ON THE CLASSIFICATION STATISTIC OF WALD

by

Mohammad Iqbal
University of North Carolina

This research was supported in part by the Office of Naval Research under Contract No. NR-042031 for research in probability and statistics at Chapel Hill. Reproduction for any purpose of the United States Government is permitted.

Institute of Statistics
Mimeograph Series No. 159
November 1956

ERRATA SHEET

Notation	92(3) will mean page 92, line 3 .
iv(18)	Replace $m = m_3$ by $w = \mid m_3 \mid$.
4(12)	Replace Z by z .
6(11)	Replace $\sqrt{\frac{N_1+N_2}{N_1N_2}}$ by $\sqrt{\frac{N_1N_2}{N_1+N_2}}$.
13(2)	Replace $m_1m_2 - m_3^2 \leq 0$ by $m_1m_2 - m_3^2 \geq 0$.
19(9)	Read $N(\varepsilon)$ as N_ε .
21(5)	Replace $(1- \frac{\alpha^2}{n}$ by $(1 - \frac{\alpha}{n})^2$.
21(6)	Replace 2β by $\alpha\beta$
22(1)	Read as $\frac{\partial^2 g}{\partial\alpha\partial\beta}$.
23(1)	From $-\frac{\alpha^2}{2n}$ to the end of line 2, is to be enclosed in square brackets.
24(6)	Put) after $\frac{h}{n^2}$.
27(4)	Put $\sim$ between I_1 and I_{11} .
29(18)	Replace $f(x)$ by $\mid f(x)\mid$ and $0 \leq c \leq \infty$ by $0 \leq c < \infty$.
48(7)	Insert a multiplier $\sqrt{2\pi}$ on the right .
51(5)	Replace 16 by 64 . (and this corresponds to $p = 2$.)
62(9),113(7)	Replace nm_3 by $\mid nm_3\mid$.
62(18), 92(3), 115(7)	Replace m_3 by $\mid m_3\mid$.
69(15)	Replace $\mid\phi(t)-\phi(t)\mid$ by $\mid\phi(t) - \widetilde{\phi(t)}\mid$.
70(14), 88(16)	Replace e^{-j} by e^{-v} .
81(9)	Put dv after $K_m(v)$.
85(7)	Read $\Psi(v) =$.
90(5)	Replace χ' by χ^2 .
103(5)	Read $\frac{1}{\eta}$ as $\frac{1}{n}$.
119(15)	Replace $V > nm_3$ by $\mid V\mid > \mid nm_3\mid$.
135(11)	Replace $e^{-\frac{1}{2}\chi^2}$ by $e^{-\frac{1}{2}\lambda^2}$.
139(18)	Replace suffices by suffixes .
140(6),(10)	Replace r by γ .
141(6)	Replace $\frac{\sigma^{ij}}{\sigma_{ij}^2}$ by $\frac{\partial^2 \sigma^{ij}}{\partial \sigma_{ij}^2}$
142(9)	Read π_i as π_1 and 145(7) Read $\sigma^{ij}_{z_i}$ as $\sigma^{ij} z_i$.

ACKNOWLEDGEMENTS

I wish to put on record my deep sense of gratitude to Professor Harold Hotelling for his inspiring guidance throughout the preparation of this work. I feel myself greatly honored on having had the privilege of working under his direction.

I am also greatly indebted to Professor R. C. Bose for the confidence derived through his encouragement at various stages.

My thanks are also due to the Fulbright Foundation in Pakistan, the Institute of International Education and the Office of Naval Research for their financial assistance which made this study possible.

The help of Mrs. Kattsoff, Mrs. Spencer and Mrs. Kiley for the careful typing of the manuscript is gratefully acknowledged.

Mohammad Iqbal

"No scientific investigation can be final;
it merely represents the most probable
conclusions which can be drawn from the
data at the disposal of the writer. A
wider range of facts or more refined
analysis, experiment and observation will
lead to new formulae and new theories.
This is the essence of scientific progress."

Karl Pearson 1898

TABLE OF CONTENTS

CHAPTER		PAGE
	ACKNOWLEDGEMENT	ii
	INTRODUCTION	vii
I.	A PROBLEM OF CLASSIFICATION CONSIDERED BY WALD	1
	1. Introduction	1
	2. Statement of the problem	3
	3. An example of its importance	5
	4. The statistic proposed by Wald	5
	5. Further work on the problem	9
II.	ON AN ASYMPTOTIC EVALUATION OF A TRIPLE INTEGRAL . . .	12
	1. Introduction	12
	2. The integral and its domain	12
	3. Order of the variables m_1, m_2 and m_3	14
	4. An important limit	18
	5. A triple integral	25
	6. The integral over D_1, an asymptotic approximation	29
	7. An upper bound to error	34
	8. The integral over D_2	43
	9. An upper bound to the value of I_2	48
	10. Comparison of I_1 and I_2	51
	11. The integral over the domain D^*	53
	12. Summary of Chapter II	57

CHAPTER PAGE

III. ON THE ASYMPTOTIC DISTRIBUTION OF WALD'S CLASSIFICATION STATISTIC 60

1. Introduction 60
2. Wald's approximate classification statistic and its moments 61
3. The asymptotic distribution of v for $p = 2m$. 66
4. An integral equation due to Wilks 75
5. A note on Bessel functions 76
6. Distribution of v for odd values of p 79
7. The use of a differential equation in the evaluation of an integral 81
8. The asymptotic distribution of v for even and odd values of p 82
9. Note on the construction of tables 88
10. Summary of Chapter III 91

IV. AN ASYMPTOTIC SERIES EXPANSION FOR THE DISTRIBUTION OF $m = m_3$ 92

1. Introduction 92
2. An asymptotic series for the distribution 93
3. The constant of integration for the first approximation 108
4. The tail areas for the first approximation . . . 110
5. Comparison with the results of Chapter III . . . 113
6. Summary of Chapter IV 115

CHAPTER PAGE

V. THE APPLICATION OF TCHEBYCHEFF-MARKOFF INEQUALITIES TO A SPECIAL CASE 116

1. Introduction 116
2. The integral over D_1 116
3. The integral over D_2 117
4. The integral over D 119
5. Moments of V 119
6. Some results due to Tschebycheff and Markoff . . 120
7. Application of Tschebycheff-Markoff theorems to this case 123

VI. NON-NULL CASE 128

1. Introduction 128
2. The joint distribution 129
3. Note on confluent hypergeometric functions . . . 130
4. An asymptotic form of $f(m_1, m_2, m_3)$ 132
5. Distribution of U for large n and $p = 1$, an independent approach 133
6. The asymptotic mean and variance of the statistic U 137
7. Correction term for the variance of the linear discriminant function 145

VII. SOME RELATED UNSOLVED PROBLEMS 148

1. On classification statistics of Wald and Anderson 148
2. The quadratic discriminators 148
3. Possibility of a different approach 150

CHAPTER PAGE

4. Efficiency 150

5. The greater mean vector 151

BIBLIOGRAPHY 152

INTRODUCTION[1]

In his paper "On a statistical problem arising in the classification of an individual into one of two groups" [50][2], the late Professor Abraham Wald made an attempt to put the theory of discriminant functions on rigorous mathematical foundations. He demonstrated by using very ingenious geometrical arguments spread out over several lemmas that a function $V = nm_3 / [(1-m_1)(1-m_2)-m_3^2]$ can be taken as the classification statistic instead of $\sqrt{\frac{N_1 N_2}{N_1+N_2}}\, U$, where

$U = \sum_{i=1}^{p} \sum_{j=1}^{p} s^{ij} z_i (\bar{y}_j - \bar{x}_j)$ is the usual discriminant function with the population parameters replaced by their sample estimates, and N_1 and N_2 are the sizes of the two samples from the two p-variate normal populations, and $n = N_1 + N_2 - 2$. Wald also obtained the joint distribution of m_1, m_2 and m_3 - $f(m_1, m_2, m_3)$.

It would be desirable to obtain, in a usable form, the distribution of V from $f(m_1, m_2, m_3)$. Such a simplification appears extremely difficult. It is related to the problem of the non-central Wishart's distribution for which T. W. Anderson and M. A. Girshick were able to obtain manageable expressions only for two or less variables. It seems that this general distribution of the discriminant function, or the classification statistic as Wald calls it,

---

[1]Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2]The numbers in square brackets refer to the bibliography listed at the end.

would involve the figurative distance, Δ, between the centers of the two populations. One approach to this highly involved distribution would be to obtain a series of powers of Δ with each coefficient involving n and p. The present work is chiefly concerned with the examination of the first term of this series with special attention to its value when n is large.

In the first chapter of the present work, a brief historical introduction to the theory of discriminant functions is followed by a mathematical formulation of the problem following Wald. The results obtained by him and also by some subsequent workers on the problem are briefly described.

The next two chapters deal with the problem of finding the distribution of V in the null case, by supposing that $n = N_1+N_2 - 2$ is a large number. Explicitly this problem can be stated as follows:

Given $f(m_1,m_2,m_3)dm_1 dm_2 dm_3 = \text{Const.} \; |M|^{\frac{p-3}{2}} |I-M|^{\frac{n-p-1}{2}} dm_1 dm_2 dm_3$

where $M = \begin{bmatrix} m_1 & m_3 \\ m_3 & m_2 \end{bmatrix}$, to find the distribution of V suitable for large n.

It will be noticed that the sample sizes, N_1 and N_2, do not separately occur in the joint distribution and the assumption that n is large, which is obviously milder than the assumption that N_1 and N_2 are both large, introduces certain simplifications. One simplification that is obtained is that the statistic itself approximates nm_3

because of the order in probability of the variables entering into the distribution. The same assumption entails simplifications both in the integrand and in the domain of integration.

In the second chapter, which can be regarded as dealing with the mathematical aspects of the problem, methods have been developed which will enable us to evaluate triple integrals giving the moments of $v = |nm_3|$. An upper bound to the error in using these simplifications is also worked out which enables us to put reliance in the approximations in suitable cases.

The third chapter deals with the asymptotic distribution problem. After finding an expression for the kth moment of v, we obtain the asymptotic distribution of v both for even and for odd values of p. For even values of p, the uniqueness of the distribution, which is obtained by the help of its moment generating function, is also established. For odd values of p, use had to be made of an integral equation due to S. S. Wilks [55]; and, because of the fact that we are considering only the principal term in the kth moment of v, the uniqueness of the result cannot be guaranteed. This section is therefore presented on a heuristic basis and has to be left for further discussion and rigorization.

In Chapter IV we have obtained an asymptotic series for the distribution of $w = |m_3|$, which is proportional to v. This is done by observing that for a fixed w the range of integration for m_1 and m_2 is a lenticular region enclosed by two hyperbolic arcs in the plane,

w = a constant. Integration is carried out over this region by using suitable transformations, and the first three terms of the asymptotic series are obtained. For the first approximation, we have also discussed the method of finding the tail areas.

Chapter V deals with the special case $N_1 + N_2 = 20$, $p = 3$. In this case, the first seven moments of V are found, and use is made of the inequalities due to Tchebycheff and Markoff in setting up bounds on probabilities of the type $P(V \geq \xi)$. These limits are rather crude due to the fact that a small number of moments is being used. The example, however, illustrates one way of proceeding to discover something about an unknown distribution when its first few moments are known.

Chapter VI contains a few remarks on the non-null case. It starts with expressing the joint distribution of m_1, m_2 and m_3 discussed by Sitgreaves [45] in another form suitable for large n. This chapter also contains a brief discussion of the asymptotic distribution of U for $p = 1$. In the next section we exemplify the differential method by finding the mean and variance of U. The concluding section of this chapter deals with finding the variance of the linear discriminant function when the sampling fluctuations of the means are taken into account. In the last chapter are listed a few unsolved problems related to the problem of classification.

CHAPTER I

A PROBLEM OF CLASSIFICATION CONSIDERED BY WALD

1. Introduction.

The problem of classification is the problem of assigning an individual (or an element), on which a set of measurements is available, to one of several groups or populations. The problem admits a simple solution when the distributions of measurements in the alternative populations are completely known or what is the same thing as saying that the sizes of the samples available from the various populations, on the basis of which we have to make a decision, tend to infinity, so that the sample estimates of the parameters tend stochastically to their population values. If, however, the samples are not large, the problem becomes rather complicated.

Research in this area of Multivariate Analysis was started with his introduction of the linear discriminant function by Sir Ronald A. Fisher [10] in 1936. The linear discriminant function is

$D = \sum_{i=1}^{p} \ell_i z_i$, in which $z = (z_1 \dots z_p)$ is the new observation, and the coefficients ℓ_i, following Fisher, can be obtained by maximizing the square of the difference of the expectations of D in the two populations divided by the standard deviation of D. The linear discriminant function provides the best solution of the problem of classification provided that,

(1) The number of alternative populations is two,

(2) The form of the distributions in both populations is multivariate normal,

(3) The parameters are all known,

(4) The covariance matrices of the two populations are equal.

