THE NON-CENTRAL WISHART DISTRIBUTION AND CERTAIN PROBLEMS OF MULTIVARIATE STATISTICS[1]

By T. W. Anderson

Cowles Commission for Research in Economics

1. Summary. The non-central Wishart distribution is the joint distribution of the sums of squares and cross-products of the deviations from the sample means when the observations arise from a set of normal multivariate populations with constant covariance matrix but expected values that vary from observation to observation. The characteristic function for this distribution is obtained from the distribution of the observations (Theorem 1). By using the characteristic functions it is shown that the convolution of several non-central Wishart distributions is another non-central Wishart distribution (Theorem 2). A simple integral representation of the distribution in the general case is given (Theorem 3). The integrand is a function of the roots of a determinantal equation involving the matrix of sums of squares and cross-products of deviations of observations and the matrix of sums of squares and cross-products of deviations of corresponding expected values.

The knowledge of the non-central Wishart distribution is applied to two general problems of multivariate normal statistics. The moments of the generalized variance, which is the determinant of sums of squares and cross-products multiplied by a constant, are given for the cases of the expected values of the variates lying on a line (Theorem 4) and lying on a plane (Theorem 5). The likelihood ratio criterion for testing linear hypotheses can be expressed as the ratio of two determinants or as a symmetric function of the roots of a determinantal equation. In either case there is involved a matrix having a Wishart distribution and another matrix independently distributed such that the sum of these two matrices has a non-central Wishart distribution. When the null hypothesis is not true the moments of this criterion are given in the non-central planar case (Theorem 6).

2. Introduction. The well-known Wishart distribution is the distribution of the sums of squares and cross-products of deviations from the sample means of observations from a multivariate normal distribution. If the expected values of the variates change from observation to observation (with the covariance matrix constant), the distribution of sums of squares and cross-products is the *non-central* Wishart distribution. This distribution has been given explicitly [1] for the simple cases of the non-central problem. If we think of the expected values of each observation as defining a point in a space of dimensionality equal to the number of variates, we can say that the cases handled are those in which the points corresponding to a sample lie on a line or

[1] Part of a thesis submitted to the Mathematics Department of Princeton University in partial fulfillment of the requirements for the degree of Doctor of Philosophy, June, 1945.

a plane. Although the explicit formulas for the distribution of higher rank are extremely complicated and have not been derived, the characteristic function is relatively simple. The distribution in general can be given in terms of a simple multiple integral.

The Wishart distribution is the basis of much of the sampling theory associated with the multivariate normal distribution. It plays a role similar to that of the χ^2-distribution in univariate normal theory. It can be used in deriving the distributions of the generalized T^2 and of the multiple correlation coefficient when all variates have a normal distribution; it is used in deriving the moments of the likelihood ratio criterion for testing the general linear hypothesis (including the test of the means of several populations being equal) as well as deriving the moments of other such criteria[2]. For the problems of the T^2 and the test of the linear hypothesis and many other problems, the non-central Wishart distribution must be substituted for the central Wishart distribution when the null hypothesis is not true. That is, the non-central distribution can be the basis of obtaining the power function for many tests in multivariate normal statistics. As an example of the application of the non-central Wishart distribution to these problems, in this paper we obtain the moments of the generalized variance and the moments of the criterion for linear hypotheses when the population means lie on a line or a plane. Applications to other problems such as testing collinearity, comparing scales of measurement, and multiple regression in time series analysis will be published in a later paper [3]. Another problem to which this non-central theory can be applied is a method of estimating the parameters of a single equation of a complete system of linear stochastic difference equations (developed by T. W. Anderson, M. A. Girshick and H. Rubin).

In [1] it was shown that one can make linear transformations on the observations which simplify the derivation of the non-central Wishart distribution in the linear and planar cases. Consider a set of N multivariate normal populations, each of p variates. Let the i-th ($i = 1, 2, \cdots, p$) variate of the α-th ($\alpha = 1, 2, \cdots, N$) population be $x_{i\alpha}$; let the mean of the variate be

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

Let the covariance matrix (of rank p) common to all N populations be

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

The probability element of the $x_{i\alpha}$ can be written as

$$| \sigma^{ij} |^{\frac{1}{2}N} (2\pi)^{-\frac{1}{2}pN} \exp [- \tfrac{1}{2} \sum_{i,j,\alpha} \sigma^{ij} (x_{i\alpha} - \mu_{i\alpha})(x_{j\alpha} - \mu_{j\alpha})] \prod_{i,\alpha} dx_{i\alpha}, \tag{2}$$

where

$$\| \sigma^{ij} \| = \| \sigma_{ij} \|^{-1}.$$

[2] See Wilks [2] for example.

The sum of squares and cross-products of deviations from the means in a sample $\{x_{i\alpha}\}$ are

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

where

$$\bar{x}_i = \frac{1}{N}\sum_{\alpha=1}^{N} x_{i\alpha}.$$

The dimensionality, say t, of the space spanned by $\| \mu_{i\alpha} \|$ is equal to the rank of

(4) $$\| \tau_{ij} \| = \| \sum_{\alpha=1}^{N} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j) \|,$$

where

$$\bar{\mu}_{i\alpha} = \frac{1}{N}\sum_{\alpha=1}^{N} \mu_{i\alpha}.$$

As a result of a linear transformation it was demonstrated that the distribution of a_{ij} is the same as that of $\sum_{\alpha=1}^{N-1} x'_{i\alpha} x'_{j\alpha}$ where the $x_{i\alpha}$ have a normal multivariate distribution with covariance matrix $\| \sigma_{ij} \|$ and expected values

$$E(x'_{i\alpha}) = \mu'_{i\alpha}$$
$$(i = 1, 2, \cdots, p; \alpha = 1, 2, \cdots, N - 1),$$

such that

$$\| \tau_{ij} \| = \| \sum_{\alpha=1}^{N-1} \mu'_{j\alpha} \|.$$

The joint distribution of a_{ij} is given for three cases:

(i) *Case* $t = 0$:

(5) $$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N-1, 0) = K_0 \, | \sigma^{ij} |^{\frac{1}{2}(N-1)} \, | a_{ij} |^{\frac{1}{2}(N-p-2)} \exp [-\tfrac{1}{2} \sum_{i,j} \sigma^{ij} a_{ij}];$$

(ii) *Case* $t = 1$:

(6) $$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N-1, 1) = K_1 \exp [-\tfrac{1}{2} \sum_{i,j} \sigma^{ij} \tau_{ij}] \, | \sigma^{ij} |^{\frac{1}{2}(N-1)} \, | a_{ij} |^{\frac{1}{2}(N-p-2)}$$
$$\times \exp [-\tfrac{1}{2} \sum_{i,j} \sigma^{ij} a_{ij}] [\sum_{i,j} a_{ij} \tau_{ij}]^{-\frac{1}{4}(N-3)} \, \mathrm{I}_{\frac{1}{2}(N-3)}(\sqrt{\textstyle\sum_{i,j} a_{ij} \tau_{ij}});$$

(iii) *Case* $t = 2$:

(7) $$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N-1, 2) = K_2 \exp [-\tfrac{1}{2} \sum_{i,j} \sigma^{ij} \tau_{ij}] \, | \sigma^{ij} |^{\frac{1}{2}(N-1)}$$
$$\times | a_{ij} |^{\frac{1}{2}(N-p-2)} \exp [-\tfrac{1}{2} \sum_{i,j} \sigma^{ij} a_{ij}] \sum_{w=0}^{\infty} \frac{(u_1 u_2)^w}{2^{2w} w! \Gamma(\frac{1}{2}[N-2] + w)}$$
$$\times (u_1 + u_2)^{-\frac{1}{2}(\frac{1}{2}[N-3]+2w)} \, \mathrm{I}_{\frac{1}{2}(N-3)+2w} (\sqrt{u_1 + u_2}),$$

where

$$K_0^{-1} = 2^{\frac{1}{2}p(N-1)} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[N-i]),$$

$$K_1^{-1} = 2^{\frac{1}{2}p(N-1)-\frac{1}{2}(N-3)} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p-1} \Gamma(\tfrac{1}{2}[N-1-i]),$$

$$K_2^{-1} = 2^{\frac{1}{2}p(N-1)-\frac{1}{2}(N-3)} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p-2} \Gamma(\tfrac{1}{2}[N-2-i]),$$

$I_n(x)$ is the Bessel function of purely imaginary argument, and u_1 and u_2 are the two non-zero roots of

$$|\mathrm{T} - \lambda A^{-1}| = 0 \tag{8}$$

(here $\mathrm{T} = \|\tau_{ij}\|$ and $A = \|a_{ij}\|$). The number $N-1$ is the *number of degrees of freedom* and t is the *rank*. The matrix $\|\sigma_{ij}\|$ we shall call the *sigma matrix*, and $\|\tau_{ij}\|$ we shall call the *means sigma matrix*.

