ARKIV FÖR MATEMATIK Band 3 nr 43

Communicated 5 June 1957 by HARALD CRAMÉR and OTTO FROSTMAN

The significance probability of the Smirnov two-sample test

By J. L. HODGES, Jr.[1]

With 3 figures in the text

1. Introduction

In 1939 N. V. Smirnov proposed the following rank-order test for the two-sample problem. Let $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$ be samples of independent observations from populations with continuous distribution functions F and G, respectively. Form from the samples the empirical distribution functions F_m and G_n; that is, $mF_m(u)$ is the number of the observations $x_1, \ldots, x_m$ which do not exceed u, with $nG_n(u)$ defined analogously. To test the hypothesis $F = G$ we use the statistic $D = \sup_u |F_m(u) - G_n(u)|$, large values of which are significant. We may without loss of generality assume $m \geqq n$.

It is clear that the signficance probability $Pr\{D \geqq d \mid F = G\}$, which we shall denote throughout by P_2, is independent of the common value of $F = G$; that is, the test (like all rank-order tests for the two-sample problem) is similar over the class of all continuous distributions. Further, the fact that $\sup_u |F_m(u) - F(u)|$ tends to 0 in probability as $m \to \infty$ implies that the test is consistent against all alternatives $F \neq G$. These properties of similarity and consistency, together with a certain mathematical elegance, give the test wide appeal to mathematical statisticians. A considerable literature has developed, the proposer of the test has been awarded a Stalin prize (Kolmogorov and Hinčin 1951), and the test has begun to appear in applied handbooks. The test is not very powerful against specific alternatives such as shift (van der Waerden 1953), but this could hardly be expected in view of its consistency.

Smirnov's test was suggested by analogy with the earlier test of Kolmogorov (1933) for the one-sample problem. In fact, Smirnov's test generalizes Kolmogorov's, for when $n \to \infty$ we may replace G_n by G, and D becomes Kolmogorov's statistic for the hypothesis that F equals a completely specified G. Thus general results on the Smirnov test usually give (by the limit passage $m \to \infty$) results on the Kolmogorov test. We shall not however attempt to discuss the significance problem for Kolmogorov's test, nor shall we take up the many variants of Smirnov's test which have been suggested.

Smirnov's test also appears in a one-sided version. We may use $D^+ = \sup_u [F_m(u) - G_n(u)]$ to test the hypothesis that $F(u) \leqq G(u)$ for all u. This form of the test is in

[1] This paper was written while the author was a fellow of the John Simon Guggenheim memorial foundation, and a guest of Stockholms högskola. It is a pleasure to record appreciation for the courtesies extended to me by the högskola and its rector, Professor Harald Cramér.