It may be remarked that Welch [53] observed that even without making any assumptions of normality or equality of covariance matrices, the problem of obtaining the best function to discriminate between two completely specified populations may be solved. He demonstrated that the desired function is simply the ratio of the two probability distributions, and the criterian level to which this function is referred is deducible either from Bayes' Theorem with given a priori probabilities or by the use of a lemma by Neyman and Pearson [32] when the errors for the two hypotheses are minimized in any given ratio. He proved that under the four assumptions stated above the function obtained in this manner is identical with the linear discriminant function.

Von Mises [31] considered the problem of classification when the number of populations is m, and showed how to subdivide the sample space into m parts so as to minimize the maximum error of misclassification.

Rao [39] gave explicit Bayes solutions with given a priori probabilities or ratios of errors for the alternative populations, and discussed the construction and use of doubtful regions and related problems.

In all these cases it is assumed that the distributions are completely specified. If, however, as will frequently be the case, one cannot justify the supposition that the distributions are completely known, and the only information at hand is what is contained in the samples available from various populations, we run into rather complicated distribution problems.

Wald [50] in 1944 set out to solve the problem of classification for the case of two alternative populations. Instead of using a distribution-free approach he simplified it further by introducing the following two restrictions:

(1) The form of the distributions is multivariate normal.

(2) The two populations have the same covariance matrix.

Though it would be desirable to solve the problem without making either of these assumptions, still one can argue that in many practical problems arising in numerous fields of scientific inquiry it is not unreasonable to make the two assumptions stated above.

In this chapter we propose to give a mathematical formulation of the problem, and to state the conclusions of Wald, and of subsequent workers on the problem.

2. Statement of the problem.

(2.1) Let
$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & \cdots & \cdots & x_{2p} \\ \vdots & & & \vdots \\ x_{N_1 1} & \cdots & \cdots & x_{N_1 p} \end{bmatrix}$$

$$(2.2) \qquad \text{and} \quad Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & \cdot & \cdots & y_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ y_{N_2 1} & \cdot & \cdots & y_{N_2 p} \end{bmatrix}$$

be two random samples from two variate normal populations π_x and π_y both having the same, though unknown, covariance matrix Σ, and different unknown mean vectors

$$\mu = (\mu_1 \; \mu_2 \; \cdots \; \mu_p) \qquad \text{and}$$

$$\nu = (\nu_1 \; \nu_2 \; \cdots \; \nu_p) \qquad \text{respectively.}$$

(2.3) Let $z = (z_1 \; z_2 \; \ldots \; z_p)$

be an observation on a new individual which is known to have come either from π_x or from π_y, but is distributed independently of both $x = (x_1 \ldots x_p)$ and $y = (y_1 \ldots y_p)$, the two sets of variates corresponding to π_x and π_y respectively.

On the basis of the information supplied by X, Y and Z the statistician has to make one of the following decisions,

(1) $d_1 : z \in \pi_x$,

(2) $d_2 : z \in \pi_y$,

such that if the probability of one type of misclassification is held fixed, the chance of second type of misclassification is minimum.

3. As an example of the importance of this problem we can consider a candidate applying for admission to an institution with certain test scores. He may have to be accepted or rejected depending on his chances of success or otherwise on the basis of the scores of candidates admitted in previous years.

4. The statistic proposed by Wald.

(4.1) Wald considered as classification statistic

$U = \sum_{i=1}^{p} \sum_{j=1}^{p} s^{ij} z_i(\bar{y}_j - \bar{x}_j)$ obtained by considering this problem as one in testing the hypothesis $H_x : z \in \pi_x$ against the alternative that

$$H_y : z \in \pi_y$$

and by replacing the population values of the parameters by their optimum estimates obtained from the samples in the statistic obtained by using the fundamental lemma of Neyman and Pearson. Thus

$$[s^{ij}] = [s_{ij}]^{-1}$$

where

$$(4.2) \qquad s_{ij} = \frac{\sum_{\alpha=1}^{N_1} (x_{i\alpha}-\bar{x}_i)(x_{j\alpha}-\bar{x}_j) + \sum_{\beta=1}^{N_2} (y_{i\beta}-\bar{y}_i)(y_{j\beta}-\bar{y}_j)}{N_1 + N_2 - 2}$$

and

$$(4.3) \qquad \bar{x}_i = \sum_{\alpha=1}^{N_1} x_{i\alpha} \Big/ N_1 \qquad \bar{y}_i = \sum_{\beta=1}^{N_2} y_{i\beta} \Big/ N_2 .$$

The statistic U can be rewritten as

(4.4) $$U = \sqrt{\frac{N_1+N_2}{N_1N_2}} \sum_i \sum_j s^{ij} z_i z_j^*$$

where

(4.5) $$z_j^* = (\bar{y}_j - \bar{x}_j) \sqrt{\frac{N_1N_2}{N_1+N_2}} \quad , \; j = 1, 2 \ldots, p.$$

and where

$$z = (z_1 \ldots z_p)$$

and

$$z^* = (z_1^* \ldots z_p^*)$$

are distributed independently of each other according to p-variate normal distributions with

$$E(z) = \begin{cases} \mu & \text{if} \quad z \in \Pi_x \\ \nu & \text{if} \quad z \in \Pi_y \end{cases}$$

and

$$E(z^*) = \sqrt{\frac{N_1+N_2}{N_1N_2}} (\nu - \mu) \quad ,$$

and with the same covariance matrix $[\sigma_{ij}]$.

Since the s_{ij} are distributed independently of the set $(z_1 \ldots z_p, z_1^* \ldots z_p^*)$, the distribution of U remains unchanged if we define

(4.6) $$s_{ij} = \sum_{\alpha=1}^{n} t_{i\alpha}^2 \Big/ n \quad ,$$

where

$$n = N_1 + N_2 - 2 \quad ;$$

and writing (s^{ij}) for $(s_{ij})^{-1}$, Wald observed that the distribution of U is the same as that of

$$(4.7) \qquad V = \sum_{i=1}^{p} \sum_{j=1}^{p} s^{ij} t_{i,n+1} \, t_{j,n+2}$$

where the probability element of $t_{i\alpha}$ is given by

$$(4.8) \qquad \frac{1}{(2\pi)^{\frac{p(n+2)}{2}}} \exp \Big[- \frac{1}{2} \Big\{ \sum_{i=1}^{p} \sum_{\alpha=1}^{n} t_{i\alpha}^2 + \sum_{i=1}^{p} (t_{i,n+1} - \quad _i)^2 + \sum_{i=1}^{p} (t_{i,n+2} - \quad _i)^2 \Big\} \Big] \prod_{i=1}^{p} \prod_{\alpha=1}^{n+2} dt_{i\alpha}$$

where

$$(4.9) \qquad \rho = (\rho_1 \, \rho_2 \; \ldots \; \rho_p)$$

$$\zeta = (\zeta_1, \zeta_2 \; \ldots \; \zeta_p)$$

are certain functions of μ_i, ν_i and σ_{ij}, $i, j = 1, 2 \ldots p$.

Here Wald introduced two sets of numbers $(u_1 \ldots u_{n+2})$ and $(v_1 \ldots v_{n+2})$ satisfying the relations

$$(4.10) \qquad \sum_{\alpha=1}^{n+2} u_\alpha^2 = \sum_{\alpha=1}^{n+2} v_\alpha^2 = 1 \quad \text{and} \quad \sum_{\alpha=1}^{n+2} u_\alpha v_\alpha = 0$$

and using a very ingenious geometrical argument, concluded that the distribution of V is the same as that of

$$\frac{nm_3}{(1-m_1)(1-m_2)-m_3^2} \tag{4.11}$$

where

$$m_1 = \sum_{\alpha=1}^{p} u_\alpha^2$$

$$m_2 = \sum_{\alpha=1}^{p} v_\alpha^2 \tag{4.12}$$

$$m_3 = \sum_{\alpha=1}^{p} u_\alpha v_\alpha$$

and the joint probability distribution of m_1, m_2 and m_3 is given by

$$f(m_1 m_2 m_3)dm_1 dm_2 dm_3 = \frac{B}{\sqrt{m_1 m_2(1-m_1)(1-m_2)}} F_p(m_1)F_p(m_2)\phi_p\left(\frac{m_3}{\sqrt{m_1 m_2}}\right) \tag{4.13}$$

$$F_{n+2-p}(1-m_1)F_{n+2-p}(1-m_2)\phi_{n+2-p}\left(\frac{-m_3}{\sqrt{(1-m_1)(1-m_2)}}\right)$$

$$e^{\frac{1}{2}(m_1 \sum_{i=1}^{p} \rho_i^2 + 2m_3 \sum_{i=1}^{p} \rho_i \zeta_i + m_2 \sum_{i=1}^{p} \zeta_i^2)} \quad .$$

$$E\left(\begin{vmatrix} r_{11} & \cdots & r_{1p} \\ \cdot & \cdots & \cdot \\ r_{p1} & \cdots & r_{pp} \end{vmatrix}\right)^{\frac{n+2-p}{2}} dm_1 dm_2 dm_3 \, .$$

where $r_{ij} = \sum_{\alpha=1}^{p} t_{i\alpha} t_{j\alpha}$ and B is the constant of integration, in the domain $0 \leq m_1 \leq 1$, $0 \leq m_2 \leq 1$, $-\sqrt{m_1 m_2} \leq m_3 \leq \sqrt{m_1 m_2}$, and zero otherwise;

where
$$F_k(t) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} t^{\frac{k-2}{2}} e^{-\frac{t}{2}} ,$$

and
$$\phi_k(t) = \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\, \Gamma(\frac{k-1}{2})} (1-t^2)^{\frac{k-3}{2}} .$$

5. Further work on the problem.

Anderson [1] considered the statistic

$$(5.1) \qquad W = \sum_i \sum_j s^{ij} z_i (\bar{y}_j - \bar{x}_j) - \frac{1}{2} \sum_i \sum_j s^{ij} (\bar{y}_j + \bar{x}_j)(\bar{y}_j - \bar{x}_j)$$

which is much like U, since it differs from U by terms independent of $z = (z_1 \, . \, . \, . \, z_p)$, the measurements observed on the new individual. He evaluated the expected value of the matrix of non-central Wishart variates occuring in the joint distribution of m_1, m_2 and m_3 in the special case when

$$(5.2) \qquad \begin{aligned} \rho &= k_1 \delta \\ \xi &= k_2 \delta \end{aligned} .$$

Sitgreaves [45] gave an analytic derivation of the distribution of W in the case considered by Anderson and also obtained exactly the constant of integration in the joint distribution of $m_1 m_2 m_3$. We shall refer to the following result from her paper in the next chapter.

$$(5.3) \qquad f(m_1 m_2 m_3)\, dm_1\, dm_2\, dm_3 =$$

$$\frac{\Gamma(\frac{n+1}{2}) e^{\frac{\lambda^2}{2}(k_1^2+k_2^2)} \left|M\right|^{\frac{p-3}{2}} \left|I - M\right|^{\frac{n-p-1}{2}}}{\Gamma(\frac{n-p+2}{2})\Gamma(\frac{n-p+1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{1}{2})}$$

$$\sum_{j=0}^{\infty} \frac{\Gamma(\frac{n+2}{2}+j)}{\Gamma(\frac{p}{2}+j)j!} \left(\frac{\lambda^2}{2}\right)^j \left(k_1^2 m_1 + 2k_1 k_2 m_3 + k_2^2 m_2\right)^j \qquad ,$$

where

$$0 \le m_1 \le 1$$

$$0 \le m_2 \le 1$$

$$\left|M\right| \ge 0 \quad \left|I - M\right| \ge 0$$

where

$$M = \begin{bmatrix} m_1 & m_3 \\ m_3 & m_2 \end{bmatrix}$$

and $\lambda = \delta' \Sigma^{-1}\delta$, and where $k_1 k_2$ are defined in (5.2).

Earlier Harter [18] in 1951 considered the joint distribution of m_1, m_2 and m_3 of (4.13) in the degenerate case $\rho_i = 0 = \zeta_i$, $i = 1, \ldots p$ and obtained the approximate distribution of m_3 in the special cases

(1-a) n even, p odd

(1-b) n even, p even .

The technique he used in deriving this distribution was essentially expanding the two binomials constituting the integrand in the joint distribution of m_1, m_2 and m_3 of (4.13) in the degenerate case, and integrating with respect to m_1 and then with respect to m_2. The number of terms in the distribution of m_3 thus obtained depends on n, which is not a small number in any practical situation. Moreover the solution thus obtained is not an asymptotic series in which the leading terms could be considered as approximating the true distribution for large n.

