SEVERAL k-SAMPLE KOLMOGOROV-SMIRNOV TESTS[1]

By W. J. Conover

Kansas State University

1. Introduction. Let the nk random variables $\{X_{ij}\}$, $i = 1, \ldots, n$, $j = 1, \cdots, k$, represent k random samples of equal size n with the common absolutely continuous distribution function $F(x)$. Order the samples within themselves, and denote the rth ordered sample by $Z_{1r} < Z_{2r} < \cdots < Z_{nr}$. Then Z_{ir} is the ith order statistic from the rth sample, and Z_{1r} will be referred to as the extreme of the rth sample. Define the empirical distribution function of the rth sample $F_r(z)$ as

$$(1.1)\qquad \begin{aligned} F_r(z) &= 0 && \text{if } z < Z_{1r}, \\ F_r(z) &= m/n && \text{if } Z_{mr} \leqq z < Z_{m+1,r}, \\ F_r(z) &= 1 && \text{if } Z_{nr} \leqq z. \end{aligned}$$

Now order the samples among themselves on the basis of their extremes. That is, let $S = \{Z_{1r}, r = 1, \cdots, k\}$ be the set of extremes from the k samples, and let $Y_{11} < Y_{12} < \cdots < Y_{1k}$ represent the set S after S is ordered. Further, let Y_{ij} represent the ith order statistic from the sample whose extreme is Y_{1j}. In other words, for each point in the sample space where Z_{1r} corresponds to Y_{1j}, the sample $(Z_{1r}, Z_{2r}, \cdots, Z_{nr})$ will be denoted by $(Y_{1j}, Y_{2j}, \cdots, Y_{nj})$. Since $Z_{1r} < Z_{2r} < \cdots < Z_{nr}$, it follows that $Y_{1j} < Y_{2j} < \cdots < Y_{nj}$. The number i is called the rank of Y_{ij} within the sample, and the number j is called the rank of the sample. Define the empirical distribution function of the sample with rank j as

$$(1.2)\qquad \begin{aligned} S_j(y) &= 0 && \text{if } y < Y_{1j}, \\ S_j(y) &= m/n && \text{if } Y_{mj} \leqq y < Y_{m+1,j}, \\ S_j(y) &= 1 && \text{if } Y_{nj} \leqq y. \end{aligned}$$

When $k = 2$, the Kolmogorov-Smirnov (two-sample) test usually involves the use of the test statistics

$$(1.3)\qquad D^+(n, n) = \sup_x [F_1(x) - F_2(x)]$$

and

$$(1.4)\qquad D(n, n) = \sup_x |F_1(x) - F_2(x)|$$

Massey (1951) tabulated the distribution of (1.4), and Birnbaum and Hall

Received 7 September 1964.

[1] Prepared with support from the Kansas Agricultural Experiment Station, Manhattan, as Contribution No. 88, Department of Statistics and Statistical Laboratory.