A BIVARIATE SIGN TEST

BY J. L. HODGES, JR.

University of California, Berkeley

1. Introduction. The sign test has proved to be a very useful means for judging the significance of treatments. Suppose that on each of n individuals (or pairs of individuals) measurements are made under two conditions, for example, before and after treatment (or on a treated and a control subject). Denote the two measurements for the ith individual (or pair of individuals) by x_i and x_i'. We formulate the null hypothesis that x_i and x_i' are identically and independently distributed, but wish to make no assumption concerning relations between the distributions of $x_1, x_2, \cdots, x_n$, nor concerning relations between those of $x_1', x_2', \cdots, x_n'$, save that each set is independent. The alternative to the null hypothesis is that the second measurements x_i' are generally shifted, with respect to the first measurements x_i, in the same direction for all (or most) of the individuals. The test is carried out by counting the number S of the differences $x_i' - x_i$ which have positive signs. Under the null hypothesis, S is binomially distributed with $p = \frac{1}{2}$, assuming there are no cases with $x_i' = x_i$, or that such cases of equality are broken randomly. Under the alternative, S would tend to have large values if the second measurements are generally increased relative to the first, small values if they are decreased. We may then reject for large S, small S, or either, according to the alternative against which we wish the test to have power. The great advantage of the test, aside from its simplicity, is the generality of the conditions under which it is valid.

The present paper proposes a bivariate analog of the two-sided sign test, which can be applied when two quantities are measured on each individual. We now have measurements x_i and y_i in a first circumstance, x_i' and y_i' in a second. Do the $4n$ measurements justify our concluding that the two circumstances differ? The null hypotheses is that the bivariate distribution for (x_i, y_i) is identical with that for (x_i', y_i'), and that these vectors are independent. The alternative of interest is that in the second circumstance the bivariate distribution has been shifted relative to the first, in generally the same direction for all individuals. The direction of this possible shift is, however, unknown.

To illustrate, suppose we measure blood pressure and blood sugar before and after treatment with a new drug on a number of individuals. We wish to know whether the drug influences these quantities, but have no preconceived notion concerning the direction or relative amount of the influence on either quantity, should it exist. The joint distribution of the quantities has an unknown form, and is presumably different in different individuals. The quantities are presumably dependent, but in an unknown way.

If we knew the direction of a possible shift, it would be easy to reduce our problem to the sign test. We could simply project the vectors of differences

Received October 14, 1954.