The latest paper in historical order of development of the theory of discrimination is that of Rao [40], in which he developed some general methods by using the ideas of sufficient statistics and fiducial probability distributions, by using which, the discrimination problem can be solved utilizing only the sample information. The distribution problems connected with the test criteria suggested in the paper have, however, yet to be tackled.

CHAPTER II

ON AN ASYMPTOTIC EVALUATION OF A TRIPLE INTEGRAL

1. Introduction.

The integral with which we shall be concerned in this chapter is the one obtained from the joint distribution of m_1, m_2 and m_3, given by Wald [50], by putting $\rho_i = 0 = \xi_i$, $i = 1,2,\ldots,p$. For the sake of convenience, we shall refer to this case as the Central Case or the Null Case. In this chapter, we shall find the value of the integral for large values of n, which is equal to $N_1 + N_2 - 2$, by introducing certain simplifications both in the integrand and in the domain of integration. Justifications shall be given for the simplifications introduced, and the final result shown to be a valid asymptotic approximation in the sense of Poincaré. Moreover, an upper bound for the error involved in the asymptotic approximation will be found.

2. The integral and its domain.

The triple integral to which we refer corresponds to

$$f(m_1 m_2 m_3)dm_1 dm_2 dm_3 = \begin{cases} (m_1 m_2 - m_3^2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}} dm_1 dm_2 dm_3 & \text{over } D \\ 0 & \text{elsewhere.} \end{cases} \tag{2.1}$$

where the domain D is defined by the following inequalities which insure a real, positive integrand in its interior.

$$(2.2) \qquad D: \begin{cases} 0 \le m_1 \le 1 \\ 0 \le m_2 \le 1 \\ m_1 m_2 - m_3^2 \le 0 \\ (1-m_1)(1-m_2) - m_3^2 \ge 0 \, . \end{cases}$$

The inequalities in (2.2) show that the domain is bounded by two right elliptical cones in three-dimensional space having vertices at (0,0,0) and (1,1,0) respectively and having a common base in the plane $m_1 + m_2 = 1$.

We define two other domains D_1 and D_2 as follows:

$$(2.3) \qquad D_1: \begin{cases} 0 \le m_1 \le 1 \\ 0 \le m_2 \le 1 \\ m_1 m_2 - m_3^2 \ge 0 \\ m_1 + m_2 \le 1 \end{cases} \qquad D_2: \begin{cases} 0 \le m_1 \le 1 \\ 0 \le m_2 \le 1 \\ (1-m_1)(1-m_2) - m_3^2 \ge 0 \\ m_1 + m_2 \ge 1 \, . \end{cases}$$

(2.4) Then it is easy to see that $D = D_1 + D_2$, except for the set of points lying on the plane $m_1 + m_2 = 1$, which are counted twice.

The truth of this statement can be seen easily by noticing that the regions defined by the two domains are the interiors of two cones, one obtainable from the other by a simple transformation and lying on opposite sides of the plane $m_1 + m_2 = 1$ in the space of three dimensions. Moreover, except for the points lying on the plane $m_1 + m_2 = 1$, the two regions are mutually exclusive because the point set corresponding to D_1 lies on the origin side, whereas the other corresponding to D_2 lies on the non-origin side. The fact that D_1 and D_2 between themselves include all the points of D can be seen by observing that

$$(2.5) \qquad \begin{matrix} m_1m_2 - m_3^2 \geq 0, \\ m_1 + m_2 \leq 1 \end{matrix} \quad \Longrightarrow \quad (1-m_1)(1-m_2) - m_3^2 \geq 0,$$

and

$$(2.6) \qquad \begin{matrix} (1-m_1)(1-m_2) - m_3^2 \geq 0, \\ m_1 + m_2 \geq 1 \end{matrix} \quad \Longrightarrow \quad m_1m_2 - m_3^2 \geq 0 \quad ,$$

where $\Longrightarrow$ is read as "imply".

As a consequence of this result, we can find the value of an integral over D by adding up its values over the two domains D_1 and D_2. The fact that points lying on the plane $m_1 + m_2 = 1$ have been taken twice would not make any difference because they form a set of Lebsegue measure zero.

3. <u>Order of the variables</u> m_1, m_2 <u>and</u> m_3.

To examine the order of the variables m_1, m_2 and m_3, we have first to define them following the original paper of Wald [50]. For the sake of clarity, therefore, we add the following paragraph.

Denote by S the 2n + 1 dimensional surface in the 2n + 4 dimensional space of the variables $u_1,\ldots u_{n+2}$, $v_1,\ldots v_{n+2}$ defined by the following equations:

$$(3.1)\qquad \begin{aligned} \sum_{\beta=1}^{n+2} u_\beta^2 = \sum_{\beta=1}^{n+2} v_\beta^2 = 1 \\ \sum_{\beta=1}^{n+2} u_\beta v_\beta = 0. \end{aligned}$$

Let $u_1\ldots u_{n+2}$, $v_1\ldots v_{n+2}$ be random variables whose joint probability distribution function is defined as follows: the point $(u_1\ldots u_{n+2}, v_1\ldots u_{n+2})$ is restricted to points of S whose probability density function is defined by

$$(3.2)\qquad \frac{dS}{\int_S dS}$$

Then for any subset A of S, the probability of A is equal to the 2n + 1 dimensional value of A divided by $\int_S dS$. It should be noted that the probability density function (3.2) is identical with the probability density function we would obtain if we were to assume that $u_1,\ldots,u_{n+2}$, $v_1,\ldots,v_{n+2}$ are independently, normally distributed with zero means and unit variances and calculate the conditional density function under the restriction that the point $(u_1,\ldots,u_{n+2},$ $v_1,\ldots,v_{n+2}$ belongs to S.

Variables m_1, m_2 and m_3 which are equal respectively to

$$\sum_{\beta=1}^{p} u_\beta^2 , \qquad \sum_{\beta=1}^{r} v_\beta^2 , \qquad \sum_{\beta=1}^{p} u_\beta v_\beta$$

can be redefined by using (2.1) as follows:

$$m_1 = \sum_{\beta=1}^{p} u_\beta^2 \Big/ \sum_{\beta=1}^{n+2} u_\beta^2$$

$$(3.3) \qquad m_2 = \sum_{\beta=1}^{p} v_\beta^2 \Big/ \sum_{\beta=1}^{n+2} v_\beta^2$$

$$m_3 = \sum_{\beta=1}^{p} u_\beta v_\beta \Bigg/ \sqrt{\sum_{\beta=1}^{n+2} u_\beta^2 \sum_{\beta=1}^{n+2} v_\beta^2} ,$$

where p is the number of variables and $n = N + N_2 - 2$ is the number of degrees of freedom.

With this explanation about the variables entering into the discussion we shall prove the following theorem:

<u>Theorem (1).</u> The variables m_1, m_2 and m_3 defined in (3.3) in terms of u_i and v_i, $i=1,2,\ldots,n+2$, which are $N(0,1)$ variates; are of order n^{-1} in the probability sense.

<u>Definition.</u> We write $x_N = 0_p [f(N)]$, and say that x_N is of probability order $0 [f(N)]$ if for each $\epsilon > 0$ there exists an $A_\epsilon > 0$ such that $[P \mid x_N \mid \leq A_\epsilon f(N)] \geq 1-\epsilon$ for all values of $N > N_0(\epsilon)$.

<u>Proof of the theorem:</u>

Since u_i and v_i, $i=1,\ldots,n+2$, are all independently, normally distributed with zero means and unit variance,

$$(3.4) \qquad f(m_i)\, dm_i = \frac{1}{\beta(p,n-p+2)} \; m_i^{p-1} (1-m_i)^{n-p+1}\, dm_i, \; i=1,2$$

since each of m_1 and m_2 is of the form $\dfrac{\chi_1^2}{\chi_1^2 + \chi_2^2}$.

Thus

$$(3.5) \qquad E(m_i) = \frac{p}{n+2} \text{ and } V(m_i) = \frac{p(n-p+2)}{(n+2)^2(n+3)} = 0\left(\frac{1}{n^2}\right);$$

and by Tchebycheff's inequality, namely

$$(3.6) \qquad P \left\{ \left| m_i - \frac{p}{n+2} \right| \geq \frac{k}{n} \right\} \leq \frac{1}{k^2},$$

it is immediately seen that for given ϵ there exist k_1 and k_2 such that

$$(3.7) \qquad P\left(\frac{k_1}{n} \leq m_i \leq \frac{k_2}{n} \right) > 1-\epsilon.$$

Hence m_1 and m_2 are of order $\frac{1}{n}$ in probability.

To see that m_3 is also of order $\frac{1}{n}$ in the probability sense, we note that

$$(3.8) \qquad m_3^2 = \frac{\left(\sum_{\beta=1}^{p} u_\beta v_\beta\right)^2}{\sum_{\beta=1}^{n+2} u_\beta^2 \sum_{\beta=1}^{n+2} v_\beta^2} \quad .$$

$$(3.9) \qquad \text{But} \quad \sum_{\beta=1}^{p} u_\beta v_\beta \leq \sum_{\beta=1}^{p} u_\beta^2 \sum_{\beta=1}^{p} v_\beta^2 \ ;$$

$$(3.10) \quad \text{therefore} \quad m_3^2 \leq \frac{\sum_{\beta=1}^{p} u_\beta^2}{\sum_{\beta=1}^{n+2} u_\beta^2} \cdot \frac{\sum_{\beta=1}^{p} v_\beta^2}{\sum_{\beta=1}^{n+2} v_\beta^2} \ .$$

$$(3.11) \qquad \text{Hence} \quad m_3^2 \leq m_1 m_2 \ \text{because of (3.3).}$$

From this, by noting that $m_1 = 0_p(\frac{1}{n})$ and $m_2 = 0_p(\frac{1}{n})$, we conclude that

$$m_3 = 0_p(\frac{1}{n}) \ .$$

4. An important limit.

In this section we shall prove a result which will be helpful in finding asymptotic values of triple integrals of the type

$$(4.1) \qquad \iiint_D m_1^{\nu_1} m_2^{\nu_2} m_3^{\nu_3} \begin{vmatrix} m_1 & m_3 \\ m_3 & m_2 \end{vmatrix}^a \begin{vmatrix} 1-m_1 & m_3 \\ m_3 & 1-m_2 \end{vmatrix}^b dm_1 dm_2 dm_3$$

where b is a large number. The result can be stated as

<u>Theorem 2.</u> If m_1, m_2, m_3 are random variables as defined in 3, each depending on n, and each being of order n^{-1} in probability, then

$$(4.2) \qquad \operatorname*{plim}_{n \to \infty} \frac{[(1-m_1)(1-m_2)-m_3^2]^n}{e^{-n(m_1+m_2)}} = 1 \ .$$

Proof: We shall replace m_1, m_2 and m_3 by $\frac{\alpha}{n}$, $\frac{\beta}{n}$ and $\frac{\gamma}{n}$ respectively. The variables α, β and γ are therefore of order one in probability. This means for a given ϵ, there exist numbers $N(\epsilon)$, $A_{1\epsilon}$, $A_{2\epsilon}$ and $A_{3\epsilon}$, such that

$$P(\alpha \geq A_{1\epsilon}) \leq \epsilon \ ,$$

$$P(\beta \geq A_{2\epsilon}) \leq \epsilon \ ,$$

$$P(\gamma \geq A_{3\epsilon}) \leq \epsilon \qquad \text{for } n \geq N_\epsilon \ .$$

If in (4.3) each of $A_{1\epsilon}$, $A_{2\epsilon}$ and $A_{3\epsilon}$ is replaced by $A_\epsilon = \max(A_{1\epsilon}, A_{2\epsilon}, A_{3\epsilon})$, the inequalities will still hold. In terms of these variables we have to show

$$(4.4) \qquad \operatorname*{plim}_{n \to \infty} \frac{\left[1 - \frac{\alpha+\beta}{n} + \frac{\alpha B-\gamma^2}{n^2}\right]^n}{e^{-\alpha-\beta}} = 1 \ .$$

That is, plim $f(\alpha,\beta,\gamma) = 1$, where

$$(4.5) \quad f(\alpha,\beta,\gamma) = e^{\alpha+\beta} \left[1 - \frac{\alpha+\beta}{n} + \frac{\alpha\beta-\gamma^2}{n^2} \right]^n .$$

