

A DISSERTATION
Presented to the
Faculty of Princeton University
**THE NON-CENTRAL WISHART DISTRIBUTION AND
ITS APPLICATION TO PROBLEMS IN
MULTIVARIATE STATISTICS**

by

Theodore W. Anderson, Jr.

A DISSERTATION
Presented to the
Faculty of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

by

Theodore W. Anderson, Jr.

THE NON-CENTRAL WISHART DISTRIBUTION
AND ITS APPLICATION TO PROBLEMS
IN MULTIVARIATE STATISTICS

	<u>Page</u>
I. Introduction.....	1
II. Simplification of the Problem of the Non- central Wishart Distribution.....	6
III. The Linear Case of the Non-central Wishart Distribution.....	18
IV. The Planar Case of the Non-central Wishart Distribution.....	24
V. The General Form of the Linear and Planar Cases.....	37
VI. The Characteristic Function of the Non- central Wishart Distribution.....	42
VII. An Integral Representation of the Non- central Wishart Distribution in the General Case.....	55
VIII. The Moments of the Generalized Variance in the Linear and Planar Non-central Cases.....	59
A. The Linear Case.....	59
B. The Two Dimensional Planar Case...	66
C. The General Planar Case.....	68
IX. The Moments of the Criterion for Testing Linear Hypotheses in the Linear and Planar Non-central Cases.....	72

	<u>Page</u>
A. The Moments of the Criterion.....	72
B. A General Class of Linear Hypotheses.....	83
C. Two Special Cases.....	91
D. Asymptotic Distribution.....	95
X. Likelihood Ratio Criteria for Testing the Dimensionality of Population Means....	97
A. Testing Collinearity.....	97
B. Testing Higher Dimensionality.....	113
C. Fisher's Criteria for Testing Dimensionality.....	137
D. Higher Dimensionality as an Alternative Hypothesis.....	138
E. Maximum Likelihood Estimates of the Lines or the Hyperplanes...	139
F. Asymptotic Distributions.....	141
XI. Distribution of the Roots of Certain Determinantal Equations.....	149
XII. Summary.....	164

THE NON-CENTRAL WISHART DISTRIBUTION AND
ITS APPLICATION TO PROBLEMS IN
MULTIVARIATE STATISTICS*

I. INTRODUCTION

The well-known Wishart distribution (expression (38)) is the joint distribution of the sums of squares and cross-products of deviations in samples drawn from a multivariate normal population under the condition that the expected value of each variate remains the same from observation to observation. The Wishart distribution can be considered as a generalization to several dimensions of the χ^2 distribution. In another sense the χ'^2 distribution, first obtained by R. A. Fisher [1], is a generalization of the χ^2 distribution. The χ'^2 distribution is the distribution of the squared deviations in samples from univariate normal populations,

* The contents of Sections II to IV have been published in the Annals of Mathematical Statistics, Vol. 15 (1944), pp. 345-357. That paper, entitled "Some Extensions of the Wishart Distribution", presents results which were obtained completely independently and simultaneously by M. R. Girshick and the present author. Some preliminary results (of Section II) were presented by Girshick at a meeting of the Institute of Mathematical Statistics at Washington, D. C., May 6, 1944.

the expected value varying from observation to observation but the variance remaining the same. In the present paper we shall consider the simultaneous distribution of the sums of squares and cross-products of deviations of observations from normal multivariate distributions, the expected values of which are not identical from observation to observation while the population variances and covariances are constant. Such a distribution could be considered as a generalization of the Wishart distribution to the non-central case, as well as a generalization of the χ'^2 distribution to several variates. 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 represent the expected values of the observations as 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. In general the non-central Wishart distribution is a Wishart distribution multiplied by a function of the variables which can be expressed as a multiple integral.

Knowledge of the non-central Wishart distribution permits finding certain other non-central distributions. In particular we shall consider the non-central cases of the generalized variance, the criterion for testing linear hypotheses, and the roots of certain determinantal equations.

Wilks [2] suggested the determinant of the sample variances and covariances as a measure of the scattering of observed points in the variate space and named this quantity the "generalized variance". He gave the moments of this quantity for any number of variates and any number of observations under the assumption that the expected value of each variate was constant from observation to observation. For certain specific pairs

of such numbers the distributions were obtained. In the present paper we give the moments for the cases of the expected values of the variates lying on a line or on a plane.

The likelihood ratio criterion for testing whether a set of multivariate samples comes from populations with the same means (assuming the variance-covariance matrices identical) has been derived by Wilks [2] and the moments are given under the null hypothesis. Hsu [3] has shown that a large class of linear hypotheses concerning the means in a set of multivariate populations can be expressed in the same terms as the above problem; hence the criterion and its moments are the same. We derive in this paper the moments of this criterion for the case of the expected values in the populations lying on a line or on a plane. This information is relevant to the power function of the criterion of the general linear hypothesis when the alternatives are means on a line or on a plane. Hsu [4] has given the asymptotic distribution of the criterion.

The obvious generalization of this test is a test of whether the population means do lie on a line or on a hyperplane of certain dimensionality. Fisher [5] on intuitive grounds has suggested a test. We obtain the likelihood ratio criterion for the hypothesis and the maximum likelihood estimates of the line (or hyperplane). In addition criteria are derived for several similar types of hypotheses with various alternative hypotheses.

The criteria are functions of the roots of certain determinantal equations involving variates distributed according to the non-central Wishart distribution. The distribution of the roots has been given in the case of the central Wishart distribution by Fisher [6] and Hsu [7] and in the linear non-central case by Roy [8]. Hsu [9] has also given the limiting distribution in the general non-central case. In the present paper we demonstrate the linear case by a different method (analytic instead of geometric) and indicate the exact distribution in the planar case.

II. SIMPLIFICATION OF THE PROBLEM OF THE NON-CENTRAL WISHART DISTRIBUTION

In this section we will state the problem of the non-central Wishart distribution and then simplify it. Consider a set of N multivariate normal populations each of p variates. Let the i -th ($i=1,2,\dots,p$) variate of the α -th ($\alpha=1,2,\dots,N$) population be $x_{i\alpha}$, the mean of this variate be

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

and let the variance-covariance matrix (of rank p) of each distribution be

$$\| E[(x_{i\alpha} - \mu_{i\alpha})(x_{j\alpha} - \mu_{j\alpha})] \| = \| \sigma_{ij} \|$$

for all α . Now consider a sample of observations $\{x_{i\alpha}\}$ one from each population.* The probability element of

* Unless the meaning is ambiguous, we shall not distinguish in notation between the stochastic variate and the observation.

the $x_{i\alpha}$ can be written as

$$(2) \quad \frac{|\sigma_{ij}|^{\frac{N}{2}}}{(2\pi)^{\frac{NP}{2}}} e^{-\frac{1}{2} \sum_{i,j,\alpha} \sigma_{ij} (x_{i\alpha} - \mu_{i\alpha})(x_{j\alpha} - \mu_{j\alpha})} \prod_{i,\alpha} dx_{i\alpha}$$

where

We wish to find the joint distribution of the quantities

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

where

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

The sample variances and covariances are multiples (by $\frac{1}{N-1}$) of the a_{ij} .

The a_{ij} may be considered as sums of squares and cross-products, for there exists a linear transformation*,

$$x'_{i\alpha} = \sum_{\beta=1}^N \theta_{\alpha\beta} x_{i\beta} \quad (i=1, 2, \dots, p),$$

* See, for example, [10].

where the matrix $|| \theta_{\alpha\beta} ||$ is orthogonal (and $\theta_{N1} = \theta_{N2} = \dots = \theta_{NN} = \frac{1}{\sqrt{N}}$), such that

$$a_{ij} = \sum_{\alpha=1}^n x'_{i\alpha} x'_{j\alpha},$$

where $n = N-1$ and

$$N \bar{x}_i \bar{x}_j = x'_{iN} x'_{jN}.$$

For a given α the $x'_{i\alpha}$ have a multivariate normal distribution with the same variances and covariances as the x 's (because of the orthogonality) and with expected values

$$E(x'_{i\alpha}) = \sum_{\beta=1}^N \theta_{\alpha\beta} \mu_{i\beta} = \mu'_{i\alpha}, \text{ say } (\alpha \neq N).$$

Let

$$\tau_{ij} = \sum_{\alpha=1}^n \mu'_{i\alpha} \mu'_{j\alpha}.$$

Then it is clear that the τ_{ij} are the same functions

of the μ 's that the a_{ij} are of the x 's, namely,

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

where

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

Now consider the two p by p matrices

$$\Sigma = \|\sigma_{ij}\|$$

(henceforth called the "sigma matrix"),

and

$$T = \|\tau_{ij}\|$$

(called the "sigma matrix of means").

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

$$(5) \quad |T - \lambda \Sigma| = 0.$$

There exists a non-singular p by p matrix

$$(6) \quad \Psi = \|\psi_{ij}\|,$$

such that*

$$(7) \quad \Psi \Sigma \Psi' = I$$

and

$$(8) \quad \Psi T \Psi' = \begin{vmatrix} \kappa_1^2 & 0 & \dots & 0 \\ 0 & \kappa_2^2 & \dots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \dots & \kappa_p^2 \end{vmatrix},$$

TW4:
lower

*See, for example, [11].

where I is the identity matrix and $\bar{T}'$ is the transpose of $\bar{T}$. Suppose the rank of T is t ; then t of the roots of (5) are non-zero and $p-t$ are zero. For the sake of convenience we shall choose $\kappa_1^2, \kappa_2^2, \dots, \kappa_t^2$ to be non-zero roots. If T is of rank t , then the means $\mu_{i\alpha}$ lie in a t dimensional sub-space of the original p dimensional space. Let us make the transformation

$$(9) \quad z_{i\alpha} = \sum_{j=1}^p \psi_{ij} x'_{j\alpha}.$$

The $z_{i\alpha}$ are normally and, because of relationship (7), independently distributed with unit variances. The mean value of $z_{i\alpha}$ is

$$E(z_{i\alpha}) = \sum_{j=1}^p \psi_{ij} \mu'_{j\alpha} = v_{i\alpha},$$

say. As a result of (8)

$$(10) \quad \sum_{\alpha=1}^n v_{i\alpha}^2 = \kappa_i^2 \quad (i=1, 2, \dots, p),$$

$$(11) \quad \sum_{\alpha=1}^n v_{i\alpha} v_{j\alpha} = 0 \quad (i \neq j).$$

In fact, $v_{1\alpha}$ is zero for $\alpha = t+1, \dots, p$. Let the new sums of squares and cross-products be

$$(12) \quad b_{ij} = \sum_{\alpha=1}^n z_{i\alpha} z_{j\alpha}.$$

We shall first find the joint distribution of the b_{ij} and then obtain the distribution of the a_{ij} by using the fact that the b_{ij} can be considered simply as a linear transformation of the a_{ij} in a $\frac{p(p+1)}{2}$ dimensional space. For we can write b_{ij} as

$$(13) \quad \begin{aligned} b_{ij} &= \sum_{\alpha=1}^n \sum_{h,k=1}^p \psi_{ih} \psi_{jk} x'_{h\alpha} x'_{k\alpha} \\ &= \sum_{h,k=1}^p \psi_{ih} \psi_{jk} a_{hk}. \end{aligned}$$

The transformation (9) is performed on all variates of each observation. The next transformation, which is the one that results in the standard form of the problem, is performed on all observations of each variate. We wish to construct the n by n matrix of this transformation

$$\Phi = \|\phi_{\alpha\beta}\|$$

in the following manner: Let

$$(14) \quad \phi_{\eta\alpha} = \frac{v_{\eta\alpha}}{k_{\eta}} \quad (\alpha = 1, 2, \dots, n; \eta = 1, 2, \dots, t).$$

In view of (10) and (11)

$$\sum_{\alpha=1}^n \phi_{\xi\alpha} \phi_{\eta\alpha} = \delta_{\xi\eta} \quad (\xi, \eta = 1, 2, \dots, t),$$

where $\delta_{\xi\eta}$ is the Kronecker delta. The remaining elements in Φ are chosen in any way that makes Φ orthogonal.

Now make the transformation

$$y_{ia} = \sum_{\beta=1}^n \phi_{\alpha\beta} z_{i\beta} \quad (i=1,2,\dots,p; \alpha=1,2,\dots,n).$$

Because $\bar{I}$ is orthogonal,

$$(15) \quad b_{ij} = \sum_{\alpha=1}^n y_{i\alpha} y_{j\alpha},$$

and the y 's are independently normally distributed. The expected value of $y_{1\alpha}$ is

$$E(y_{1\alpha}) = \sum_{\beta=1}^n \phi_{\alpha\beta} v_{1\beta}.$$

For $i=t+1, \dots, p$ this is zero because $v_{1\beta}$ is zero for such values of i . For $i \leq t$, we have by

$$\begin{aligned} E(y_{1\alpha}) &= \kappa_i \sum_{\beta=1}^n \phi_{\alpha\beta} \phi_{i\beta} \quad (i=1,2,\dots,t) \\ &= \kappa_i \delta_{\alpha i} // \end{aligned}$$

by the orthogonality of $\bar{I}$. In other words, the expected value of each $y_{1\alpha}$ is zero except for t of the

variates, namely,

$$E(y_{\eta\eta}) = \kappa_{\eta} \quad (\eta = 1, 2, \dots, t).$$

Now the problem can be stated in this form: Find the distribution of b'_{1j} (given by (15)) when the probability density function of the y 's (in the standard form) is

$$(16) \quad \frac{1}{(2\pi)^{\frac{pn}{2}}} e^{-\frac{1}{2} \sum_{i=1}^p \sum_{\alpha=1}^n (y_{i\alpha} - \kappa_i \delta_{i\alpha})^2},$$

where $\kappa_1^1, \kappa_2^2, \dots, \kappa_t^t$ are different from zero and the other κ 's are zero.

The solution of our problem is a multiple integral of $t.p$ variables. Let

$$b'_{ij} = \sum_{\alpha=t+1}^n y_{i\alpha} y_{j\alpha}.$$

Since the b'_{1j} have the Wishart distribution with $n-t$ degrees of freedom (we assume $n \geq t+p$) we can write the joint probability density function of the b'_{1j} and y_1

($i=1,2,\dots,p$; $\eta=1,2,\dots,t$) as

$$\frac{|b'_{ij}|^{\frac{n-p-t-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p b'_{ii}}}{2^{\frac{p(n-t)}{2}} \pi^{\frac{p(p-1)}{2}} \prod_{i=1}^p \Gamma\left(\frac{n-t+1-i}{2}\right)} \times \frac{e^{-\frac{1}{2} \sum_{i=1}^p \sum_{\eta=1}^t (y_{i\eta} - \kappa_{\eta} \delta_{i\eta})^2}}{(2\pi)^{\frac{pt}{2}}}$$

Considering the equations

$$b_{ij} = b'_{ij} + \sum_{\eta=1}^t y_{i\eta} y_{j\eta}$$

as a transformation of the b'_{ij} , we immediately obtain the joint probability element of the b 's and the $y_{i\eta}$ ($i=1,2,\dots,p$; $\eta=1,2,\dots,t$) as

$$(17) \frac{e^{-\frac{1}{2} \sum_{\eta=1}^t \kappa_{\eta}^2} |b_{ij} - \sum_{\eta=1}^t y_{i\eta} y_{j\eta}|^{\frac{n-p-t-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p b_{ii} + \sum_{\eta=1}^t \kappa_{\eta} y_{\eta\eta}}}{2^{\frac{pn}{2}} \pi^{\frac{p(p-1)}{2} + \frac{pt}{2}} \prod_{i=1}^p \Gamma\left(\frac{n-t+1-i}{2}\right) \prod_{i=1}^p \prod_{j \leq i} db_{ij} \prod_{i=1}^p \prod_{\eta=1}^t dy_{i\eta}}$$

To find the distribution of the b_{ij} we must integrate out the $y_{i\eta}$, where the range of integration is such that the matrix

$$\| b_{ij} - \sum_{\eta=1}^t y_{i\eta} y_{j\eta} \|$$

is positive semi-definite. For $t = 1$ or 2 , we can integrate (17) and express the results in a convenient form. However, for higher values of t the integration affords considerable difficulty and has not been done for the general case. However, in §7 we give an integral representation of the general non-central Wishart distribution which shows more clearly the form of the function.

In terms of geometry the case $t = 1$ is the case in which the expected values of the observations lie on a line in the p dimensional space. In the case of $t = 2$, similarly the expected values lie in a plane in this space. Hence, we shall call these two cases the linear and planar cases, respectively. Indeed, t will be called the rank of the distribution.

The κ 's can be given a geometric interpretation. Suppose the matrix Σ is the identity; that is, the original distributions are spherical. If $t = 1$, κ_1^2 is then simply the sum of squared deviations of the means (of the populations) along the line. If $t = 2$, κ_1^2 is the sum of squared deviations along one line in the plane, and κ_2^2 is the sum in the orthogonal direction where these two directions are chosen so that the sum of cross-products vanishes. For general t , the κ^2 's are multiples by $(N-1)$ of the variances in certain orthogonal directions. If the matrix Σ is not the identity, the above remarks hold true in the metric induced by Σ in the p -space.

III. THE LINEAR CASE OF THE NON-CENTRAL WISHART DISTRIBUTION

In the linear case there is one root of the equation (5) which is not equal to zero, that is, there is simply one κ in the distribution (17) and one set of y 's, namely y_{1i} ($i=1,2,\dots,p$). The problem is to integrate the y_{1i} over the range for which the matrix

$$\| b_{ij} - \gamma_i \gamma_j \|$$

is positive semi-definite; the integrand we are interested in is

$$(18) \quad | b_{ij} - \gamma_i \gamma_j |^{\frac{n-p-2}{2}} e^{+\kappa \gamma_i} \prod_{i=1}^p d\gamma_i,$$

where the subscript "1" has been dropped from the κ , and γ_{11} and the part of (14) not involving the y 's has been neglected. The determinant in expression (18) can be expressed as*

$$| b_{ij} - \gamma_i \gamma_j | = | b_{ij} | (1 - \sum_{i,j=1}^p b^{ij} \gamma_i \gamma_j),$$

where

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

* For example, see [12], p. 237.

The inverse exists with probability one because the probability is zero that $||b_{1j}||$ is singular. There is a linear transformation

$$\gamma_i = \sum_{j=1}^p q_{ij} u_j ,$$

such that

$$\sum_{i,j=1}^p b^{ij} \gamma_i \gamma_j = \sum_{j=1}^p u_j^2$$

and

$$\kappa \gamma_i = \lambda u_i ,$$

where λ^2 is the one non-zero root of the equation

$$(19) \quad |\kappa^2 E_{11} - \lambda B^{-1}| = 0 ,$$

where $B^{-1} = ||b^{ij}||$ and E_{11} is the matrix with unity in the upper left hand corner and zeros elsewhere. This fact is a result of the well-known theorem concerning diagonalization of pairs of quadratic forms.* The Jacobian of this transformation is

$$|q_{ij}| = |b_{ij}|^{\frac{1}{2}},$$

and the range of integration is $\sum_{i=1}^p u_i^2 \leq 1$.

The integrand is transformed into

$$|b_{ij}|^{\frac{n-p-1}{2}} \left(1 - \sum_{i=1}^p u_i^2\right)^{\frac{n-p-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p u_i^2} \prod_{i=1}^p du_i.$$

Now let

$$u_1 = \sin w$$

$$u_i = \cos w v_{i-1} \quad (i=2, 3, \dots, p),$$

* The transformation is the so-called "regression transformation". See [13].

The Jacobian of this transformation is $\cos^p w$ and

$$1 - \sum_{i=1}^p u_i^2 = \cos^2 w \left(1 - \sum_{i=1}^{p-1} v_i^2\right).$$

The integration is over the ranges of $-\pi/2 \leq w \leq \pi/2$ and $\sum_{i=1}^{p-1} v_i^2 \leq 1$. We integrate the following expression

$$|b_{ij}|^{\frac{n-p-1}{2}} \left[\left(1 - \sum_{i=1}^{p-1} v_i^2\right)^{\frac{n-p-2}{2}} \prod_{i=1}^{p-1} dv_i \right] \left\{ \cos^{n-2} w e^{i \sin w} dw \right\}.$$

The integral of the quantity within the brackets is simply a Dirichlet integral [14] and its value is

$$\frac{\Gamma\left(\frac{n-p}{2}\right) \pi^{\frac{p-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}.$$

The integral of the expression within the braces is a multiple of a Bessel function of purely imaginary argument

[15, p. 79]; that is,

$$\frac{\Gamma(\frac{n-1}{2}) \sqrt{\pi}}{(1/2)^{\frac{n-2}{2}}} I_{\frac{n-2}{2}}(1).$$

Hence, the integral of (18) is

$$(20) \quad |b_{ij}|^{\frac{n-p-1}{2}} \Gamma(\frac{n-p}{2}) \pi^{\frac{p}{2}} (1/2)^{-\frac{n-2}{2}} I_{\frac{n-2}{2}}(1).$$

Multiplying equation (19) by the matrix $||b_{1j}||$ one can easily show that the non-zero root, λ^2 , is simply $\kappa^2 b_{11}$.

