ON THE EFFICIENCY OF OPTIMAL NONPARAMETRIC PROCEDURES IN THE TWO SAMPLE CASE[1]

By Piotr Witold Mikulski[2]

University of California, Berkeley

1. Summary. A series of papers has been published recently dealing with the efficiency of nonparametric procedures in testing statistical hypotheses. A frequently discussed problem is that of the efficiency of nonparametric procedures compared to some parametric methods in the two sample case, when the hypothesis tested asserts no shift versus the alternative that two samples are drawn from two populations with distributions differing only by the location parameter.

Hodges and Lehmann in [5] compared the Wilcoxon test with the t-test for this case. Chernoff and Savage [3] have proved that the Fisher-Yates test compared to the t-test has Pitman's efficiency exceeding one, with equality sign achieved only if the underlying distribution is normal. In [3] it has also been shown that under mild regularity restrictions the optimal nonparametric procedure as compared to the best parametric procedure (in the sense of the likelihood ratio test) has Pitman's efficiency equal to unity assuming that the underlying distribution is known. Also in [3] the authors implicitly stated the following question: "Suppose we construct two tests for the two sample problem, one parametric and one nonparametric for some fixed distribution believed to occur in investigated populations. How does Pitman's efficiency behave if the true distribution departs from the assumed one?" The present investigation deals with this particular problem.

It turns out that among a class of distributions satisfying some regularity conditions, the normal is the only one possessing the property proved in [3].

This investigation was suggested by Professor E. L. Lehmann to whom I would like to express my gratitude for stating the problem and for all his valuable comments.

2. Assumptions, definitions and notation. Let $X_1 X_2 \cdots X_n \; Y_1 Y_2 \cdots Y_m$ be independent random variables such that $\Pr\{X_k \leqq z\} = K\{z + (1 - \lambda)\Delta\}$ for $k = 1, 2, \cdots, n$ and $\Pr\{Y_j \leqq z\} = K(z - \lambda\Delta)$ for $j = 1, 2, \cdots, m$, with $K(\cdot)$ being a Lebesgue absolutely continuous distribution function, $\lambda = n/(n + m)$ and $\Delta \geqq 0$ an unknown parameter.

For the purpose of constructing test procedures for the hypothesis $H: \Delta = 0$

Received September 7, 1961; July 16, 1962.

[1] This paper was prepared with the partial support of the National Science Foundation, Grant G-14648 and Office of Ordnance Research Grant DA-ARO (D)-31-124-G83. This research forms a part of a dissertation submitted to the University of California, Berkeley, in partial fulfillment of the requirements for the Ph.D. degree.

[2] Now at the University of Maryland.