Let $\kappa_1^2, \kappa_2^2, \cdots, \kappa_p^2$ be the real, non-negative roots of the determinantal equation

$$|\mathrm{T} - \lambda\Sigma| = 0 \tag{9}$$

(where $\Sigma = \|\sigma_{ij}\|$). There is a non-singular $p \times p$ matrix Ψ $(= \|\psi_{ij}\|)$ such that

$$\Psi\Sigma\Psi' = I \tag{10}$$

and

$$\Psi T\Psi' = \|\kappa_i^2\delta_{ij}\| \tag{11}$$

(where I is the identity; Ψ' is the transpose of Ψ and $\delta_{ij} = 1$ for $i = j$ and 0 for $i \neq j$). Then the quantities

$$b_{ij} = \sum_{h,k=1}^{p} \psi_{ih}\psi_{jk}a_{hk} \tag{12}$$

have the distribution $W(b_{ij}, \delta_{ij}, \kappa_i^2\delta_{ij}; p, n, t)$ where $n = N - 1$ and $\kappa_i^2 = 0$ for $i = t+1, t+2, \cdots, p$). This is the same distribution that would be derived if the b_{ij} were defined by

$$b_{ij} = \sum_{\alpha=1}^{n} y_{i\alpha}y_{j\alpha}, \tag{13}$$

where the distribution of the $y_{i\alpha}$ is

$$(2\pi)^{-\frac{1}{2}pn} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p}\sum_{\alpha=1}^{n}(y_{i\alpha} - \kappa_i\delta_{i\alpha})\right]^2. \tag{14}$$

This simplified distribution of the observations has been called the canonical form.

3. The characteristic function of the non-central Wishart distribution. We shall find the characteristic function of the a_{ii} and $2a_{ij}$ $(i \neq j)$ as defined in (3). We first obtain the characteristic function of the b_{ii} and $2b_{ij}$ $(i \neq j)$ as defined in (13) and then perform a linear transformation to obtain the characteristic function of the a_{ij}. The characteristic function of the b_{ii} and $2b_{ij}$ $(i \neq j)$ is defined as

$$E\left(\exp\left[i \sum_{i,j=1}^{p} b_{ij}\theta_{ij}\right]\right), \tag{15}$$

where

$$\theta_{ij} = \theta_{ji}$$

and i in the exponent is the imaginary quantity.

We can write (15) as

$$E\left(\exp\left[i \sum_{i,j=1}^{p} \sum_{\alpha=1}^{n} y_{i\alpha}\, y_{j\alpha}\, \theta_{ij}\right]\right)$$
$$= (2\pi)^{-\frac{1}{2}pn} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left[-\tfrac{1}{2} \sum_{i=1}^{p} \sum_{\alpha=1}^{n} (y_{i\alpha} - \kappa_i\, \delta_{i\alpha})^2 + i \sum_{i,j=1}^{p} \sum_{\alpha=1}^{n} y_{i\alpha}\, y_{j\alpha}\, \theta_{ij}\right]$$
$$\times \prod_{i=1}^{p} \prod_{\alpha=1}^{n} dy_{i\alpha}\,.$$

Let us first integrate the $y_{i\alpha}$ for $i = 1, 2, \cdots, p$ and $\alpha = t + 1, t + 2, \cdots, n$; that is, make the integration

$$(2\pi)^{-\frac{1}{2}p(n-t)} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}$$
$$\cdot \exp\left[-\frac{1}{2} \sum_{i=1}^{p} \sum_{\alpha=t+1}^{n} y_{i\alpha}^2 + i \sum_{i,j=1}^{p} \sum_{\alpha=1}^{n} y_{i\alpha} y_{j\alpha} \theta_{ij}\right] \prod_{i=1}^{p} \prod_{\alpha=t+1}^{n} dy_{i\alpha}.$$

This is, however, the characteristic function of a Wishart distribution with $n - t$ degrees of freedom [4], namely

$$| \delta_{ij} - 2i\theta_{ij} |^{-\frac{1}{2}(n-t)}. \tag{16}$$

Now we must make the integration

$$(2\pi)^{-\frac{1}{2}pt} \exp\left[-\frac{1}{2} \sum_{\eta=1}^{t} \kappa_\eta^2\right] \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}$$
$$\cdot \exp\left[-\frac{1}{2} \sum_{i=1}^{p} \sum_{\eta=1}^{t} y_{i\eta}^2 + i \sum_{i,j=1}^{p} \sum_{\eta=1}^{t} y_{i\eta} y_{j\eta} \theta_{ij} + \sum_{\eta=1}^{t} y_{\eta\eta}\, \kappa_\eta\right] \times \prod_{i=1}^{p} \prod_{\eta=1}^{t} dy_{i\eta}. \tag{17}$$

There is a $p \times p$ matrix $G = \| g_{ij} \|$ such that

$$\sum_{h,k=1}^{p} g_{hi}\, d_{kh}\, g_{kj} = \delta_{ij},$$

where

$$d_{kh} = \delta_{kh} - 2i\theta_{kh}.$$

Let us make the transformation

$$y_{i\eta} = \sum_{h=1}^{p} g_{ih} z_{h\eta} + d^{i\eta} \kappa_\eta,$$

where

$$\| d^{ih} \| = \| d_{ih} \|^{-1}.$$

Then the exponent of (17) within the integral sign is

$$-\frac{1}{2}\left\{\sum_{i=1}^{p} \sum_{\eta=1}^{t} z_{i\eta}^2 - \sum_{\eta=1}^{t} d^{\eta\eta} \kappa_\eta^2\right\},$$

and the Jacobian of the transformation is

$$| d_{i\eta} |^{-\frac{1}{2}t}.$$

Hence, the integral of (17) is

$$(18) \qquad | d^{ij} |^{\frac{1}{2}t} \exp\left[-\frac{1}{2}\left(\sum_{\eta=1}^{t} \kappa_\eta^2 - \sum_{\eta=1}^{t} d^{\eta\eta} \kappa_\eta^2\right)\right].$$

This result is obviously true if the θ_{ij} are pure imaginary and sufficiently small so that $\| d_{ij} \| = \| \delta_{ij} - 2i\theta_{ij} \|$ (which is real in this case) is positive definite. For all complex θ_{ij} in a neighborhood of the origin (17) converges because the real part of $\| d_{ij} \|$ is positive definite. Similarly the integral of the derivative with respect to θ_{ij} of the integrand converges for θ_{ij} in this neighborhood. It follows that the (complex) derivative of the characteristic function exists in this neighborhood because the derivative of the integrand is measurable and is absolutely integrable. Therefore, the characteristic function is analytic in a neighborhood of the origin. From this it follows that the characteristic function is analytic in an open set containing the flat space of real θ_{ij}. By analytic continuation, then, (18) is the value of (17) in the open set containing real θ_{ij}. The characteristic function (15) is the product of (16) and (18). Accordingly, we have the result that the characteristic function of the b_{ii} and $2b_{ij}$ ($i \neq j$) defined by (13) is

$$(19) \qquad | d^{ij} |^{\frac{1}{2}n} \exp\left[-\frac{1}{2}\left(\sum_{\eta=1}^{t} \kappa_\eta^2 - \sum_{\eta=1}^{t} d^{\eta\eta} \kappa_\eta^2\right)\right].$$

It is clear that if $\kappa_\eta = 0$ (for all η), this function reduces to the characteristic function of the Wishart distribution with n degrees of freedom, namely,

$$(20) \qquad | \delta_{ij} - 2i\theta_{ij} |^{-\frac{1}{2}n}.$$

It is interesting to note that (19) factors into two parts, one of which is (20) and the other is

$$(21) \qquad \exp\left[-\frac{1}{2}\left(\sum_{\eta=1}^{t} \kappa_\eta^2 - \sum_{\eta=1}^{t} d^{\eta\eta} \kappa_\eta^2\right)\right].$$

The distribution function similarly factors into two parts, one of which is the Wishart distribution, whose characteristic function is (20). Thus the non-central Wishart distribution function is the convolution of a function (central Wishart distribution) and another (the transform of (21) the first of which is a factor of this same non-central Wishart distribution.