The probability element of the b_{1j} in the linear case then is

$$(21) \quad \frac{e^{-\frac{1}{2}\kappa^2} |b_{ij}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p b_{ii}}}{2^{\frac{pn}{2} - \frac{n-2}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p-1} \Gamma(\frac{n-i}{2})} (\kappa^2 b_{11})^{-\frac{n-2}{4}} I_{\frac{n-2}{2}}(\kappa \sqrt{b_{11}}) \prod_{i=1}^p \prod_{j=i+1}^p db_{ij}$$

In § 5 we shall give the distribution in terms of the original variables, namely, the a_{ij} .

IV. THE PLANAR CASE OF THE NON-CENTRAL WISHART DISTRIBUTION

The case of two non-zero roots of equation (5) can be handled by continuing the process of integration of § 3 another step. The essential problem is the integration of

$$(22) \left| b_{ij} - \sum_{\eta=1}^2 \gamma_{i\eta} \gamma_{j\eta} \right|^{\frac{n-p-3}{2}} e^{K_1 \gamma_{11} + K_2 \gamma_{22}} \prod_{i=1}^p \prod_{\eta=1}^2 \gamma_{i\eta}$$

over the range of y 's for which the matrix

$$\left\| b_{ij} - \sum_{\eta=1}^2 \gamma_{i\eta} \gamma_{j\eta} \right\|$$

is positive semi-definite. The integration is done in two stages, first with respect to the y_{11} , then with

respect to the y_{21} . Let

$$\tilde{b}_{ij} = b_{ij} - \gamma_{2i} \gamma_{2j},$$

$$\tilde{n} = n - 1,$$

$$\tilde{\kappa} = \kappa_1,$$

$$\tilde{\gamma}_i = \gamma_{i1}.$$

Omitting for the time being

$$(23) \quad e^{\kappa_2 \gamma_{22}} \prod_{i=1}^p d\gamma_{i2},$$

we can write the first stage of the integration of (22) as

$$(24) \quad \int |\tilde{b}_{ij} - \tilde{\gamma}_i \tilde{\gamma}_j|^{\frac{\tilde{n}-p-2}{2}} e^{\tilde{\kappa} \tilde{\gamma}_1} \prod_{i=1}^p d\tilde{\gamma}_i$$

over the range $||\tilde{b}_{1j} - \tilde{y}_1 \tilde{y}_j||$ positive semi-definite. However, the only difference between this and the integration of (18), is that we are now writing all variables with " $\sim$ " signs. Hence, the integral of (24) is (20) will all variables written with " $\sim$ " signs. Changing back again to our other variables of § 4 and inserting again (23), we can write the first stage of the integration of (22) as

$$(25) \quad \Gamma\left(\frac{n-p-1}{2}\right) \pi^{\frac{p}{2}} |b_{1j} - y_{12} y_{j2}|^{\frac{n-p-2}{2}} e^{K_2 y_{12}} \\ \cdot \left[\frac{K_1^2 (b_{11} - y_{12}^2)}{4} \right]^{-\frac{n-3}{4}} I_{\frac{n-3}{2}} \left(\sqrt{K_1^2 b_{11} - K_1^2 y_{12}^2} \right) \prod_{i=1}^p dy_{i2}.$$

Now we must integrate (25) with respect to the y_{12} over the range $||b_{1j} - y_{12} y_{j2}||$ positive semi-definite. The

determinant in (25) can be written as

$$|b_{ij}| \left(1 - \sum_{i,j=1}^p b^{ij} \gamma_{i2} \gamma_{j2}\right).$$

There is a transformation*

$$\gamma_{i2} = \sum_{j=1}^p g_{ij} s_j,$$

(26)

such that

$$\sum_{i,j=1}^p b^{ij} \gamma_{i2} \gamma_{j2} = \sum_{j=1}^p s_j^2,$$

(27)

$$K_1^2 \gamma_{12}^2 = f^2 s_1^2,$$

$$K_2 \gamma_{22} = d_1 s_1 + d_2 s_2,$$

* Again the "regression transformation" [13] is used.

where f^2 is the one non-zero root of the equation

$$(28) \quad |\kappa_1^2 E_{11} - \lambda B^{-1}| = 0,$$

where B^{-1} and E_{11} are used as in equation (19). Since this equation is similar to (17) we know that $f^2 = \kappa_1^2 b_{11}$. The values of d_1 and d_2 will be considered later. This result is deduced from an extension of the theorem concerning the diagonalization of pairs of quadratic forms which can be demonstrated as follows.

The theorem on diagonalization of quadratic forms states that there is a transformation

$$y_{i2} = \sum_{j=1}^p h_{ij} t_j,$$

such that

$$\sum_{i,j=1}^p b^{ij} y_{i2} y_{j2} = \sum_{j=1}^p t_j^2,$$

$$\kappa_1^2 y_{12}^2 = f^2 t_1^2,$$

where f^2 is the non-zero root of (28). Then we have

$$k_2 y_{22} = k_2 \sum_{j=1}^p h_{2j} t_j.$$

By the same theorem we have a transformation of the last $(p-1)$ variates

$$t_j = \sum_{i=2}^p f_{ji} u_i, \quad j = 2, \dots, p,$$

such that

$$\sum_{j=2}^p t_j^2 = \sum_{i=2}^p u_i^2$$

and

$$k_2 \sum_{j=2}^p h_{2j} t_j = c u_2.$$

Hence, the transformation (26) can be written as

$$y_{i2} = h_{i1} t_1 + \sum_{j,k=2}^p h_{ij} f_{jk} u_k,$$

where $t_1 = s_1$ and $u_k = s_k$ ($k=2,3,\dots,p$). Then $d_1 = k_2 h_{21}$ and $d_2 = c$. The Jacobian of the transformation (26) is

$$|g_{ij}| = |b_{ij}|^{\frac{1}{2}}$$

and the range of integration is $\sum_{i=1}^p s_i^2 \leq 1$. The integrand (25) is now changed to

$$\Gamma\left(\frac{n-p-1}{2}\right) \pi^{\frac{p}{2}} |b_{ij}|^{\frac{n-p-1}{2}} \left(1 - \sum_{i=1}^p s_i^2\right)^{\frac{n-p-2}{2}} e^{d_1 s_1 + d_2 s_2} \\ \cdot \left[\frac{f^2(1-s_i^2)}{4}\right]^{-\frac{n-3}{4}} I_{\frac{n-3}{2}}\left(\sqrt{f^2(1-s_i^2)}\right) \prod_{i=1}^p ds_i.$$

Next the following transformation is made:

$$s_1 = \sin w_1$$

$$s_2 = \cos w_1 \sin w_2$$

$$s_i = \cos w_1 \cos w_2 \dots \cos w_{i-2} \quad (i=3,4,\dots,p).$$

The Jacobian is $\cos^p w_1 \cos^{p-1} w_2$, and

$$1 - \sum_{i=1}^p s_i^2 = \cos^2 w_1 \cos^2 w_2 \left(1 - \sum_{i=1}^{p-2} v_i^2\right).$$

We now integrate

$$(29) \quad \Gamma\left(\frac{n-p-1}{2}\right) \pi^{\frac{p}{2}} |b_{ij}|^{\frac{n-p-1}{2}} \left[\left(1 - \sum_{i=1}^{p-2} v_i^2\right)^{\frac{n-p-2}{2}} \prod_{i=1}^{p-2} dv_i \right] \\ \cdot \left\{ \cos^{n-2} w_1 \cos^{n-3} w_2 e^{d_1 \sin w_1 + d_2 \cos w_1 \sin w_2} \cdot \left(\frac{f}{2} \cos w_1\right)^{-\frac{n-3}{2}} I_{\frac{n-3}{2}}(f \cos w_1) dw_1 dw_2 \right\}.$$

The integral of the expression within the brackets is another Dirichlet integral; its value is

$$\frac{\Gamma\left(\frac{n-p}{2}\right) \pi^{\frac{p-2}{2}}}{\Gamma\left(\frac{n-2}{2}\right)}.$$

Similar to § 3 the integration of the quantity within the braces with respect to w_2 is

(30)

The only function of w_1 in (30) that is not even is $\sin w_1$. Hence, the integral of (30) over half the range of w_1 , namely, is the same as the integral of (30) over if is replaced by . Therefore, the integral of (30) can be transformed into an integration over the range by replacing by

Since [15, p. 54]

$$\sqrt{2\pi x} I_{-\frac{1}{2}}(x) = \sinh x,$$

where $I_{\nu}(x)$ is the Bessel function of purely imaginary argument, we can write the integral of the above equation

as

$$\Gamma\left(\frac{n-2}{2}\right) \sqrt{\pi} \left(\frac{d_2}{2}\right)^{-\frac{n-3}{2}} \left(\frac{f}{2}\right)^{-\frac{n-3}{2}} \sqrt{d_1} \sqrt{2\pi}$$

$$\int_0^{\frac{\pi}{2}} I_{-\frac{1}{2}}(d_1 \sin w_1) I_{\frac{n-3}{2}}(d_2 \cos w_1) I_{\frac{n-3}{2}}(f \cos w_1) \sin^{\frac{1}{2}} w_1 \cos w_1 dw_1.$$

By using a Sonine finite integral for Bessel functions [15, p. 377] we can write this integral as

$$(31) \sqrt{\pi} 2^{\frac{n-2}{2}} \int_0^{\pi} \frac{I_{\frac{n-2}{2}}(\sqrt{d_1^2 + d_2^2 + f^2 - 2d_2 f \cos u})}{(d_1^2 + d_2^2 + f^2 - 2d_2 f \cos u)^{\frac{n-2}{4}}} \sin^{\frac{n-3}{2}} u du.$$

Letting $d_1^2 + d_2^2 + f^2 = x$ and $d_2 f = y$ and using an expansion formula for Bessel functions [15, p.140] we can write (31) as

$$\sqrt{\pi} 2^{\frac{n-2}{2}} \int_0^{\pi} \sum_{\gamma=0}^{\infty} \frac{(y^2)^{\gamma} \cos^{\gamma} u}{2^{2\gamma} \gamma! \Gamma(\frac{n-1}{2} + \gamma)} x^{-\frac{1}{2}(\frac{n-2}{2} + \gamma)} I_{\frac{n-2}{2} + \gamma}(\sqrt{x}) \sin^{n-3} u \, du.$$

Since the integral of $\cos^{\gamma} u \sin^{n-3} u$ where γ is odd is zero, the result of the integration (using the duplication formula" for Γ functions and letting $\gamma = 2\omega$) is

$$\Gamma(\frac{n-2}{2}) \pi 2^{\frac{n-2}{2}} \sum_{\omega=0}^{\infty} \frac{(y^2)^{\omega} x^{-\frac{1}{2}(\frac{n-2}{2} + 2\omega)}}{2^{2\omega} \omega! \Gamma(\frac{n-1}{2} + \omega)} I_{\frac{n-2}{2} + 2\omega}(\sqrt{x}).$$

From the relationship (27) it is clear that the equation

$$(32) \quad |\kappa_i^2 \delta_{ij} - \lambda b^{ij}| = 0 \quad (\kappa_i^2 = 0 \text{ for } i=3,4,\dots,p)$$

is transformed by (26) into

$$(33) \quad \begin{vmatrix} f^2 + d_1^2 - \lambda & d_1 d_2 & 0 & \dots & 0 \\ d_1 d_2 & d_2^2 - \lambda & 0 & \dots & 0 \\ 0 & 0 & -\lambda & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & -\lambda \end{vmatrix} = 0.$$

Equation (32) is also equivalent to

$$(34) \quad |\kappa_i \kappa_j b_{ij} - \lambda \delta_{ij}| = 0.$$

Hence, x and y^2 , which are the sum and product, respectively, of the non-zero roots of the two equivalent equations (33) and (34), are given by

$$\begin{aligned} x^2 &= f^2 + d_1^2 + d_2^2 = \kappa_1^2 b_{11} + \kappa_2^2 b_{22}, \\ y^2 &= f^2 d_2^2 = \kappa_1^2 \kappa_2^2 (b_{11} b_{22} - b_{12}^2). \end{aligned}$$

In view of this result we can now write the integration of (29) as

$$\Gamma\left(\frac{n-p-1}{2}\right) \Gamma\left(\frac{n-p}{2}\right) \pi^p 2^{\frac{n-2}{2}} |b_{ij}|^{\frac{n-p-1}{2}}$$

$$\sum_{\omega=0}^{\infty} \frac{[K_1^2 K_2^2 (b_{11} b_{22} - b_{12}^2)]^{\omega}}{2^{2\omega} \omega! \Gamma\left(\frac{n-2}{2} + \omega\right)} (K_1^2 b_{11} + K_2^2 b_{22})^{-\frac{1}{2}(\frac{n-2}{2} + 2\omega)}$$

$$\cdot I_{\frac{n-2}{2} + 2\omega} \left(\sqrt{K_1^2 b_{11} + K_2^2 b_{22}} \right).$$

Finally, by multiplying in what was left out of (22) we obtain the integral of (17), which is the solution to the problem as stated in the standard form:

(35)

$$\begin{aligned}
(35) \quad & \frac{e^{-\frac{1}{2}(k_1^2 + k_2^2)} |b_{ij}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p b_{ii}}}{2^{\frac{pn}{2} - \frac{n-2}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p-2} \Gamma\left(\frac{n-1-i}{2}\right)} \\
& \cdot \sum_{\omega=0}^{\infty} \frac{[k_1^2 k_2^2 (b_{11} b_{22} - b_{12}^2)]^{\omega}}{2^{2\omega} \omega! \Gamma\left(\frac{n-1}{2} + \omega\right)} (k_1^2 b_{11} + k_2^2 b_{22})^{-\frac{1}{2}\left(\frac{n-2}{2} + 2\omega\right)} \\
& \cdot I_{\frac{n-2}{2} + 2\omega} \left(\sqrt{k_1^2 b_{11} + k_2^2 b_{22}} \right),
\end{aligned}$$

V. THE GENERAL FORM OF THE LINEAR AND PLANAR CASES

To answer the problem as stated originally it is necessary to make the transformation (13) and obtain the distribution in terms of the a_{ij} for the linear and planar cases.

It is clear that equation (32) is equivalent to

$$(36) \quad |T - \lambda A^{-1}| = 0,$$

where $A^{-1} = ||a_{ij}||^{-1}$, for T is the transform (by (6)) of $||k_i^2 \delta_{ij}||$ and A^{-1} is the transform of $||b^{ij}||$. The sum and product of the non-zero roots, which are the arguments of the infinite series in (31), remain unchanged.

Since the quantity $\sum_{i=1}^p \kappa_i^2$ is the sum of the roots of (5) it can be expressed as

$$\sum_{i=1}^p \kappa_i^2 = \sum_{i,j=1}^p \sum_{\alpha=1}^N \sigma_{ij}^{\alpha} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)$$

where $||\sigma_{ij}^{\alpha}|| = ||\sigma_{ij}^{\alpha}||^{-1}$ since (5) is equivalent to

$$|\Sigma T - \lambda I| = 0$$

and

$$\text{tr}(\Sigma T) = \text{tr}(\sum ||\mu_{i\alpha} - \bar{\mu}_i|| \cdot ||\mu_{j\alpha} - \bar{\mu}_j||) = \text{tr}(\sum ||\mu_{i\alpha} - \bar{\mu}_i|| \sum ||\mu_{j\alpha} - \bar{\mu}_j||).$$

Furthermore, we have

$$|b_{ij}| = |\Sigma| |a_{ij}|,$$

$$\sum_{i=1}^p b_{ii} = \text{tr} ||b_{ij}|| = \text{tr}(\Psi ||a_{ij}|| \Psi') = \sum_{i,j=1}^p \sigma_{ij}^{\alpha} a_{ij}.$$

Moreover, the absolute value of the Jacobian of the transformation* (12) is

$$(37) \quad |J| = |\Psi|^{p+1} = |\Sigma|^{-\frac{1}{2}(p+1)}.$$

* One method of demonstrating this fact is to apply (9) to centrally distributed variates and compare the Wishart distribution of the transformed variates with the Wishart distribution of the original variates. The transformation (9) induces a transformation (13) also in this case.

By making these substitutions into (21) and (35) we have the following theorem:

Theorem 1.

Given N multivariate normal populations each of p variates with identical variance-covariance matrices $|| \sigma_{ij} ||$ and with expected values of the pN variates $x_{1\alpha}$ equal to $\mu_{1\alpha}$. Let a_{1j} be defined by equation (3); let the rank of the matrix $|| \tau_{ij} ||$, defined by (4) be t, and suppose $N \geq p + t$.

(1) When $t = 0$, the joint probability density function of the a_{1j} is given by the Wishart distribution,

$$(38) \quad W(a_{ij}, \sigma^{ij}, \tau_{ij}; p, N-1, 0) \\ = \frac{|\sigma^{ij}|^{\frac{N-1}{2}} |a_{ij}|^{\frac{N-p-2}{2}} e^{-\frac{1}{2} \sum_{i,j=1}^p \sigma^{ij} a_{ij}}}{2^{\frac{p(N-1)}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma(\frac{N-i}{2})}.$$

(11) When $t = 1$, the joint probability density

function of the a_{ij} is

$$W(a_{ij}, \sigma^{ij}, \tau_{ij}, p, N-1, 1)$$

$$(39) \quad = \frac{e^{-\frac{1}{2} \sum_{i,j=1}^p \sum_{\alpha=1}^N \sigma^{ij} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)} |\sigma^{ij}|^{\frac{N-1}{2}} |a_{ij}|^{\frac{N-p-2}{2}}}{2^{\frac{p(N-1)-N-3}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p-1} \Gamma\left(\frac{N-1-i}{2}\right)}$$

$$e^{-\frac{1}{2} \sum_{i,j=1}^p \sigma^{ij} a_{ij} \left[\sum_{i,j=1}^p \sum_{\alpha=1}^N a_{ij} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j) \right]^{\frac{N-3}{4}}}$$

$$I_{\frac{N-3}{2}} \left(\sqrt{\sum_{i,j=1}^p \sum_{\alpha=1}^N a_{ij} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)} \right)$$

(iii) When $t = 2$, the joint probability density function of the a_{ij} is given by

$$\begin{aligned}
 & W(a_{ij}, \sigma^{ij}, \tau_{ij}, p, N-1, 2) \\
 &= \frac{e^{-\frac{1}{2} \sum_{i,j=1}^p \sum_{\alpha=1}^N \sigma^{ij} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)} |\sigma^{ij}|^{\frac{N-1}{2}} |a_{ij}|^{\frac{N-p-2}{2}}}{2^{\frac{p(N-1)}{2} - \frac{N-3}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p-2} \Gamma\left(\frac{N-2-i}{2}\right)} \\
 (40) \quad & e^{-\frac{1}{2} \sum_{i,j=1}^p \sigma^{ij} a_{ij}} \sum_{w=0}^{\infty} \frac{(u_1^2 u_2^2)^w (u_1^2 + u_2^2)^{-\frac{1}{2}(\frac{N-3}{2} + 2w)}}{2^{2w} w! \Gamma\left(\frac{N-2}{2} + w\right)}
 \end{aligned}$$

$$I_{\frac{N-3}{2} + 2w} \left(\sqrt{u_1^2 + u_2^2} \right),$$

where u_1^2 and u_2^2 are the two non-zero roots of (36).