To show this we consider the function

$$(4.6) \qquad g(\alpha,\beta,\gamma) = \log f(\alpha,\beta,\gamma)$$

and expand it by Taylor's Theorem, with a remainder after two terms, namely

$$(4.7) \quad g(\alpha,\beta,\gamma) = g(0,0,0) + \left(\alpha \frac{\partial}{\partial\alpha} + \beta \frac{\partial}{\partial\beta} + \gamma \frac{\partial}{\partial\gamma}\right) g(0,0,0)$$

$$+ \left(\alpha \frac{\partial}{\partial\alpha} + \beta \frac{\partial}{\partial\beta} + \gamma \frac{\partial}{\partial\gamma}\right)^2 g(\theta,\phi,\psi) ,$$

where

$$0 < \theta < \alpha$$

$$0 < \phi < \beta$$

$$0 < \psi < \gamma .$$

We have

$$g(\alpha,\beta,\gamma) = \alpha + \beta + n \log \left[1 - \frac{\alpha+\beta}{n} + \frac{\alpha\beta-\gamma^2}{n^2} \right] ,$$

so that

$$\frac{\partial g}{\partial \alpha} = 1 - \frac{1 - \beta/n}{(1 - \frac{\alpha}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}} \quad ,$$

$$\frac{\partial g}{\partial \beta} = 1 - \frac{1 - \frac{\alpha}{n}}{(1 - \frac{\alpha}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}} \quad ,$$

$$\frac{\partial g}{\partial \gamma} = \frac{-2\frac{\gamma}{n}}{(1 - \frac{\alpha}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}}$$

Also

$$\frac{\partial^2 g}{\partial \alpha^2} = \frac{-\frac{1}{n}(1 - \frac{\beta^2}{n})}{\left[(1 - \frac{2}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} \quad ,$$

$$\frac{\partial^2 g}{\partial \beta^2} = \frac{-\frac{1}{n}\,(1 - \frac{\alpha^2}{n}}{\left[(1 - \frac{\alpha}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} \quad ,$$

$$\frac{\partial^2 g}{\partial \gamma^2} = \frac{\frac{2}{n}\left[1 - \frac{\alpha}{n} - \frac{\beta}{n} + \frac{2\beta + \gamma^2}{n^2}\right]}{\left[(1 - \frac{\alpha}{n})(1 - \frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} \quad ,$$

$$\frac{\partial^2 g}{\partial\alpha\partial\beta} = \frac{-\gamma^2/n^3}{\left[(1-\frac{\alpha}{n})(1-\frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} ,$$

$$\frac{\partial^2 g}{\partial\alpha\partial\gamma} = \frac{(1-\frac{\beta}{n})(-\frac{2\gamma}{n^2})}{\left[(1-\frac{\alpha}{n})(1-\frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} ,$$

$$\frac{\partial^2 g}{\partial\beta\partial\gamma} = \frac{(1-\frac{\alpha}{n})(-\frac{2\gamma}{n^2})}{\left[(1-\frac{\alpha}{n})(1-\frac{\beta}{n}) - \frac{\gamma^2}{n^2}\right]^2} .$$

Now $g, \frac{\partial g}{\partial\alpha}, \frac{\partial g}{\partial\beta}, \frac{\partial g}{\partial\gamma}$ are all zero for $(\alpha,\beta,\gamma) = (0,0,0)$.

Thus (4.7) gives

(4.8) $$g(\alpha,\beta,\gamma) = R_2 = \frac{\alpha^2}{2} g_{\alpha\alpha} + \frac{\beta^2}{2} g_{\beta\beta} + \frac{\gamma^2}{2} g_{\gamma\gamma}$$

$$+ \alpha\beta\, g_{\alpha\beta} + \beta\gamma\, g_{\beta\gamma} + \gamma\alpha\, g_{\gamma\alpha}$$

where the value of each of the derivatives involved is calculated at (θ,ϕ,ψ). This gives

$$(4.9) \quad R_2 = \frac{1}{\left[1 - \frac{\theta}{n} - \frac{\phi}{n} + \frac{\theta\phi - \psi^2}{n^2}\right]^2} \quad - \frac{\alpha^2}{2n}\left(1 - \frac{\phi}{n}\right)^2 - \frac{\beta}{2n}\left(1 - \frac{\phi}{n}\right)^2 +$$

$$\frac{\gamma^2}{n}\left(1 - \frac{\theta}{n} - \frac{\phi}{n} + \frac{\theta\phi - \psi^2}{n^2}\right) - \frac{\alpha\beta\psi^2}{n^3} - 2\,\frac{\beta\gamma\psi}{n^2}\left(1 - \frac{\theta}{n}\right) - \frac{2\gamma\alpha\psi}{n^2}\left(1 - \frac{\phi}{n}\right) \quad .$$

Using the inequalities of (4.3) in (4.9) we get

$$(4.10) \qquad R_2 \quad < \max \left\{ \frac{\frac{A^2}{n} + \frac{A^2}{n^2} + \frac{5A^4}{n^3}}{\left(1 - \frac{2A}{n} - \frac{A^2}{n^2}\right)^2} \qquad \text{and} \right.$$

$$\left. \frac{\frac{A^2}{n} + \frac{6A^3}{n^2} + \frac{2A^4}{n^3} + \frac{A^4}{n^4}}{\left(1 - \frac{2A}{n} - \frac{A^2}{n^2}\right)^2} \right\}$$

in the probability sense; where the two expressions inside the brackets in (4.10) are calculated from (4.9) by considering the fact that R_2 may be positive or negative. Since A is finite and independent of n , therefore both the expressions tend to zero as n tends to infinity or

$$\underset{n \longrightarrow \infty}{\text{Plim}} \; R_2 = 0$$

Hence

$$\operatorname{Plim} e^{\alpha+\beta}\left[1 - \frac{\alpha+\beta}{n} + \frac{\alpha\beta-\gamma^2}{n^2}\right]^n = \operatorname{Plim} e^{R_2} = 1$$

That is, $\left[1 - \frac{\alpha+\beta}{n} + \frac{\alpha\beta-\gamma^2}{n^2}\right]^n$ is asymptotically equivalent to $e^{-\alpha-\beta}$ in the probability sense. Note: An alternative proof of the statement (4.2) shall be provided if we are able to prove that

$$(4.12) \qquad \operatorname*{plim}_{n \to \infty} \left[n \log \left(1 - \frac{c}{n} + \frac{h}{n^2} + c\right)\right] = 0,$$

where c stands for $\alpha + \beta$, and $h = \alpha\beta - \gamma^2$ and α, β and γ are restricted by the conditions (4.3).

Since $0 \leq 1 - m_1 - m_2 + m_1 m_2 - m_3^2 \leq 1$, $0 \leq (\frac{c}{n} - \frac{h}{n^2}) \leq 1$.

It is easy to verify that

$$(4.13) \qquad \frac{-x}{n-x} \leq \log \left(1 - \frac{x}{n}\right) \leq -\frac{x}{n} - \frac{x^3}{2n^2} \quad ,$$

in which the lower bound is written by observing that

$$-\log\left(1 - \frac{x}{n}\right) = \frac{x}{n} + \frac{x^2}{2n^2} + \frac{x^3}{3n^3} + \dots \leq \frac{x}{n} + \left(\frac{x}{n}\right)^2 + \left(\frac{x}{n}\right)^3 + \dots,$$

and this series adds up to $\frac{x}{n-x}$.

The use of (4.13) gives

$$(4.14)\quad \frac{-cn+h}{n^2-cn+h} \leq \log\left(1-\frac{c}{n}+\frac{h}{n^2}\right) \leq -\frac{c}{n}+\frac{h}{n^2}-\frac{c^2}{2n^2} \quad .$$

Thus

$$\frac{hn-c^2n+ch}{n^2-cn+h} \leq \left[n\log\left(1-\frac{c}{n}+\frac{h}{n^2}\right)+c\right] \leq \frac{h}{n}-\frac{c^2}{2n}$$

and both the limits coverge to zero because of the restrictions (4.3). This proves the statement (4.12) and hence the Theorem.

5. A triple integral.

In this and the remaining sections of this chapter, we shall confine ourselves to the study of the integral

$$(5.1)\quad I = \iiint_D (m_1m_2 - m_3^2)^{\frac{p-3}{2}} \left[(1-m_1)(1-m_2)-m_3^2\right]^{\frac{n-p-1}{2}} dm_1dm_2dm_3 \ ,$$

where D is defined in (2.2) . To find an asymptotic approximation to the value of I we first write

$$(5.2)\qquad I = I_1 + I_2 \ ,$$

where I_1 and I_2 denote the values of the integral over D_1 and D_2. To find I_1, we shall first evaluate the integral

$$(5.3) \quad I_{11} = \iiint_{D_1} (m_1 m_2 - m_3^2)^{\frac{p-3}{2}} e^{-\frac{n}{2}(m_1+m_2)} dm_1 dm_2 dm_3 ,$$

and then find an upper bound to

$$(5.4) \qquad E = I_1 - I_{11} ;$$

that is, an upper bound for the error committed in replacing the the factor $[(1-m_1)(1-m) - m_3^2]^{\frac{n-p-1}{2}}$ in the integrand by $e^{-\frac{n}{2}(m_1 + m_2)}$. Using (5.3) and (5.4), we can state that

$$(5.5) \qquad I_{11} - E \leq I_1 \leq I_{11} + E.$$

It will then be demonstrated that both the error committed in approximating I_1 by I_{11} and the value of I_2 are negligible as compared to the least possible to value of I_1. Mathematically

$$\lim_{n \to \infty} \frac{E}{I_{11} - E} = 0$$

and

$$(5.7) \qquad \lim_{n \longrightarrow \infty} \frac{I_2}{I_{11} - E} = 0 .$$

As a consequence of (5.5) and (5.6) we can write

$$(5.8) \qquad I_1 \quad I_{11} ,$$

and as a result of (5.1), (5.7) and (5.8) we can write

$$(5.9) \qquad I \sim I_{11} .$$

This will be the general line of argument to be followed in obtaining an asymptotic approximation for the value of I.

It would appear from Theorem 2, proved in Section 4, that the approximation $e^{-\frac{n}{2}(m_1+m_2)}$ for $\left[(1-m_1)(1-m_2)-m_3^2\right]^{\frac{n-p-1}{2}}$ is valid only in the domain $D^* \subset D_1$ which is such that throughout D^* the variables m_i, i=1,2 and 3 are $O_p(\frac{1}{n})$. We shall, however, work in terms of the division D_1 and D_2 of the total domain and use the exponential approximation over the whole of D_1 because of certain simplifications which result. Justification of the results thus obtained is provided by two factors:

(1) The integrand shows that almost the whole of the density is concentrated in that part of the domain D_1 which is close to the

origin. In fact, if we define a domain $D^* \subset D$ by the inequalities

$$(5.10) \qquad \begin{aligned} m_1 m_2 - m_3^2 &\geq 0 \\ m_1 + m_2 &\leq \frac{A}{n} , \end{aligned}$$

then it is shown in section 11 that D^* contains almost the whole of the density. This is probably the main reason why the exponential approximation, which is true over the domain D^*, gives close results.

(2) The discussion on the upper bound to error given in section 7 actually proves that the loss of accuracy in using $e^{-\frac{n}{2}(m_1+m_2)}$ instead of $[(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}}$ is negligible when n is large.

It may also be remarked at this point that the exact value of I is known from Sitgreaves [45]. It would appear obvious, therefore, that the asymptotic value of I could be obtained from the one given by Sitgreaves by using Stirling's approximation to $\Gamma(x)$. This would no doubt hold true provided we were interested merely in the asymptotic value of I. The reason, however, for our following an independent approach is that we are interested in finding the solution to a distribution problem. The techniques and simplifications used in the asymptotic evaluation of I, which emerge mainly as a result of the supposition that <u>n tends to infinity</u>, will be used in evaluating the limiting moments of a certain statistic, to be called Wald's approximate classification statistic. This distribution problem will be our subject of discussion in chapter III.

