A k-SAMPLE EXTENSION OF THE ONE-SIDED TWO-SAMPLE SMIRNOV TEST STATISTIC[1]

BY W. J. CONOVER

Kansas State University

1. Introduction. Let $F_{1,n}(x)$ and $F_{2,m}(x)$ be the empirical cdf's (cumulative distribution functions) of the random samples $(X_{1,1}, \cdots, X_{1,n})$ and $(X_{2,1}, \cdots, X_{2,m})$ drawn from populations with continuous cdf's $F_1(x)$ and $F_2(x)$, respectively. The one-sided, two sample Smirnov test statistic $D^+(n, m) = \sup_x (F_{1,n}(x) - F_{2,m}(x))$ provides a test of $H_0: F_1 = F_2$ that is consistent against all alternatives of the type $H_1: F_1(x) > F_2(x)$ for some x. A k-sample extension of $D^+(n, m)$ would be useful in situations where the experimenter has k populations, $k > 2$, and may legitimately assume, from biological or other nonmathematical considerations, that $F_1(x) \geqq F_2(x) \geqq \cdots \geqq F_k(x)$ for all x. Such k sample extensions have been obtained, in various forms, by Ozols (1956), Darling (1957), David (1958), Kiefer (1959), Dwass (1960), Birnbaum and Hall (1960), and Conover (1965), among others. However, each extension, except that of Conover (1965), fails to furnish the small sample distribution function of a test statistic that may be used for all $k \geqq 2$.

This paper furnishes the small sample, and asymptotic, distribution functions of a k-sample extension of $D^+(n, n)$, valid for all $k \geqq 2$, but restricted to equal sample sizes for all k samples. Such a restriction is not surprising since the distribution function in the two-sample case is still not known for all cases of $n \neq m$. Consistency is discussed in Section 5.

2. Principal results. While the interesting result is the corollary, the more general theorem is no more difficult to prove, although the notation may appear cumbersome. For $i = 1, 2, \cdots, k$, let $F_{i,n_i}(x)$ be the empirical cdf of a random sample $(X_{i,1}, \cdots, X_{i,n_i})$ of size n_i drawn from a population with the continuous cdf $F_i(x)$. Let $I_i(x) = n_i F_{i,n_i}(x)$ and let c_i' represent the smallest integer not less than c_i. Let

$$(2.1) \quad P_k^* = P(\sup_x (I_i(x) - I_{i+1}(x)) < c_i ; i = 1, 2, \cdots, k-1).$$

It is assumed hereafter that $H_0: F_1 \equiv F_2 \equiv \cdots \equiv F_k$ is true.

THEOREM.

$$(2.2) \qquad P_k^* = |A^{k \times k}|$$

where $|A^{k\times k}|$ *is the determinant of the* $k \times k$ *matrix* A *whose elements are*

$$a_{ij} = 0 \quad \text{if} \quad n_j - \textstyle\sum_{\alpha=1}^{i-1} c_\alpha' + \sum_{\beta=1}^{j-1} c_\beta' < 0$$
$$= n_j!/(n_j - \textstyle\sum_{\alpha=1}^{i-1} c_\alpha' + \sum_{\beta=1}^{j-1} c_\beta')! \text{ otherwise.}$$

Received 20 February 1967.

[1] Prepared with support from the Kansas Agricultural Experiment Station, Manhattan as Contribution No. 110, Department of Statistics and Statistical Laboratory.