VI. THE CHARACTERISTIC FUNCTION OF THE NON-CENTRAL WISHART DISTRIBUTION

We shall find the characteristic function of the a_{11} and $2a_{1j}$ ($1 \neq j$) as defined in (3). We first obtain the characteristic function of the b_{11} and $2b_{1j}$ ($1 \neq j$) as defined in (15) and then perform a linear transformation to obtain the characteristic function of the a_{1j} . The characteristic function of the b_{11} and $2b_{1j}$ ($1 \neq j$) is defined as

$$(41) \quad E \left(e^{i \sum_{i,j=1}^p b_{ij} \theta_{ij}} \right),$$

where

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

and i in the exponent is the imaginary quantity. Since

$$b_{ij} = \sum_{\alpha=1}^n \gamma_{i\alpha} \gamma_{j\alpha},$$

where the $y_{i\alpha}$ have the normal multivariate distribution specified in (16), we can write (41) as

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

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

$$\frac{1}{(2\pi)^{\frac{p(n-t)}{2}}} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \sum_{i=1}^p \sum_{\alpha=t+1}^n y_{i\alpha}^2 + i \sum_{i,j=1}^p \sum_{\alpha=t+1}^n y_{i\alpha} y_{j\alpha} \theta_{ij}} \prod_{i=1}^p \prod_{\alpha=t+1}^n dy_{i\alpha}.$$

This is, however, the characteristic function of a Wishart distribution with $n-t$ degrees of freedom [16], namely,

$$(42) \quad |\delta_{ij} - 2i\theta_{ij}|^{-\frac{n-t}{2}}.$$

Now we must integrate

$$(43) \frac{e^{-\frac{1}{2} \sum_{\eta=1}^t K_{\eta}^2}}{(2\pi)^{\frac{1}{2} p t}} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \sum_{i=1}^p \sum_{\eta=1}^t \gamma_{i\eta}^2 + i \sum_{i,j=1}^p \sum_{\eta=1}^t \gamma_{i\eta} \gamma_{j\eta} \theta_{ij} + \sum_{\eta=1}^t \gamma_{\eta\eta} K_{\eta}} \prod_{i=1}^p \prod_{\eta=1}^t d\gamma_{i\eta}.$$

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

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

where

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

Then the exponent of (43) within the integral sign is

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

and the Jacobian of the transformation is

$$|d_{ih}|^{-\frac{t}{2}}$$

Hence, the integral of (43) is

$$(44) \quad \frac{e^{-\frac{1}{2} \left(\sum_{\eta=1}^t k_{\eta}^2 - \sum_{\eta=1}^t d^{\eta\eta} k_{\eta}^2 \right)}}{| \delta_{ij} - 2i \theta_{ij} |^{\frac{1}{2}t}},$$

where

$$\|\delta^{ij}\| = \|\delta_{ij} - 2i\theta_{ij}\|^{-1}.$$

The characteristic function is the product of (42) and (44). Accordingly, we have the result that the characteristic function of the b_{ii} and $2b_{ij}$ ($i \neq j$) defined by (15) is

$$(45) \quad \frac{e^{-\frac{1}{2} \left(\sum_{\eta=1}^t \kappa_{\eta}^2 - \sum_{\eta=1}^t d^{\eta\eta} \kappa_{\eta}^2 \right)}}{\left| \delta_{ij} - 2i\theta_{ij} \right|^{\frac{1}{2}n}}.$$

where $\|\delta^{ij}\|$ is defined by (45).

It is clear that if $\kappa_{\eta} = 0$, this function reduces to the characteristic function of the Wishart distribution

with n degrees of freedom, namely,

$$(46) \quad | \delta_{ij} - 2i\theta_{ij} |^{-\frac{1}{2}n}.$$

It is interesting to note that (45) factors into two parts, one of which is (46) and the other is

$$(47) \quad e^{-\frac{1}{2} \left(\sum_{\gamma=1}^t K_{\gamma}^2 - \sum_{\gamma=1}^t d^{\gamma\gamma} K_{\gamma}^2 \right)}.$$

The distribution function similarly factors into two parts, one of which is the Wishart distribution, whose characteristic function is (46). This is unusual, for here the non-central Wishart distribution function is the convolution of a function (central Wishart distribution) and another (the transform of (47)) 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

$$\frac{e^{-\frac{1}{2}(\kappa_1^2 - d'' \kappa_1^2)}}{|\delta_{ij} - 2i\theta_{ij}|^{\frac{1}{2}n_1}} \cdot \frac{e^{-\frac{1}{2}(\kappa_2^2 - d^{22} \kappa_2^2)}}{|\delta_{ij} - 2i\theta_{ij}|^{\frac{1}{2}n_2}},$$

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 (16). Let

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

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

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''_{1j} has the non-central Wishart distribution with n_2 degrees of freedom and parameter k_2^2 in the direction of the second coordinate axis. Since

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

the distribution of the b_{1j} is a convolution of the distributions of b'_{1j} and b''_{1j} . In general the non-central distribution is the convolution of t distributions of the linear case (provided $n \geq t_{p+t}$).

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'_{1j} and b''_{1j} have non-central Wishart distributions with parameter $k_1'^2$ and $k_1''^2$ in the direction of the first coordinate axes and n_1 and n_2 degrees of freedom. The characteristic functions are

$$\frac{e^{-\frac{1}{2}(k_1'^2 - d'' k_1'^2)}}{|d_{ij}|^{\frac{1}{2} n_1}}$$

and

$$\frac{e^{-\frac{1}{2}(\kappa_1''^2 - d'' \kappa_1''^2)}}{|d_{ij}|^{\frac{1}{2}n_2}}$$

The product is

$$\frac{e^{-\frac{1}{2}(\kappa_1^2 - d'' \kappa_1^2)}}{|d_{ij}|^{\frac{1}{2}n}}$$

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

Now let us deduce the characteristic function of the a_{11} and $2a_{1j}$ ($1 \neq j$). Since by (13) the b 's are transforms of the a 's we can write $a_{1j} = \sum_{h,k=1}^p \psi^{ih} \psi^{jk} b_{hk}$. Then

$$(48) E\left(e^{i \sum_{i,j=1}^p a_{ij} \phi_{ij}}\right) = E\left(e^{i \sum_{i,j,k}^p \phi_{ij} \psi^{ih} \psi^{jk} b_{hk}}\right)$$

where $\phi_{1j} = \phi_{j1}$. If we define

$$(49) \quad \theta_{hk} = \sum_{i,j=1}^p \phi_{ij} \psi^{ih} \psi^{jk},$$

then (48) can be derived by substituting (49) in (45).

Let

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

Then

$$\|\phi_{ih}\| = D = \Phi^{-1} (\Sigma^{-1} - 2i\Phi) \Phi^{-1}$$

and

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

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

as

$$e^{-\frac{1}{2} \{ \text{tr}(\Phi T \Phi') - \text{tr}[\Phi (\Sigma^{-1} - 2i\Phi)^{-1} \Phi' \Phi T \Phi'] \}}$$

$$\left\{ |\Phi^{-1}| \cdot |\Sigma^{-1} - 2i\Phi| \cdot |\Phi^{-1}| \right\}^{\frac{1}{2}n}$$

using (7) and (8). The denominator is

$$\left\{ |\Phi'| |\Phi| \right\}^{-\frac{1}{2}n} \left| \Sigma^{-1} - 2i\Phi \right|^{\frac{1}{2}n},$$

and the numerator can be written as *

$$e^{-\frac{1}{2} \left\{ \text{tr}(M' \Phi' \Phi M) - \text{tr}[M' \Phi' \Phi (\Sigma^{-1} - 2i\Phi)^{-1} \Phi' \Phi M] \right\}}$$

where

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

and

$$M' M = T.$$

We may summarize in the following theorem:

Theorem 2.

Given a_{1j} ($1, j = 1, 2, \dots, p$) defined by (3) where
the $x_1 \sim (1 = 1, 2, \dots, p; \alpha = 1, 2, \dots, N)$ are distributed

* The result follows from the fact that $\text{tr}(AB) = \text{tr}(BA)$.

according to (2), the characteristic function of a_{1j} and $2a_{1j}$ ($1 \neq j$) is

$$\begin{aligned}
 & E \left(e^{i \sum_{i,j=1}^p a_{ij} \phi_{ij}} \right) \\
 (50) = & \frac{|\sigma_{ij}|^{\frac{1}{2}(N-1)} e^{-\frac{1}{2} \sum_{i,j=1}^p \sum_{\alpha=1}^N \sigma_{ij} (\mu_{i\alpha} - \bar{\mu}_i)(\mu_{j\alpha} - \bar{\mu}_j)}}{|\tilde{\sigma}_{ij}|^{\frac{1}{2}(N-1)}} \\
 & \cdot e^{\frac{1}{2} \sum_{\substack{i,j=1 \\ h,k}}^p \sum_{\alpha=1}^N \sigma^{hi} \tilde{\sigma}_{ij} \sigma^{jk} (\mu_{h\alpha} - \bar{\mu}_h)(\mu_{k\alpha} - \bar{\mu}_k)}
 \end{aligned}$$

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 (50) 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 observation from a non-central normal population. We can state this in a theorem.

Theorem 3.

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

VII. AN INTEGRAL REPRESENTATION OF THE NON-CENTRAL WISHART DISTRIBUTION IN THE GENERAL CASE

The integral of (17) which is the non-central Wishart distribution, can be indicated in another way that reveals more clearly the *form* of the non-central Wishart distribution. We shall now show that it is the product of a Wishart distribution and a symmetric function of the non-zero roots of the determinantal equation

$$(51) \quad |k_i^2 \delta_{ij} - \lambda b^{ij}| = 0.$$

Let

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

$$Y = \| y_{in} \|,$$

$$K = \| \kappa_i \delta_{ij} \|.$$

Then (17) can be written as

$$(52) C e^{-\frac{1}{2} \text{tr} B} |B|^{\frac{1}{2}(n-p-t-1)} |I - Y'B^{-1}Y|^{\frac{1}{2}(n-p-t-1)} e^{\text{tr}(KY)} dB dY,$$

where

$$(53) C = \frac{2^{\frac{1}{2}pn} \pi^{\frac{1}{4}p(p-1) + \frac{1}{2}pt} \prod_{i=1}^p \Gamma(\frac{1}{2}[n-t+1-i])}{},$$

and where dB represents $\prod_{i=1}^p \prod_{j=i}^p db_{ij}$ and $dY = \prod_{i=1}^p \prod_{j=1}^t dy_{ij}$.

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

$$G' B^{-1} G = I,$$

$$K G = D = \| u_i \delta_{ij} \|,$$

where the u_i^2 are roots of (51). Then make the transformation

$$Y = G Z,$$

where

$$Z = \| z_{i\eta} \|.$$

The Jacobian of this transformation is

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

Now (52) can be written as

$$(54) \quad C e^{-\frac{1}{2} \text{tr} B} |B|^{\frac{1}{2}(n-p-1)} \times |I - Z' Z|^{\frac{1}{2}(n-p-t-1)} e^{\text{tr}(DZ)} dB dZ.$$

The first part of (54) is, except for a constant factor, a Wishart distribution of n degrees of freedom. The integral of the second part is obviously a symmetric function of the u_1 . In terms of the a_{1j} the u_1^2 are simply the roots of (36). We can sum these results in a theorem.

Theorem 4.

Given a sample of observations $\{x_{i\alpha}\}$ ($i=1, 2, \dots, p; \alpha=1, 2, \dots, N$) 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

$$W(a_{ij}, \sigma^{ij}, \tau_{ij}; p, N-1, t)$$

$$= C |\sigma^{ij}|^{\frac{1}{2}(N-1)} |a_{ij}|^{\frac{1}{2}(N-p-2)} e^{-\frac{1}{2} \sum_{ij=1}^p \sigma^{ij} a_{ij}}$$

$$\cdot \int |\delta_{\eta\zeta} - \sum_{i=1}^p z_{i\eta} z_{i\zeta}|^{\frac{1}{2}(N-p-t-2)} e^{\sum_{\eta=1}^t u_{\eta} z_{\eta\eta}} \prod_{i=1}^p \prod_{\eta=1}^t d|z_{i\eta}|$$

integrated over

$$\| \delta_{\eta\zeta} - \sum_{i=1}^p z_{i\eta} z_{i\zeta} \|$$

positive where C and τ_{ij} are defined by (53) and (4), respectively, and where u_n^2 are the t non-zero roots of (36).

VIII. THE MOMENTS OF THE GENERALIZED VARIANCE IN THE LINEAR AND PLANAR NON-CENTRAL CASES

A. The Linear Case.

The generalized variance which is the determinant of the variances and covariances is a measure of the spread of the observations. If one thinks of the n observations (reduced by the mean) of each variate as a vector in n -space, the generalized variance is 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 ^{of the squared volumes} the sum of all possible parallelotopes that can be joined by choosing as the p principal edges p of the n sample vectors.

In this section we consider the moments of the general-

ized 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 origin. For convenience we shall assume that the distribution of the observations $\{y_{i\alpha}\}$ is according to (15). This will give as great 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 determinant $|b_{ij}|$, where

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

is simply a multiple (by $|\Sigma| (N-1)$) of $|\frac{a_{ij}}{N-1}|$, where

$$\|a_{ij}\| = \left\| \sum_{\alpha=1}^N (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) \right\|.$$

Let us first consider the linear case, i.e., $\kappa = \kappa_1 \neq 0$ and $\kappa_i = 0$ ($i=2, \dots, p$) in (15). 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

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

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 sample variance of the i -th variate and $||r_{ij}||$ is the matrix of sample correlation coefficients. The Jacobian of this transformation is

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

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

$$(55) \quad \frac{e^{-\frac{1}{2}k^2 - \sum_{i=1}^p s_i^2} \prod_{i=1}^p (s_i^2)^{\frac{n}{2}-1}}{2^{\frac{1}{2}pn} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n-i])} |r_{ij}|^{\frac{n-p-1}{2}}$$

$$\sum_{\alpha=0}^{\infty} \frac{(k^2 s_i^2)^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}n+\alpha)} \prod_{i=1}^p d(s_i^2) \prod_{i=1}^p \prod_{j=i+1}^p d(r_{ij}).$$

It is clear from (55) 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 ($i=2,3,\dots,p$) is

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

which is simply the χ^2 -distribution.

The h-th moment of s_1^2 ($h=2,3,\dots,p$) is

$$E[(s_1^2)^h] = \frac{2^h \Gamma(\frac{1}{2}n+h)}{\Gamma(\frac{1}{2}n)}.$$

The probability element of s_1^2 is

$$(56) \quad \frac{e^{-\frac{1}{2}k^2}(s_1^2)^{\frac{1}{2}n-1} e^{-\frac{1}{2}s_1^2}}{2^{\frac{1}{2}n}} \sum_{\alpha=0}^{\infty} \frac{(k^2 s_1^2)^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}n+\alpha)} d(s_1^2).$$

This is the χ'^2 -distribution (non-central χ^2 -distribution) which was first given by Fisher [1]. Applying term-by-term integration (the series converges properly) we get the moment

$$E[(s_1^2)^h] = 2^h e^{-\frac{1}{2}k_1^2} \sum_{\alpha=0}^{\infty} \frac{(k^2)^{\alpha} \Gamma(\frac{1}{2}n+h+\alpha)}{2^{\alpha} \alpha! \Gamma(\frac{1}{2}n+\alpha)}.$$

The probability element of the r_{1j} is the well known dis-

tribution of correlation coefficients,

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

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 (permissible) range of the r_{ij} , we have as a consequence the h -th moment of the determinant (since n is arbitrary).

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

Hence, the h -th moment of $|a_1 a_j r_{1j}|$ is

$$(57) \quad 2^{ph} \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n-i] + h)}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n-i])} e^{-\frac{1}{2}k^2} \sum_{\alpha=0}^{\infty} \frac{k^{2\alpha} \Gamma(\frac{1}{2}n + h + \alpha)}{2^\alpha \alpha! \Gamma(\frac{1}{2}n + \alpha)}.$$

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

Theorem 5.

If the quantities a_{ij} ($i, j = 1, 2, \dots, p$) have the distribution

$$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, n, 1)$$

defined by (39), then the moments of $|a_{ij}|$ are given by

$$E(|a_{ij}|^h) = |\sigma_{ij}|^h 2^{ph} e^{-\frac{1}{2}k^2} \frac{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n-i] + h)}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n-i])} \sum_{\alpha=0}^{\infty} \frac{k^{2\alpha} \Gamma(\frac{1}{2}n + h + \alpha)}{2^\alpha \alpha! \Gamma(\frac{1}{2}n + \alpha)}.$$

If $k^2 = 0$, expression (57) clearly reduces to the moment given by Wilks [2]

$$(58) \quad 2^{ph} \frac{\prod_{i=1}^p \Gamma\left(\frac{1}{2}[n+1-i] + h\right)}{\prod_{i=1}^p \Gamma\left(\frac{1}{2}[n+1-i]\right)}.$$

The expression (57) 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 $e^{\frac{1}{2}k^2}$.

B. The Two Dimensional Planar Case.

Next we shall treat the planar case for two dimensions. Suppose that $k_i^2 \neq 0$ ($i=1,2$), $k_i^2 = 0$ ($i=3,4,\dots,p$), the probability density function of b_{11} , b_{12} , and b_{22} is

$$\frac{e^{-\frac{1}{2}(k_1^2 + k_2^2) - \frac{1}{2} \sum_{i=1}^2 b_{ii}}}{2^n \sqrt{\pi}} (b_{11} b_{22} - b_{12}^2)^{\frac{1}{2}(n-3)}$$

$$\sum_{\alpha, \beta=0}^{\infty} \frac{[k_1^2 k_2^2 (b_{11} b_{22} - b_{12}^2)]^{\alpha} (k_1^2 b_{11} + k_2^2 b_{22})^{\beta}}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[n-1]+\alpha) \Gamma(\frac{1}{2}n+2\alpha+\beta)}$$

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

$$(59) \quad \frac{e^{-\frac{1}{2}(k_1^2 + k_2^2)}}{2^n \sqrt{\pi}} (s_1^2)^{\frac{1}{2}n-1} (s_2^2)^{\frac{1}{2}n-1} (1-r^2)^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}(s_1^2 + s_2^2)}$$

$$\sum_{\alpha, \beta=0}^{\infty} \frac{(k_1^2 k_2^2 s_1^2 s_2^2)^{\alpha} (1-r^2)^{\alpha} (k_1^2 s_1^2 + k_2^2 s_2^2)^{\beta} d(s_1^2) d(s_2^2) dr}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[n-1]+\alpha) \Gamma(\frac{1}{2}n+2\alpha+\beta)}$$

We wish to find $E \{ [s_1^2 s_2^2 (1-r^2)]^h \}$. Let us first multiply (59) by $(1-r^2)^h$ and integrate from -1 to +1.

We then obtain

$$2^{-h} e^{-\frac{1}{2}(k_1^2 + k_2^2)} (s_1^2)^{\frac{1}{2}n-1} (s_2^2)^{\frac{1}{2}n-1} e^{-\frac{1}{2}(s_1^2 + s_2^2)}$$

$$\sum_{\alpha, \beta=0}^{\infty} \frac{(k_1^2 k_2^2 s_1^2 s_2^2)^{\alpha} (k_1^2 s_1^2 + k_2^2 s_2^2)^{\beta} \Gamma(\frac{1}{2}[n-1]+h+\alpha) d(s_1^2) d(s_2^2)}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[n-1]+\alpha) \Gamma(\frac{1}{2}n+2\alpha+\beta) \Gamma(\frac{1}{2}n+h+\alpha)}$$

Next we multiply by $(s_1^2)^h (s_2^2)^h$, set $(k_1^2 s_1^2 + k_2^2 s_2^2)^\beta / \beta!$ equal to

$$\sum_{\beta_1 + \beta_2 = \beta} \frac{(k_1^2 s_1^2)^{\beta_1} (k_2^2 s_2^2)^{\beta_2}}{\beta_1! \beta_2!}$$

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

$$E([b_{11} b_{22} - b_{12}^2]^h) \\ (60) = 2^{2h} e^{-\frac{1}{2}(k_1^2 + k_2^2)} \sum_{\alpha, \beta_1, \beta_2=0}^{\infty} \frac{(k_1^2)^{\alpha+\beta_1} (k_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}n+h+\alpha+\beta_1) \Gamma(\frac{1}{2}n+h+\alpha+\beta_2)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n-1]+\alpha) \Gamma(\frac{1}{2}n+2+\alpha+\beta_1+\beta_2)} \\ \cdot \frac{\Gamma(\frac{1}{2}[n-1]+h+\alpha)}{\Gamma(\frac{1}{2}n+h+\alpha)}$$

which is the expected value we are seeking.

Clearly this reduces to a special case of (57) if k_2^2 is set equal to zero.

