A COMPARISON OF ALTERNATIVE TESTS OF SIGNIFICANCE FOR THE PROBLEM OF m RANKINGS[1]

By Milton Friedman

A paper published in 1937 [2] suggested that the consilience of a number of sets of ranks can be tested by computing a statistic designated χ_r^2. A mathematical proof by S. S. Wilks demonstrated that the distribution of χ_r^2 approaches the ordinary χ^2 distribution as the number of sets of ranks increases. The rapidity with which this limiting distribution is approached was investigated by obtaining the exact distributions of χ_r^2 for a number of special cases. It was concluded that "when the number of sets of ranks is moderately large (say greater than 5 for four or more ranks) the significance of χ_r^2 can be tested by reference to the available χ^2 tables" [2, p. 695]. The use of the normal distribution was recommended when the number of ranks in each set is large, but the number of sets of ranks is small, although no rigorous justification of this procedure was presented.

Except for the few special cases for which exact distributions were given, the paper did not provide a test of significance for data involving less than six sets of ranks and a small or moderate number of ranks in each set. This important gap has now been filled by M. G. Kendall and B. Babington Smith [1]. In addition, they furnish a somewhat more exact test of significance for tables of ranks for which the earlier article recommended the use of the χ^2 distribution.

Kendall and Smith use a different statistic, W, defined as χ_r^2 divided by its maximum value, $m(n - 1)$, where n is the number of items ranked, and m the number of sets of ranks.[2] The new statistic (independently suggested by W. Allen Wallis [3] who terms it the rank correlation ratio and denotes it by η_r^2) is thus not fundamentally different from χ_r^2. A more radical innovation is the improvement in the test of significance that they suggest. Instead of testing χ_r^2 by reference to the χ^2 distribution for $n - 1$ degrees of freedom, Kendall and Smith, generalizing from the first four moments of W, recommend that the significance of W be tested by reference to the analysis of variance distribution (Fisher's z-distribution) with $z = \frac{1}{2}\log_e\left(\frac{(m-1)W}{1-W}\right)$, $n_1 = (n-1) - \frac{2}{m}$, $n_2 = (m-1)\left[(n-1) - \frac{2}{m}\right]$. For small values of m and n, they introduce con-

[1] The author is indebted to Mr. W. Allen Wallis for valuable criticism and to Miss Edna R. Ehrenberg for computational assistance.

[2] This is Kendall and Smith's notation which will be used in the present paper. The original paper [2] designated the number of items ranked by p, and the number of sets of ranks by n.