ON SOME QUESTIONS CONNECTED WITH TWO-SAMPLE TESTS OF SMIRNOV TYPE

I. VINCZE
MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES

1. Introduction

1.1. In the following we shall consider some questions concerning the comparison of two samples. The test around which our investigations will center is the Kolmogorov-Smirnov two-sample test, restricted always to the case of equal sample sizes.

In the first part we shall treat the power function for certain alternatives and make some remarks on the efficiency of the test considered in the case of small samples. In the second part some remarks will be given on distributions and limiting distributions occurring in connection with the treated problems. The investigations given here are closely connected with the author's work presented at the Fourth Berkeley Symposium.

1.2. For diminishing the difference in efficiency between parametric and nonparametric tests, the author has in his papers [9], [11] proposed the use of a pair of statistics instead of one statistic. In consequence of the Neyman-Pearson lemma, this results, for given alternatives, in a better test than the one based on either single test statistic. We apply the maximum deviation of the two empirical distribution functions as the first statistic, which ensures the asymptotic consistency of the test. Then we can add to this for several types of alternatives a suitable corresponding pair, for example, the first maximum index, the number of intersections, the Galton statistic, and so on. In order to examine the increase in the efficiency of the two-sample Smirnov test, we shall treat the situation in the case of a special alternative, for which the computation is relatively easy.

In our treatment we make use of the power functions of the original test and of the two-statistic test as well. The power function can be constructed easily in case of a (continuous) alternative containing piecewise linear parts. With such alternatives we can approximate any given alternative. Following Z. W. Birnbaum [1], these kinds of alternatives (for instance, the maximum and minimum alternatives) were treated by many authors in the one-sample case. As we shall see in section 1, this power function can be easily obtained in the two-sample case for all tests for which the distribution of the test statistic under null hypothesis is known; the idea used is the extension of the method used by