In the planar case the characteristic function can be written as

$$\frac{\exp\left[-\frac{1}{2}(\kappa_1^2 - d^{11}\kappa_1^2)\right]}{|\delta_{ij} - 2i\theta_{ij}|^{\frac{1}{2}n_1}} \cdot \frac{\exp\left[-\frac{1}{2}(\kappa_2^2 - d^{22}\kappa_2^2)\right]}{|\delta_{ij} - 2i\theta_{ij}|^{\frac{1}{2}n_2}},$$

where $n_1 + n_2 = n$. From this fact it is clear that the distribution for the planar case (if $n \geq 2p + 2$) is a convolution of two distributions each of the linear case.

This deduction can also be made from the distribution (14). Let

$$b'_{ij} = y_{i1}y_{j1} + \sum_{\alpha=3}^{n_1+1} y_{i\alpha}y_{j\alpha},$$

$$b''_{ij} = y_{i2}y_{j2} + \sum_{\alpha=n_1+2}^{n} y_{i\alpha}y_{j\alpha}.$$

Then it is clear that the b'_{ij} has the non-central Wishart distribution with n_1 degrees of freedom and parameter κ_1^2 in the direction of the first coordinate axis, while the b''_{ij} has the non-central Wishart distribution with n_2 degrees of freedom and parameter κ_2^2 in the direction of the second coordinate axis. Since

$$b_{ij} = b'_{ij} + b''_{ij},$$

the distribution of the b_{ij} is a convolution of the distributions of b'_{ij} and b''_{ij}. In general the non-central distribution is the convolution of t distributions of the linear case (provided $n \geq tp + t$).

It is easy to show that if one has two (or more) non-central Wishart distributions of rank 1 with parameters in the same direction, the convolution is again a non-central Wishart distribution with parameter in the same direction. Suppose b'_{ij} and b''_{ij} have non-central Wishart distributions with parameter κ'^2_1 and κ''^2_1 in the direction of the first coordinate axes and n_1 and n_2 degrees of freedom respectively. The characteristic functions are

$$|d^{ij}|^{\frac{1}{2}n_1} \exp\left[-\tfrac{1}{2}(\kappa_1'^2 - d^{11}\kappa_1'^2)\right]$$

and

$$|d^{ij}|^{\frac{1}{2}n_2} \exp\left[-\tfrac{1}{2}(\kappa_1''^2 - d^{11}\kappa_1''^2)\right].$$

The product is

$$|d^{ij}|^{\frac{1}{2}n} \exp\left[-\tfrac{1}{2}(\kappa_1^2 - d^{11}\kappa_1^2)\right],$$

where $n = n_1 + n_2$ and $\kappa_1^2 = \kappa'^2_1 + \kappa''^2_1$.

Now let us deduce the characteristic function of the a_{ii} and $2a_{ij}$ $(i \neq j)$. Since by (12) the b's are transforms of the a's we can write $a_{ij} = \sum_{h,k=1}^{p} \psi^{ih}\psi^{jk} b_{hk}$.

Then

$$E\left(\exp\left[i \sum_{i,j=1}^{p} a_{ij}\phi_{ij}\right]\right) = E\left(\exp\left[i \sum_{i,j=1}^{p} \phi_{ij}\psi^{ih}\psi^{jk} b_{hk}\right]\right), \tag{22}$$

where $\phi_{ij} = \phi_{ji}$. If we define

$$\theta_{hk} = \sum_{i,j=1}^{p} \phi_{ij}\psi^{ih}\psi^{jk}, \tag{23}$$

then (22) can be derived by substituting (23) in (19).

Let

$$\Phi = \|\phi_{ij}\|.$$

Then

$$\|d_{ij}\| = D = \Psi'^{-1}(\Sigma^{-1} - 2i\Phi)\Psi^{-1}$$

and

$$D^{-1} = \Psi(\Sigma^{-1} - 2i\Phi)^{-1}\Psi'.$$

The characteristic function of the a's then can be written as

$$\frac{\exp\left[-\frac{1}{2}\{tr(\Psi T\Psi') - tr[\Psi(\Sigma^{-1} - 2i\Phi)^{-1}\Psi'\Psi T\Psi']\}\right]}{\{|\Psi'^{-1}| \cdot |\Sigma^{-1} - 2i\Phi| \cdot |\Psi^{-1}|\}^{\frac{1}{2}n}}$$

using (10) and (11). The denominator is

$$\{|\Psi'| \cdot |\Psi|\}^{-\frac{1}{2}n} |\Sigma^{-1} - 2i\Phi|^{\frac{1}{2}n}$$

and the numerator can be written as[3]

$$\exp\left[-\tfrac{1}{2}\{\mathrm{tr}\,(M'\Psi'\Psi M) - \mathrm{tr}\,[M'\Psi'\Psi(\Sigma^{-1} - 2i\Phi)^{-1}\Psi'\Psi M]\}\right]$$

where

$$M = \|\mu_{i\alpha} - \mu_i\|$$

and

$$M'M = T.$$

We may summarize in the following theorem:

THEOREM 1. *Given* a_{ij} $(i, j = 1, 2, \cdots, p)$ *defined by* (3) *where the* $x_{i\alpha}$ $(i = 1, 2, \cdots, p;\ \alpha = 1, 2, \cdots, N)$ *are distributed according to* (2), *the characteristic function of* a_{ij} *and* $2a_{ij}$ $(i \neq j)$ *is*

$$E\left(\exp\left[i \sum_{i,j=1}^{p} a_{ij}\phi_{ij}\right]\right)$$

$$= \frac{|\sigma^{ij}|^{\frac{1}{2}(N-1)} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{p}\sum_{\alpha=1}^{N}\sigma^{ij}(\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)\right]}{|\tilde{\sigma}^{ij}|^{\frac{1}{2}(n-1)}} \tag{24}$$

$$\cdot \exp\left[\frac{1}{2}\sum_{\substack{i,j=1\\h,k=1}}^{p}\sum_{\alpha=1}^{N}\sigma^{hi}\tilde{\sigma}_{ij}\sigma^{jk}(\mu_{h\alpha} - \bar{\mu}_h)(\mu_{k\alpha} - \bar{\mu}_k)\right],$$

[3] The result follows from the fact that tr(AB) = tr(BA).

where

$$\| \tilde{\sigma}_{ij} \|^{-1} = \| \tilde{\sigma}^{ij} \| = \| \sigma^{ij} - 2i\phi_{ij} \|$$

and

$$\phi_{ij} = \phi_{ji} .$$

Suppose we have two sets of quantities a'_{ij} and a''_{ij} each set of which is distributed according to a non-central Wishart distribution with sigma matrix $\| \sigma^{ij} \|$, one having n' degrees of freedom (or n''), means sigma matrix τ'_{ij} (or τ''_{ij}) of rank t' (or t''). Consideration of the characteristic functions (24) shows that

$$a_{ij} = a'_{ij} + a''_{ij}$$

has a non-central Wishart distribution with matrix $\| \sigma^{ij} \|$, $n' + n''$ degrees of freedom and a matrix

$$\| \tau_{ij} \| = \| \tau'_{ij} \| + \| \tau''_{ij} \|.$$

The rank of the distribution is equal to the rank of $\| \tau_{ij} \|$. This result can also be deduced from the representation of a'_{ij} and a''_{ij} in terms of observations from non-central normal populations. It is a straightforward generalization of the same result for central Wishart distributions.

THEOREM 2. *The convolution of two or more non-central Wishart distributions with identical sigma matrices is a non-central Wishart distribution with means sigma matrix equal to the sum of the means sigma matrices of the components.*

4. An integral representation of the non-central Wishart distribution in the general case. It was shown in [1] that

$$W(b_{ij}, \delta_{ij}, \kappa_i^2 \delta_{ij}; p, n, t) = Ce^{-\frac{1}{2}trB} \int | B - YY' |^{\frac{1}{2}(n-p-t-1)} e^{tr(K'Y)} \, dY$$

where

$$C^{-1} = e^{\frac{1}{2}tr(K'K)} 2^{\frac{1}{2}pn} \pi^{\frac{1}{4}p(p-1)+\frac{1}{2}pt} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n - t + 1 - i]),$$

$$dB = \prod_{i=1}^{p} \prod_{j=i}^{p} db_{ij},$$

$$dY = \prod_{i=1}^{p} \prod_{\eta=1}^{t} dy_{i\eta},$$

$$B = \| b_{ij} \|,$$

$$Y = \| y_{i\eta} \| \qquad (\eta = 1, 2, \cdots, t),$$

$$K = \| \kappa_i \delta_{i\eta} \| \qquad (\eta = 1, 2, \cdots, t),$$

and the integration is on Y over the range $\| B - YY' \|$ positive semi-definite. This is equivalent to

$$(25) \qquad Ce^{-\frac{1}{2}trB} | B |^{\frac{1}{2}(n-p-t-1)} \int | I - Y'B^{-1}Y |^{\frac{1}{2}(n-p-t-1)} e^{tr(K'Y)} \, dY.$$

The integration is over the range of Y for which $|| I - Y'B^{-1}Y ||$ is positive semi-definite.

