Paper given at the Washington, D. C. meeting of the Institute on 29 December 1959.