ON MULTISTAGE ESTIMATION[1]

By J. R. Blum and Judah Rosenblatt[2]

University of New Mexico

1. Introduction. Let $\mathfrak{F}$ be a family of distribution functions and let $\theta(\cdot)$ be a real-valued functional defined on $\mathfrak{F}$. In this paper we shall be concerned for the most part with the problem of finding a confidence interval of preassigned length and confidence for $\theta(F)$ based on a sample from $F \varepsilon \mathfrak{F}$. For simplicity of notation we will assume that $\mathfrak{F}$ consists of univariate distributions. It will be clear that this restriction is not necessary.

Apparently Dantzig [3] was the first to point out that many such problems cannot be solved in a single stage of estimation; i.e. it is impossible to prescribe an integer n and give a confidence interval of preassigned length and confidence based on a sample of size n. Bahadur and Savage [2] showed that if $\mathfrak{F}$ is the class of all distributions for which the mean exists, it is impossible to obtain a confidence interval of prescribed length and confidence for the mean even with a purely sequential scheme. Intuitively this follows from the fact that no matter what data have been observed, there can exist a "small spike" close to $\pm \infty$ which affects the mean, but is not likely to affect the data. Farrell [6] showed that a purely sequential scheme is both necessary and sufficient for estimation of the median within the class of distributions possessing a unique median. This is plausible because to pin down the median, the sample median must be closely surrounded by sufficiently many other observations. Though with probability one this will occur, the necessary sample size is not determinable if it has not yet occurred. Farrell's results are actually considerably deeper, since he obtains the order of magnitude of the minimum expected sample size as the density at the median becomes small.

The earliest result of a positive nature is the paper by Stein [9] who gave a two-stage sampling procedure for estimation of the mean of a normal distribution with unknown variance. Graybill [7] gave sufficient conditions for two-stage estimation in certain parametric cases, while Weiss [10] showed that a two-stage scheme suffices for estimation of quantiles when $\mathfrak{F}$ is the class of unimodal distributions. Birnbaum and Healy [3] considered the problem of two-stage unbiased point estimation with fixed variance. Abbott and Rosenblatt [1] gave sufficient conditions for two-stage estimability with one observation on the first stage. A number of other papers treat these and related problems, e.g. Matthes [8].

Unless otherwise specified we shall assume throughout that there is available a sequence $X_1, X_2, \cdots$ of independent random variables with common distri-

Received September 29, 1962.

[1] Research supported by the National Science Foundation, Grant NSF G-21219.

[2] Sandia Corporation. Work performed under the auspices of the United States Atomic Energy Commission.