6. The integral over D_1, an asymptotic approximation.

$$\text{(6.1)} \quad \text{Let } I_1 = \iiint\limits_{D_1} (m_1 m_2 - m_3^2)^{\frac{p-3}{2}} \left[(1-m_1)(1-m_2)-m_3^2\right]^{\frac{n-p-1}{2}} dm_1 dm_2 dm_3,$$

where D_1 is defined by the following inequalities:

$$\text{(6.2)} \quad \begin{aligned} m_1 m_2 - m_3^2 &\geq 0 \\ m_1 + m_2 &\leq 1 \\ 0 \leq m_1 &\leq 1 \\ 0 \leq m_2 &\leq 1 \ ; \end{aligned}$$

and where $m_i = 0_p(\frac{1}{n})$ for i=1,2, and 3. We shall replace the second factor in (6.1) by $e^{-\frac{n}{2}(m_1+m_2)}$, but the operation of integration after this replacement needs some justification. There is no loss of generality if we consider a similar univariate case and prove that it is possible to replace a binomial raised to a large power by an exponential factor to which it increases. We shall state this result formally as

Lemma. Let f(x) be a function of the real variable x, such that

$$f(x) \leq A \quad \text{if } 0 \leq x \leq c, \quad \text{where } 0 \leq c \leq \infty.$$

Then

$$(6.3) \qquad \lim_{n \to \infty} \frac{\int_0^c (1 - \frac{x}{n})^n f(x)\, dx}{\int_0^c e^{-x} f(x)\, dx} = 1.$$

Proof

$$(6.4) \qquad \text{Let } I_d = \int_0^c f(x) \left[(1 - \frac{x}{n})^n - e^{-x} \right] dx.$$

Then (6.3) states that $\lim_{n \to \infty} \left| I_d \right| = 0.$

It is known that

$$(6.5) \qquad 0 \leq e^{-x} - (1 - \frac{x}{n})^n \leq \frac{x^2 e^{-x}}{n} .$$

(See, for instance, Whittacker and Watson [54], page 242). Therefore, using this and condition (1), we get

$$(6.6) \qquad I_d \leq A \int_0^c \frac{x^2 e^{-x}}{n} \leq A \int_0^\infty \frac{x^2 e^{-x}}{n} .$$

This shows that for all c

$$(6.7) \qquad I_d < A \frac{\Gamma(c)}{n} .$$

The quantity on the right hand side of (6.7) is positive and tends to zero as n increases. Hence (6.3) is established.

We shall rewrite (6.1) as

$$(6.8) \qquad I_{11} = \iiint_{D_1} (m_1 m_2 - m_3^2)^{\frac{p-3}{2}} e^{\frac{n}{2}(m_1+m_2)} dm_1 dm_2 dm_3$$

and evaluate I_{11}, an asymptotic approximation to I_1. This will be followed by a Section on an upper bound to the error in using I_{11} in place of I_1.

To integrate with respect to m_3 we first put

$$(6.9) \qquad m_3 = (m_1 m_2 t^*)^{1/2} \quad .$$

Thus

$$(6.10) \qquad I_{11} = \iiint_{t^*=0}^{1} (m_1 m_2)^{\frac{p-2}{2}} t^{*\,-\frac{1}{2}} (1-t^*)^{\frac{p-3}{2}} e^{-\frac{n}{2}(m_1+m_2)} dt^* dm_1 dm_2 .$$

Integration with respect to t^* gives

$$(6.11) \qquad I_{11} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})} \iint_{0 \le m_1 + m_2 \le 1} (m_1 m_2)^{\frac{p-2}{2}} e^{-\frac{n}{2}(m_1+m_2)} dm_1 dm_2 .$$

Making the transformation

$$(6.12) \qquad \begin{aligned} m_1 &= z \cos^2 \Theta , \\ m_2 &= z \sin^2 \Theta , \end{aligned}$$

we have

$$\frac{\partial(m_1 m_2)}{\partial(z,\Theta)} = 2z \sin \Theta \cos \Theta ;$$

and so (6.11) becomes

$$(6.13)\quad I_{11} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})}\, 2 \int_{z=0}^{1} \int_{\Theta=0}^{\pi/2} z^{p-1} e^{-\frac{n}{2}z} \cos^{p-1}\Theta \sin^{p-1}\Theta \, dz d\Theta .$$

Integration with respect to Θ yields

$$(6.14)\quad I_{11} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})} \; \frac{\Gamma(\frac{p}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{z=0}^{1} e^{-\frac{n}{2}z} z^{p-1} \, dz ;$$

and putting $\frac{n}{2}z = t$ we obtain the form

$$(6.15)\quad I_{11} = \left(\frac{2}{n}\right)^{p} \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{0}^{\frac{n}{2}} e^{-t} t^{p-1} \, dt .$$

Now for large n it is well known that

$$(6.16)\quad \int_{0}^{\frac{n}{2}} e^{-t} t^{p-1} \, dt \sim \int_{0}^{\infty} e^{-t} t^{p-1} \, dt ,$$

$$(6.17)\quad (6.15) \text{ gives } I_{11} \sim \left(\frac{2}{n}\right)^{p} \Gamma(\tfrac{1}{2})\, \Gamma(\tfrac{p-1}{2})\, \Gamma(\tfrac{p}{2}) .$$

This further simplifies to

$$(6.18)\quad I_{11} \sim \frac{4\pi(p-2)!}{n^{p}} .$$

To be more exact we can write

$$(6.19) \quad \int_0^{\frac{n}{2}} e^{-t}\, t^{p-1}\, at = \int_0^{\infty} e^{-t}\, t^{p-1}\, dt - \int_{\frac{n}{2}}^{\infty} e^{-t}\, t^{p-1}\, dt ,$$

and successive integration by parts shows that the right hand side of (6.16) reduces to

$$(6.20) \qquad \Gamma(p) - (\tfrac{n}{2})^{p-1} \frac{1}{e^{n/2}} - \frac{n}{2}^{p-2} \frac{p-1}{e^{n/2}} - \ldots,$$

which can be written as $\Gamma(p) - O(\frac{n^{p-1}}{e^{n/2}})$. Since $\frac{n^{p-1}}{e^{n/2}}$ tends to zero as n tends to infinity,

$$(6.21) \qquad I_{11} \sim \frac{4\pi\ (p-2)!}{n^p} \text{, and } I_{11} = \frac{4\pi(p-2)!}{n^p}(1+\varepsilon_n) .$$

We will show in Section 7 that the error in taking I_{11} as an approximation to the value of I_1 is negligible in comparison to the value of I_1.

It may be remarked in passing that the integral occuring in (6.11) could also be evaluated by usin g Dirichlet's formula [54, p. 258] namely

$$\int\int \cdots \int t_1^{\alpha_1-1} t_2^{\alpha_2-1} \cdots t_n^{\alpha_n-1} f(t_1+t_2\ldots+t_n)dt_1\ldots dt_n$$

(6.22)

$$= \frac{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\ldots\,\Gamma(\alpha_n)}{\Gamma(\alpha_1+\alpha_2+\ldots\alpha_n)} \int_0^1 z^{\sum_{i=1}^{n}\alpha_i - 1} f(z)dz \ .$$

7. An Upper Bound to Error.

In this section we shall consider the following problem: How much error is committed by replacing

$$(1 - m_1 - m_2 + m_1m_2 - m_3^2)^{\frac{n-p-1}{2}}$$

by

$$e^{-\frac{n}{2}(m_1+m_2)}$$

in the triple integral (6.1) over the domain D_1? We shall consider two separate cases,

(A) When $(1 - m_1 - m_2 + m_1m_2 - m_3^2)^{\frac{n-p-1}{2}} > e^{-\frac{n}{2}(m_1+m_2)}$

and

(B) $e^{-\frac{n}{2}(m_1+m_2)} > (1 - m_1 - m_2 + m_1m_2 - m_3^2)^{\frac{n-p-1}{2}}$,

and find an upper bound to error in both cases. The larger of these shall ultimately be taken as the upper bound.

Case A. Let I_{d_1} denote the difference

$$(7.1)\quad \iiint_{D_1} (m_1m_2-m_3^2)^{\frac{p-3}{2}} \left\{ [1 - m_1 - m_2 + m_1m_2 - m_3^2]^{\frac{n-p-1}{2}} - e^{-\frac{n}{2}(m_1+m_2)} \right\} dm_1dm_2dm_3 \ .$$

I_{d_1} will not decrease if in the factor $[1 - m_1 - m_2 + m_1m_2 - m_3^2]^{\frac{n-p-1}{2}}$ we omit m_3^2 and replace m_1m_2 by $(\frac{m_1+m_2}{2})$. Making these changes, we get

$$(7.2)\quad I_{d_1} \leq \iiint_{D_1} (m_1m_2-m_3^2)^{\frac{p-3}{2}} [(1 - \frac{m_1+m_1}{2})^{n-p-1} -e^{-\frac{n}{2}(m_1+m_2)}] dm_1dm_2dm_3.$$

The variable m_3 can be integrated out by using the transformation

$$m_3 = (m_1m_2t^*)^{1/2}, \text{ and then}$$

$$(7.3) \qquad I_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})} \iint (m_1 m_2)^{\frac{p-2}{2}} \left[(1 - \frac{m_1 + m_2}{2})^{n-p-1} - e^{-\frac{n}{2}(m_1+m_2)} \right] dm_1 dm_2 \quad .$$

Using (6.22) in the double integral involved in (7.3), we get

$$(7.4) \qquad I_{d_1} \leq \frac{\Gamma(\frac{1}{2})\ \Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{t=0}^{1} z^{p-1} \left[(1 - \frac{z}{2})^{n-p-1} - e^{-\frac{n}{2} z} \right] dz \ .$$

Replacing $^{z}/_{2}$ by t, (7.4) becomes

$$(7.5) \qquad I_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)}\ 2^p \int_{t=0}^{\frac{1}{2}} t^{p-1} \left[(1-t)^{n-p-1} - e^{-nt} \right] dt \ .$$

The expression

$$(7.6) \qquad (1 - t)^{n-p-1} - e^{-nt}$$

can be written as

$$(7.7) \qquad (1 - t)^{n-p-1} - \bar{e}^{(n-p-1)t} \ . \ \bar{e}^{(p+1)t} \ ,$$

can be written as

$$(7.7) \qquad (1 - t)^{n-p-1} - e^{-(n-p-1)t} \cdot e^{-(p+1)t} ,$$

which can further be written as

$$(7.8) \qquad (1 - t)^{n'} - e^{-n't}\left[1 - (p+1) + \frac{(p+1)t^2}{2!} - \frac{(p+1)t^3}{3!} + \dots\right] ;$$

and using the fact that $1 - y \leq e^{-y}$ we find as a first approximation that

$$(7.9) \qquad (1 - t)^{n'} - e^{-nt} \leq (1 - t)^{n'} - e^{-n't} + e^{-n't}(p+1)t .$$

This reduces to

$$(7.10) \qquad (1 - t)^{n'} - e^{-nt} \leq e^{-n't}(p + 1)t$$

by using the well known inequality

$$(7.11) \qquad 0 \leq e^{-n't} - (1 - t)^{n'} \leq n't^2 e^{-n't} .$$

Using (7.10) in (7.5), we get

$$(7.12) \qquad I_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \; 2^p \int_{t=0}^{\frac{1}{2}} t^p \; (p+1)e^{-(n-p-1)t} \, dt \; .$$

The integral in (7.12) can be simplified by replacing $(n-p-1)t$ by w and extending the upper limit of integration for w to infinity instead of $\frac{n-p-1}{2}$.

This will only increase the upper bound for I_{d_1} , and we get a simpler result, namely

$$(7.13) \qquad I_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \; \frac{2^p(p+1)}{(n-p-1)^{p+1}} \int_{w=0}^{\infty} w^p \, e^{-w} \, dw \; .$$

On simplifying (7.13) we obtain

$$(7.14) \qquad I_{d_1} \leq \frac{4\pi(p+1)!}{(p-1)} \cdot \frac{1}{(n-p-1)^{p+1}} \; ,$$

which gives an upper bound to error in case A.