There is a p by p matrix $H = || h_{ij} ||$ such that

$$H'B^{-1}H = I$$

$$H'K = W = || w_i \delta_{i\eta} ||,$$

where w_i^2 are the roots of

$$| \kappa_i^2 \delta_{ij} - \lambda b^{ij} | = 0,$$

$$|| b^{ij} || = || b_{ij} ||^{-1}.$$

Then make the transformation to $Z = || z_{i\eta} ||$ by

$$Y = HZ.$$

The Jacobian of the transformation is

$$| H |^t = | B |^{\frac{1}{2}t}.$$

Then (25) can be written as

$$(26) \qquad Ce^{-\frac{1}{2}trB} | B |^{\frac{1}{2}(n-p-1)} \int | I - Z'Z |^{\frac{1}{2}(n-p-t-1)} e^{trW('Z)} \, dZ.$$

Partition

$$Z = \begin{Vmatrix} Z^1 \\ Z_1 \end{Vmatrix}$$

such that Z_1 is square ($t \times t$). Let $I - Z_1'Z_1 = E'E$, (in terms of Z_1), where E is specified uniquely and consider the transformation of variables from Z_2 to V defined by

$$Z_2 = VE.$$

Then (26) can be written as

$$Ce^{-\frac{1}{2}trB} | B |^{\frac{1}{2}(n-p-1)} \, dB \int | I - Z_1'Z_1 |^{\frac{1}{2}(n-2t-1)} e^{tr(W_1'Z_1)} \, dZ_1$$

$$\cdot \int | I - V'V |^{\frac{1}{2}(n-p-t-1)} \, dV_1$$

where

$$W_1 = || w_\eta \delta_{\eta\xi} || \qquad (\eta, \xi = 1, 2, \cdots, t).$$

The first integration is over the range $(I - Z_1'Z_1)$ positive semi-definite and the second is over $(I - V'V)$ positive semi-definite. The value of the second integral is

$$\int |I - V'V|^{\frac{1}{2}(n-p-t-1)}\, dV = \frac{\pi^{\frac{1}{2}t(p-t)} \prod_{i=1}^{p-t} \Gamma(\frac{1}{2}[n - 2t + 1 - i])}{\prod_{i=1}^{p-t} \Gamma(\frac{1}{2}[n - t + 1 - i])}.$$

Hence (26) can be written as

$$C_1 e^{-\frac{1}{2}trB} |B|^{\frac{1}{2}(n-p-1)} \int |I - Z_1'Z_1|^{\frac{1}{2}(n-2t-1)} e^{tr(W_1'Z_1)}\, dZ_1 \tag{27}$$

with

$$\begin{aligned} C_1^{-1} &= e^{\frac{1}{2}tr(K'K)} 2^{\frac{1}{2}pn} \pi^{\frac{1}{4}p(p-1)+\frac{1}{2}t^2} \\ &\times \prod_{i=1}^{t} \Gamma(\tfrac{1}{2}[n - t + 1 - i]) \prod_{i=1}^{p-t} \Gamma(\tfrac{1}{2}[n - t + 1 - i]). \end{aligned} \tag{28}$$

The first part of (27) is, except for a constant factor, a central Wishart distribution with n degrees of freedom. The integral of the second part is obviously a symmetric function of the w_i'. In terms of the a_{ij} the w_i^2 are simply the roots of (8). We can sum these results in a theorem.

THEOREM 3. *Given a sample of observations* $\{x_{i\alpha}\}$ $(i = 1, 2, \cdots, p;\ \alpha = 1, 2, \cdots, N)$ *distributed according to* (2), *the probability density function of the sums of squares and cross products of deviations from the sample means defined by* (3) *is*

$$\begin{aligned} W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N - 1, t) &= C_1 |\sigma^{ij}|^{\frac{1}{2}(N-1)} |a_{ij}|^{\frac{1}{2}(N-p-2)} \\ &\cdot \exp\left[-\frac{1}{2}\sum_{i,j} \sigma^{ij} a_{ij}\right] \int \left|\delta_{\eta\xi} - \sum_{i=1}^{t} z_{i\eta} z_{i\xi}\right|^{\frac{1}{2}(N-2t-2)} \\ &\cdot \exp \sum_{i=1}^{t} w_i z_{ii} \prod_{\eta,\xi=1}^{t} dz_{\eta\xi} \end{aligned}$$

integrated over

$$\left\| \delta_{\eta\xi} - \sum_{i=1}^{t} z_{i\eta} z_{i\xi} \right\|$$

positive semi-definite where C_1 $(n = N - 1)$ *and* τ_{ij} *are defined by* (28) *and* (4), *respectively, and where* w_η^2 *are the t non-zero roots of* (8).

5. The moments of the generalized variance in the linear and planar non-central cases.

5.1. *The linear case.* The generalized variance, which is the determinant of the variances and covariances,[4] is a measure of the spread of the observations. If one thinks of the N observations of each variate as a vector in N-space with

[4] This definition of Wilks [5] was made in terms of variances and covariances defined by a_{ij}/N (from equation (3)). Since we consider $a_{ij}/(N-1)$ to be the variances and covariances we define $|a_{ij}/(N-1)|$ as the generalized variance.

origin at the sample mean, the generalized variance is proportional to the square of the volume of the p dimensional parallelotope which is defined by these vectors as principal edges. Another geometric interpretation can be given in terms of the p-dimensional variate space. The generalized variance is proportional to the sum of the squared volumes of all possible parallelotopes that can be joined by choosing as the p principal edges p of the N sample vectors (origin at the sample mean).

In this section we consider the moments of the generalized variance when the distributions of the observations are non-central multivariate normal. In terms of the first geometric representation this means that the center of one or more of the vector distributions is different from the others. For convenience we shall assume that the distribution of the observations $\{y_{i\alpha}\}$ is according to (14). This will give as much generality as if we treated observations $\{x_{i\alpha}\}$ having the distribution (2). Moreover, we shall consider the determinant of sums of squares and cross-products instead of the determinant of variances and covariances. It is clear that the determinantant $|b_{ij}|$, defined by (13), is simply a multiple (by $|\Sigma|(N-1)^p$) of $\left|\frac{a_{ij}}{N-1}\right|$, defined by (3).

Let us first consider the linear case, i.e., $\kappa = \kappa_1 \neq 0$ and $\kappa_i = 0$ $(i = 2, \cdots, p)$ in (14). The first of the p vectors is centered on the first coordinate axis, not at the origin. Then the probability density function of the b_{ij} is

$$\frac{e^{-\frac{1}{2}\kappa^2}|b_{ij}|^{\frac{1}{2}(n-p-1)}\exp\left[-\frac{1}{2}\sum_{i=1}^{p} b_{ii}\right]}{2^{\frac{1}{2}pn}\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p-1}\Gamma(\frac{1}{2}[n-i])}\sum_{\alpha=0}^{\infty}\frac{(\kappa^2 b_{11})^\alpha}{2^{2\alpha}\alpha!\Gamma(\frac{1}{2}n+\alpha)}. \tag{29}$$

We wish to find the moments $E(|b_{ij}|^h)$. Let

$$b_{ij} = s_i s_j r_{ij}.$$

Then s_i^2 is the sum of squares of the i-th variate and $\|r_{ij}\|$ is the matrix of sample correlation coefficients. The Jacobian of this transformation (to s_i^2, r_{ij}) is

$$(s_1^2)^{\frac{1}{2}(p-1)}(s_2^2)^{\frac{1}{2}(p-1)}\cdots(s_p^2)^{\frac{1}{2}(p-1)}.$$

The probability element of the s^2's and r's is

$$\frac{\exp\left[-\frac{1}{2}\kappa^2-\frac{1}{2}\sum_{i=1}^{p}s_i^2\right]\prod_{i=1}^{p}(s_i^2)^{\frac{1}{2}n-1}}{2^{\frac{1}{2}pn}\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p-1}\Gamma(\frac{1}{2}[n-i])}|r_{ij}|^{\frac{1}{2}(n-p-1)} \tag{30}$$
$$\times\sum_{\alpha=0}^{\infty}\frac{(\kappa^2 s_1^2)^\alpha}{2^{2\alpha}\alpha!\Gamma(\frac{1}{2}n+\alpha)}\prod_{i=1}^{p}d(s_i^2)\prod_{i=1}^{p}\prod_{j=i+1}^{p}dr_{ij}.$$

