COMPARISON OF THE NORMAL SCORES AND WILCOXON TESTS

J. L. HODGES, JR. AND E. L. LEHMANN
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction

For testing the equality of two distributions F and G on the basis of samples $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ from these distributions, a number of procedures are available. If the tests are to be powerful against shift alternatives given by

$$G(y) = F(y - \theta), \tag{1.1}$$

the most commonly proposed tests are

(a) Student's t test;

(b) Wilcoxon's two-sample test based on the sum $s_1 + \cdots + s_n$ of ranks of the Y's;

(c) The Normal scores test. This test has been proposed in two asymptotically equivalent versions, the test statistic in both cases being of the form

$$h(s_1) + \cdots + h(s_n) \tag{1.2}$$

with large values significant against the alternatives $\theta > 0$.

(i) The function

$$h(s) = E(W^{(s)}), \tag{1.3}$$

where $W^{(1)} < \cdots < W^{(m+n)}$ are the order statistics of a sample of size $m + n$ from a standard normal distribution was introduced by Fisher and Yates in the introduction to table XX of [3], who also gave a table of (1.3). These authors propose replacing the variables X_i and Y_j in the t-statistic by the function (1.3) of their ranks and applying to these values the usual analysis of variance, which amounts to using as critical value that appropriate to the t-test. The corresponding rank test (in which the critical value is obtained from the distribution of ranks rather than that of t), was proposed by Hoeffding [5] and discussed further by Terry [8], who also gave a table of percentage points.

(ii) The closely related function

$$h(s) = \Phi^{-1}\left(\frac{s}{m + n + 1}\right), \tag{1.4}$$

where Φ denotes the cumulative distribution function of the standard normal

This research was sponsored in part by the Office of Naval Research under Contracts. Nonr-222(43) and 2842(00).