C. The General Planar Case.

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 (16) 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 parallelepiped which can be expressed as

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

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

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

If the y's had simply the distribution

$$(61) \quad \frac{1}{(2\pi)^{\frac{1}{2}pn}} e^{-\frac{1}{2} \sum_{i=1}^p \sum_{a=1}^n y_{ia}^2}$$

then the h-th moment of $|b_{ij}|$ would be (58) and the h-th moment of

$$V_1^2 V_2^2 \sin^2 \theta_1 = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

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 $v_3, v_4, \dots, v_p$ and $\theta_2, \dots, \theta_{p-1}$ is the same whether the y's are distributed according to (16) or (61), we have

$$(62) E[(v_3 \dots v_p \sin \theta_2 \dots \sin \theta_{p-1})^{2h}] = 2^{h(p-2)} \frac{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i] + h)}{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i])}$$

Multiplying (60) by (62) we obtain the h-th moment of

$|b_{ij}|$, namely,

$$E(|b_{ij}|^h) = 2^{hp} e^{-\frac{1}{2}(k_1^2 + k_2^2)} \frac{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i] + h)}{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i])}$$

$$\cdot \sum_{\alpha, \beta_1, \beta_2=0}^{\infty} \frac{(k_1^2)^{\alpha+\beta_1} (k_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}n+h+\alpha+\beta_1)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n-1]+\alpha)}$$

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

This result may be summarized as follows:

Theorem 6.

Let the probability density function of the quantities

a_{ij} ($i, j=1, 2, \dots, p$) be

$$W(a_{ij}, \sigma_{ij}, \tau_{ij}; p, n, 2)$$

defined by (38). Then the h-th moment of $|a_{ij}|$ is

$$E(|a_{ij}|^h) = |\sigma_{ij}|^h 2^{hp} e^{-\frac{1}{2}(K_1^2 + K_2^2)} \frac{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i] + h)}{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+1-i])}$$

$$\sum_{\alpha, \beta_1, \beta_2=0}^{\infty} \frac{(K_1^2)^{\alpha+\beta_1} (K_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}n+h+\alpha+\beta_1) \Gamma(\frac{1}{2}n+h+\alpha+\beta_2) \Gamma(\frac{1}{2}[n-1]+h+\alpha)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n-1]+\alpha) \Gamma(\frac{1}{2}n+2\alpha+\beta_1+\beta_2) \Gamma(\frac{1}{2}n+h+\alpha)}$$

IX. THE MOMENTS OF THE CRITERION FOR TESTING LINEAR HYPOTHESES IN THE LINEAR AND PLANAR NON-CENTRAL CASES

A. 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 B we shall indicate the linear hypothesis 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 $z_1 \alpha$ ($i=1,2,\dots,p$; $\alpha=1,2,\dots,n$) and $y_1 \gamma$ ($i=1,2,\dots,p$; $\gamma=1,2,\dots,m$) have the probability element

(64)

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

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

$$\frac{|\sigma_{ij}|^{\frac{n+m}{2}}}{(2\pi)^{\frac{p(n+m)}{2}}} e^{-\frac{1}{2} \left\{ \sum_{i,j=1}^p \sum_{\alpha=1}^n \sigma_{ij} z_{i\alpha} z_{j\alpha} + \sum_{i,j=1}^p \sum_{\gamma=1}^m \sigma_{ij} (y_{i\gamma} - \mu_{i\gamma})(y_{j\gamma} - \mu_{j\gamma}) \right\}} \prod_{i=1}^p \prod_{\alpha=1}^n dz_{i\alpha} \prod_{i=1}^p \prod_{\gamma=1}^m dy_{i\gamma}$$

Let

(65)

$$a_{ij} = \sum_{\gamma=1}^m y_{i\gamma} y_{j\gamma},$$

$$(66) \quad b_{ij} = \sum_{\alpha=1}^n z_{i\alpha} z_{j\alpha},$$

$$(67) \quad c_{ij} = a_{ij} + b_{ij}.$$

Then the likelihood ratio criterion for testing the hypothesis, called by Hsu [4] the Wilks-Lawley criterion, is the $\frac{n+m}{2}$ power of

$$(68) \quad W = \frac{|b_{ij}|}{|c_{ij}|}.$$

Under the null hypothesis the b_{ij} have a Wishart distribution with n degrees of freedom, and the a_{ij} are

are independently distributed such that c_{1j} has a Wishart distribution with $n+m$ degrees of freedom. Wilks [2] 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 (64) where the rank of $||\mu_{i\gamma}||$ is 2, i.e., the planar case. Under this assumption the b_{1j} have a Wishart distribution with n degrees of freedom, the a_{1j} are independently distributed in such a way that the c_{1j} 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

$$(69) \quad \left| \sum_{\gamma=1}^m \mu_{i\gamma} \mu_{j\gamma} - \lambda \sigma_{ij} \right| = 0.$$

It is clear that W is unchanged if σ^{1j} is set equal to

δ_{ij} . Then the c_{ij} are distributed according to (35) with $n+m$ ^{degrees of freedom.} $\frac{p}{1-p}$. The moments will be obtained by a method similar to that used by Wilks [2].

Let the expected value given by (63) be

$$(70) E(|c_{ij}|^h) = K(n+m, h, p, \kappa_i^2),$$

which is a constant depending on $n+m$, h , p , κ_1^2 , and κ_2^2 . If $D(a_{ij})$ represents the distribution function of the a_{ij} , one can write (70) as

$$(71) K(n+m, h, p, \kappa_i^2) = \frac{1}{2^{\frac{1}{2}pn} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i])} \int |c_{ij}|^h |b_{ij}|^{\frac{1}{2}(n-p-1)} e^{-\frac{1}{2} \sum_{i=1}^p b_{ii}} D(a_{ij}) \prod_{i=1}^p \prod_{j=i}^p db_{ij} dA,$$

where dA is the volume element of the a_{ij} , and

where the integration is over the entire ranges of the b_{ij} and a_{ij} . Equation (71) holds since the c 's are functions of the b 's and a 's. Multiplying (71) by

$$(72) \quad 2^{\frac{pn}{2}} \prod_{i=1}^p \Gamma\left(\frac{n+1-i}{2}\right),$$

then replacing n by $n+2g$ and dividing by (72) again, we obtain

$$(73) \quad K(n+m+2g, h, p, k_i^2) \frac{2^{\frac{p(n+2g)}{2}} \prod_{i=1}^p \Gamma\left(\frac{n+2g+1-i}{2}\right)}{2^{\frac{pn}{2}} \prod_{i=1}^p \Gamma\left(\frac{n+1-i}{2}\right)}$$

$$= \frac{1}{2^{\frac{pn}{2}} \pi^{\frac{p(p-1)}{2}} \prod_{i=1}^p \Gamma\left(\frac{n+1-i}{2}\right)}$$

$$\cdot \int |c_{ij}|^h |b_{ij}|^{\frac{n+2g-p-1}{2}} e^{-\frac{1}{2} \sum_{i=1}^p b_{ii}} D(a_{ij}) \prod_{i=1}^p \prod_{j=i}^p db_{ij} dA.$$

By definition the right hand side of (73) is the expected value of $|c_{1j}|^h |b_{1j}|^g$. Hence,

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

In this expression it is permissible to set h equal to $-g$ (n could have been replaced by $n+2g$ in (71) to insure the argument of each Γ function being positive). Then we have

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

Finally, the g -th moment is

$$E(W^g) = e^{-\frac{1}{2}(K_1^2 + K_2^2)} \frac{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+m+1-i]) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i] + g)}{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+m+1-i] + g) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i])}$$

(74)

$$\sum_{\alpha, \beta_1, \beta_2=0}^{\infty} \frac{(K_1^2)^{\alpha+\beta_1} (K_2^2)^{\alpha+\beta_2} \Gamma(\frac{1}{2}[n+m]+u+\beta_1) \Gamma(\frac{1}{2}[n+m]+\alpha+\beta_2) \Gamma(\frac{1}{2}[n+m-1]+\alpha)}{2^{2\alpha+\beta_1+\beta_2} \alpha! \beta_1! \beta_2! \Gamma(\frac{1}{2}[n+m-1]+\alpha+g) \Gamma(\frac{1}{2}[n+m]+\alpha+\beta_1+\beta_2+g) \Gamma(\frac{1}{2}[n+m]+\alpha)}$$

We can summarize in the following theorem:

Theorem 7.

Let $z_1 \propto (i=1, 2, \dots, p; \alpha=1, 2, \dots, n)$ and $y_1 \propto (i=1, 2, \dots, p; \gamma=1, 2, \dots, m)$ have (64) as a joint distribution. Define a_{ij} , b_{ij} and c_{ij} by (65), (66), and (67), respectively. Let K_1^2 and K_2^2 be the non-zero roots of (69). Then the g -th moment of W , defined by (68), is (74).

Expression (74) 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 moments of W in the linear case is given by

$$E(W^g) = e^{-\frac{1}{2}\kappa_1^2} \frac{\prod_{i=2}^p \Gamma(\frac{1}{2}[n+m+1-i]) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i] + g)}{\prod_{i=2}^p \Gamma(\frac{1}{2}[n+m+1-i] + g) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i])}$$

(75)

$$\sum_{\beta_1=0}^{\infty} \frac{(\kappa_1^2)^{\beta_1} \Gamma(\frac{1}{2}[n+m] + \beta_1)}{2^{\beta_1} \beta_1! \Gamma(\frac{1}{2}[n+m] + \beta_1 + g)}.$$

For $\kappa_1^2 = 0$, (75) reduces to the expression given for the moments under the null hypothesis, namely,

$$(76) E(W^g) = \frac{\prod_{i=1}^p \Gamma(\frac{1}{2}[n+m+1-i]) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i] + g)}{\prod_{i=1}^p \Gamma(\frac{1}{2}[n+m+1-i] + g) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i])}.$$

Wilks [17] 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 $K_1^2 \neq 0$) the distribution functions are yet more involved. Hence, we shall not write any explicitly.

For most purposes, alternative hypotheses of the means μ_i being on a line (i.e., of rank one) are sufficiently general. In any particular case, one can compute from (75) numerical values for several moments and then fit an appropriate distribution function. If one wishes to consider alternative hypotheses of rank 2, one can use (74) and similarly compute numerical values for moments. The series in either (75) or (74) converge rapidly. To construct an approximate power function for linear alternatives, say, one would fit distribution functions for several values of K_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

$$(77) \quad |a_{ij} - \lambda b_{ij}| = 0.$$

It follows that

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

Then W can be written as

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

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

The distribution of the roots of (77) is given in Section XI for a_{ij} of rank p in either the linear or planar case. One could obtain the probability of W not exceeding a given value by integrating the λ 's over the proper range.

B. A General Class of Linear Hypotheses.

A large class of hypotheses concerning the means of multivariate normal variates can be put in the form of H_0 . Hsu [3] has demonstrated that the general regression problem can be put in this form. Suppose that

$X_{i\alpha}$ ($i=1,2,\dots,p$; $\alpha=1,2,\dots,N$) follow a multivariate normal distribution with variance-covariance matrix $\|\sigma_{ij}\|$, and let the expected value of $X_{i\alpha}$ be

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

where the q by N matrix

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

is of rank q . Let H_1 be the hypothesis that

$$H_1: B_1 = \|\beta_{iu}\| = 0 \quad (i=1,2,\dots,p; u=1,2,\dots,m \leq q)$$

with the w 's known. Then Hsu has shown that there is a transformation of the x 's into the y 's and z 's which is dependent on the w 's such that the distribution of the x 's

is transformed into (64) times a distribution of (q-m) extraneous variables and such that H_1 is transformed into H_0 .

The hypothesis H_1 has great generality. For example, H_1 includes the hypothesis that q populations (of p variates) known to have the same set of variances and covariances have also a common set of means. In Hsu's paper this is defined as

$$W_{r\alpha} = \begin{cases} 1 & \text{when } N_1 + \dots + N_{r-1} + 1 \leq \alpha \leq N_1 + \dots + N_r \\ 0 & \text{otherwise} \end{cases} \quad \left. \begin{matrix} \\ \end{matrix} \right\} r=1, \dots, q-1,$$

$$W_{q\alpha} = 1, \quad \alpha=1, \dots, N,$$

where $N = N_1 + N_2 + \dots + N_q$ (N_r is the size of the r-th sample).

It is convenient to use matrix notation here.

Let

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

$$W_2 = \| w_{r\alpha} \| \quad (r=1,2,\dots,q; \alpha=1,2,\dots,N),$$

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

Then

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}.$$

Then the likelihood ratio criterion for H_1 is*

$$(78) \quad \Lambda = \frac{\begin{vmatrix} \bar{X} \bar{X}' & \bar{X} W' \\ W \bar{X}' & W W' \end{vmatrix} \begin{vmatrix} W_2 W_2' \end{vmatrix}}{\begin{vmatrix} W W' & \bar{X} \bar{X}' & \bar{X} W_2' \\ W_2 \bar{X}' & W_2 W_2' \end{vmatrix}} \cdot \frac{1}{2} N$$

The following identity

$$\begin{vmatrix} P & Q \\ R & S \end{vmatrix} = |S| \cdot |P - Q S^{-1} R|,$$

* This follows from well-known regression theory. See [12], for instance.

which results from

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} I & 0 \\ -S^{-1}R & I \end{pmatrix} = \begin{pmatrix} P - QS^{-1}R & Q \\ 0 & S \end{pmatrix},$$

may be applied to (78) to obtain

$$(79) \lambda = \frac{|\bar{X}\bar{X}' - \bar{X}W'(WW')^{-1}W\bar{X}'|^{\frac{1}{2}N}}{|\bar{X}\bar{X}' - \bar{X}W_2'(W_2W_2')^{-1}W_2\bar{X}'|^{\frac{1}{2}N}}.$$

If $WW' = EE'$, where E has zeros below the main diagonal, let $\Gamma_1 = E^{-1}W$. Then $\Gamma_1\Gamma_1' = I$.

We can choose Γ_2 (of $N-q$ rows) so that

$$\Gamma = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}$$

is orthogonal. Then Γ is the matrix of Hsu's transformation, that is,

$$Y = \| y_{i\gamma} \| = X \Gamma_1'$$

$$Z = \| z_{i\alpha} \| = X \Gamma_2'$$

Hsu [3] shows that the $y_{i\gamma}$ ($i=1,2,\dots,p$; $\gamma=m+1,\dots,q$) have means unspecified by H_1 and hence can be omitted from further consideration. We must have $N-q = n$.

This transformation transforms the right hand side of (79) into

$$\Lambda = \frac{|b_{ij}|^{\frac{1}{2}N}}{|c_{ij}|^{\frac{1}{2}N}}.$$

That is, the criterion for testing H_1 is equivalent to that for testing H_0 .*

The equation (69) can be written in terms of Z , B , and W as

$$(81) \quad |B_1 W_1 (I - W_2' (W_2 W_2') W_2) W_1' B_1' - \lambda \Sigma| = 0$$

for $m < q$. If $m = q$, (169) becomes

$$(82) \quad |B_1 W W_1' B_1' - \lambda \Sigma| = 0.$$

In (81) and (82) there are no more non-zero roots than the rank of B . It is clear that the roots of (81) (or (82)) depend on the matrix W as well as B . The distri-

* λ is the $N/2$ power of (68) instead of the $(n+m)/2$ power.

bution of Λ 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 k_1^2 or on k_1^2 and k_2^2 , and hence, on the distribution of the elements of W as well as the value of B .

C. Two Special Cases.

Suppose $m = 1$. Let $k_1^2 = k_2^2$; of course, $k_2^2 = 0$. Then (75) is

$$(83) e^{-\frac{1}{2}k^2} \frac{\Gamma(\frac{1}{2}[n+1-p]+g)}{\Gamma(\frac{1}{2}[n+1-p])} \sum_{\beta=0}^{\infty} \frac{(k^2)^{\beta} \Gamma(\frac{1}{2}[n+1]+\beta)}{2^{\beta} \beta! \Gamma(\frac{1}{2}[n+1]+\beta+g)}.$$

Expression (82) can be given the following integral representation:

$$\frac{e^{-\frac{1}{2}k^2}}{\Gamma(\frac{1}{2}[n+1-p])} \int_0^1 Y^{\frac{1}{2}(n-p-1)+g} \sum_{\beta=0}^{\infty} \frac{(k^2)^{\beta} \Gamma(\frac{1}{2}[n+1]+\beta) (1-Y)^{\beta+\frac{1}{2}p-1}}{2^{\beta} \beta! \Gamma(\frac{1}{2}p+\beta)} dY.$$

This is simply the definition of the g-th moment of Y, namely, $E(Y^g)$, where Y has the distribution

$$\frac{e^{-\frac{1}{2}k^2} Y^{\frac{1}{2}(n-p-1)}}{\Gamma(\frac{1}{2}[n-p+1])} \sum_{\beta=0}^{\infty} \frac{(k^2)^{\beta} \Gamma(\frac{1}{2}[n+1]+\beta)(1-Y)^{\beta+\frac{1}{2}p-1}}{2^{\beta} \beta! \Gamma(\frac{1}{2}p+\beta)}$$

Hotelling's generalized T^2 [18] is a rational function of Y; it is,

$$T^2 = \frac{1-Y}{Y}$$

$$Y = \frac{1}{1+T^2}$$

Hence, the distribution of T^2 when the null hypothesis is false is

$$(84) \frac{e^{-\frac{1}{2}k^2} (1+T^2)^{-\frac{1}{2}(n-p-3)}}{\Gamma(\frac{1}{2}[n-p+1])} \sum_{\beta=0}^{\infty} \frac{(k^2)^{\beta} \Gamma(\frac{1}{2}[n+1]+\beta)}{2^{\beta} \beta! \Gamma(\frac{1}{2}p+\beta)} \left(\frac{T^2}{1+T^2} \right)^{\beta+\frac{1}{2}p-1}$$

Here

$$f = n$$

note: here used
+ for T^2

-92-

$$f_1 = k = p$$

$$f_2 = f + 1 - k = n + 1 - p$$

$$e^{-\frac{k^2}{2}} \sum_{\beta=0}^{\infty} \frac{(\frac{k^2}{2})^{\beta}}{\beta!} \frac{\Gamma(\frac{n+1}{2}+\beta)}{\Gamma(\frac{p}{2}+\beta) \Gamma(\frac{n+1-p}{2})} \frac{\lambda^{\frac{p}{2}-1+\beta}}{(1+T^2)^{\frac{n+1}{2}+\beta}}$$

This result has been obtained by Hsu [19] by another method. The function (84) has been studied by Tang [20] in connection with the power functions of the analysis of variance tests. Here n is the number of degrees of freedom.

Suppose $p = 1$ and $m = q$. Then

$$A^{\frac{2}{N}} = \frac{\begin{vmatrix} \sum_{\alpha=1}^N \chi_{\alpha}^2 & \sum_{\alpha=1}^N \chi_{\alpha} w_{1\alpha} & \dots & \sum_{\alpha=1}^N \chi_{\alpha} w_{q\alpha} \\ \sum_{\alpha=1}^N \chi_{\alpha} w_{1\alpha} & \sum_{\alpha=1}^N w_{1\alpha}^2 & \dots & \sum_{\alpha=1}^N w_{1\alpha} w_{q\alpha} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{\alpha=1}^N \chi_{\alpha} w_{q\alpha} & \sum_{\alpha=1}^N w_{1\alpha} w_{q\alpha} & \dots & \sum_{\alpha=1}^N w_{q\alpha}^2 \end{vmatrix}}{\begin{vmatrix} \sum_{\alpha=1}^N \chi_{\alpha}^2 & \sum_{\alpha=1}^N w_{1\alpha}^2 & \dots & \sum_{\alpha=1}^N w_{1\alpha} w_{q\alpha} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{\alpha=1}^N w_{1\alpha} w_{q\alpha} & \dots & \dots & \sum_{\alpha=1}^N w_{q\alpha}^2 \end{vmatrix}}$$

is one minus the square of the multiple correlation coefficient. Then k^2 is the root of

$$\sum_{\alpha=1}^N \left(\sum_{u=1}^q \beta_u w_{u\alpha} \right)^2 - \lambda \sigma^2 = 0,$$

where σ^2 is the variance of X . We shall assume that the $w_{u\alpha}$ ($u=1,2,\dots,q$; $\alpha=1,2,\dots,n$) are fixed. Then the moments of $1-R^2$, where R is the multiple correlation coefficient, are given by

$$(85) E[(1-R^2)^\beta] = \frac{e^{-\frac{1}{2}k^2} (1-R^2)^{\frac{N-q}{2}-1}}{\Gamma(\frac{N-q}{2})} \sum_{\beta=0}^{\infty} \frac{(k^2)^\beta \Gamma(\frac{N}{2} + \beta)}{2^\beta \beta! \Gamma(\frac{N}{2} + \beta)}$$

Expression (85) can be put into the form of an integral, that is,

$$E[(1-R^2)^\beta] = \frac{e^{-\frac{1}{2}k^2}}{\Gamma(\frac{N-q}{2})} \int_0^1 \sum_{\beta=0}^{\infty} \frac{(k^2)^\beta \Gamma(\frac{1}{2}N + \beta) (1-R^2)^{\frac{1}{2}(N-q) + \beta - 1}}{2^\beta \beta! \Gamma(\frac{1}{2}q + \beta)} d(R^2)$$

It follows that the (symmetric) distribution of R (which is continuous) is

$$\frac{e^{-\frac{1}{2}k^2} (1-R^2)^{\frac{1}{2}(N-1)-1}}{\Gamma(\frac{1}{2}[N-1])} \sum_{\beta=0}^{\infty} \frac{(k^2)^{\beta} \Gamma(\frac{1}{2}N+\beta) R^{q+2\beta-1}}{2^{\beta} \beta! \Gamma(\frac{1}{2}q+\beta)}$$

There is a difference between this result and the one given by Fisher [1] because here the w 's are fixed instead of being normally distributed variates.