Case B. Now suppose

$$e^{-\frac{n}{2}(m_1+m_2)} > [1 - m_1 - m_2 + m_1m_2 - m_3^2]^{\frac{n-p-1}{2}},$$

and let

$$(7.15)\quad I_{d_1} = \iiint_{D_1} (m_1m_2-m_3^2)^{\frac{p-3}{2}} \left\{ e^{-\frac{n}{2}(m_1+m_2)} - [1-m_1-m_2+m_1m_2-m_3^2]^{\frac{n-p-1}{2}} \right\} \times dm_1dm_2dm_3 .$$

Omitting the factor $m_1m_2-m_3^2$, which is non-negative in the domain of integration, one can write that

$$(7.16)\quad I_{d_1} \leq \iiint_{D_1} (m_1m_2-m_3^2)^{\frac{p-3}{2}} [e^{-\frac{n}{2}(m_1+m_2)} - (1-m_1-m_2)^{\frac{n-p-1}{2}}] dm_1dm_2dm_3.$$

Integrating out m_3 by the same transformation as was used in Case A, we get

$$(7.17)\quad I^*_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})} \iint\limits_{0 \leq m_1+m_2 \leq 1} (m_1 m_2)^{\frac{p-2}{2}} \left[e^{-\frac{n}{2}(m_1+m_2)} - (1-m_1-m_2)^{\frac{n-p-1}{2}} \right] \times dm_1 dm_2$$

Using (6.22) on the double integral involved in (7.17), we obtain

$$(7.18)\quad I^*_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{z=0}^{1} z^{p-1} \left[e^{-\frac{n}{2}z} - (1-z)^{\frac{n-p-1}{2}} \right] dz .$$

Write

$$(7.19)\quad e^{-\frac{n}{2}z} - (1-z)^{\frac{n-p-1}{2}} = e^{-\frac{p+1}{2}} e^{-\frac{n'z}{2}} - (1-z)^{\frac{n'}{2}} ,$$

where $n' = n-p-1$. Then

$$(7.20)\quad e^{-\frac{n}{2}z} - (1-z)^{\frac{n-p-1}{2}} = e^{-\frac{n'z}{2}} \left[1 - \frac{p+1}{2}z + \left(\frac{p+1}{2}\right)^2 \frac{t^2}{2!} \cdots \right] - (1-z)^{\frac{n'}{2}} .$$

Since $e^{-y} \leq 1$, as a first approximation, we can write

$$(7.21)\quad e^{-\frac{n}{2}z} - (1-z)^{\frac{n-p-1}{2}} \leq e^{-\frac{n'}{2}} - (1-z)^{\frac{n'}{2}} \quad ,$$

and the use of (7.11) gives

$$(7.22)\quad e^{-\frac{n}{2}z} - (1-z)^{\frac{n-p-1}{2}} \leq \frac{n'}{2}z^2 \, e^{-\frac{n'}{2}z} \quad ,$$

Replacing n' by n-p-1, and using this result in (7.18), we get

$$(7.23)\quad I^*_{d_1} \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{z=0}^{1} z^{p+1} \, \frac{n-p-1}{2} \, e^{-\frac{n-p-1}{2}z} \, dz \quad .$$

We put $\frac{n-p-1}{2}z = w$ to get

$$(7.24)\quad I^*_{d_1} \leq \frac{(2)^{p+1}\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\,\Gamma(\frac{p}{2})}{(n-p-1)^{p+1}\quad\Gamma(p)} \int_{w=0}^{\frac{n-p-1}{2}} w^{p+1} \, e^{-w} \, dw \quad .$$

Extending the range of the integral to infinity and integrating, we get

$$(7.25)\quad I^*_{d_1} \leq \left(\frac{2}{n-p-1}\right)^{p+1} \frac{\pi(p-2)!}{2^{p-2}\Gamma(p)}\, \Gamma(p+2) = \frac{8\pi(p+1)!}{(p-1)(n-p-1)^{p+1}} .$$

The larger of the bounds I_{d_1} and $I^*_{d_1}$ namely

$$(7.26)\qquad \frac{8\pi(p+1)!}{(p-1)(n-p-1)^{p+1}}$$

can therefore be taken as an upper bound to the error involved in replacing the factor raised to a high power in the integrand by an exponential factor. It may be remarked here that (7.26) gives only a first approximation for the upper bound to error, and that a closer bound would be obtained if we considered four terms in the expansion of $e^{-(p+1)t}$ in (7.8), and three terms in the expansion of $e^{-\frac{p+1}{2}z}$ in (7.20). Needless to say, we can get closer and closer bounds by considering a larger number of terms in (7.8) and (7.20). It should be noted that the result (7.26) enables us to put greater confidence in our approximation of the value of I, which is of order $\frac{1}{n^p}$; because (7.26) asserts that the maximum error committed by supposing that I_1 is approximated by I_{11} is of order $\frac{1}{n^{p+1}}$; and therefore negligible for large values of n.

The bound $\frac{8\pi(p+1)!}{p-1} \frac{1}{(n-p-1)^{p+1}}$ can be rewritten as $\frac{R_{D_1}}{n^{p+1}}$

by using the inequality $\frac{1}{(n-p-1)^{p+1}} < \frac{2}{n^{p+1}}$ for large n.

Thus

(7.27) $$\text{"Error} < \frac{16\pi(p+1)!}{p-1} \frac{1}{n^{p+1}} = \frac{R_{D_1}}{n^{p+1}} \text{" say.}$$

As a result of the discussion in Sections 6 and 7 we can write a formal proof of

Theorem 3.

$$I_1 \sim \frac{4\pi(p-2)!}{n^p} \quad .$$

Proof:

From the results of Sections 6 and 7 we can write

(7.28) $$I_1 = \frac{4\pi(p-2)!}{n^p}(1+\varepsilon_n) + \frac{J_{D_1}}{n^{p+1}}$$

where $\left| J_{D_1} \right| < R_{D_1} = \frac{16\pi(p+1)!}{p-1}$ and $\lim_{n \to \infty} (\varepsilon_n) = 0$.

Multiplying both sides of (7.28) by n^p, and taking the limit as n tends to infinity, of the right hand side in

(7.29) $$n^p I_1 = 4\pi(p-2)! + \frac{J_{D_1}}{n} + 4\pi(p-2)!\ \varepsilon_n$$

the truth of Theorem 3 is established.

8. The integral over D_2.

We now consider the integral

$$(8.1)\quad I_2 = \iiint_{D_2} (m_1m_2-m_3^2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}} dm_1dm_2dm_3,$$

where D_2 is defined by the inequalities

$$(8.2)\quad \begin{aligned} (1-m_1)(1-m_2) - m_3^2 &\geq 0 \\ m_1 + m_2 &\geq 1 \\ 0 \leq m_1 &\leq 1 \\ 0 \leq m_2 &\leq 1 \end{aligned} \quad .$$

To integrate (8.1) with respect to m_3, we make the transformation

$$(8.3)\quad m_3 = [(1-m_1)(1-m_2)t]^{\frac{1}{2}}$$

so that $dm_3 = \frac{1}{2} [(1-m_1)(1-m_2)]^{\frac{1}{2}} t^{-\frac{1}{2}} dt.$

Then

$$*8.4)\quad I_2 = \iiint_{t=0}^{1} [m_1m_2-(1-m_1)(1-m_2)t]^{\frac{p-3}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p}{2}} t^{-\frac{1}{2}} (1-t)^{\frac{n-p-1}{2}} dt\, dm_1\, dm_2.$$

If we notice that

$$F(a,b,c,x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \Theta^{b-1}(1-\Theta)^{c-b-1}(1-\Theta x)^{-a}\, d\Theta,$$

(8.5) provided that $|x| < 1$, we can rewrite (8.4) as

$$(8.6)\quad I_2 = \iint\limits_{m_1+m_2\geq 1} (m_1m_2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p}{2}} \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{n-p+2}{2})}$$

$$\cdot\ F\left[-\frac{p-3}{2}, \frac{1}{2}, \frac{n-p+2}{2}, \frac{(1-m_1)(1-m_2)}{m_1m_2}\right] dm_1\, dm_2.$$

This step is justified by the fact that

$$(8.7)\qquad \frac{(1-m_1)(1-m_2)}{m_1m_2} < 1, \quad \text{since } m_1+m_2 \geq 1$$

in the domain under consideration; except that on the surface of the plane $m_1+m_2 = 1$ we have an equality sign in (8.7). But the omission of the point set determined by the plane $m_1+m_2 = 1$ does not alter the value of I_2, since it forms a set of measure zero.

$$(8.8)\qquad \text{Since } F(a,b,c,x) = 1 + \frac{ab}{c}x + \frac{a(a+1)b(b+1)}{2!}x^2 + \ldots$$

where, for the hypergeometric function involved in (8.6), we have $c = \frac{n-p+2}{2}$, we see that F gives a rapidly convergent series suitable for asymptotic purposes. Since $a = -\frac{p-3}{2}$ and is non-positive for $p \geq 3$, it gives rise to a terminating series if p is odd and is ≥ 3. In all the other cases we shall have an infinite series in which the rth term is of order $\frac{1}{n^{r-1}}$. Hence for asymptotic purposes the first term in the expansion of $F(a,b,c,x)$, namely unity, will provide a reasonably good approximation.

With these considerations in view, we can rewrite (8.6) as

$$(8.9)\quad I_2 = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{n-p+2}{2})}\left\{ \iint\limits_{m_1+m_2\geq 1} (m_1m_2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p}{2}} dm_1 dm_2 \right.$$

$$\left. - \frac{p-3}{2(n-p+2)} \iint\limits_{m_1+m_2\geq 1} (m_1m_2)^{\frac{p-5}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p+2}{2}} dm_1\, dm_2 + \cdots \right\} .$$

Each of the double integrals involved in (8.9) can be changed to a repeated integral and evaluated. As an example, we consider the first one, namely

$$(8.10)\quad \iint\limits_{m_1+m_2\geq 1} (m_1m_2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p}{2}} dm_1\, dm_2 .$$

This can be written as

$$(8.11)\quad \int_{m_1=0}^{1} m_1^{\frac{p-3}{2}} (1-m_1)^{\frac{n-p}{2}} \int_{m_2=1-m_1}^{1} m_2^{\frac{p-3}{2}} (1-m_2)^{\frac{n-p}{2}} dm_1 dm_2 .$$

Integration by parts shows that

$$(8.12)\quad \int_{1-m_1}^{1} m_2^{\frac{p-3}{2}} (1-m_2)^{\frac{n-p}{2}} dm_2 = \frac{2}{n-p+2} (1-m_1)^{\frac{p-3}{2}} m_1^{\frac{n-p+2}{2}} +$$

$$\frac{p-3}{n-p+2} \cdot \frac{2}{n-p+4} \cdot (1-m_1)^{\frac{p-5}{2}} m_1^{\frac{n-p+4}{2}} +$$

$$\cdot \quad \frac{p-3}{n-p+2} \cdot \frac{p-5}{n-p+4} \cdot \frac{2}{n-p+6} \cdot (1-m_1)^{\frac{p-7}{2}} m_1^{\frac{n-p+6}{2}} + \ldots \quad .$$

Using (8.12) in (8.11) we get the series:

$$(8.13)\quad \frac{2}{n-p+2} \int_{m_1=0}^{1} m_1^{\frac{n-1}{2}} (1-m_1)^{\frac{n-3}{2}} dm_1 + \frac{p-3}{n-p+2} \cdot \frac{2}{n-p+4} \int_0^1 m_1^{\frac{n+1}{2}} (1-m_1)^{\frac{n-5}{2}} dm_1$$

$$+ \frac{p-3}{n-p+2} \cdot \frac{p-5}{n-p+4} \cdot \frac{2}{n-p+6} \int_0^1 m_1^{\frac{n+3}{2}} (1-m_1)^{\frac{n-7}{2}} dm_1 + \ldots \quad .$$

Writing the values of the integrals involved above, (8.13) gives rise to the series:

$$(8.14)\quad \frac{2}{n-p+2} \; \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{n-1}{2})}{\Gamma(n)} + \frac{p-3}{n-p+2} \; \frac{2}{n-p+4} \; \frac{\Gamma(\frac{n+3}{2})\Gamma(\frac{n-3}{2})}{\Gamma(n)}$$

$$+ \frac{p-3}{n-p+2} \cdot \frac{p-5}{n-p+4} \cdot \frac{2}{n-p+6} \cdot \frac{\Gamma(\frac{n+5}{2})\Gamma(\frac{n-5}{2})}{\Gamma(n)} + \ldots \quad .$$

Similarly we can find the series expansions for the values of the remaining double integrals involved in (8.9). Using (8.14) as the value of (8.10), and similar values for the other integrals (8.9) gives

$$(8.15)\quad I_2 = \frac{\pi(n-p)!}{(n-2)!(n-p+2)2^{n-p}} \left[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-p+2}{2})}\right]^2 + O(\frac{1}{n^3 2^n}) \quad .$$

Use can be made of Stirling's approximation to the value of $\Gamma(x)$, namely

$$(8.16)\qquad \Gamma(x) = e^{-x}\, x^{x-\frac{1}{2}} \left[1 + \frac{1}{12x} + \ldots \right] .$$

Equation (8.15) shows that the principal term in the value of I_2 is of order $\frac{1}{n^2 2^n}$, where the actual value of I_2 differs from the principal term by terms which are of order $\frac{1}{n^3 2^n}$ and higher.