It is clear from (30) that the s_i^2 are distributed independently and that the set r_{ij} have a joint distribution independent of the s_i^2's. Hence

$$E(|b_{ij}|^h) = E(|s_i s_j r_{ij}|^h) = \prod_{i=1}^{p} E[(s_i^2)^h] \cdot E(|r_{ij}|^h).$$

The probability element of s_i^2 $(i = 2, 3, \cdots, p)$ is

$$\frac{e^{-\frac{1}{2}s_i^2}(s_i^2)^{\frac{1}{2}n-1}}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)} d(s_i^2),$$

which is simply the χ^2-distribution. The h-th moment of s_i^2 $(i = 2, 3, \cdots, p)$ is

$$E[(s_i^2)^h] = \frac{2^h\Gamma(\frac{1}{2}n + h)}{\Gamma(\frac{1}{2}n)}.$$

The probability element of s_1^2 is

$$\frac{e^{-\frac{1}{2}\kappa^2}(s_1^2)^{\frac{1}{2}n-1}e^{-\frac{1}{2}s_1^2}}{2^{\frac{1}{2}n}} \sum_{\alpha=0}^{\infty} \frac{(\kappa^2 s_1^2)^\alpha}{2^{2\alpha}\alpha!\Gamma(\frac{1}{2}n + \alpha)} d(s_1^2). \tag{31}$$

This is the χ'^2-distribution (non-central χ^2-distribution) which was given by Fisher [6]. Applying term-by-term integration (the series converges properly) we get the h-th moment

$$E[(s_1^2)^h] = 2^h e^{-\frac{1}{2}\kappa^2} \sum_{\alpha=0}^{\infty} \frac{(\kappa^2)^\alpha\Gamma(\frac{1}{2}n + h + \alpha)}{2^\alpha\alpha!\Gamma(\frac{1}{2}n + \alpha)}.$$

The probability element of the r_{ij} is the well known distribution of correlation coefficients,

$$\frac{\Gamma^{p-1}(\frac{1}{2}n) \, |r_{ij}|^{\frac{1}{2}(n-p-1)}}{\pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i])} \prod_{i=1}^{p} \prod_{j=i+1}^{p} dr_{ij}.$$

Since

$$\int \frac{|r_{ij}|^{\frac{1}{2}(n-p-1)}}{\pi^{\frac{1}{4}p(p-1)}} \prod_{i=1}^{p} \prod_{j=i+1}^{p} dr_{ij} = \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i])}{\Gamma^{p-1}(\frac{1}{2}n)},$$

where the integration is over the entire (permissable) range o the r_{ij}, we have as a consequence the h-th moment of the determinant (since n is arbitrary)

$$E(|r_{ij}|^h) = \frac{\Gamma^{p-1}(\frac{1}{2}n)}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i])} \int \frac{|r_{ij}|^{\frac{1}{2}(n-p-1)+h}}{\pi^{\frac{1}{4}p(p-1)}} \prod_{i=1}^{p} \prod_{j=i+1}^{p} dr_{ij}$$

$$= \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i] + h)\Gamma^{p-1}(\frac{1}{2}n)}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i])\Gamma^{p-1}(\frac{1}{2}n + h)}.$$

Hence, the h-th moment of $| s_i s_j r_{ij} |$ is

$$(32) \qquad 2^{ph} \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n + 2h - i])}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n - i])} e^{-\frac{1}{2}\kappa^2} \sum_{\alpha=0}^{\infty} \frac{\kappa^{2\alpha}\Gamma(\frac{1}{2}n + h + \alpha)}{2^\alpha \alpha! \Gamma(\frac{1}{2}n + \alpha)}.$$

Let us summarize this in a theorem for the a_{ij}.

THEOREM 4. *If the quantities a_{ij} $(i, j = 1, 2, \cdots, p)$ have the distribution $W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N - 1, 1)$ defined by (6), then the moments of $| a_{ij} |$ are given by*

$$E(| a_{ij} |^h) = | \sigma^{ij} |^h 2^{ph} e^{-\frac{1}{2}\kappa^2} \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[N - 1 - i] + h)}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[N - 1 - i])} \sum_{\alpha=0}^{\infty} \frac{\kappa^{2\alpha}\Gamma(\frac{1}{2}[N - 1] + h + \alpha)}{2^\alpha \alpha! \Gamma(\frac{1}{2}[N - 1] + \alpha)},$$

where κ^2 is the non-zero root of (9).

The h-th moment of the generalized variance $| a_{ij}/(N - 1) |$ is obtained by dividing the above expression by $(N - 1)^{ph}$.

If $\kappa^2 = 0$, expression (32) clearly reduces to the moment given by Wilks [5]

$$(33) \qquad 2^{ph} \frac{\prod_{i=1}^{p} \Gamma(\frac{1}{2}[n + 1 - i] + h)}{\prod_{i=1}^{p} \Gamma(\frac{1}{2}[n + 1 - i])}.$$

The expression (32) gives the moments of the generalized variance when the means of the observations are not fixed, but lie on a line. The distribution of $| b_{ij} |$ is not a simple function even in the central case. However, in any particular case one could find the first few moments of $| b_{ij} |$ and fit a distribution function. It is to be noted that the convergence of the series is nearly as rapid as that for $e^{\frac{1}{2}\kappa^2}$.

5.2. *The planar case.* Next we shall treat the planar case for two dimensions. Suppose that $\kappa_i^2 \neq 0$ $(i = 1, 2)$. The probability density function of b_{11}, b_{12}, and b_{22} is

$$(34) \qquad \frac{\exp\left[-\frac{1}{2}(\kappa_1^2 + \kappa_2^2) - \frac{1}{2}\sum_{i=1}^{2} b_{ii} \right]}{2^n \sqrt{\pi}} (b_{11} b_{22} - b_{12}^2)^{\frac{1}{2}(n-3)} \times \sum_{\alpha,\beta=0}^{\infty} \frac{[\kappa_1^2 \kappa_2^2 (b_{11} b_{22} - b_{12}^2)]^\alpha (\kappa_1^2 b_{11} + \kappa_2^2 b_{22})^\beta}{2^{4\alpha+2\beta}\alpha!\beta!\Gamma(\frac{1}{2}[n - 1] + \alpha)\Gamma(\frac{1}{2}n + 2\alpha + \beta)}.$$

Let $b_{11} = s_1^2$, $b_{22} = s_2^2$, and $b_{12} = s_1 s_2 r$. The Jacobian is $s_1 s_2$. The probability element of s_1^2, s_2^2 and r is

$$(35) \qquad \frac{e^{-\frac{1}{2}(\kappa_1^2+\kappa_2^2)}}{2^n \sqrt{\pi}} (s_1^2)^{\frac{1}{2}n-1} (s_2^2)^{\frac{1}{2}n-1} (1 - r^2)^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}(s_1^2+s_2^2)} \times \sum_{\alpha,\beta=0}^{\infty} \frac{(\kappa_1^2 \kappa_2^2 s_1^2 s_2^2)^\alpha (1 - r^2)^\alpha (\kappa_1^2 s_1^2 + \kappa_2^2 s_2^2)^\beta \, d(s_1^2)\, d(s_2^2)\, dr}{2^{4\alpha+2\beta}\alpha!\beta!\Gamma(\frac{1}{2}[n - 1] + \alpha)\Gamma(\frac{1}{2}n + 2\alpha + \beta)}.$$

We wish to find $E\{[s_1^2 s_2^2(1 - r^2)]^h\}$. Let us first multiply (35) by $(1 - r^2)^h$ and integrate from -1 to $+1$. We then obtain

$$2^{-n} e^{-\frac{1}{2}(\kappa_1^2+\kappa_2^2)} (s_1^2)^{\frac{1}{2}n-1} (s_2^2)^{\frac{1}{2}n-1} e^{-\frac{1}{2}(s_1^2+s_2^2)}$$
$$\times \sum_{\alpha,\beta=0}^{\infty} \frac{(\kappa_1^2 \kappa_2^2 s_1^2 s_2^2)^\alpha (\kappa_1^2 s_1^2 + \kappa_2 s_2^2)^\beta \Gamma(\frac{1}{2}[n-1] + h + \alpha)\, d(s_1^2)\, d(s_2^2)}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[n-1] + \alpha) \Gamma(\frac{1}{2}n + 2\alpha + \beta) \Gamma(\frac{1}{2}n + h + \alpha)}.$$