Fisher gave both

D. Asymptotic Distribution.

Hsu [4] has given the asymptotic distribution of W . Suppose that the $w_{u\alpha}$ are such that the following function of the transformed constants

$$\mathcal{I}_n = \sum_{i,j=1}^p \sum_{\gamma=1}^m \mu_{i\gamma} \mu_{j\gamma}$$

tends to the limit $\mathcal{I}_0$ as n tends to infinity. Then the limiting distribution of $x = -n \log W$ (which equals $-2 \log \Lambda$ where Λ is the likelihood ratio criterion) is

$$\frac{e^{-\frac{1}{2}\mathbb{F}_0} \chi^{\frac{1}{2}pm-1} e^{-\frac{1}{2}\chi^2}}{2^{\frac{1}{2}pm}} \sum_{\alpha=0}^{\infty} \frac{\mathbb{F}_0^{\alpha} \chi^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}pm+\alpha)}.$$

That is, it is the χ'^2 distribution with pm degrees of freedom and parameter $\mathbb{F}_0$.

Hsu demonstrates this by first showing that the Lawley criterion

$$V = \sum_{i,j=1}^p b^{ij} a_{ij}$$

multiplied by n has the above χ'^2 -distribution and that nV and $-n \log W$ approach each other (in the probability sense).

If the $w_{u\alpha}$ are statistical variates (viz. with given continuous distribution function) it is easy to apply Wald's theorem [21] on the asymptotic distribution of the likelihood ratio criterion to obtain the same result.

X. LIKELIHOOD RATIO CRITERIA FOR TESTING THE DIMENSIONALITY OF POPULATION MEANS

A. Testing Collinearity.

The likelihood ratio criterion considered in Section IX can be used to test, on the basis of a sample, the hypothesis that the sets of means in several multivariate populations are identical assuming that the matrices of variances and covariances are the same. If this hypothesis is false, one may wish to determine whether there is a relationship between the sets of means and what this relationship is. The set of means of a p -variate population may be considered as the coordinates of a point in a p -dimensional space. One may then state the question of relationship between sets of means in geometrical terms. The problem is: Given a set of samples each from a multivariate normal population, what is the dimensionality of the linear space determined by the points representing the means of the populations under the assumption that the matrices of variances and covariances in the several normal multivariate populations

are the same? In Section A we develop a criterion for testing whether these points lie on a line; in Section B we obtain criteria for testing any specified dimensionality. The alternative hypotheses are that the means lie anywhere in the p -dimensional space. Criteria for testing one dimensionality against a specified greater dimensionality are given in Section D. The criteria obtained are the likelihood ratio criteria.*

Let Ω be the set of admissible hypotheses which contains m multivariate normal populations with the same variance-covariance matrix. Symbolically Ω is defined as

$$\Omega : \|\sigma_{\dot{1}}^{\dot{1}}\| = \|\sigma_{\dot{2}}^{\dot{2}}\| = \dots = \|\sigma_{\dot{m}}^{\dot{m}}\| = \|\sigma^{\dot{1}}\| \text{ positive definite,}$$

$$-\infty < \mu_i^q < +\infty \quad q = 1, 2, \dots, m; \quad i = 1, 2, \dots, p,$$

where the probability density function of the q -th popu-

* See [12], p. 150, for example.

lation is

$$(86) \frac{|\sigma_{ig}^{-1}|^{-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}} e^{-\frac{1}{2} \sum_{i,j=1}^p \sigma_g^{ij} (x_i^g - \mu_i^g)(x_j^g - \mu_j^g)}$$

Let ω be the hypothesis we test which contains m multivariate normal populations whose means line on a line through $(\mu_1, \mu_2, \dots, \mu_p)$ with direction numbers λ_i , (where the μ 's and λ 's are not specified). We can denote the hypothesis as

$\omega: \|\sigma_1^{ij}\| = \|\sigma_2^{ij}\| = \dots = \|\sigma_m^{ij}\| = \|\sigma^{ij}\|$ positive definite;

$$-\infty < \mu_i < +\infty$$

$$i = 1, 2, \dots, p,$$

$$-\infty < \lambda_i < +\infty$$

$$i = 1, 2, \dots, p,$$

$$-\infty < x_g < +\infty$$

$$g = 1, 2, \dots, m,$$

with

$$(87) \quad \sum_{q=1}^m n^q \gamma^q = 0$$

and

$$(88) \quad \sum_{q=1}^m n^q (\gamma^q)^2 = 1.$$

The number of observations in the q -th sample is n^q . The probability density function of the q -th population is

$$\frac{|\sigma_{ij}^q|^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}} e^{-\frac{1}{2} \sum_{i,j=1}^p \sigma_{ij}^q (\chi_i^q - \mu_i - \lambda_i \gamma^q)(\chi_j^q - \mu_j - \lambda_j \gamma^q)}$$

In other words, the hypothesis is

$$\mu_i^q = \mu_i + \lambda_i \gamma^q.$$

Let us consider m samples $\{x_{i\alpha}^q\}$ ($\alpha = 1, 2, \dots, n^q$) from the m populations. The associated probability density function is

$$(89) P_N = \frac{|\sigma_{ij}|^{\frac{1}{2}N}}{(2\pi)^{\frac{1}{2}PN}} e^{-\frac{1}{2} \sum_{i,j=1}^p \sum_{q=1}^m \sum_{\alpha=1}^{n^q} \sigma^{ij} (x_{i\alpha}^q - \mu_i^q)(x_{j\alpha}^q - \mu_j^q)}$$

where $N = \sum_{q=1}^m n^q$. We maximise this function with respect to the Ω set. Differentiating with respect to μ_i^q and setting the resulting expression equal to zero, we obtain

$$\frac{\partial P_N}{\partial \mu_i^q} = P_N \sum_{j=1}^p \sum_{\alpha=1}^{n^q} \sigma^{ij} (x_{j\alpha}^q - \mu_j^q) = 0, \quad \begin{matrix} i=1, 2, \dots, p, \\ q=1, 2, \dots, m. \end{matrix}$$

Since $||\sigma^{ij}||$ is the matrix of each set of p equations in the p unknowns $\sum_{\alpha=1}^{n^q} (x_{j\alpha}^q - \mu_j^q)$ and is non-singular, we

have

$$\sum_{\alpha=1}^{n^q} (\chi_{j\alpha}^q - \mu_j^q) = 0.$$

Hence, we choose $\hat{\mu}_j^q = \bar{\chi}_j^q = \frac{1}{n^q} \sum_{\alpha=1}^{n^q} \chi_{j\alpha}^q$.

Next we differentiate (89) with respect to σ_{ij}^{-1} , obtaining

$$\frac{\partial P_N}{\partial \sigma_{ij}} = P_N \left[N \sigma_{ij} - \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i^q)(\chi_{j\alpha}^q - \mu_j^q) \right] = 0, \quad i \neq j,$$

and

$$\frac{\partial P_N}{\partial \sigma_{ii}} = P_N \left[\frac{N}{2} \sigma_{ii} - \frac{1}{2} \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i^q)^2 \right] = 0,$$

where

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

Therefore, we choose as the maximum likelihood estimate of σ_{ij}

$$\hat{\sigma}_{ij} = \frac{1}{N} \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \hat{\mu}_i^q)(\chi_{j\alpha}^q - \hat{\mu}_j^q) = \frac{1}{N} a_{ij}, \quad \text{say.}$$

Then the maximum of the likelihood function (89) over Ω is

$$P_N(\Omega_{\max}) = \frac{|a_{ij}|^{-\frac{1}{2}N}}{(2\sigma)^{\frac{1}{2}PN}} N^{\frac{1}{2}PN} e^{-\frac{1}{2}PN}.$$

Now let us maximise (89) with respect to the w set. First we differentiate with respect to μ_1 and set the result equal to zero; that is,

$$\frac{\partial P_N}{\partial \mu_i} = P_N \sum_{q=1}^m \sum_{j=1}^p \sum_{\alpha=1}^{n^q} \sigma_{ij}^q (\chi_{j\alpha}^q - \mu_j - \gamma^q \lambda_i) = 0.$$

Since $|\sigma_{ij}^q| \neq 0$, we have

$$(90) \quad \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{j\alpha}^q - \mu_j - \gamma^q \lambda_j) = 0.$$

Since

$$\sum_{q=1}^m n^q \gamma^q = 0,$$

the summation (90) gives as the maximum likelihood estimate

$$\hat{\mu}_j = \bar{X}_j = \frac{1}{N} \sum_{q=1}^m \sum_{\alpha=1}^{n^q} \chi_{j\alpha}^q.$$

Next we differentiate with respect to λ_1 , obtaining,

$$\frac{\partial P_N}{\partial \lambda_i} = P_N \sum_{q=1}^m \sum_{j=1}^p \sum_{\alpha=1}^{n^q} \sigma_{ij}^q (\chi_{j\alpha}^q - \mu_j - \gamma^q \lambda_j) \gamma^q = 0.$$

Since $|\sigma^{1j}| \neq 0$, we have

$$\sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\gamma^q \chi_{j\alpha}^q - \gamma^q \mu_j - (\gamma^q)^2 \lambda_j) = 0.$$

In view of (87) and (88), we find

$$\hat{\lambda}_j = \sum_{q=1}^m \eta^q \gamma^q \bar{x}_j^q.$$

Next we differentiate P_N with respect to σ^{1j} , obtaining,

$$\frac{\partial P_N}{\partial \sigma_{ij}^1} = P_N \left[N \sigma_{ij}^1 - \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i) (\chi_{j\alpha}^q - \mu_j) \right] = 0, \quad i \neq j,$$

and

$$\frac{\partial P_N}{\partial \sigma_{ii}^1} = P_N \left[\frac{N}{2} \sigma_{ii}^1 - \frac{1}{2} \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i)^2 \right] = 0.$$

Hence, we choose

$$N \hat{\sigma}_{ij} = \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \bar{\chi}_i - \gamma^q \sum_{r=1}^m n^r \gamma^r \bar{\chi}_i^r) (\chi_{j\alpha}^q - \bar{\chi}_j - \gamma^q \sum_{r=1}^m n^r \gamma^r \bar{\chi}_j^r) \\ = \tilde{a}_{ij} - \left(\sum_{r=1}^m n^r \gamma^r \bar{\chi}_i^r \right) \left(\sum_{q=1}^m n^q \gamma^q \bar{\chi}_j^q \right),$$

where

$$\tilde{a}_{ij} = \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \bar{\chi}_i) (\chi_{j\alpha}^q - \bar{\chi}_j) = a_{ij} + b_{ij}$$

and

$$b_{ij} = \sum_{q=1}^m n^q (\bar{\chi}_i^q - \bar{\chi}_i) (\bar{\chi}_j^q - \bar{\chi}_j).$$

We note, incidentally, that for a given q $\sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \bar{\chi}_i^q) (\chi_{j\alpha}^q - \bar{\chi}_j^q)$ has a Wishart distribution with $n^q - 1$ degrees of freedom (for $n^q > p$). So $||a_{1j}||$ has a Wishart distribution with $N - m$ degrees of freedom.

Next we maximize P_N with respect to γ^r , that is, we

maximize

$$\frac{\left| \tilde{a}_{ij} - \left(\sum_{r=1}^m n^r \gamma^r \bar{x}_i^r \right) \left(\sum_{q=1}^m n^q \gamma^q \bar{x}_j^q \right) \right|^{\frac{1}{2}N}}{(2\pi)^{\frac{1}{2}PN}} N^{\frac{1}{2}PN} e^{-\frac{1}{2}PN}.$$

To find the $\hat{\gamma}^q$ we must minimise

$$\begin{aligned} & \left| \tilde{a}_{ij} - \left(\sum_{r=1}^m n^r \gamma^r \bar{x}_i^r \right) \left(\sum_{q=1}^m n^q \gamma^q \bar{x}_j^q \right) \right| \\ (91) \quad & = \left| \tilde{a}_{ij} \right| \cdot \left[1 - \sum_{i,j=1}^p \tilde{a}_{ij} \left(\sum_{r=1}^m n^r \gamma^r \bar{x}_i^r \right) \left(\sum_{q=1}^m n^q \gamma^q \bar{x}_j^q \right) \right], \end{aligned}$$

that is, we must maximize

$$Q = \sum_{i,j=1}^p \sum_{r,q=1}^m \tilde{a}_{ij} \bar{x}_i^r \bar{x}_j^q n^r n^q \gamma^r \gamma^q,$$

subject to the restrictions

$$\sum_{q=1}^m n^q \gamma^q = 0,$$

$$\sum_{q=1}^m n^q (\gamma^q)^2 = 1.$$

Let

$$f = Q - \nu \sum_{q=1}^m n^q \gamma^q - \mu \left(\sum_{q=1}^m n^q (\gamma^q)^2 - 1 \right),$$

where ν and μ are Lagrange multipliers. Differentiating f with respect to γ^r and setting the result equal to zero, we obtain

$$(92) \quad \frac{\partial f}{\partial \gamma^r} = 2 \sum_{i,j=1}^p \sum_{q=1}^m \tilde{a}^{ij} \tilde{\chi}_i^r \tilde{\chi}_j^q n^r n^q \gamma^q - \nu n^r - 2\mu n^r \gamma^r = 0.$$

Summing (92) with respect to r , in view of (87) we obtain

$$N \cdot 2 \sum_{i,j=1}^p \sum_{q=1}^m \tilde{a}^{ij} \bar{x}_i \bar{x}_j^q n^q \gamma^q - \nu N = 0$$

Hence, (92) can be written as

$$2 \sum_{i,j=1}^p \sum_{q=1}^m \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) \bar{x}_j^q n^r n^q \gamma^q - 2\mu n^r \gamma^r = 0,$$

or, in view of (87), dividing by 2, we have

$$(93) \sum_{i,j=1}^p \sum_{q=1}^m \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) n^r n^q \gamma^q - \mu n^r \gamma^r = 0.$$

Multiplying (93) by γ^p and summing, we obtain

$$\sum_{r,q=1}^m \sum_{i,j=1}^p \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) n^r \gamma^r n^q \gamma^q - \mu = 0.$$

Substituting this value of μ in (93) we obtain the set of equations for γ^q ,

$$(94) \quad \sum_{q=1}^m \left\{ \sum_{i,j=1}^p \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) \sqrt{n^r} \sqrt{n^q} \right\} (\sqrt{n^q} \gamma^q) - (\sqrt{n^r} \gamma^r) \left[\sum_{r,q=1}^m \left\{ \sum_{i,j=1}^p \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) \sqrt{n^r} \sqrt{n^q} \right\} (\sqrt{n^r} \gamma^r) (\sqrt{n^q} \gamma^q) \right]$$

For a solution $(\sqrt{n^r} \gamma^r)$ it is necessary that the determinant of the matrix of the coefficients be zero. Such a condition is simply that the quantity in brackets in (94) is a root of

$$(95) \quad \left| \sum_{i,j=1}^p \tilde{a}^{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) \sqrt{n^r} \sqrt{n^q} - \phi \delta^{rq} \right| = 0.$$

Since Q can be written as

$$\sum_{r,q=1}^m \left\{ \sum_{i,j=1}^p \tilde{a}_{ij} (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) \sqrt{n^r} \sqrt{n^q} \right\} (\sqrt{n^r} \delta^r) (\sqrt{n^q} \delta^q),$$

its maximum is equal to a root of (95). Because we want Q maximum, we take the largest root, say, ϕ_1 , for Q . However, the non-zero roots of (95) are also the non-zero roots of

$$(96) \quad \left| \sum_{q=1}^m n^q (\bar{x}_i^q - \bar{x}_i) (\bar{x}_j^q - \bar{x}_j) - \phi \tilde{a}_{ij} \right| = 0.$$

This fact is proved in the Lemma of (B). The equations (95) and (96) are similar to (111) and (113).

Equation (96) can be written as

$$(97) \quad |b_{ij} - \phi(a_{ij} + b_{ij})| = 0.$$

The likelihood ratio criterion is then (by use of (91))

$$\Lambda = \frac{P_N(\omega_{\max})}{P_N(\Omega_{\max})} = \frac{|a_{ij}|^{\frac{1}{2}N}}{\{ |a_{ij} + b_{ij}|(1 - \phi_1) \}^{\frac{1}{2}N}},$$

where ϕ_1 is the largest root of the equation (97). The test for collinearity of means may be taken as

$$\Lambda^{\frac{2}{N}} = \frac{|a_{ij}|}{|a_{ij} + b_{ij}|(1 - \phi_1)}.$$

Let $\theta = \frac{\phi}{1 - \phi}$

Then

$$\Lambda^{\frac{2}{N}} = \frac{|a_{ij}|(1 + \theta_1)}{|a_{ij} + b_{ij}|},$$

where θ_1 is the largest

root of

$$(98) \quad |b_{ij} - \theta a_{ij}| = 0,$$

There exists a non-singular matrix H , since $\|a_{ij}\|$ is positive definite, so

$$H \|a_{ij}\| H' = I, \quad \text{the identity, and}$$

$$H \|b_{ij}\| H' = \begin{vmatrix} \theta_1 & 0 & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \theta_p \end{vmatrix} \quad \text{where } \theta_1, \dots, \theta_p$$

are the roots in descending order of (98). The test function then is

$$\begin{aligned} \Lambda &= \left\{ \frac{|H| \cdot |a_{ij}| \cdot |H'| (1+\theta_1)}{|H| \cdot |a_{ij} + b_{ij}| \cdot |H'|} \right\}^{\frac{1}{2}N} \\ &= \frac{(1+\theta_1)^{\frac{1}{2}N}}{\begin{vmatrix} 1+\theta_1 & 0 & \dots & 0 \\ 0 & 1+\theta_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1+\theta_p \end{vmatrix}^{\frac{1}{2}N}} \\ &= \prod_{i=2}^p (1+\theta_i)^{-\frac{1}{2}N} \end{aligned}$$

B. Testing Higher Dimensionality.

Let us consider the general problem of testing on the basis of samples whether the means of m populations in p variates lie on an h dimensional hyperplane. The hypothesis to be tested is whether the mean of the i -th variate in the q -th population can be written as

$$\mu_i^q = \mu_i + \sum_{\beta=1}^h \gamma_{\beta}^q \lambda_{i\beta}, \quad i=1, 2, \dots, p; q=1, 2, \dots, m.$$

