ON THE DISTRIBUTION OF WILKS' STATISTIC FOR TESTING THE INDEPENDENCE OF SEVERAL GROUPS OF VARIATES

BY A. WALD[1] AND R. J. BROOKNER[1]

Columbia University

1. **Introduction.** We consider p variates $x_1, x_2, \cdots, x_p$ which have a joint normal distribution. Let the variates be divided into k groups; group one containing $x_1, x_2, \cdots, x_{p_1}$, group two containing $x_{p_1+1}, x_{p_1+2}, \cdots, x_{p_2}$, etc. We are interested in testing the hypothesis that the set of all population correlation coefficients between any two variates which belong to different groups is zero

Wilks[2] has derived, by using the Neyman-Pearson likelihood ratio criterion, a statistic based on N independent observations on each variate with which one may test this hypothesis. Let $\| r_{ij} \|$ be the matrix of sample correlation coefficients; Wilks' statistic, λ, is the ratio of the determinant of the p-rowed matrix of sample correlations to the product of the p_1-rowed determinant of correlations of the variates of group one, the $(p_2 - p_1)$-rowed determinant of correlations of the second group, etc. That is

$$\lambda = \frac{| r_{ij} |}{| r_{\alpha_1\beta_1} | \cdot | r_{\alpha_2\beta_2} | \cdots | r_{\alpha_k\beta_k} |}$$

where $| r_{\alpha_i\beta_i} |$ is the principal minor of $| r_{ij} |$ corresponding to the ith group.

In order to use the test, the distribution function of λ must be known. Wilks has shown that in certain cases the exact distribution is a simple elementary function; in other cases it is an elementary function, but one which is rather unwieldy and which does not lend itself readily to practical use. It is our purpose in this paper (1) to show a method by which the exact distribution can be explicitly given as an elementary function for a certain class of groupings of the variates, and (2) to give an expansion of the exact cumulative distribution function in an infinite series which is applicable to any grouping.

2. **The exact distribution of λ.** By the method to be described, the exact distribution of λ can be found when the numbers of variates in the groups are such that there are an odd number in at most one group. If the number of variates is small, say at most eight, the method will increase only slightly the list of distribution functions that Wilks gives in his paper.

[1] Research under a grant-in-aid of the Carnegie Corporation of New York.

[2] S. S. Wilks, "On the independence of k sets of normally distributed statistical variables," *Econometrica*, Vol. 3 (1935), pp. 309-326. Other references to Wilks in this paper except where otherwise noted are to this publication.