NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

A CRITERION FOR TESTING THE HYPOTHESIS THAT TWO SAMPLES ARE FROM THE SAME POPULATION

By W. J. Dixon

1. Introduction. The purpose of this paper is to consider a criterion for testing the hypothesis that two samples have been drawn from populations with the same distribution function, assuming only that the cumulative distribution function common to the two populations is continuous. Let the two samples, O_n and O_m, be of size n and m respectively. We may assume $n \leq m$ without loss of generality. Suppose the elements $u_1, \cdots, u_n$ of O_n are arranged in order from the smallest to the largest, that is, $u_1 < u_2 < \cdots < u_n$. These may be represented as points along a line. The elements of 0_m represented as points on the same line are then divided into $(n + 1)$ groups by the first sample, O_n. Let m_1 be the number of points having a value less than u_1, m_i the number lying between u_i and u_{i+1}, $(i = 1, 2, \cdots, n)$ and m_{n+1} the number greater than u_n, $(m_{n+1} = m - m_1 - m_2 - \cdots - m_n)$. The criterion here proposed is[1]

$$C^2 = \sum_{i=1}^{n+1} \left(\frac{1}{n+1} - \frac{m_i}{m} \right)^2. \tag{1}$$

[1] A similar criterion

$$d^2 = \sum_{i=0}^{n} \left(\frac{i}{n} - \frac{\sum_0^i n_j}{n} \right)^2$$

for two samples of the same size was investigated (unpublished) by A. M. Mood. He found the mean and variance to be

$$E(d^2) = \frac{2n+1}{3n}, \qquad \sigma_{d^2}^2 = \frac{8(n-1)(2n+1)}{45n^2}.$$

It can be seen that this is the sum of the squares of the differences between the ordinates of the two cumulative sample distributions calculated at the jumps of the first sample distribution.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org