9. <u>An upper bound to the value of I_2.</u>

We can start from (8.6) and write

$$(9.1)\quad I_2 = \iint\limits_{m_1+m_2\geq 1} (m_1 m_2)^{\frac{p-3}{2}} \left[(1-m_1)(1-m_2)\right]^{\frac{n-p}{2}} \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{n-p+2}{2})} \,.$$

$$F(-\frac{p-3}{2}, \frac{1}{2}, \frac{n-p+2}{2}, \frac{(1-m_1)(1-m_2)}{m_1 m_2})\, dm_1\, dm_2 \;.$$

where it is known that

$$\frac{(1-m_1)(1-m_2)}{m_1 m_2} \leq 1. \tag{9.2}$$

The maximum value of the hypergeometric series involved will correspond to the case in which $\frac{(1-m_1)(1-m_2)}{m_1 m_2} = 1$, and in that case, using the formula

$$F(a,b,c,1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \quad , \text{ we get} \tag{9.3}$$

$$I_2 \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} \iint\limits_{m_1+m_2\geq 1} (m_1 m_2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)]^{\frac{n-p}{2}} \, dm_1 \, dm_2 \; . \tag{9.4}$$

$$\text{Since} \quad (1-m_1)(1-m_2) < (1 - \frac{m_1+m_2}{2})^2 \; , \tag{9.5}$$

(9.4) can be rewritten as

$$I_2 \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} \iint\limits_{1\leq m_1+m_2\leq 2} (m_1 m_2)^{\frac{p-3}{2}} (1 - \frac{m_1+m_2}{2})^{n-p} \, dm_1 \, dm_2 \; . \tag{9.6}$$

Transformation

$$\begin{aligned} m_1 &= z \cos^2 \theta \\ m_2 &= z \sin^2 \theta \end{aligned} \tag{9.7}$$

reduces (9.6) to

$$(9.8)\quad I_2 \leq 2\,\frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} \int_{z=1}^{2}\int_{\Theta=0}^{\frac{\pi}{2}} z^{p-2}(1-\tfrac{z}{2})^{n-p} \sin^{p-2}\Theta \cos^{p-2}\Theta \, d\Theta \, dz \, .$$

By using the formula

$$(9.9)\quad \int_{0}^{\frac{\pi}{2}} \sin^{a-1}\Theta \cos^{b-1}\Theta \, d\Theta = \frac{1}{2}\,\frac{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})}{\Gamma(\frac{a+b}{2})} \qquad \text{for}$$

integration with respect to Θ, (9.8) reduces to

$$(9.10)\quad I_2 \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}\,\frac{\Gamma(\frac{p-1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)} \int_{z=1}^{2} z^{p-2}(1-\tfrac{z}{2})^{n-p}\, dz .$$

This inequality is the same as

$$(9.11)\quad I_2 \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}\,\frac{\Gamma(\frac{p-1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)}\, 2^{p-1} \int_{w=\frac{1}{2}}^{1} w^{p-2}(1-w)^{n-p}\, dw .$$

Observing that

$$\int_{\frac{1}{2}}^{1} w^{p-2}\,(1-w)^{n-p}\, dw = \frac{1}{2^{n-1}(n-p+1)} - \frac{p-2}{n-p+1}\int_{\frac{1}{2}}^{1} w^{p-3}(1-w)^{n-p+1}\, dw,$$

we have

(9.12) $$I_2 \leq \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})} \frac{\Gamma(\frac{p-1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)\ \ 2^{n-p}} \cdot \frac{1}{(n-p+1)}$$

Inequality (9.12) shows that we can find a number R_{D_2} such that

(9.13) $$I_2 \leq \frac{R_{D_2}}{n^{3/2}2^n} \quad .$$

Slight simplification would indicate that if n is so large that Stirling's approximation for $\Gamma(n)$ is valid then $R_{D_2} = 16$ would give a liberal upper bound.

10. Comparison of I_1 and I_2.

In Section 8 we proved that

(10.1) $$I_2 = \frac{c}{n^2 2^n} + O(\frac{1}{n^3 2^n}), \qquad \text{where } c$$

is some constant, and in Section 9 we established that

(10.2) $$I_2 < \frac{R_{D_2}}{n^{3/2}2^n} \quad .$$

A comparison of these results with the value of I_1 namely

(10.3) $$I_1 = \frac{4\pi(p-2)!}{n^p}(1+a_n) + \frac{J_{D_1}}{n^{p+1}}$$

where J_{D_1} is a certain constant less in absolute value than another known constant which is independent of n, shows that

$$(10.4) \qquad \lim_{n \to \infty} \frac{I_2}{I_1} = 0 \quad .$$

This statement follows from the obvious fact that 2^n tends to infinity more rapidly than n^p where p is finite. This means that the relative contribution of the domain D_2 to the value of the Integral I carried over D is negligible in the limit.

Theorem (4).

$$(10.5) \qquad I \sim \frac{4\pi(p-2)!}{n^p} \; .$$

Proof.

From (5.2) $I = I_1 + I_2$. Using (10.4) we have $I \sim I_1$ and from theorem 3, $I_1 \sim \frac{4\pi(p-2)!}{n^p}$. Hence $\frac{4\pi(p-2)!}{n^p}$, which can also be written as $\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2}) \; (\frac{2}{n})^p$, is an asymptotic approximation to the value of I.

As a further check of the correctness of our approximation I_{11}, we can compare it with the exact value of the integral referred to in Section 5. That value can be written as

$$(10.6) \qquad I_s = \frac{4\pi(p-2)!}{n(n-1)\ldots(n-p+1)} \quad ,$$

where the subscript 's' is for the author of the formula. Comparing it with I_{11} written in (10.5), we have

$$\lim_{n \to \infty} \frac{I_{11}}{I_s} = \lim_{n \to \infty} \frac{n(n-1)\ldots(n-p+1)}{n^p} = 1.$$

Hence our approximation is asymptotically equivalent to I_s, the exact value, in the sense of Poincaré [13].

11. <u>The integral over the domain D^*.</u>

Domain D^* was defined in Section 5 as that subset of D_1 in which $m_i = O_p(\frac{1}{n})$ for i=1,2 and 3. Since $-\sqrt{m_1 m_2} \leq m_3 \leq \sqrt{m_1 m_2}$, one way of characterizing this domain would be to say that D^* corresponds to the inequalities

(11.1)
$$m_1 m_2 - m_3^2 \geq 0$$
$$0 \leq m_1 + m_2 \leq \frac{A}{n},$$

where A is a finite number, independent of n.

We can evaluate the integral over D^* as follows: Let

(11.2)
$$I^* = \iiint_{D^*} (m_1 m_2 - m_3^2)^{\frac{p-3}{2}} e^{-\frac{n}{2}(m_1+m_2)} \, dm_1 \, dm_2 \, dm_3.$$

Integration with respect to m_3 by the usual transformation gives

$$(11.3) \qquad I^* = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})} \iint\limits_{0 \leq m_1+m_2 \leq A/n} (m_1 m_2)^{\frac{p-2}{2}} e^{-\frac{n}{2}(m_1+m_2)} \, dm_1 \, dm_2 \ .$$

Putting $m_1 = z \cos^2\Theta$, and

$m_2 = z \sin^2\Theta$

and integrating out Θ, we get

$$(11.4) \qquad I^* = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \int_{z=0}^{\frac{A}{n}} z^{p-1} e^{-\frac{n}{2}z} \, dz \ .$$

Substituting w for $\frac{n}{2}$ z, (12.4) can be written as

$$(11.5) \qquad I^* = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \left(\frac{2}{n}\right)^p \int_{w=0}^{\frac{A}{2}} w^{p-1} e^{-w} \, dw.$$

Thus

$$(11.6) \qquad I^* = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p}{2})}{\Gamma(p)} \left(\frac{2}{n}\right)^p \left[\Gamma(p) - \int_{\frac{A}{2}}^{\infty} w^{p-1} e^{-w} \, dw \right],$$

which on further simplification gives

$$(11.7) \qquad I^* = \frac{4\pi(p-2)!}{n^p} - \frac{4\pi}{p-1} \frac{1}{n^p} \int_{\frac{A}{2}}^{\infty} w^{p-1} e^{-w} \, dw \ .$$

Comparing this value with the exact value I_s we have

$$\text{(11.8)} \qquad \lim_{n \to \infty} \frac{I^*}{I_s} = 1 - \frac{1}{(p-1)!} \int_{\frac{A}{2}}^{\infty} w^{p-1} e^{-w} \, dw,$$

which is also = $\lim_{n \to \infty} \frac{I^*}{I_{11}}$.

Since A might be a large number though not of the order of n, the term

$$\text{(11.9)} \qquad \frac{1}{(p-1)!} \int_{\frac{A}{2}}^{\infty} w^{p-1} e^{-w} \, dw$$

shall be small compared to 1; e.g., for p=3 and A=200, we get

$$\frac{1}{(p-1)!} \int_{\frac{A}{2}}^{\infty} w^{p-1} e^{-w} \, dw = e^{-100} [100^2 + 200 + 2] ,$$

which will give a small fraction, and the fact is established that almost the whole of the density is concentrated in the domain D^* near the origin. Equation (11.8) would indicate that even for A as small as 10, and p = 3 say D^* accounts for more than 99 per cent of the density. In practice, however, A can be taken larger, consistent with (11.1).

Another point needing clarification is the use of the exponential approximation over the domain $D_1 - D^*$. At this stage the justification is provided by the upper bound to error involved in using $e^{-\frac{n}{2}(m_1+m_2)}$ instead of $[(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}}$ inside D_1, which was worked out in Section 7. The upper bound to error for D_1 was found to be $\frac{R_{D_1}}{n^{p+1}}$. A closer bound can be worked out for the domain $D_1 - D^*$, and it can be shown that it is a constant times the same upper bound multiplied by an integral of the type $\int_0^{A/2} w^p e^{-w} dw$.

This upper bound can be obtained by following the same lines as those followed in Section 7.

12. Summary of Chapter II.

In this chapter we have considered the asymptotic evaluation of the integral

$$(12.1) \qquad I = \iiint_D (m_1m_2-m_3^2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}} dm_1dm_2dm_3,$$

where D is determined by the fact that both factors involved in (12.1) are non-negative, and $0 \leq m_i \leq 1$, i=1,2. Two simplifications used in the evaluation of I are:

(1) D can be split up into two domains, D_1 and D_2, by the plane $m_1+ m_2= 1$. The contribution due to the domain D_2, for which $m_1+ m_2 \geq 1$ is negligible in the limit, in comparison with that of D_1.

(2) The integral over D_1 is evaluated by replacing the factor $[(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}}$ by $e^{-\frac{n}{2}(m_1+m_2)}$. The justification for the approximation thus obtained is provided partly by the probability order of the variables, and partly by the bounds to error found subsequently.

With these simplifications it is proved that

$$(12.2) \qquad I \sim 4\pi(p-2)! \frac{1}{n^p} ,$$

and that the exact value of I can be written as

$$(12.3) \qquad I = \frac{4\pi(p-2)!}{n^p}(1+\epsilon_n) + \frac{J_{D_1}}{n^{p+1}} + \frac{J_{D_2}}{n^2 2^n} + O\left(\frac{1}{n^3 2^n}\right)$$

where the second term is the remaining contribution due to D_1, and the remaining terms give the integral over D_2. Bounds have been found, (7.27), for J_{D_1} and, (9.13), for the integral over D_2. These have been shown to be negligible as compared to the principal term in the value of I, giving $\frac{4\pi(p-2)!}{n^p}$ as an asymptotic approximation to the value of I.

CHAPTER III

ON THE ASYMPTOTIC DISTRIBUTION OF WALD'S CLASSIFICATION STATISTIC IN THE NULL CASE

1. Introduction.

We are dealing with the problem of classifying an individual into one of two groups or populations such that the information regarding the two populations is based on two samples of sizes N_1 and N_2 respectively. One may be called upon to consider the following three situations:

(A) N_1 and N_2 large,

(B) $N_1 + N_2$ or $n(= N_1 + N_2 - 2)$ large,

(C) N_1 and N_2 small.

The study of case A is equivalent to the study of a linear function of normal variates, that is, treating the statistic U, defined in Chapter I, or the linear discriminant function, as normally distributed with means and covariance matrix replaced by their sample estimates to get the mean and variance of the approximating normal distribution. This case has been completely exploited by several workers in this field.

The results available in case C have been summarized in Sections 4 and 5 of Chapter I. The difficulties involved in obtaining the exact sampling distribution of $V = \dfrac{nm_3}{(1-m_1)(1-m_2)-m_3^2}$ from the joint distribution of m_1, m_2 and m_3 being substantial, it makes sense to ask whether it would be possible to get the distribution of V in case B. Obviously the results obtained would not be as exact as one would like to have, but they should be better than the large sample normal approximation of case A. It is thus in the sense of large n that we shall use the words "asymptotic" and "limiting", and it should be noted that the assumption n large is less restrictive than the assumption N_1 and N_2 both large.

In this chapter we shall find the asymptotic moments of a statistic v which will be called Wald's approximate classification statistic, and then use those moments to find the limiting distribution of v, in the null case, separately for even and for odd values of p.

2. Wald's approximate classification statistic and its moments.

From Chapter I we recall that Wald expressed the statistic ultimately as a function of three variables, and stated that

$$V = \frac{nm_3}{(1-m_1)(1-m_2)-m_3^2} \tag{2.1}$$

can be considered as the classification statistic.

We can rewrite V as $nm_3 [1-m_1-m_2+m_1m_2-m_3^2]^{-1}$, and so

$$V = nm_3(1 + \varepsilon) \tag{2.2}$$

by Section 2 of Chapter II. Thus, by a convergence theorem due to Kolmogoroff [25], the distribution of V can be well approximated by the distribution of nm_3 as stated by Wald. There is no loss of generality in considering

$$(2.3) \qquad v = | nm_3 |$$

as the statistic instead of nm_3. We shall refer to this as the approximate classification statistic of Wald, as against the exact statistic V suitable for small samples.