Next we multiply by $(s_1^2)^h (s_2^2)^h$, set $(\kappa_1^2 s_1^2 + \kappa_2^2 s_2^2)^\beta / \beta!$ equal to

$$\sum_{\beta_1+\beta_2=\beta} \frac{(\kappa_1^2 s_1^2)^{\beta_1} (\kappa_2^2 s_2^2)^{\beta_2}}{\beta_1! \beta_2!},$$

and integrate s_1^2 and s_2^2 from 0 to ∞. We obtain

$$E([b_{11}b_{22} - v_{12}^2]^h)$$

$$= 2^{2h} \exp[-\tfrac{1}{2}(\kappa_1^2 + \kappa_2^2)] \tag{36}$$
$$\times \sum_{\alpha,\beta_1,\beta_2=0}^{\infty} \frac{(\kappa_1^2)^{\alpha+\beta_1} (\kappa_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}n + h + \alpha + \beta_1) \Gamma(\frac{1}{2}n + h + \alpha + \beta_2)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n-1] + \alpha) \Gamma(\frac{1}{2}n + 2a + \beta_1 + \beta_2)}$$
$$\times \frac{\Gamma(\frac{1}{2}[n-1] + h + \alpha)}{\Gamma(\frac{1}{2}n + h + \alpha)},$$

which is the expected value we are seeking.

Clearly this reduces to a special case of (32) if κ_2^2 is set equal to zero.

Now we consider the planar case in p dimensions. Geometrically we have p vectors in n-space. If the $\{y_{i\alpha}\}$ are distributed according to (14) the mean point (i.e., center of distribution) of the first two vectors is different from the origin, but the mean point of each of the other $p - 2$ vectors is the origin. The vectors are distributed independently. The determinant

$$|b_{ij}| = \left| \sum_{\alpha=1}^{n} y_{i\alpha} y_{j\alpha} \right|$$

is the square of the volume of the parallelopiped which can be expressed as

$$v_1 v_2 \cdots v_p \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{p-1},$$

where v_i is the length of the i-th vector and θ_i is the angle between the $(i + 1)$-st vector and the flat space determined by the first i vectors. The distribution of $v_3, \cdots, v_p$ and $\theta_2, \cdots, \theta_{p-1}$ is statistically independent of v_1, v_2, and θ_1; for no matter what the plane of the first two vectors is, the conditional distribution of the other variables is the same. Hence

$$E(|b_{ij}|^h) = E[(v_1 v_2 \sin\theta_1)^{2h}] \cdot E[(v_3 v_4 \cdots v_p \sin\theta_2 \cdots \sin\theta_{p-1})^{2h}].$$

If the y's had simply the distribution

$$\frac{1}{(2\pi)^{\frac{1}{2}pn}} \exp\left[-\tfrac{1}{2} \sum_{i=1}^{p} \sum_{\alpha=1}^{n} y_{i\alpha}^2 \right], \tag{37}$$

then the h-th moment of $| b_{ij} |$ would be (33), and the h-th moment of

$$v_1^2 v_2^2 \sin^2 \theta_1 = \begin{vmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{vmatrix}$$

would be

$$2^{2h} \frac{\prod_{i=1}^{2} \Gamma(\frac{1}{2}[n+1-i]+h)}{\prod_{i=1}^{2} \Gamma(\frac{1}{2}[n+1-i])}.$$

Since the distribution of $v_3, v_4, \cdots, v_p$ and $\theta_2, \cdots, \theta_{p-1}$ is the same whether the y's are distributed according to (14) or (37), we have

$$E[(v_3 \cdots v_p \sin \theta_2 \cdots \sin \theta_{p-1})^{2h}] = 2^{h(p-2)} \frac{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[n+2h+1-i])}{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[n+1-i])}. \tag{38}$$

Multiplying (36) by (38) we obtain the h-th moment of $| b_{ij} |$, namely,

$$E(| b_{ij} |^h) = 2^{hp} \exp[-\tfrac{1}{2}(\kappa_1^2 + \kappa_2^2)] \frac{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[n+2h+1-i])}{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[n+1-i])}$$

$$\times \sum_{\alpha,\beta_1,\beta_2=0}^{\infty} \frac{(\kappa_1^2)^{\alpha+\beta_1} (\kappa_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}n - h + \alpha + \beta_1)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n-1]+\alpha)}$$

$$\frac{\Gamma(\frac{1}{2}n + h + \alpha + \beta_2)\Gamma(\frac{1}{2}[n-1] + h + \alpha)}{\Gamma(\frac{1}{2}n + 2\alpha + \beta_1 + \beta_2)\Gamma(\frac{1}{2}n + h + \alpha)}.$$

This result may be summarized as follows:

THEOREM 5. *Let the probability density function of the quantities a_{ij} $(i, j = 1, 2, \cdots, p)$ be*

$$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, N-1, 2)$$

defined by (7). *Then the h-th moment of $| a_{ij} |$ is*

$$E(| a_{ij} |^h) = | \sigma^{ij} |^h 2^{hp} \exp[-\tfrac{1}{2} \kappa_1^2 + \kappa_2^2] \frac{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[N-i]+h)}{\prod_{i=3}^{p} \Gamma(\frac{1}{2}[N-i])}$$

$$(39) \quad \sum_{\alpha,\beta_1,\beta_2=0}^{\infty} \frac{(\kappa_1^2)^{\alpha+\beta_1}(\kappa_2^2)^{\alpha+\beta_2}\Gamma(\frac{1}{2}[N-1]+h+\alpha+\beta_1)\Gamma(\frac{1}{2}[N-1]+h+\alpha+\beta_2)}{2^{2\alpha+\beta_1+\beta_2}\alpha!\,\beta_1!\,\beta_2!\,\Gamma(\frac{1}{2}[N-2]+\alpha)\Gamma(\frac{1}{2}[N-1]+2\alpha+\beta_1+\beta_2)} \times \frac{\Gamma(\frac{1}{2}[N-2]+h+\alpha)}{\Gamma(\frac{1}{2}[N-1]+h+\alpha)},$$

with κ_i^2 *defined by* (9).

The h-th moment of the generalized variance $|a_{ij}/(N-1)|$ is obtained by dividing the above expression by $(N-1)^{ph}$. This formula holds for all $h > -\frac{1}{2}(N-p)$.

6. The moments of the criterion for testing linear hypothesis in the linear and planar non-central cases.

6.1. *The moments of the criterion.* There are several linear hypotheses concerning the means of multivariate normal populations that can be included in a general formulation of the problem. We shall first of all consider a simple case of a linear hypothesis and find the moments of the criterion under linear and planar alternatives. In Section 6.2 we shall indicate some linear hypotheses that can be reduced to this simple case. Regression problems and the problem of equality of means in several populations (studied by Wilks) are included.

Suppose the variates $z_{i\alpha}$ $(i = 1, 2, \cdots, p; \alpha = 1, 2, \cdots, n)$ and $y_{i\gamma}$ $(i = 1, 2, \cdots p; \gamma = 1, 2, \cdots, q)$ have the probability element

$$(40) \quad \frac{|\sigma^{ij}|^{\frac{1}{2}(n+q)}}{(2\pi)^{\frac{1}{2}p(n+q)}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{p}\sum_{\alpha=1}^{n}\sigma^{ij}z_{i\alpha}z_{j\alpha}\right] \exp\left[-\frac{1}{2}\sum_{i,j=1}^{p}\sum_{\gamma=1}^{q}\sigma^{ij}(y_{i\gamma}-\mu_{i\gamma})(y_{j\gamma}-\mu_{j\gamma})\right]\prod_{i=1}^{p}\prod_{\alpha=1}^{n}dz_{i\alpha}\prod_{i=1}^{p}\prod_{\gamma=1}^{q}dy_{i\gamma}.$$

Let us consider the hypothesis H_0 that the means of mp y's are zero, namely,

$$H_0: \mu_{i\gamma} = 0 \quad (i = 1, 2, \cdots, p; \gamma = 1, 2, \cdots, m)$$

Let

$$(41) \quad a_{ij} = \sum_{\gamma=1}^{m} y_{i\gamma}\alpha_{j\gamma},$$

$$(42) \quad b_{ij} = \sum_{\alpha=1}^{n} z_{i\alpha}z_{j\alpha},$$

$$(43) \quad c_{ij} = a_{ij} + b_{ij}.$$

Then the likelihood ratio criterion for testing H_0, called by Hsu [8] the Wilks-Lawley hypothesis, is the $\frac{1}{2}(n+q)$ power of

$$(44) \quad W = \frac{|b_{ij}|}{|c_{ij}|}.$$

Under the null hypothesis the b_{ij} have a Wishart distribution with n degrees of freedom, and the a_{ij} are distributed independently of b_{ij} such that c_{ij} has a Wishart distribution with $n + m$ degrees of freedom. Wilks [5] has given the moments of W and in some special cases the distribution of W.