We form the likelihood ratio criterion defining the sets Ω and ω as follows:

$$\Omega: \|\sigma_1^q\| = \|\sigma_2^q\| = \dots = \|\sigma_m^q\| = \|\sigma^q\|, \text{ positive definite,}$$

$$-\infty < \mu_i^q < +\infty, \quad i=1, 2, \dots, p; q=1, 2, \dots, m.$$

$$\omega: \|\sigma_1^q\| = \|\sigma_2^q\| = \dots = \|\sigma_m^q\| = \|\sigma^q\|, \text{ positive definite,}$$

$$-\infty < \mu_i < +\infty, \quad i=1, 2, \dots, p,$$

$$-\infty < \lambda_{i\beta} < +\infty, \quad i=1, 2, \dots, p; \beta=1, 2, \dots, h;$$

$$-\infty < \gamma_{\beta}^q < +\infty, \quad q=1, 2, \dots, p; \beta=1, 2, \dots, h;$$

where

$$(100) \quad \sum_{q=1}^m n^q \gamma_{\beta}^q = 0,$$

$$\sum_{q=1}^m n^q \gamma_{\beta}^q \gamma_{\epsilon}^q = \delta_{\beta\epsilon},$$

and

$$(101) \quad \sum_{i,j=1}^p \sigma_{ij} \lambda_{i\beta} \lambda_{j\epsilon} = 0, \quad \beta \neq \epsilon.$$

The maximum of P_N over the Ω set is, as before,

$$(102) \quad P_N(\Omega_{\max}) = \frac{|a_{ij}|^{\frac{1}{2}N}}{(2\pi)^{\frac{1}{2}PN}} N^{\frac{1}{2}PN} e^{-\frac{1}{2}PN}.$$

To maximize with respect to ω we find $\hat{\mu}_i$, $\hat{\lambda}_{i\beta}$, $\hat{\sigma}_{ij}$, and γ_β^q , respectively. Differentiating with respect to μ_i , we have

$$\frac{\partial P_N}{\partial \mu_i} = P_N \sum_{q=1}^m \sum_{j=1}^p \sum_{\alpha=1}^{n^q} \sigma_{ij}^q (\chi_{j\alpha}^q - \mu_j - \sum_{\beta=1}^h \gamma_\beta^q \lambda_{i\beta}) = 0.$$

Therefore,

$$\sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{j\alpha}^q - \mu_j - \sum_{\beta=1}^h \gamma_\beta^q \lambda_{i\beta}) = 0.$$

Since $\sum_{q=1}^m n^q \gamma_\beta^q = 0$, we have

$$\text{and} \quad \sum_{q=1}^m \sum_{\alpha=1}^{n^q} \chi_{j\alpha}^q = N \mu_j,$$

(103)

$$\hat{\mu}_j = \bar{x}_j$$

Next, we differentiate with respect to $\lambda_{i\beta}$ and obtain

$$\frac{\partial P_N}{\partial \lambda_{i\beta}} = P_N \sum_{q=1}^m \sum_{j=1}^p \sum_{\alpha=1}^{n^q} \sigma^{ij} (\chi_{j\alpha}^q - \mu_j - \sum_{\epsilon=1}^h \gamma_{\epsilon}^q \lambda_{j\epsilon}) \gamma_{\beta}^q = 0.$$

Therefore,

$$\sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\gamma_{\beta}^q \chi_{j\alpha}^q - \gamma_{\beta}^q \mu_j - \gamma_{\beta}^q \sum_{\epsilon=1}^h \gamma_{\epsilon}^q \lambda_{j\epsilon}) = 0.$$

Since $\sum_{q=1}^m n^q \gamma_{\epsilon}^q \gamma_{\beta}^q = \delta_{\epsilon\beta}$, we have

$$\sum_{q=1}^m \sum_{\alpha=1}^{n^q} \gamma_{\beta}^q \chi_{j\alpha}^q = \sum_{q=1}^m n^q \gamma_{\beta}^q \gamma_{\beta}^q \lambda_{j\beta},$$

$$\hat{\lambda}_{j\beta} = \sum_{q=1}^m n^q \gamma_{\beta}^q \bar{\chi}_j^q.$$

It will be shown later that (101) is satisfied by $\hat{\lambda}_{j\beta}$ because of the properties of $\hat{\gamma}_{\beta}^q$.

Differentiating with respect to σ_{ij} , we obtain

$$\frac{\partial P_N}{\partial \sigma_{ij}} = P_N \left\{ N \sigma_{ij} - \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i^q)(\chi_{j\alpha}^q - \mu_j^q) \right\} = 0, \quad i \neq j$$

and

$$\frac{\partial P_N}{\partial \sigma_{ii}} = P_N \left\{ \frac{N}{2} \sigma_{ii} - \frac{1}{2} \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \mu_i^q)^2 \right\} = 0.$$

Therefore

$$\begin{aligned} N \hat{\sigma}_{ij} &= \sum_{q=1}^m \sum_{\alpha=1}^{n^q} (\chi_{i\alpha}^q - \bar{x}_i - \sum_{\beta=1}^h \gamma_{\beta}^q \chi_{\alpha}^q \sum_{r=1}^m n^r \gamma_{\beta}^r \bar{x}_i^r) (\chi_{j\alpha}^q - \bar{x}_j - \sum_{\beta=1}^h \gamma_{\beta}^q \chi_{\alpha}^q \sum_{r=1}^m n^r \gamma_{\beta}^r \bar{x}_j^r) \\ &= \tilde{a}_{ij} - \sum_{\beta=1}^h \left(\sum_{r=1}^m n^r \gamma_{\beta}^r \bar{x}_i^r \right) \left(\sum_{q=1}^m n^q \gamma_{\beta}^q \bar{x}_j^q \right) \\ &= a_{ij}, \text{ say.} \end{aligned}$$

$$a_{ij} + b_{ij} \quad -117-$$

$$\sum (\chi_{i\alpha}^q - \bar{x}_i^q)(\chi_{j\alpha}^q - \bar{x}_j^q) + \sum n^q (\bar{x}_i^q - \bar{x}_i)(\bar{x}_j^q - \bar{x}_j)$$

The maximum of P_N with respect to the γ_β^q under restriction (100) is the minimum of

$$f = |a'_{ij}| + \sum_{\beta=1}^h \nu_\beta \left(\sum_{q=1}^m n^q \gamma_\beta^q \right) + \sum_{\substack{\beta, \epsilon=1 \\ \beta \leq \epsilon}}^h \mu_{\beta\epsilon} \left(\sum_{q=1}^m n^q \gamma_\beta^q \gamma_\epsilon^q - \delta_{\beta\epsilon} \right),$$

where ν_β and $\mu_{\beta\epsilon}$ are Lagrange multipliers. Differentiating with respect to γ_β^q and setting the result equal to zero, we obtain

$$(104) \quad -2 \sum_{r=1}^m \sum_{i,j=1}^p A'_{ij} n^q \bar{x}_i^q n^r \bar{x}_j^r \gamma_\beta^r + n^q \nu_\beta + 2 \sum_{\epsilon=\beta}^h \mu_{\beta\epsilon} n^q \gamma_\epsilon^q = 0,$$

where A'_{ij} is the cofactor of a'_{ij} in $|a'_{ij}|$. Summing (104) with respect to q we find

$$-2N \sum_{r=1}^m \sum_{i,j=1}^p A'_{ij} \bar{x}_i^r \bar{x}_j^r n^r \gamma_\beta^r + N \nu_\beta = 0.$$

Substituting in (104), we obtain, noting that $\sum_{q=1}^m n^q \gamma_{\beta}^q = 0$,

$$-2 \sum_{r=1}^m \sum_{i,j=1}^p A'_{ij} (\bar{x}_i^q - \bar{x}_i) (\bar{x}_j^r - \bar{x}_j) n^q n^r \gamma_{\beta}^r + 2 \sum_{\epsilon=1}^h \mu_{\beta \epsilon} n^q \gamma_{\epsilon}^q = 0.$$

Thus the problem reduces to maximizing

$$(105) \left| \tilde{a}_{ij} - \sum_{\beta=1}^h \left[\sum_{r=1}^m n^r \gamma_{\beta}^r (\bar{x}_i^r - \bar{x}_i) \right] \left[\sum_{q=1}^m n^q \gamma_{\beta}^q (\bar{x}_j^q - \bar{x}_j) \right] \right|$$

under the restriction $\sum_{r=1}^m n^r \gamma_{\beta}^r \gamma_{\epsilon}^r = \delta_{\beta \epsilon}$.

We need the following lemma.

Lemma

$$\text{Let } d_{ij} = \sum_{r=1}^m \chi_i^r \chi_j^r \quad (i, j=1, 2, \dots, p),$$

where $||d_{ij}||$ is of rank $\min(m-1, p)$ and let

$$a_{ij} = c_{ij} + d_{ij}$$

where $||c_{ij}||$ is positive definite and symmetric. Then the maximum of

$$\phi = | a_{ij} - \sum_{\beta=1}^h \left(\sum_{r=1}^m \gamma_r^\beta \chi_{ir} \right) \left(\sum_{q=1}^m \gamma_q^\beta \chi_{jq} \right) |$$

($h \leq \min(m-1, p)$) over variations of the set $|\gamma_q^\beta|$ subject to the restrictions

$$(106) \quad \sum_{r=1}^m \gamma_r^\alpha \gamma_r^\beta = \delta^{\alpha\beta}$$

(where $\delta^{\alpha\beta}$ is the Kronecker delta) is

$$|a_{ij}| \prod_{\beta=1}^h (1 - \psi_\beta),$$

where the ψ_β are the h largest roots of

$$|d_{ij} - \psi_\beta a_{ij}| = 0.$$

Proof: Consider

$$\omega = \phi + \sum_{\substack{\alpha, \beta=1 \\ \alpha \leq \beta}}^h \mu_{\alpha\beta} \left(\sum_{r=1}^m \gamma_r^\alpha \gamma_r^\beta - \delta^{\alpha\beta} \right),$$

where the μ 's are Lagrange multipliers.

Differentiating ω with respect to γ_r^β and setting equal to zero we obtain

$$(107) \quad \frac{\partial \omega}{\partial \gamma_r^\beta} = -2 |b_{ij}| \sum_{q=1}^m \sum_{i,j=1}^p b^{ij} x_i^r x_q^\beta x_j^q + 2 \sum_{\alpha=1}^h \mu_{\alpha\beta} \gamma_r^\alpha = 0,$$

where b^{ij} is an element of the inverse of

$$\|b_{ij}\| = \left\| a_{ij} - \sum_{\beta=1}^h \left(\sum_{r=1}^m \gamma_r^\beta x_i^r \right) \left(\sum_{q=1}^m \gamma_q^\beta x_j^q \right) \right\|.$$

Multiplied (107) by γ_r^γ ($\gamma \leq \beta$) and summing on r , we obtain

$$-2 |b_{ij}| \sum_{q,r=1}^m \sum_{i,j=1}^p b^{ij} x_c^r \gamma_r^\beta x_j^q \gamma_q^\gamma + 2 \mu_{\gamma\beta} = 0 \quad (\gamma \leq \beta).$$

This equation can be solved for $\mu_{\gamma\beta}$.

Then (107) can be written as

$$(108) \quad \sum_{q=1}^m D^{rq} \gamma_q^\beta = \sum_{\alpha=1}^\beta \gamma^\alpha \left(\sum_{s,q=1}^m D^{sq} \gamma_s^\beta \gamma_q^\alpha \right),$$

where

$$D^{rq} = \sum_{i,j=1}^p b^{ij} x_c^r x_j^q,$$

and is of rank $\min(m-1, p)$. Multiplying (108) by y_r^γ
 ($\gamma > \beta$) and summing over r , we obtain

$$\sum_{r,q=1}^m D^{rq} y_q^\beta y_r^\gamma = \sum_{\alpha=1}^{\beta} \delta^{\gamma\alpha} \left(\sum_{s,q=1}^m D^{sq} y_s^\beta y_q^\alpha \right)$$

for $\gamma > \beta$ and for $\alpha \leq \beta$, that is

$$(109) \quad \sum_{r,q=1}^m D^{rq} y_q^\beta y_r^\gamma = 0$$

for $\gamma > \beta$. Since

$$\sum_{r,q=1}^m D^{rq} y_q^\beta y_r^\gamma = \sum_{r,q=1}^m D^{rq} y_q^\gamma y_r^\beta,$$

we have

$$\sum_{r,q=1}^m D^{rq} y_q^\beta y_r^\gamma = 0$$

for $\gamma \neq \beta$. Then (108) can be written as

$$(110) \quad \sum_{q=1}^m D^{rq} \gamma_q^\beta = \gamma_r^\beta \left(\sum_{s,q=1}^m D^{sq} \gamma_s^\beta \gamma_q^\beta \right).$$

It is clear that γ_r^β are the characteristic vectors of $||D^{rq}||$ and

$$\sum_{r,q=1}^m D^{rq} \gamma_r^\beta \gamma_q^\beta$$

is the β -th characteristic root of $||D^{rq}||$; i.e., root of

$$(111) \quad |D^{rq} - \lambda \delta^{rq}| = 0,$$

say λ^β . Then (109) can be written as

$$(112) \quad \sum_{q=1}^m D^{rq} \gamma_q^\beta = \lambda^\beta \gamma_r^\beta.$$

Next consider the pair of matrices $||b_{ij}||$ and $||\sum_{r=1}^m \chi_i^r \chi_j^r||$. Let $|w_\beta^i|$ be the i -th coordinate of the β -th vector associated with the equation

$$(113) \quad \left| \sum_{r=1}^m \chi_i^r \chi_j^r - \theta b_{ij} \right| = 0;$$

that is,

$$(114) \quad \theta_\beta \sum_{j=1}^p b_{ij} w_\beta^j = \sum_{j=1}^p \sum_{r=1}^m \chi_i^r \chi_j^r w_\beta^j,$$

where θ_β is a root of (113) (and $\sum_{i,j=1}^p b_{ij} w_\alpha^i w_\beta^j = \delta_{\alpha\beta}$).

Multiplying (114) by b^{ik} and summing with respect to i , we have

$$(115) \quad \theta_\beta w_\beta^k = \sum_{i,j=1}^k \sum_{r=1}^m b^{ik} \chi_i^r \chi_j^r w_\beta^j.$$

Multiplying (115) by χ_k^q and sum on k ; we have

$$(116) \quad \theta_\beta \left(\sum_{k=1}^p \chi_k^q w_\beta^k \right) = \sum_{r=1}^m \sum_{i,h=1}^p b^{ih} \chi_i^r \chi_h^q \left(\sum_{j=1}^p \chi_j^r w_\beta^j \right).$$

Comparing (116) with (112) we see that

$$(117) \quad \gamma_q^\beta = \sum_{k=1}^p \chi_k^q w_\beta^k$$

and

$$\lambda^\beta = \theta_\beta.$$

Hence, from the characteristic vectors and roots associated with (111) we can go to those of (113).

Next we show that the w_β^k are the vectors associated

with

$$(118) \quad \left| \sum_{r=1}^m x_i^r x_j^r - \psi a_{ij} \right| = 0.$$

From (114) we have that

$$\begin{aligned} \theta_Y \sum_{j=1}^p a_{ij} w_Y^j &= \theta_Y \sum_{j=1}^p \sum_{\beta=1}^h \left(\sum_{r=1}^m \gamma_r^\beta x_i^r \right) \left(\sum_{q=1}^m \gamma_q^\beta x_j^q \right) w_Y^j \\ &= \sum_{j=1}^p \sum_{r=1}^m x_i^r x_j^r w_Y^j ; \end{aligned}$$

that is, (in view of (117) and (106))

$$\theta_Y \sum_{j=1}^p a_{ij} w_Y^j = \theta_Y \sum_{\beta=1}^h \sum_{r=1}^m \gamma_r^\beta x_i^r \delta^{\beta Y} = \sum_{j=1}^p \sum_{r=1}^m x_i^r x_j^r w_Y^j.$$

Transposing (and letting $y_r^\beta = \sum_{j=1}^p x_j^r w_{\beta}^j$) we obtain

$$(119) \quad \frac{\theta_r}{1+\theta_r} \sum_{j=1}^p a_{ij} w_{\gamma}^j = \sum_{j=1}^p \sum_{r=1}^m x_i^r x_j^r w_{\gamma}^j.$$

Hence, we also have the result that the roots ψ_{β} of (118) are

$$\psi_{\beta} = \frac{\theta_{\beta}}{1+\theta_{\beta}} = \frac{\lambda^{\beta}}{1+\lambda^{\beta}}.$$

Hence, the λ^{β} can be defined in terms of a_{1j} and x_1^r as well as the y_r^{β} . Although the y_r are in the D^{rq} of (112), we nevertheless obtain a perfectly valid determination of the y_r^{β} . It remains to demonstrate the value of ϕ .

We have

$$\phi = \left| a_{ij} - \sum_{\beta=1}^h \left(\sum_{r=1}^m \sum_{k=1}^p x_i^r x_k^r w_{\beta}^h \right) \left(\sum_{q=1}^m \sum_{k=1}^p x_j^q x_k^q w_{\beta}^k \right) \right|$$

$$= \left| a_{ij} - \sum_{\beta=1}^h \frac{\left(\sum_{k=1}^p a_{ik} w_{\beta}^k \right) \left(\sum_{k=1}^p a_{jk} w_{\beta}^k \right)}{\left(\frac{1 + \lambda^{\beta}}{\lambda^{\beta}} \right)^2} \right|$$

by (119).

Let

$$(120) \quad a_{ij} = \sum_{k=1}^p f_{ik} f_{jk} \quad . \quad \text{Then}$$

$$(121) \quad \phi = |f_{ik}|^2 \cdot \left| \delta_{ij} - \sum_{\beta=1}^h \frac{\left(\sum_{k=1}^p f_{ik} w_{\beta}^k \right) \left(\sum_{k=1}^p f_{jk} w_{\beta}^k \right)}{\left(\frac{1 + \lambda^{\beta}}{\lambda^{\beta}} \right)^2} \right|.$$

However, $\|f_{1j}\|$ can be chosen so as to make

$$(122) \quad \left\| \frac{\sum_{k=1}^p f_{ki} w_{\beta}^k}{\frac{1 + \lambda^{\beta}}{\lambda^{\beta}}} \right\| = G, \text{ say } (i \text{ is the row index, } \beta \text{ is the column index),}$$

diagonal. Hence (121) is

$$(123) \quad \phi = |a_{ij}| \cdot |I - GG'|.$$

But the elements of $G'G$ must be the same as those of GG' except for zeros. From (122) we have

$$\begin{aligned} G'G &= \left\| \sum_{i,j,h=1}^p \frac{f_{hi} f_{ji} w_{\beta}^h w_{\alpha}^j}{\frac{1+\lambda^{\beta}}{\lambda^{\beta}} \cdot \frac{1+\lambda^{\alpha}}{\lambda^{\alpha}}} \right\| \\ &= \left\| \sum_{h,j=1}^p \frac{a_{hj} w_{\beta}^h w_{\alpha}^j}{\frac{1+\lambda^{\beta}}{\lambda^{\beta}} \cdot \frac{1+\lambda^{\alpha}}{\lambda^{\alpha}}} \right\| \end{aligned}$$

by (120). Using the fact that the w 's are characteristic vectors we have

$$(124) \quad G' G = \left\| \frac{\delta^{\alpha\beta}}{\frac{1+\lambda^\beta}{\lambda^\beta}} \right\|$$

Since G is diagonal $G G'$ must be diagonal with, say, the upper left hand h by h minor simply (124). Then we have

$$|I - G G'| = \prod_{\beta=1}^h \left(\frac{1}{1+\lambda^\beta} \right) = \prod_{\beta=1}^h (1 - \psi_\beta).$$

Hence, from (123)

$$\phi = |a_{ij}| \prod_{\beta=1}^h (1 - \psi_\beta).$$

Thus the lemma is proved.

It follows from applying the above lemma that the maximum of (105) is

$$|\tilde{a}_{ij}| = \prod_{i=1}^h (1 - \psi_i),$$

where ψ_i ($i=1, 2, \dots, h$) are the h largest roots of

$$(126) \quad \left| \sum_{r=1}^m (\bar{x}_i^r - \bar{x}_i)(\bar{x}_j^r - \bar{x}_j) n^r - \psi_i \tilde{a}_{ij} \right| = 0.$$

Now we can demonstrate that (101) is fulfilled.

Equation (109) written in terms of a_{ij} , $\sqrt{n^r} \hat{g}_\beta^r$
and $\sqrt{n^r} (\bar{x}_i^r - \bar{x}_i)$ is

$$(127) \quad \sum_{r,q=1}^m \sum_{i,j=1}^p a_{ij} (\bar{x}_i^r - \bar{x}_i)(\bar{x}_j^q - \bar{x}_j) n^r \hat{g}_\beta^r n^q \hat{g}_\epsilon^q = 0$$

for $\beta \neq \epsilon$. Since we can write $\hat{\lambda}_{i\beta}$ as

$$\hat{\lambda}_{i\beta} = \sum_{r=1}^m n^r \hat{g}_\beta^r (\bar{x}_i^r - \bar{x}_i),$$

equation (127) is

$$\sum_{i,j=1}^p a_{ij} \hat{\lambda}_{i\beta} \hat{\lambda}_{j\epsilon} = 0 = \int_{\beta \in \Theta_{\beta}}$$

for $\beta \neq \epsilon$. Because a^{1j}/N is the maximum likelihood estimate of σ^{1j} (i.e., $\hat{\sigma}^{1j} = a^{1j}/N$), the restrictions of (101) are satisfied.

