SOME EXTENSIONS OF THE WISHART DISTRIBUTION[1]

T. W. ANDERSON AND M. A. GIRSHICK

Princeton University and Columbia University

1. Introduction. The well-known Wishart distribution is the distribution of the variances and covariances of a sample drawn from a multivariate normal population assuming that the expected value of each variate remains the same from observation to observation. For problems such as testing collinearity [1], comparing scales of measurement [2], and multiple regression in times series analysis [3], it is desirable to have the distribution of sample variances and covariances for observations, the expected values of which are not all identical. Such a distribution could be considered as a generalization to several variates of the χ'^2 (non-central χ^2) distribution, as well as a generalization of the Wishart distribution to the non-central case. In this paper we shall discuss the general problem of finding the distribution in question and shall derive this distribution for two particular cases. We shall start out with the problem in its most general form and as a result of linear transformations express the distribution as a certain multiple integral.

We can think of the expected values of the observations as defining points in a space of dimension equal to the number of variates. If these points lie on a line, the non-central Wishart distribution is essentially the Wishart distribution multiplied by a Bessel function; if the points lie in a plane, it is a Wishart distribution multiplied by an infinite series of Bessel functions. For higher dimensionality the integration of the multiple integral becomes extremely troublesome; it has not been possible yet to express the general integration in a concise form. These results are summarized precisely at the end of the paper.

2. Reduction to canonical form. Consider a set of N multivariate normal populations each of p variates. Let the ith $(i = 1, 2, \cdots, p)$ variate of the αth $(\alpha = 1, 2, \cdots, N)$ population be $x_{i\alpha}$; let the mean of this variate be

$$E(x_{i\alpha}) = \mu_{i\alpha} \quad (i = 1, 2, \cdots, p; \alpha = 1, 2, \cdots, N); \tag{1}$$

and let the variance-covariance matrix (of rank p) common to all N distributions be

$$\| E[(x_{i\alpha} - \mu_{i\alpha})(x_{j\alpha} - \mu_{j\alpha})] \| = \| \sigma_{ij} \| \quad (\alpha = 1, 2, \cdots, N).$$

Now consider a sample of observations $\{x_{i\alpha}\}$ one from each population.

The purpose of this paper is to find the joint distribution of the quantities

$$a_{ij} = \sum_{\alpha=1}^{N} (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j), \tag{2}$$

[1] The results given below were arrived at independently by the two authors. Some preliminary results were given before the Institute of Mathematical Statistics at Washington, D. C., May 6, 1944, by Girshick.