We shall now obtain the moments of W for distributions specified by (40) where the rank of $\| \mu_{i\gamma} \|$ ($\gamma = 1, 2, \cdots, m$) is 2, i.e., the planar case. Under this assumption the b_{ij} have a Wishart distribution with n degrees of freedom, the a_{ij} are independently distributed in such a way that the c_{ij} have a non-central Wishart distribution with $n + m$ degrees of freedom. Let κ_1^2 and κ_2^2 be the non-zero roots of

$$(45) \qquad \left| \sum_{\gamma=1}^{m} \mu_{ij}\mu_{j\gamma} - \lambda\sigma_{ij} \right| = 0.$$

It is clear that the distribution of W is unchanged if σ^{ij} is set equal to δ_{ij}. Furthermore, we can take $\tau_{ij} = \kappa_i^2\delta_{ij}$ then the c_{ij} are distributed according to $W(c_{ij}, \delta_{ij}, \kappa_j^2\delta_{ij}; p, n, 2)$ with $n + m$ degrees of freedom. The moments will be obtained by a method similar to that used by Wilks [5].

Let the expected value given by (39) be

$$(46) \qquad E(|c_{ij}|^h) = K(n + m, h, p, \kappa_i^2),$$

which is a constant depending on $n + m, h, p, \kappa_1^2$, and κ_2^2. If $D(a_{ij})$ represents the distribution function of the a_{ij}, one can write (46) as

$$(47) \qquad K(n+m, h, p, \kappa_i^2) = \frac{1}{2^{\frac{1}{2}pn}\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p}\Gamma(\frac{1}{2}[n+1-i])}$$
$$\int |c_{ij}|^h |b_{ij}|^{\frac{1}{2}(n-p-1)} \exp[-\tfrac{1}{2}\sum b_{ii}]\, D(a_{ij})] \prod_{i=1}^{p}\prod_{j=i}^{p} db_{ij}\, dA$$

where dA is the volume element of the a_{ij}, and where the integration is over the entire (permissable) ranges of the b_{ij} and a_{ij}. Equation (47) holds since the c's are functions of the b's and a's. Multiplying (47) by

$$(48) \qquad 2^{\frac{1}{2}pn}\prod_{i=1}^{p}\Gamma(\tfrac{1}{2}[n+1-i]),$$

then replacing n by $g + 2$ and dividing by (48) again, we obtain

$$(49) \qquad K(n+m+2g, h, p, \kappa_i^2) \frac{2^{\frac{1}{2}p(n+2g)}\prod_{i=1}^{p}\Gamma(\frac{1}{2}[n+1-i]+g)}{2^{\frac{1}{2}pn}\prod_{i=1}^{p}\Gamma(\frac{1}{2}[n+1-i])}$$
$$= \frac{1}{2^{\frac{1}{2}pn}\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p}\Gamma(\frac{1}{2}[n+1-i])}$$
$$\cdot\int |c_{ij}|^h |b_{ij}|^{\frac{1}{2}(n+2g-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} b_{ii}]\, D(a_{ij})\prod_{i=1}^{p}\prod_{i=j}^{p} db_{ij}\, dA]\right].$$

By definition the right hand side of (49) is the expected value of $|c_{ij}|^h \, |b_{ij}|^g$. Hence

$$E(|c_{ij}|^h \, |b_{ij}|^g) = K(n + m + 2g, h, p, \kappa_i^2) \frac{2^{pg} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i] + g)}{\prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i])}$$

In this expression it is permissible to set h equal to $-g$ (n could have been replaced by $n + 2g$ in (47) to insure the argument of each Γ function being positive). Then we have

$$\begin{aligned} E(W^g) &= E(|c_{ij}|^{-g} \, |b_{ij}|^g) \\ &= K(n + m + 2g, -g, p, \kappa_i^2) \frac{2^{gp} \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i] + g)}{\prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i])}. \end{aligned}$$

Finally, the g-th moment is

$$\begin{aligned} E(W^g) &= \exp[-\tfrac{1}{2}(\kappa_1^2 + \kappa_2^2)] \frac{\prod_{i=3}^{p} \Gamma(\tfrac{1}{2}[n + m + 1 - i]) \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i] + g)}{\prod_{i=3}^{p} \Gamma(\tfrac{1}{2}[n + m + 1 - i] + g) \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i])} \\ &\quad \cdot \sum_{\alpha, \beta_1, \beta_2 = 0}^{\infty} \left[\frac{(\kappa_1^2)^{\alpha + \beta_1} (\kappa_2^2)^{\alpha + \beta_2} \Gamma(\tfrac{1}{2}[n + m] + \alpha + \beta_1) \Gamma(\tfrac{1}{2}[n + m] + \alpha + \beta_2)}{2^{2\alpha + \beta_1 + \beta_2} \alpha! \beta_1! \beta_2! \Gamma(\tfrac{1}{2}[n + m - 1] + g + \alpha) \Gamma(\tfrac{1}{2}[n + m] + \alpha)} \right. \\ &\quad \left. \cdot \frac{\Gamma(\tfrac{1}{2}[n + m - 1] + \alpha)}{\Gamma(\tfrac{1}{2}[n + m] + g + 2\alpha + \beta_1 + \beta_2)} \right]. \end{aligned} \tag{50}$$

We can summarize in the following theorem:

THEOREM 6. *Let* $z_{i\alpha}$ $(i = 1, 2, \cdots, p\colon \alpha = 1, 2, \cdots, n)$ *and* $y_{i\gamma}$ $(i = 1, 2, \cdots, p; \gamma = 1, 2, \cdots, q)$ *have* (40) *as a joint distribution. Define* a_{ij}, b_{ij} *and* c_{ij} *by* (41), (42), *and* (43), *respectively. Let* κ_1^2 *and* κ_2^2 *be the non-zero roots of* (45). *Then the g-th moment of W, defined by* (44), *is* (50). Expression (50) gives the moments of W in the planar case. The linear case is a special case of the planar case, that is, it is the planar case for $\kappa_2^2 = 0$. The g-th moment of W in the linear case is given by

$$\begin{aligned} E(W^g) &= \exp[-\tfrac{1}{2}\kappa_1^2] \frac{\prod_{i=2}^{p} \Gamma(\tfrac{1}{2}[n + m + 1 - i]) \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i] + g)}{\prod_{i=2}^{p} \Gamma(\tfrac{1}{2}[n + m + 1 - i] + g) \prod_{i=1}^{p} \Gamma(\tfrac{1}{2}[n + 1 - i])} \\ &\quad \times \sum_{\beta_1 = 0}^{\infty} \frac{(\kappa_1^2)^{\beta_1} \Gamma(\tfrac{1}{2}[n + m] + \beta_1)}{2^{\beta_1} \beta_1! \Gamma(\tfrac{1}{2}[n + m] + g + \beta_1)}. \end{aligned} \tag{51}$$

For $\kappa_1^2 = 0$, (51) reduces to the expresion given for the moments under the null hypothesis.

Wilks [7] has given the distribution of W under the null hypothesis for several special cases (i.e., certain pairs of n and p). In general, however, the distribution function is too complicated to write down explicitly. When the null hypothesis is not satisfied (i.e., at least one $\kappa_1^2 \neq 0$) the distribution functions are yet more involved. Hence, we shall not write any explicitly.

Hsu [8] has given the asymptotic distribution of W. Suppose that

$$\Psi_n = \sum_{i,j=1}^{p} \sum_{\gamma=1}^{m} \mu_{i\gamma}\mu_{j\gamma}$$

tends to the limit Ψ_0 as n tends to infinity (if the μ's are functions of n). Then the limiting distribution of $x = -(n + q) \log W$ (which equals $-2 \log \Lambda$, where Λ is the likelihood ratio criterion) is

$$2^{-\frac{1}{2}pm} e^{-\frac{1}{2}\Psi_0} x^{\frac{1}{2}pm-1} e^{-\frac{1}{2}x^2} \sum_{\alpha=0}^{\infty} \frac{\Psi_0^\alpha x^\alpha}{2^{2\alpha}\alpha!\,\Gamma(\frac{1}{2}pm + \alpha)}. \tag{52}$$

That is, it is the χ'^2 distribution with pm degrees of freedom and parameter Ψ_0.

