EFFECT OF NON-NORMALITY ON STEIN'S TWO SAMPLE TEST

By G. P. Bhattacharjee

Indian Institute of Technology, Kharagpur

1. Introduction. Student's t test is used to test the hypothesis about the mean of a normal population when the variance is not known; the power of this test being dependent on the unknown variance. It is shown by Dantzig (1940) that for a sample of fixed size there does not exist a test for Student's hypothesis whose power is independent of the variance. Stein (1945) gave a two sample test for a linear hypothesis whose power is independent of the unknown variance. He used it (i) to test the hypothesis about the mean of a normal population and (ii) to estimate the mean by a confidence interval of predetermined length with a given confidence coefficient. As in other tests of significance the basic assumptions in Stein's test is the normality of the parent population. This assumption of normality may not hold good in practice, and hence the validity of normal theory Stein's test for non-normal populations should be examined.

The effect of non-normality on Student's test has been investigated, among others, by Pearson and Adyanthaya (1929), Bartlett (1935), Geary (1936), Gayen (1949), Ghurye (1949) and Srivastava (1958). Pearson and Adyanthaya (1929) have shown by some experimental investigation that the effect of skewness and kurtosis of the parent population on Student's t may be considerable. Bartlett (1935) confirmed Pearson's results theoretically by obtaining an approximate distribution of t in non-normal samples. Assuming the parent population to be represented by the first two terms of an Edgeworth series, Geary (1936) obtained the approximate distribution of t. Gayen (1949) considered the effect of both skewness (λ_3) and kurtosis (λ_4) by using the first four terms of the Edgeworth series as the frequency function of the population to derive the distribution of t. A theoretical study on the effect of non-normality on the power of the t test was first made by Ghurye (1949) by considering the first two terms of the Edgeworth series and later Srivastava (1958) extended this work by considering the effects of λ_4 and λ_3^2. In a recent paper, Bhattacharjee and Nagendra (1964) have studied the effect of non-normality on the Wald sequential test for mean. This will be of particular interest as Stein's test can be considered as a special case of sequential test.

In this paper, the effect of non-normality on Stein's two sample scheme is investigated by deriving the distribution of Stein's t for non-normal populations represented by the first four terms of an Edgeworth series. The power function of Stein's test and the confidence level of the fixed length confidence interval are also obtained.

2. Stein's two sample scheme. Stein's (1945) two sample procedure for (i) testing the mean of a normal population with unknown variance and (ii) for

Received 24 March 1964; revised 7 July 1964.