We have, therefore, the maximum of P_N over ω , namely,

$$(128) \quad P_N(\omega_{max}) = \frac{|_{\omega} a_{ij}|^{-\frac{1}{2}N}}{(2\pi)^{\frac{1}{2}pN}} N^{\frac{1}{2}pN} e^{-\frac{1}{2}pN},$$

where $|_{\omega} a_{ij}|$ is simply (125). The likelihood ratio criterion is the ratio of (128) and (102), namely

$$\Lambda = \frac{P_N(\omega_{max})}{P_N(\Omega_{max})} = \frac{|a_{ij}|^{\frac{1}{2}N}}{\left\{ |a_{ij} + b_{ij}| \prod_{i=1}^n (1 - \phi_i) \right\}^{\frac{1}{2}N}}.$$

There is a transformation $\|c_{ij}\|$ such that

$$\|a_{ij} + b_{ij}\| = \|c_{ij}\| \cdot \|c_{ij}\|'$$

and

$$\|a_{ij}\| = \|c_{ij}\| \cdot \|(1 - \psi_i) \delta_{ij}\| \cdot \|c_{ij}\|'$$

Therefore we can use as our test criterion

$$\Lambda = \prod_{i=h+1}^p (1 - \psi_i)^{\frac{1}{2}N}$$

where the ψ_i are the $(p-h)$ smallest roots of (126).

In terms of the $(p-h)$ smallest roots θ_i of

$$(129) \quad |b_{ij} - \theta a_{ij}| = 0,$$

the criterion is

$$\lambda = \prod_{i=h+1}^p (1+\theta_i)^{-\frac{1}{2}N} \cdot \frac{|a_{ij}|}{|b_{ij}|}$$

One can give a geometric interpretation of this criterion. The matrix $\|a_{ij}\|$ represents the ellipsoid of the scatter of points about their means. It is an estimate of the population variance-covariance matrix. The matrix $\|b_{ij}\|$ represents the sample estimate of the ellipsoid of the scatter of the means of the populations. The largest root of (129) measures the greatest difference between these two ellipsoids: The h largest roots measure the greatest h dimensional difference between the two. That is, the h corresponding vectors span an h dimensional hyperplane; the area on this hyperplane between the intersections with the ellipsoids is a maximum (over all h dimensional hyperplanes). Under the null hypothesis the h roots should absorb most of the scatter due to the true means being on an h -dimensional hyperplane. The remaining $p-h$ roots

simply measure the scatter due to error. However, if the null hypothesis is not true, the $p-h$ smallest roots also include scatter due to the means. The test determines whether this scatter is due only to "error".

Let us summarize the results of Sections A and B.

Theorem 8..

Suppose a sample of n^r ($r=1,2,\dots,m$) observations is drawn from each of the m normal multivariate populations with probability density functions (86). Under the assumption that the variance-covariance matrix for each population is the same, the likelihood ratio criterion for testing the hypothesis that the means of the populations lie in an h dimensional hyperplane (against the alternative hypothesis of the means lying anywhere) is

$$\Lambda = \prod_{i=h+1}^p (1 + \theta_i)^{-\frac{1}{2}N},$$

where the θ_i ($i=h+1,\dots,p$) are the $p-h$ smallest roots (including zeros) of equation (129).

C. Fisher's Criteria for Testing Dimensionality.

R. A. Fisher [5] has intuitively obtained a test function for the hypothesis that the means lie on an h dimensional hyperplane. It can be readily shown (as Hsu has pointed out [9]) that Fisher's criterion is the sum of the $p-h$ smallest roots of

$$(130) \quad \left| \sum_{r=1}^m n^r (\bar{x}_i^r - \bar{x}_i)(\bar{x}_j^r - \bar{x}_j) - \lambda s_{ij} \right| = 0,$$

where

$$s_{ij} = \frac{1}{N-m} a_{ij}$$

Each set of characteristic vectors γ_{β}^r (in our notation) he calls "components" or "comparisons".

Fisher suggested that his criterion was approximately distributed as χ^2 with $(p-h)(m-1-h)$ degrees of freedom. Hsu [22] has proved that as N approaches infinity with

the ratios n^r/N constant the limiting distribution of the quantity

$$\sum_{i=h+1}^p \lambda_i$$

is the χ^2 distribution with $(p-h)(m-1-h)$ degrees of freedom.

D. Higher Dimensionality as an Alternative Hypothesis.

Suppose we wish to test the hypothesis that the means of the m populations lie on an f dimensional hyperplane assuming that they lie on a g ($f < g$) dimensional hyperplane. In this case the ω set is defined in (99) (100) and (101) with h equal to f . The Ω set is the same with h equal to g . Then $P_N(\omega \text{ max})$ is (128) with h equal to f and $P_N(\Omega \text{ max})$ is the same with h equal to g . The ratio is

$$\Lambda = \prod_{i=f+1}^g (1 + \theta_i)^{-\frac{1}{2}N}$$

where the θ_i are the roots of (129).

If one wishes to test whether the means lie at a point assuming they lie on a line, the test function is

simply

$$(1 + \theta_1)^{-\frac{1}{2} N}$$

If one wishes to test linearity assuming coplanarity, one uses

$$(1 + \theta_2)^{-\frac{1}{2} N}$$

as the criterion.

Fisher's intuitive approach would lead to the general test function (f dimension hyperplane against a g dimensional hyperplane) of

$$\sum_{i=f+1}^g \lambda_i$$

where the λ 's are roots of (130).

E. Maximum Likelihood Estimates of the Lines or Hyperplanes.

If one tests for collinearity and finds that the hypothesis is not contradicted, one may wish to estimate the line and the position of the mean points on it. In general one may wish to estimate the h dimensional hyper-

plane and the position of the mean points upon it. These estimates are the maximum likelihood estimates. First one obtains w_{β}^i for each β ($\beta = 1, 2, \dots, h$) as the coordinates of the characteristic vectors associated with the h largest roots of (129). Then (using (117)) we obtain the $\sqrt{n^r} \hat{\gamma}_{\beta}^r$ from the equations

$$\sqrt{n^r} \hat{\gamma}_{\beta}^r = \sum_{i=1}^p w_{\beta}^i \sqrt{n^r} (\bar{x}_i^r - \bar{x}_i).$$

The direction ratios of the lines spanning the h dimensional hyperplane in p dimensions are given by

$$\hat{\lambda}_{j\beta} = \sum_{r=1}^m n^r \hat{\gamma}_{\beta}^r \bar{x}_j^r$$

$$= \sum_{r=1}^m \sum_{i=1}^p n^r (\bar{x}_i^r - \bar{x}_i) (\bar{x}_j^r - \bar{x}_j) w_{\beta}^i$$

$$= \sum_{i=1}^p b_{ij} w_{\beta}^i$$

(not b_{ij} of lemma)

$$= \psi_{\beta} \sum_{i=1}^p a_{ij} w_{\beta}^i$$

Finally the $\hat{\mu}_j$ are given by (103). From these estimates $(\hat{\mu}_j, \hat{\lambda}_{j\beta}, \hat{\gamma}_{\beta}^r)$ one can estimate the mean μ_j^r of the j -th characteristic of the r -th population under the null hypothesis.

F. Asymptotic Distributions.

Hsu [9] has given the limiting distribution of the $\hat{\theta}_i$. Suppose there are h κ^2 's different from zero and suppose l_1 is the minimum of $m-1$ and p . Let the multiplicities of the κ^2 's be μ_s ($s=1, 2, \dots, \nu$). Then let

$$a_0 = 0, a_1 = \mu_1, a_s = \mu_1 + \dots + \mu_s \quad (s=1, 2, \dots, \nu),$$

$$a_\nu = h.$$

Let

$$\lambda_s = \kappa_s^2, \quad i = a_{s-1} + 1, \dots, a_s; \quad s=1, 2, \dots, \nu.$$

Now the λ 's are distinct ($\lambda_1 > \lambda_2 > \dots > \lambda_v$). Let

$$\xi_i = \frac{\sqrt{N}(\theta_i - \lambda_s)}{\sqrt{2\lambda_s^2 + 4\lambda_s}}, \quad i = a_{s-1} + 1, \dots, a_s; s = 1, 2, \dots, v;$$

$$\xi_i = N\theta_i, \quad i = h+1, \dots, l_1.$$

The remaining θ 's are zero. Define M and q^r so that

$n^r = M q^r$ ($r=1, 2, \dots, m$), where $N = M \sum_{r=1}^m q^r$. Then, as shown by Hsu, the limiting distribution of the ξ 's as $M \rightarrow \infty$ is represented by the density

$$D(\xi_1, \dots, \xi_{a_1}) \cdot D(\xi_{a_1+1}, \dots, \xi_{a_2}) \dots D(\xi_{a_{v-1}+1}, \dots, \xi_h)$$

(130)

$$\cdot D(\xi_{h+1}, \dots, \xi_{l_1}),$$

where

$$D(x_1, \dots, x_m) = \frac{e^{-\frac{1}{2} \sum_{i=1}^m x_i^2} \prod_{i=1}^p \prod_{j=i+1}^p (x_i - x_j)}{2^{\frac{1}{2}m} \prod_{i=1}^m \Gamma(\frac{1}{2}i)} \\ (+\infty > x_1 \geq x_2 \geq \dots \geq x_m > -\infty),$$

and

$$(131) D_1(\xi_{h+1}, \dots, \xi_{l_1}) = \frac{\pi^{\frac{1}{2}(l_1-h)} e^{-\frac{1}{2} \sum_{j=h+1}^{l_1} \xi_j} \prod_{j=h+1}^{l_1} \xi_j^{\frac{1}{2}(l_2-l_1-1)} \prod_{i=h+1}^{l_1} \prod_{j=i+1}^{l_1} (\xi_i - \xi_j)}{2^{\frac{1}{2}(p-h)(n_1-h)} \prod_{i=1}^{l_1-h} \Gamma(\frac{1}{2}[l_2-l-i]+1) \prod_{i=1}^{l_1-h} \Gamma(\frac{1}{2}i)},$$

where $n_1 = m-1$ and $l_2 = \max(m-1, k)$.

We shall now show that $-2 \log \Lambda$ is asymptotically distributed like χ^2 where

$$(132) \Lambda = \prod_{j=h+1}^{l_1} \left(1 + \frac{\xi_j}{N}\right)^{-\frac{1}{2}N}$$

(since the last $p-1$ θ_1 are zero). Let $F_N(\xi)$ be the cumulative distribution function of the ξ 's and let $F(\xi)$ be the cumulative of the probability density function (130). Then Hsu's theorem can be stated as

$$\lim F_N(\xi) = F(\xi)$$

Let

$$g_N(\xi) = e^{it(-2 \log 1)}$$

This can be written as

$$\begin{aligned} g_N(\xi) &= e^{it N \sum_{j=h+1}^{l_1} \log \left(1 + \frac{\xi_j}{N}\right)} \\ &= \prod_{j=h+1}^{l_1} e^{\log \left(1 + \frac{\xi_j}{N}\right) N i t} \\ &= \prod_{j=h+1}^{l_1} \left(1 + \frac{\xi_j}{N}\right)^{N i t} \end{aligned}$$

It is clear that

$$\lim_{N \rightarrow \infty} g_N(\xi) = \prod_{j=h+1}^{l_1} e^{i t \xi_j} = e^{i t \sum_{j=h+1}^{l_1} \xi_j} = g(\xi), \text{ say.}$$

The convergence is uniform in any finite region of the ξ 's.

Given an $\epsilon > 0$, there is a region R in the space of the ξ such that

$$\int_{\bar{R}} dF(\xi) < \frac{\epsilon}{2},$$

where $\bar{R}$ is the complement of R . From the convergence of the $F_N(\xi)$ we have for $N > N_0$ sufficiently large,

$$\left| \int_{\bar{R}} dF_N(\xi) - \int_{\bar{R}} dF(\xi) \right| < \frac{\epsilon}{2};$$

that is, for $N > N_0$, we have

$$\int_{\bar{R}} dF_N(\xi) < \epsilon.$$

Hence, we can confine our attention to R since with

$|g_N(\xi)| = 1$ the error may be made arbitrarily small.

Young [23] proves the following theorem for a real variable:

Suppose that $f_n(\chi)$ tends uniformly to a continuous $f(\chi)$ in (a,b) and that $g_n(\chi)$ tends to a limit $g(\chi)$ in an everywhere dense set of points E which includes a and b ; suppose further that the total variation of g_n is less than K , independent of n . Then

$$\int_a^b f_n d g_n \rightarrow \int_a^b f d g ,$$

provided that these integrals exist and that $g(\chi)$ is completed so as to be a function of bounded variation in (a,b) .

The theorem can easily be shown to hold in the case that χ is a vector. Since $e^{ia} = \cos a + i \sin a$, the integrations we make may all be broken up into two parts so that each integration is of a real variable. This theorem applied to our problem gives

$$\lim_{N \rightarrow \infty} \int_R g_N(\xi) d F_N(\xi) = \int_R g(\xi) d F(\xi).$$

It then follows that

$$\phi_N(t) = \int g_N(\xi) dF_N(\xi)$$

converges to

$$\phi(t) = \int g(\xi) dF(\xi)$$

for all t .

Since $dF(\xi)$ factors (according to (130)) and since each factor integrates to unity, $\phi(t)$ can be written as

$$\int_0^\infty \int_0^{\xi_{l-1}} e^{it \sum_{j=h+1}^{l-1} \xi_j} D_1(\xi_{h+1}, \dots, \xi_{l-1}) d\xi_{l-1} \dots d\xi_{h+1}.$$

This is the characteristic function of $\sum_{j=h+1}^{l-1} \xi_j$ where $D_1(\xi_j)$ is the distribution of the ξ_j ($j = h+1, \dots, l-1$). Since

Hsu [22] has shown that $\sum_{j=h+1}^{l_1} \xi_j$ has the χ^2 distribution with $r = (l_1 - h)(l_2 - h)$ degrees of freedom, we have

$$\phi(t) = (1 - 2it)^{-\frac{1}{2}(l_1 - h)(l_2 - h)}$$

Since $\phi(t)$ is continuous at $t = 0^*$, the distribution of $-2 \log \Lambda$ must converge to that of χ_r^2 . We have the following theorem:

Theorem 2.

The asymptotic distribution of $-2 \log \Lambda$, where Λ is the likelihood ratio criterion for testing on the basis of m samples whether the means of m normal multivariate populations in p variates with identical variance-covariance matrices lie on an h dimensional hyperplane, is the χ^2 distribution with $(p-h)(m-h-1)$ degrees of freedom.

The above demonstration indicates that Fisher's test function is equivalent to the likelihood ratio criterion in the limit. Hence, the asymptotic properties of the likelihood ratio criterion that apply to $-2 \log \Lambda$ also

* See Curtiss [24], for example.

apply to Fisher's criterion.

XI. DISTRIBUTION OF THE ROOTS OF CERTAIN DETERMINANTAL EQUATIONS

In this section we consider the problem of determining the distribution of the roots which are used in the likelihood ratio criteria. The distribution of the roots can be derived from the non-central Wishart distribution for the linear and planar cases. For the case of the means of the populations all being at one point, the distribution of the roots has been given by Hsu [7].

Let us put the problem in a general form: Given the symmetric matrix $||a_{ij}||$ whose elements follow a Wishart distribution with n degrees of freedom and of order p , ($p \leq n$), and with sigma matrix $||\sigma_{ij}||$ and given $||b_{ij}||$ whose elements follow a non-central Wishart distribution with m degrees of freedom and of order p ($p \leq m$), with sigma matrix $||\sigma_{ij}||$, and with sigma

matrix of means $||\tau_{ij}||$ of rank h , find the distribution of the roots of

$$(133) \quad |b_{ij} - \lambda a_{ij}| = 0.$$

The transformation $||\psi_{ij}||$ defined in (6) when applied to $||a_{ij}||$ and $||b_{ij}||$, leave the roots of (133) invariant, for

$$(134) \quad \left| \sum_{i,j=1}^p \psi_{hi} a_{ij} \psi_{kj} \lambda - \sum_{i,j=1}^p \psi_{hi} b_{ij} \psi_{kj} \right| \\ = |\psi_{ij}| \cdot |a_{ij} \lambda - b_{ij}| \cdot |\psi_{ij}|,$$

Hence, one loses no generality by assuming

$$||\sigma_{ij}|| = ||\delta_{ij}||,$$

$$\| \tau_{ij} \|_2 = \left\| \begin{pmatrix} K_1^2 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & K_2^2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & & & & \\ 0 & 0 & \dots & K_h^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \right\|_2.$$

We obtain the solution of this problem for $h=1$ and indicate how it is solved in any particular case for $h=2$.

The joint probability density function of a_{ij} and b_{ij} in the linear case ($h=1$) is

$$(135) K |a_{ij}|^{\frac{1}{2}(n-p-1)} |b_{ij}|^{\frac{1}{2}(n-p-1)} e^{-\frac{1}{2} \sum_{i=1}^p (a_{ii} + b_{ii})} \sum_{\alpha=0}^{\infty} \frac{(K_1^2 b_{11})^\alpha}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}n + \alpha)},$$

where

$$K = \frac{1}{2^{\frac{1}{2}p(n+m)} \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i]) \prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[m-i])}.$$

There is a matrix $|| w_{ij} ||$ such that

$$a_{ij} = \sum_{k=1}^p w_{ik} w_{jk},$$

$$b_{ij} = \sum_{k=1}^p w_{ik} \lambda_k w_{jk},$$

where λ_k is a root of (134). The Jacobian of this transformation (from $2 \cdot \frac{p(p+1)}{2}$ variables to $(p + p^2)$ variables) has been given by Hsu [7]. It is

$$2^p |w_{ij}|^{p+2} \prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j),$$

where the λ 's are numbered in descending order of magnitude.* The transformation is determined uniquely except that all the signs in one column may be changed. Hence, we require

$$0 \leq w_{ij} < \infty.$$

* The probability is zero that two roots are equal.

The resulting distribution is

$$K 2^p |w_{ij}|^{n+m-p} \prod_{i=1}^p \lambda_i^{\frac{1}{2}(m-p-1)} e^{-\frac{1}{2} \sum_{i,j=1}^p w_{ij}^2 (1+\lambda_j)}$$

$$\prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j) \sum_{\alpha=0}^{\infty} \frac{\left(K_1^2 \sum_{j=1}^p w_{ij}^2 \lambda_j \right)^{\alpha}}{2^{2\alpha} \alpha! \Gamma\left(\frac{1}{2}m + \alpha\right)}.$$

Let

$$\sqrt{1+\lambda_k} w_{ik} = u_{ik}.$$

The Jacobian is

$$\prod_{j=1}^p (1+\lambda_j)^{\frac{1}{2}p}.$$

The distribution of u 's and λ 's is

$$K 2^p |u_{ij}|^{n+m-p} \prod_{j=1}^p (1+\lambda_j)^{-\frac{1}{2}(n+m)} \prod_{j=1}^p \lambda_j^{\frac{1}{2}(m-p-1)} \quad (136)$$

$$e^{-\frac{1}{2} \sum_{i,j=1}^p u_{ij}^2} \prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j) \sum_{\alpha=0}^{\infty} \frac{\left(K_1^2 \sum_{j=1}^p u_{ij}^2 \frac{\lambda_j}{1+\lambda_j} \right)^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}m+\alpha)}.$$

Next we shall find the integral of (136) with respect to u_{ij} .