For most purposes, alternative hypotheses of the means being on a line (i.e., of rank one) are sufficiently general. In any particular case, one can compute from (51) mumerical values for several moments and then fit an appropriate distribution function. If one wishes to consider alternative hypotheses of rank two, one can use (50) and similarly compute numerical values for moments. The series in either (51) or (50) converge rapidly. To construct an approximate power function for linear alternatives, say, one would fit distribution functions for several values of κ_1^2 and find the desired percentage levels.

There is a matrix $\| d_{ij} \|$ such that

$$\| b_{ij} \| = \| d_{ij} \| \cdot \| d_{ij} \|'$$

and

$$\| a_{ij} \| = \| d_{ij} \| \cdot \| \lambda_j \delta_{ij} \| \cdot \| d_{ij} \|',$$

where the λ's are roots of

$$| a_{ij} - \lambda b_{ij} | = 0 \tag{53}$$

It follows that

$$\| c_{ij} \| = \| d_{ij} \| \cdot \| (1 + \lambda_j)\delta_{ij} \| \cdot \| d_{ij} \|'.$$

Then W can be written as

$$W = \frac{| d_{ij} | \cdot | d_{ij} |'}{| d_{ij} | \cdot | (1 + \lambda_j)\delta_{ij} | \cdot | d_{ij} |'}$$

$$= \frac{1}{\prod_{j=1}^{p} (1 + \lambda_j)}. \tag{54}$$

The distribution of the roots of (53) in the linear case has been given by Roy [9] for a_{ij} o dimensionality p.[5] The distribution in the planar case has been indicated by Anderson [3]. One could obtain the probability of W not exceeding a given value by integrating the λ's over the proper range.

6.2. *Examples of the general linear hypothesis.* A number of hypotheses concerning the expected values of variates with multivariate normal distributions can be put into the form of H_0. The equivalence of the hypotheses is demonstrated by means of linear transformations.

As an example consider the hypothesis H_1 that the means of several normal multivariate populations are equal when the respective covariance matrices are equal. Let $x_{i\alpha}^u$ be the α-th ($a = 1, 2, \cdots, N^u$) observation of the i-th ($i = 1, 2, \cdots, p$) variate in the u-th ($u = 1, 2, \cdots, U$) population. Let

$$E(x_{i\alpha}^u) = \mu_i^u \qquad (i = 1, 2, \cdots, p) \quad (u = 1, 2, \cdots, U), \tag{55}$$

and let the covariance matrix be $\|\sigma_{ij}\|$. Then the hypothesis is

$$H_1 : \mu_i^u = \mu_i \qquad (i = 1, 2, \cdots, p) \quad (u = 1, 2, \cdots, U). \tag{56}$$

For testing this hypothesis let

$$b_{ij} = \sum_{u=1}^{U} \sum_{\alpha=1}^{N^u} (x_{i\alpha}^u - \bar{x}_i^u)(x_{j\alpha}^u - \bar{x}_j^u), \tag{57}$$

$$a_{ij} = \sum_{u=1}^{U} N^u(\bar{x}_i^u - \bar{x}_i)(\bar{x}_j^u - \bar{x}_j), \tag{58}$$

where

$$\bar{x}_i^u = \frac{1}{N^u} \sum_{\alpha=1}^{N^u} x_{i\alpha}^u, \qquad \bar{x}_i = \frac{1}{N} \sum_{u=1}^{u} \sum_{\alpha=1}^{N^u} x_{i\alpha}^u, \qquad N = \sum_{u=1}^{u} N^u. \tag{59}$$

The b_{ij} have $n = N - U$ degrees of freedom and $c_{ij} = a_{ij} + b_{ij}$, have $N - 1$ degrees of freedom. Then the $N/2$ root of the likelihood ratio criterion for H_1 is W defined by (44). For this case equation (45) is

$$\left| \sum_{u=1}^{u} N^u(\mu_i^u - \bar{\mu}_i)(\mu_j^u - \bar{\mu}_j) - \lambda\sigma_{ij} \right| = 0,$$

where

$$\bar{\mu}_i = \frac{1}{N} \sum_{u=1}^{u} N^u \mu_i^u.$$

[5] Roy erroneously claims his distribution to hold for the planar case and higher rank.

Hsu has demonstrated that the general regression problem can be put into the form of H_0. Suppose that $x_{i\alpha}$ ($i = 1, 2, \cdots, p$; $\alpha = 1, 2, \cdots, N$) follow a multivariate normal distribution with covariance matrix $\| \sigma_{ij} \|$, and let the expected value of $x_{i\alpha}$ be

$$E(x_{i\alpha}) = \sum_{r=1}^{q} \beta_{ir} w_{r\alpha} \qquad (q \leq N - p),$$

where the q by N matrix

$$W = \| w_{r\alpha} \|$$

is of rank q. Let H_2 be the hypothesis that

$$H_2 : B_1 = \| \beta_{iu} \| = 0 \qquad (i = 1, 2, \cdots, p; u = 1, 2, \cdots, m \leq q)$$

with the w's known. Let

$$W_1 = \| w_{u\alpha} \| \qquad (u = 1, 2, \cdots, m; \alpha = 1, 2, \cdots, N)$$

$$W_2 = \| w_{r\alpha} \| \qquad (r = m + 1, \cdots, 9; \alpha = 1, 2, \cdots, N),$$

$$X = \| x_{i\alpha} \| \qquad (i = 1, 2, \cdots, p; \alpha = 1, 2, \cdots, N).$$

Let

$$\| b_{ij} \| = XX' - XW'(WW')^{-1}WX$$

$$\| c_{ij} \| = XX' - XW_2'(W_2W_2')^{-1}W_2X'.$$

(with $\| c_{ij} \| = XX'$ if $W_2 = 0$). Then the likelihood ratio criterion for H_2 is the $N/2$-th power of W, defined by (44).

The equation (45) can be written in terms of Z, B_1, and W as

$$(60) \qquad | B_1W_1(I - W_2'(W_2W_2')^{-1}W_2)W_1'B_1' - \lambda\Sigma | = 0$$

for $m < q$. If $m = q$, (45) becomes

$$(61) \qquad | B_1WW'B_1' - \lambda\Sigma | = 0.$$

In (60) and (61) there are no more non-zero roots than the rank of B_1. It is clear that the roots of (60) (or (61)) depend on the matrix W as well as B_1. The distribution of Λ the likelihood ratio criterion under the null hypothesis does not depend on the distribution of the matrix W (if W is not constant). However, the distribution when the null hypothesis is not satisfied does depend on κ_1^2 or on κ_1^2 and κ_2^2, and hence, on the distribution of the elements of W as well as the value of B_1.

The special case of H_0 for $m = q = 1$ gives as the likelihood ratio criterion as a function of Hotelling's generalized T^2. From the moments indicated in (50) we can deduce the distribution of T^2 when the null hypothesis is not true [3]. This result has been obtained by Hsu [10] by another method.

The author is indebted to Professor S. Bochner, Mr. H. Rubin, Professors C. L. Siegel, J. W. Tukey and S. S. Wilks for suggestions concerning this paper.

REFERENCES

[1] T. W. Anderson and M. A. Girshick, "Some extensions of the Wishart distribution, *Annals of Math. Stat.*, Vol. 15 (1944), pp. 345–357.
[2] S. S. Wilks, *Mathematical Statistics*, Princeton University Press, 1943.
[3] T. W. Anderson, *The Non-central Wishart Distribution and Its Application to Problems in Multivariate Statistics*, unpublished thesis, Library, Princeton University, 1945.
[4] J. Wishart and M. S. Bartlett, "The generalized product moment distribution," *Proc. Camb. Phil. Soc.*, Vol. 29 (1933), pp. 260–270.
[5] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika*, Vol. 24 (1932), pp. 471–94.
[6] R. A. Fisher, "The general sampling distribution of the multiple correlation coefficient," *Proc. Roy. Soc. A.*, Vol. 121 (1928), pp. 654–673.
[7] S. S. Wilks, "On the dependence of K sets of normally distributed statistical variables," *Econometrica*, Vol. 3 (1935), pp. 309–326.
[8] P. L. Hsu, "On generalized analysis of variance (I)," *Biometrika* 31 (1940), pp. 221–237.
[9] S. N. Roy, "Analysis of variance for multivariate populations: the sampling distribution of the requisite p-statistics on the null and non-null hypothesis," *Sankhyā*, Vol. 6, (1941), pp. 35–50.
[10] P. L. Hsu, "Notes on Hotelling's generalized T," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 231–243.