SOME PRINCIPLES OF THE THEORY OF TESTING HYPOTHESES[1]

By E. L. Lehmann

University of California, Berkeley

Introduction:

1. The likelihood ratio principle. The development of a theory of hypothesis testing (as contrasted with the consideration of particular cases), may be said to have begun with the 1928 paper of Neyman and Pearson [16]. For in this paper the fundamental fact is pointed out that in selecting a suitable test one must take into account not only the hypothesis but also the alternatives against which the hypothesis is to be tested, and on this basis the likelihood ratio principle is proposed as a generally applicable criterion. This principle has proved extremely successful; nearly all tests now in use for testing parametric hypotheses are likelihood ratio tests, (for an extension to the non-parametric case see [33]), and many of them have been shown to possess various optimum properties.

At least in the parametric case the likelihood ratio test has a number of desirable properties. Among these we mention:

(i) Frequently it is easy to apply and leads to a definite and reasonable test.

(ii) If the sample size is large, and if certain regularity conditions are satisfied an approximate solution can be given for the distribution problems that arise in the determination of size and power of the test (Wilks [32], Wald [25]). In fact, if the likelihood ratio is denoted by λ, $-2 \log \lambda$ approximately has a central χ^2-distribution under the hypothesis, a non-central χ^2-distribution under the alternatives. The number of degrees of freedom in these distributions equal the number of constraints imposed by the hypothesis.

(iii) As was shown by Wald [25], under certain restrictions the likelihood ratio test possesses various pleasant large sample properties.

In view of this, one may feel that the likelihood ratio principle, although perhaps not always leading to the optimum test, is completely satisfactory, and that a more systematic study of the problem of test selection is not necessary. Unfortunately, against the pleasant properties just mentioned there stands a very unpleasant one. Cases exist, in which the likelihood ratio test is not only unsatisfactory but worse than useless, and hence the likelihood ratio principle is not reliable. Examples of this kind were constructed independently by H. Rubin and C. Stein; the following is Stein's example.

---

[1] Parts of this paper were presented in an invited address at the meeting of the Institute of Mathematical Statistics on Dec. 30, 1948, in Cleveland, Ohio.