The determinant $|u_{ij}|$ is the volume spanned by the vectors $U_i (i_1, \dots, i_p)$. In terms of the lengths v_i and the angles ϕ_i (ϕ_i is the angle between the flat space of the first i vectors and the $i+1$ st vector) the determinant is

$$v_1 v_2 \dots v_p \sin \phi_1 \dots \sin \phi_{p-1}.$$

The problem of integrating (136) is equivalent to finding the $(n+m-p)/2$ -moment of $|\sum_{k=1}^p u_{ik} u_{jk}|$ given that the distribution of u_{ij} is

$$C e^{-\frac{1}{2} \sum_{i,j=1}^p u_{ij}^2} \sum_{\alpha=0}^{\infty} \frac{\left(\kappa^2 \sum_{j=1}^p u_{ij}^2 \frac{\lambda_j}{1+\lambda_j} \right)^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2} m + \alpha)},$$

where C is the proper normalizing constant. Using the argument of Section VII, C we have

$$E(|u_{ij}|^g) = E(v_1^g) \cdot E([v_2 \dots v_p \sin \phi_1 \dots \sin \phi_{p-1}]^g).$$

The same argument leading to (62) yields

$$\begin{aligned} (137) E([v_2 \dots v_p \sin \phi_1 \dots \sin \phi_{p-1}]^g) \\ = 2^{\frac{1}{2} g(p-1)} \frac{\prod_{i=2}^p \Gamma(\frac{1}{2} [p+1-i] + \frac{g}{2})}{\prod_{i=2}^p \Gamma(\frac{1}{2} [p+1-i])}. \end{aligned}$$

Dividing (137) by the normalizing constant (not involved in $E(v, \mathbf{g})$) we have

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (v_2 \dots v_p \sin \phi_1 \dots \sin \phi_{p-1})^g e^{-\frac{1}{2} \sum_{j=2}^p \sum_{i=1}^p u_{ij}^2} \prod_{i=2}^p \prod_{j=1}^p du_{ij}.$$

To obtain $E(v, \mathbf{g})$ we must integrate

$$(138) \quad C' \left(\sum_{j=1}^p u_{ij}^2 \right)^{\frac{1}{2}g} e^{-\frac{1}{2} \sum_{j=1}^p u_{ij}^2} \sum_{\alpha=0}^{\infty} \frac{\left(k_1^2 \sum_{j=1}^p u_{ij}^2 \frac{\lambda_j}{1+\lambda_j} \right)^{\alpha}}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}m+\alpha)} \prod_{j=1}^p du_{ij},$$

where C' is the normalizing constant ($0 \leq u_{ij} < \infty$).

Expand $\sum_{j=1}^p u_{ij}^2 \frac{\lambda_j}{1+\lambda_j}$ by the multinomial rule

and let $u_{ij}^2 = x_j$ ($du_{ij} = \frac{1}{2} x_j^{-\frac{1}{2}} dx_j$) . Then

(138) becomes

$$\frac{C'}{2^p} \left(\sum_{j=1}^p x_j \right)^{\frac{g}{2}} e^{-\frac{1}{2} \sum_{j=1}^p x_j} \sum_{\alpha_1, \dots, \alpha_p=0}^{\infty} \frac{(k_1^2)^{\sum_{j=1}^p \alpha_j} \prod_{j=1}^p \left(x_j \frac{\lambda_j}{1+\lambda_j} \right)^{\alpha_j} \prod_{j=1}^p (x_j^{-\frac{1}{2}} dx_j)}{2^{2 \sum \alpha_j} \Gamma(\frac{1}{2}m + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \alpha_j!}.$$

Let

$$\sum_{j=1}^p x_j = y$$

$$x_j = y y_j, \quad j = 2, 3, \dots, p.$$

Then the integration becomes

$$(139) \quad \frac{C'}{2^p} \int_0^1 \int_0^{1-\sum_{i=2}^{p-1} y_i} \int_0^\infty e^{-\frac{1}{2}y} \sum_{\alpha_1, \dots, \alpha_p=0}^\infty \left[\frac{(K_1^2)^{\sum_{j=1}^p \alpha_j} y^{\frac{p}{2} + \sum_{j=1}^p \alpha_j + \frac{1}{2}p-1}}{2^{2\sum \alpha_j} \Gamma(\frac{1}{2}m + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \alpha_j!} \right. \\ \left. \cdot \prod_{j=1}^p \left(\frac{\lambda_j}{1+\lambda_j} \right)^{\alpha_j} \prod_{j=2}^p y_j^{\alpha_j - \frac{1}{2}} \left(1 - \sum_{j=2}^p y_j \right)^{\alpha_1 - \frac{1}{2}} \right] dy dy_p \dots dy_2.$$

Using the integral formula for the gamma function and the Dirichlet integral, we find (139) to be

$$(140) \quad C' 2^{\frac{1}{2}(p-p)} \sum_{\alpha_1, \dots, \alpha_p=0}^\infty \left[\frac{(K_1^2)^{\sum_{j=1}^p \alpha_j} \Gamma(\frac{1}{2}[p+p] + \sum_{j=1}^p \alpha_j)}{2^{2\sum \alpha_j} \Gamma(\frac{1}{2}m + \sum_{j=1}^p \alpha_j)} \right. \\ \left. \cdot \frac{\prod_{j=1}^p \Gamma(\alpha_j + \frac{1}{2})}{\prod_{j=1}^p \alpha_j! \Gamma(\frac{1}{2}p + \sum_{j=1}^p \alpha_j)} \prod_{j=1}^p \left(\frac{\lambda_j}{1+\lambda_j} \right)^{\alpha_j} \right].$$

The equality between (138) and (140) is not disturbed by letting $C' = 1$. Hence, we have that ($g = n+m-p$)

$$\int |u_{ij}|^{n+m-p} e^{-\frac{1}{2} \sum_{j=1}^p u_{ij}^2} \sum_{\alpha \geq 0} \frac{(K_1^2 \sum_{j=1}^p u_{ij}^2 \frac{\lambda_j}{1+\lambda_j})^\alpha}{2^{2\alpha} \alpha! \Gamma(\frac{1}{2}m + \alpha)} \prod_{j=1}^p du_{ij}$$

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

$$\cdot \sum_{\alpha_1, \dots, \alpha_p \geq 0} \frac{(K_1^2)^{\sum_{j=1}^p \alpha_j} \Gamma(\frac{1}{2}[n+m] + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \Gamma(\alpha_j + \frac{1}{2})}{2^{\sum \alpha_j} \Gamma(\frac{1}{2}m + \sum_{j=1}^p \alpha_j) \Gamma(\frac{1}{2}p + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \alpha_j!} \prod_{j=1}^p \left(\frac{\lambda_j}{1+\lambda_j}\right)^{\alpha_j}.$$

Finally, the distribution of the λ 's for the case $h = 1$ is

$$(141) \quad \frac{e^{-\frac{1}{2}K_1^2} \prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[n+m-i]) \prod_{i=1}^p (1+\lambda_i)^{-\frac{1}{2}(n+m)} \prod_{i=1}^p \lambda_i^{\frac{1}{2}(m-p-1)}}{\prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[m-i]) \prod_{i=1}^p \Gamma(\frac{1}{2}[n+1-i]) \prod_{i=1}^{p-1} \Gamma(\frac{1}{2}[p-i])}$$

$$\cdot \prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j) \sum_{\alpha_1, \dots, \alpha_p \geq 0} \frac{(K_1^2)^{\sum_{j=1}^p \alpha_j} \Gamma(\frac{1}{2}[n+m] + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \Gamma(\alpha_j + \frac{1}{2})}{2^{\sum \alpha_j} \Gamma(\frac{1}{2}m + \sum_{j=1}^p \alpha_j) \Gamma(\frac{1}{2}p + \sum_{j=1}^p \alpha_j) \prod_{j=1}^p \alpha_j!} \prod_{j=1}^p \left(\frac{\lambda_j}{1+\lambda_j}\right)^{\alpha_j}.$$

From this distribution one can in any special instance find the probability that the likelihood ratio criterion for collinearity does not exceed a certain value by integration of (141) over the proper range. Essentially the same distribution has been obtained by Roy [8] by geometric methods. Roy, however, did not evaluate the constant.

The result given by Roy is claimed by him to hold for the general case, that is, for h K 's different from 0, with $1 \leq h \leq p$. His distribution has only one population parameter, namely, $\sum_{i=1}^h K_i^2$. It can easily be shown, however, that Roy's result does not hold in the generality he claims. For if n and m approach infinity the roots properly normalized (i.e., multiplied by $\frac{n+m}{m}$) must converge stochastically to the population roots. The distribution of Roy has only one population parameter appearing; hence for $h \geq 2$ his distribution does not give proper limiting results. This argument shows that Roy's results hold only for $h = 1$. The flaw in Roy's argument seems to be in the geometric reasoning. The correct result for $h = 2$ is indicated in the

next paragraph.

We shall indicate how the distribution of these roots in the planar case ($h = 2$) may be found. Expression (135) is replaced by the following probability density function of a_{ij} and b_{ij} :

$$K_2 |a_{ij}|^{\frac{1}{2}(n-p-1)} |b_{ij}|^{\frac{1}{2}(m-p-1)} e^{-\frac{1}{2} \sum_{i=1}^p (a_{ii} + b_{ii})}$$

$$\sum_{\alpha, \beta=0}^{\infty} \frac{[K_1^2 K_2^2 (b_{11} b_{22} - b_{12}^2)]^{\alpha} (K_1^2 b_{11} + K_2^2 b_{22})^{\beta}}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[n-1] + \alpha) \Gamma(\frac{1}{2}m + 2\alpha + \beta)},$$

where

$$K_2 = \frac{e^{-\frac{1}{2}(K_1^2 + K_2^2)}}{2^{\frac{1}{2}p(n+m)} \pi^{\frac{1}{2}p(p-1)} \prod_{i=2}^p \Gamma(\frac{1}{2}[n+1-i]) \prod_{i=3}^p \Gamma(\frac{1}{2}[m+1-i])}$$

Then (136) is replaced by the following joint distribution of u_{ij} and λ_i :

$$(142) K_2 2^p \prod_{i=1}^p (1+\lambda_i)^{-\frac{1}{2}(n+m)} \prod_{i=1}^p \lambda_i^{\frac{1}{2}(m-p-1)} \prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j)$$

$$\cdot \left[|u_{ij}|^{n+m-p} e^{-\frac{1}{2} \sum_{i,j=1}^p u_{ij}^2} \sum_{\alpha, \beta=0}^{\infty} \left\{ \frac{[K_1^2 K_2^2 \{ \sum_{i,j=1}^p u_{ij}^2 \xi_j \sum_{i,j=1}^p u_{ij}^2 \xi_j \}}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma(\frac{1}{2}[m-1]+\alpha)} - \left(\sum_{j=1}^p u_{1j} u_{2j} \xi_j \right)^2 \}^{\alpha} (K_1^2 \sum_{i,j=1}^p u_{ij} \xi_j + K_2^2 \sum_{i,j=1}^p u_{2j} \xi_j)^{\beta} \right\} \right] \frac{1}{\Gamma(\frac{1}{2}m + 2\alpha + \beta)}$$

where

$$\xi_j = \frac{\lambda_j}{1+\lambda_j}.$$

We now integrate what is within the large brackets of (142) in two steps. First of all we have

$$(143) \int (v_3 \dots v_p \sin \phi_2 \dots \sin \phi_{p-1})^{n+m-p} e^{-\frac{1}{2} \sum_{i=3}^p \sum_{j=1}^p u_{ij}^2} \prod_{i=3}^p \prod_{j=1}^p du_{ij} \\ = (2\pi)^{\frac{1}{2}p(p-2)} 2^{\frac{1}{2}(n+m-p)(p-2)} \frac{\prod_{i=3}^p \Gamma(\frac{1}{2}[n+m+1-i])}{\prod_{i=3}^p \Gamma(\frac{1}{2}[p+1-i])}$$

by the same reasoning leading to equation (137). Since

$$V_1^2 V_2^2 \sin^2 \phi = \begin{vmatrix} \sum_{j=1}^p u_{1j}^2 & \sum_{j=1}^p u_{1j} u_{2j} \\ \sum_{j=1}^p u_{1j} u_{2j} & \sum_{j=1}^p u_{2j}^2 \end{vmatrix},$$

we must finally integrate ($0 \leq u_{1j} < \infty$, $-\infty < u_{2j} < \infty$)

$$e^{-\frac{1}{2} \left(\sum_{j=1}^p u_{1j}^2 + \sum_{j=1}^p u_{2j}^2 \right)} \left[\sum_{j=1}^p u_{1j}^2 \sum_{j=1}^p u_{2j}^2 - \left(\sum_{j=1}^p u_{1j} u_{2j} \right)^2 \right]^{\frac{1}{2} (n+m-p)}$$

(144)

$$\sum_{\alpha, \beta=0}^{\infty} \frac{(K_1^2 K_2^2)^{\alpha} \left[\sum_{j=1}^p u_{1j}^2 \xi_j \sum_{j=1}^p u_{2j}^2 \xi_j - \left(\sum_{j=1}^p u_{1j} u_{2j} \xi_j \right)^2 \right]^{\alpha} \left[K_1^2 \sum_{j=1}^p u_{1j}^2 \xi_j + K_2^2 \sum_{j=1}^p u_{2j}^2 \xi_j \right]^{\beta}}{2^{4\alpha+2\beta} \alpha! \beta! \Gamma\left(\frac{1}{2} [m-1] + \alpha\right) \Gamma\left(\frac{1}{2} m + 2\alpha + \beta\right)},$$

The distribution of the λ 's is the product of (143),

$$K_2 2^p \prod_{i=1}^p (1 + \lambda_i)^{-\frac{1}{2} (n+m)} \prod_{i=1}^p \lambda_i^{\frac{1}{2} (m+p-1)} \prod_{i=1}^p \prod_{j=i+1}^p (\lambda_i - \lambda_j)$$

and the integral of (144). In general the integration is too complicated to write down in a convenient form. In any particular case the integration could be carried through to obtain the desired distribution. The limiting distribution in this planar case ($h = 2$), as well as in general, has been obtained by Hsu and is stated in Section X, F.

XII. SUMMARY

In the present paper the non-central Wishart distribution is derived for the rank one and rank two cases. Knowledge of this distribution is applied to such problems of multivariate statistics as finding the moments of the generalized variance and the moments of the criterion for testing linear hypotheses in the non-central cases. From the non-central Wishart distribution the distribution of the roots of certain determinantal equations is derived. Likelihood ratio criteria are developed for testing dimensionality of population means.

The non-central Wishart distribution is the joint distribution of the sums of squares and cross-products of deviations (from the sample means) of observations from normal multivariate distributions, the expected values of which are not identical from observation to observation, but the population variances and covariances are. The non-central Wishart distributions are given explicitly for the cases of the means lying on a line

or on a plane, when considered as points in the space of the variates of the populations. In the linear case the distribution is essentially a Wishart distribution multiplied by a Bessel function; in the planar case the distribution is a Wishart distribution multiplied by an infinite series of Bessel functions (Theorem 1). In general the non-central Wishart distribution is the product of a Wishart distribution and a function of the roots of a determinantal equation involving the matrix of squares and cross-products and the matrix of population means (Theorem 4). The characteristic function of the distribution is given (Theorem 2). It is shown that the convolution of two non-central Wishart distributions with the same variance-covariance matrices is a non-central Wishart distribution (Theorem 3).

The same ^{p1} generalized variance is defined as the determinant of the sample variances and covariances. The moments of the generalized variance in the cases where the means of the observations lie on a line or on a plane are derived from the corresponding non-central Wishart distribution^s. In the linear case the moments

are an infinite series involving the variance of the population means which is the root of a determinantal equation (Theorem 5). In the planar case the moments are a triple infinite series involving the two non-zero roots of a certain population determinantal equation (Theorem 6).

The likelihood ratio criterion for testing the hypothesis that the means of a set of populations are identical, given that the population variances and covariances are the same, is a ratio of two determinants. One is the determinant of variates distributed according to a Wishart distribution and the other is the determinant of the same matrix augmented by a set of non-centrally distributed variates such that the new elements have a non-central Wishart distribution when the hypothesis is false. The moments of this criterion derived from knowledge of the non-central Wishart distribution are infinite series when the means are on a line and triple infinite series when the means are on a plane (Theorem 7). The asymptotic distribution of the criterion when the null hypothesis is false is

a non-central χ^2 distribution.

The likelihood ratio criterion is developed for testing on the basis of a set of m samples, each drawn from a normal multivariate population, the hypothesis that the means of the set of populations in p variates lie on an h dimensional hyperplane given that the matrices of variances and covariances are identical. The criterion is a power of the product of 1 plus the $p-h$ smallest roots of the determinantal equation involving the sample estimate of the matrix of variances and covariances and the sums of squares and cross-products of deviations of sample means (Theorem 8).

no P The maximum likelihood estimates of the hyperplanes and the positions of the population means on the planes are given. The asymptotic distribution of the criterion is shown to be the χ^2 -distribution (Theorem 9).

The distribution of the roots involved in the test of dimensionality can be derived in the cases of the population means being on a line or a plane. The distribution is given explicitly for the linear case. For the planar case, the distribution is indicated as

a certain definite integral.

The author wishes to express his appreciation to Professor S. S. Wilks who has been of great help in many ways in writing this dissertation. Conversations with Professors S. Bochner, S. S. Chern, and J. W. Tukey have also assisted the author in the solution of special problems in this paper.

REFERENCES

- [1] R. A. FISHER, "The general sampling distribution of the multiple correlation coefficient," Proc. Roy. Soc. A., Vol. 121 (1928), pp. 654-673.
- [2] S. S. WILKS, "Certain generalizations in the analysis of variance," Biometrika, Vol. 24 (1932), pp. 471-94.
- [3] P. L. HSU, "Canonical reduction of the general regression problem," Ann. of Eug. Vol. 11, (1941), pp. 42-46.
- [4] P. L. HSU, "On generalized analysis of variance (I)," Biometrika 31 ('40), pp. 221-237.
- [5] R. A. FISHER, "The statistical utilization of multiple measurements," Ann. of Eug., Vol. 8 (1938), pp. 376-386.
- [6] R. A. FISHER, "The sampling distribution of some statistics obtained from non-linear equations," Ann. of Eug., Vol. 9 (1939), pp. 238-249.
- [7] P. L. HSU, "On the distribution of roots of certain determinantal equations," Ann. of Eug., Vol. 9, (1939), pp. 250-258.
- [8] S. N. ROY, "Analysis of variance for multi-variate populations: the sampling distribution of the requisite p-statistics on the null and non-null hypothesis," Sankhya, Vol. 6, (1941), pp. 35-50.
- [9] P. L. HSU, "On the limiting distribution of roots of a determinantal equation," Jour. London Math. Soc., Vol. 16 (1941) pp. 183-194.
- [10] H. HOTELLING, "Relations between two sets of variates," Biometrika, Vol. 28 (1936) pp. 336-337.

- [11] M. BOCHER, Introduction to Higher Algebra, (1907), MacMillan Company, New York, pp. 163-170.
- [12] S. S. WILKS, Mathematical Statistics. P. U. Press
- [13] W. G. MADOW, "Contributions to the theory of multivariate statistical analysis," Trans. of the Amer. Math. Soc., Vol. 44 (1938), p. 476.
- [14] E. GOURSAT (E. R. HEDRICK), A Course in Mathematical Analysis, (1904), Ginn and Company, Cambridge, Mass., p. 308.
- [15] G. N. WATSON, A Treatise on the Theory of Bessel Functions, (1944), MacMillan Company, New York.
- [16] J. WISHART and M. S. BARTLETT, "The generalized product moment distribution," Proc. Camb. Phil. Soc., Vol. 29 (1933), pp. 260-270.
- [17] S. S. WILKS, "On the dependence of K sets of normally distributed statistical variables," Econometrica, Vol. 3 (1935), pp. 309-326.
- [18] H. HOTELLING, "The generalization of Student's ratio," Ann. of Math. Stat., Vol. 2 (1933) pp. 360-378.
- [19] P. L. HSU, "Notes on Hotelling's generalized T," Ann. of Math. Stat., Vol. 9 (1938), pp. 231-243.
- [20] P. C. TANG, "The power function of the analysis of variance tests," Stat. Res. Mem. 2 (1938), pp. 126-149.
- [21] A. WALD, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Trans. of the Amer. Math. Soc., Vol. 54 (1943), pp. 426-482.

- [22] P. L. HSU, "On the problem of rank and the limiting distribution of Fisher's test function," Ann. of Eug., Vol. 11 (1941), pp. 39-41.
- [23] L. C. YOUNG, "Limits of Stieljes integrals," Jour. of the London Math. Soc., Vol. 9 (1934), pp. 119-126.
- [24] J. H. CURTISS, "A note on the theory of moment generating functions," Ann. of Math. Stat., Vol. 13 (1942), pp. 430-433.