ON ANALYSIS OF VARIANCE FOR THE K-SAMPLE PROBLEM[1]

By Dana Quade

University of North Carolina

1. Introduction. Suppose we have $k \geqq 2$ random samples, possibly multivariate, where X_{ij} is the jth observation in the ith sample for $1 \leqq i \leqq k$, $1 \leqq j \leqq n_i$, and there are $N = \sum n_i$ observations in all. Let the distribution function of X_{ij} be G_i. The null hypothesis to be tested is

$$H_0 : G_1 \equiv G_2 \equiv \cdots \equiv G_k .$$

We shall be particularly concerned with the large sample situation where $N \rightarrow \infty$ and $n_i/N \rightarrow p_i > 0$ for $1 \leqq i \leqq k$. Define

$$G = \sum_{i=1}^{k} p_i G_i ;$$

then our null hypothesis can be reexpressed as

$$H_0 : G_i \equiv G, \qquad i = 1, 2, \cdots, k.$$

The general approach we have in mind is an extended version of one-way analysis of variance. Let $f_N(x_1, x_2, \cdots, x_N)$ be a function of N arguments which is symmetric in the last $(N - 1)$ of them. Next, corresponding to each observation X_{ij} define a score

$$y_{ij} = f_N(X_{ij}, [(N - 1)\ X\text{'s} \quad \text{other than} \quad X_{ij}]).$$

Then we perform an ordinary analysis of variance based on the scores: that is, we calculate

$$F = (N - k) \sum n_i(\bar{y}_i - \bar{\bar{y}})^2/(k - 1) \sum \sum (y_{ij} - \bar{y}_i)^2,$$

where $\bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$ and $\bar{\bar{y}} = N^{-1} \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}$, and test H_0 by referring F to the F-distribution with $(k - 1, N - k)$ degrees of freedom.

In the remainder of this paper we present conditions under which such a test will be asymptotically valid, and we show that several of the tests proposed in the literature are essentially of this type.

2. The null-hypothesis distribution. In accordance with Chernoff and Teicher [2], we shall say that n random variables are interchangeable if their joint distribution function is symmetric. Then clearly the scores as defined above are interchangeable random variables if the null hypothesis is true. Now, by a completely straightforward extension of Theorem 1 of Chernoff and Teicher, using the multivariate extension of the Wald-Wolfowitz-Noether limit theorem, we can obtain the following:

Theorem 1. *For every sufficiently large N let $\{Z_{ij}^{(N)}\}$ be a set of N interchangeable*

Received 21 October 1965.

[1] Supported by the National Institutes of Health Grant No. GM-10397.