THE BEHAVIOR OF SOME STANDARD STATISTICAL TESTS UNDER NONSTANDARD CONDITIONS

HAROLD HOTELLING
UNIVERSITY OF NORTH CAROLINA

1. Introduction and summary

In the burst of new statistical developments that has followed the work of Sir Ronald Fisher in introducing and popularizing methods involving exact probability distributions in place of the old approximations, questions as to the effects of departures from the assumed normality, independence, and uniform variance have often been subordinated. It is true that much recognition has been given to the existence of serial correlation in time series, with the resultant vitiation of statistical tests carried out in the absence of due precautions, and to some other special situations, such as manifestly unequal variances in least-square problems. Wassily Hoeffding has established some general considerations on the role of hypotheses in statistical decisions [21]. Also, there have been many studies of distributions of the Student ratio, the sample variance, variance ratio, and correlation coefficient in samples from nonnormal populations. (Some are cited at the end of this paper.) These efforts have encountered formidable mathematical difficulties, since the distribution functions sought cannot usually, except in trivial cases, be easily specified or calculated in terms of familiar or tabulated functions. Because of these difficulties, mathematics has in some such studies been supplemented or replaced by experiment ([20], [38], [54], and others); or, as in some important work, approximations for which definite error bounds are not at hand have been used.

Practical statisticians have tended to disregard nonnormality, partly for lack of an adequate body of mathematical theory to which an appeal can be made, partly because they think it is too much trouble, and partly because of a hazy tradition that all mathematical ills arising from nonnormality will be cured by sufficiently large numbers. This last idea presumably stems from central limit theorems, or rumors or inaccurate recollections of them.

Central limit theorems have usually dealt only with linear functions of a large number of variates, and under various assumptions have proved convergence to normality as the number increases. For a large but fixed number the approxima-

This research was carried out in part under Office of Naval Research Project NR 042 031; Contract Nonr-855(06) and its predecessor contract.