2A. Limiting moments of the statistic.

As a first step in finding ν_k, the kth moment about the origin, we shall discuss briefly the value of the integral

$$(2.3) \quad I^{(k)} = \iiint_D nm_3{}^k (m_1m_2-m_3^2)^{\frac{p-3}{2}} [(1-m_1)(1-m_2)-m_3^2]^{\frac{n-p-1}{2}} dm_1 dm_2 dm_3 \ .$$

If we recall the discussion about the domain D from Section 2, Chapter II, it can be easily seen that the integral can be written as the sum of two integrals over the interiors of the two cones defined by D_1 and D_2. Thus (2.3) can be written as

$$(2.4) \qquad I^{(k)} = I_1^{(k)} + I_2^{(k)}$$

where $I_1^{(k)}$ and $I_2^{(k)}$ denote the values of the integrals over the two cones D_1 and D_2 .

Define

$$(2.5) \qquad I_{11}^{(k)} = n^k \iiint_{D_1} m_3{}^k (m_1m_2-m_3^2)^{\frac{p-3}{2}} e^{-\frac{n}{2}(m_1+m_2)} dm_1 dm_2 dm_3 \ .$$

By the procedure followed in Section 6, Chapter II, we get

$$(2.6) \qquad I_{11}^{(k)} = 2^k \Gamma(\frac{k+1}{2}) \Gamma(\frac{p-1}{2}) \Gamma(\frac{k+p}{2}) (\frac{2}{n})^p [1+\varepsilon]$$

which, for k = 0, gives I_{11} of Chapter II.

By following the methods of Section 7 and 9 of Chapter II, we can show that the upper bound to the error in estimating $I_1^{(k)}$ by $I_{11}^{(k)}$ is of order $\frac{1}{n^{p+1}}$, and an upper bound to the value of $I_2^{(k)}$ is of order $\frac{1}{n^{3/2} 2^n}$.

Thus we can write

$$(2.7) \qquad I^{(k)} = 2^k (\frac{2}{n})^p \, \Gamma(\frac{k+1}{2}) \Gamma(\frac{k+p}{2}) \Gamma(\frac{p-1}{2}) \, [1+\varepsilon] + I_{D_1} + I_{D_2}$$

where

$$(2.8) \qquad I_{D_1} = I_1^{(k)} - I_{11}^{(k)} \quad ,$$

and

$$(2.9) \qquad I_{D_2} = I_2^{(k)} \quad .$$

It should be noted that it is the upper bounds to, and not the exact values of, I_{D_1} and I_{D_2} that are known; and to avoid duplication in their derivation, since they are obtained in exactly the same way as similar bounds were found in Chapter II, we write the results. They are

$$(2.10) \qquad I_{D_1} \leq \Gamma(\frac{k+1}{2}) \Gamma(\frac{p-1}{2}) \Gamma(\frac{k+p}{2}) (k+p)(k+p+1) (\frac{2}{n-p+1})^{k+p+1} n^k$$

and

$$\text{(2.11)} \qquad I_{D_2} \leq \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{n-2}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{p-1}{2})\, n^k}{\Gamma(\frac{n+k-1}{2})\Gamma(p-1)\; 2^{n+k-p}(n-p+k+1)} \; .$$

It is easy to see from (2.7), (2.10) and (2.11) that

$$\text{(A)} \qquad \lim_{n \to \infty} \frac{I_{D_1}}{\min I^{(k)}} = 0 \; , \quad \text{and}$$

$$\text{(B)} \qquad \lim_{n \to \infty} \frac{I_{D_2}}{\min I^{(k)}} = 0 \; ,$$

showing thereby that I_{D_1} and I_{D_2} are negligible in comparison with the principal term in the value of $I^{(k)}$.

Dividing (2.7) by I_{11} we get the expression for the asymptotic moments, namely

$$\text{(2.12)} \qquad \nu_k = 2^k \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+p}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})} [1+\varepsilon] + R_{D_1}(k,n) + R_{D_2}(k,n) \; ,$$

where

$$\text{(2.13)} \qquad R_{D_1}(k,n) = \frac{I_{D_1}}{I_{11}} \leq \frac{\gamma(\frac{k+1}{2})\Gamma(\frac{k+p}{2})(k+p)(k+p+1)}{\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})} \cdot \frac{n^{p+k} 2^{k+1}}{(n-p+1)^{k+p+1}}$$

and

$$\text{(2.14)} \qquad R_{D_2}(k,n) = \frac{I_{D_2}}{I_{11}} \leq \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{n-2}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{n+k-1}{2})\Gamma(p-1)2^{n+k}\; \Gamma(\frac{p}{2})\Gamma(\frac{1}{2})} \cdot \frac{n^{p+k}}{(n-p+k+1)} .$$

We shall rewrite (2.13) and (2.14) as

$$R_{D_1}(k,n) \leq \frac{1}{n}R_1(k,n) \qquad (2.15)$$

and

$$R_{D_2}(k,n) \leq \frac{1}{n}R_2(k,n) \qquad (2.16)$$

,

where

$$R_1(k,n) = \Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{k+p}{2}\right)(k+p)(k+p+1)\,\frac{n^{p+k+1}2^{k+1}}{(n-p-1)^{k+p+1}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p}{2}\right)} \qquad (2.17)$$

and

$$R_2(k,n) = \frac{\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}{\Gamma\left(\frac{n+k-1}{2}\right)\Gamma(p-1)2^{n+k}\quad\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{1}{2}\right)}\cdot\frac{n^{p+k+1}}{(n-p+k+1)} \qquad (2.18)$$

.

We will also write (2.12) as

$$\nu_k = \tilde{\nu}_k(1+\varepsilon)+R_{D_1}(k,n)+R_{D_2}(k,n) \qquad (2.19)$$

.

We will refer to $\tilde{\nu}_k$ as the principal term in the value of ν_k, because, as can be easily verified,

$$\lim_{n \to \infty} \frac{R_{D_1}(k,n)}{\tilde{\nu}_k} = 0 \qquad (2.20)$$

$$\lim_{n \to \infty} \frac{R_{D_2}(k,n)}{\tilde{\nu}_k} = 0 \qquad (2.21)$$

.

To conclude this section, therefore, we can state that (2.19) gives the kth moment ν_k; and, because of (2.20) and (2.21), we can write

$$\nu_k \sim \tilde{\nu}_k \quad . \tag{2.22}$$

3. <u>The asymptotic distribution of</u> v <u>for</u> p = 2m .

In this section we shall find the asymptotic distribution of v for even values of p. By applying the general result we shall also explicitly obtain the distribution for p = 2, 4 and 6.

<u>Lemma 5.1</u>.

$$R_1(k,n) \quad < M_1(k) = \Gamma(\tfrac{k+1}{2})\Gamma(\tfrac{k+p}{2})(k+p)(k+p+1)\,2^{p+2k+2} \tag{3.1}$$

and

$$R_2(k,n) \quad < M_2(k) = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)\Gamma(\frac{p}{2})\Gamma(\frac{1}{2})} \quad \text{for large } n \ , \tag{3.2}$$

where $R_1(k,n)$ and $R_2(k,n)$ are defined in (2.17) and (2.18) .

Proof:

$$R_1(k,n) = \Gamma(\tfrac{k+1}{2})\Gamma(\tfrac{k+p}{2})\,\frac{(k+p)(k+p+1)\quad n^{p+k+1}2^{k+1}}{\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})(n-p+1)^{k+p+1}} \quad . \tag{3.3}$$

The maximum value of $\frac{n}{(n-p-1)}$ is 2 for $n \geq 2p+2$, and $\Gamma(\frac{1}{2})\Gamma(\frac{p}{2}) \geq 1$, therefore the truth of (3.1) is established.

To prove (3.2) we consider

$$(3.4)\quad R_2(k,n) = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)\Gamma(\frac{p}{2})\Gamma(\frac{1}{2})} \; \frac{\Gamma(\frac{n-2}{2})}{\Gamma\,\frac{n+k-1}{2}} \; \frac{n^{p+k+1}}{2^{n+k}\quad(n-p+k+1)} .$$

We first note that $\frac{1}{n-p+k+1} < \frac{2}{n}$ for all large n; and since $\frac{n^{p+k}}{2^{n+k-1}} < 1$ for all large n and $\frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n+k-1}{2})} < 1$ for all $n \geq 5$, (3.2) follows.

Lemma 5.2

The series

$$(3.5)\qquad (1)\quad \Sigma \frac{t^k}{k!} M_1(k)$$

and

$$(3.6)\qquad (2)\quad \Sigma \frac{t^k}{k!} M_2(k)$$

are both convergent.

Proof:

Let u_k denote the kth term of the series. For the series (3.5), we have

$$(3.7)\qquad u_k = \frac{t^k}{k!}\,\Gamma(\frac{k+1}{2})\Gamma(\frac{k+p}{2})(k+p)(k+p+1)\,2^{p+2k+2}$$

The ratio

$$(3.8)\qquad \frac{u_k}{u_{k+1}} = \frac{k+1}{4t}\;\frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+p}{2})(k+p)}{\Gamma(\frac{k+2}{2})\Gamma(\frac{k+p+1}{2})(k+p+2)} .$$

Using Stirling's approximation to factorials, we have

(3.9) $$\lim_{k \to \infty} \frac{u_k}{u_{k+1}} = \lim_{k \to \infty} \frac{k+1}{4t} \left(\frac{2}{k}\right).$$

This simplifies to

(3.10) $$\lim_{k \to \infty} \frac{u_k}{u_{k+1}} = \frac{1}{2t} .$$

The ratio test states that the behavior of a series $\sum_{k=0}^{\infty} u_k$ is determined by the following formula:

(3.11) If $\lim_{k \to \infty} \frac{u_k}{u_{k+1}} = c$
$$\begin{cases} \text{Series converges if } c > 1 , \\ \text{The test fails if } c = 1 , \\ \text{Series diverges if } c < 1 . \end{cases}$$

Application of this formula shows that series (3.5) converges if $t < 1/2$.

Consider now the series (3.6). Since

(3.12) $$u_k = \frac{t^k}{k!} \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(p-1)\Gamma(\frac{p}{2})\Gamma(\frac{1}{2})} ,$$

(3.13) $$\frac{u_k}{u_{k+1}} = \frac{k+1}{t} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k+2}{2})} .$$

(3.14) $$\lim_{k \to \infty} \frac{u_k}{u_{k+1}} = \infty ,$$

so that the series (3.6) converges for all values of t . In particular, therefore, we can say that for $t < \frac{1}{2}$ both the series (3.5) and

(3.6) are convergent.

In

$$\nu_k = \tilde{\nu}_k [1+\varepsilon] + R_{D_1}(k,n) + R_{D_2}(k,n) \tag{3.15}$$

there are three error terms if we approximate ν_k by $\tilde{\nu}_k$. Since the other two are negligible in comparison to the upper bound to $R_{D_1}(k,n)$, it will be enough to consider the contribution of this to $\phi(t)$ the moment generating function of ν_k.

We define

$$\tilde{\phi}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \tilde{\nu}_k \quad , \tag{3.16}$$

then by (2.19)

$$\phi(t) = \tilde{\phi}(t) + \sum_k \frac{t^k}{k!} R_{D_1}(k,n) + \varepsilon_n \tag{3.17}$$

where ε_n is the contribution due to other error terms and is easily seen to be an infinitesimal of an order higher than that of $\frac{1}{n}$.

By virtue of Lemmas 1 and 2, we write (2.17) as

$$|\tilde{\phi}(t) - \phi(t)| \leq \frac{1}{n} \sum_{k=0}^{\infty} \frac{t^k}{k!} R_1(k,n) + \varepsilon_n \tag{3.18}$$

uniformly for all $|t| < |T_0| < \frac{1}{2}$.

Therefore by Paul Levy's theorem [9, p. 96]

$$|F_n(v) - F(v)| < \frac{c}{n} \quad , \tag{3.20}$$

where $F_n(v)$ denotes the sequence of cumulative distribution functions corresponding to $\phi_n(t)$, and $F(v)$ corresponds to $\tilde{\phi}(t)$. Thus we have proved the following theorem:

Theorem 5.

If $F_n(v)$ is the sequence of cumulative distribution functions corresponding to ν_k for large values of n, and $F(v)$ is the distribution function corresponding to $\tilde{\nu}_k$, then given ε, there exists an N_ε, such that $\left| F_n(v) - F(v) \right| < \varepsilon$ for $n > N_\varepsilon$.

Theorem 6.

When p, the number of variables, is even, the asymptotic distribution of v is given by

$$f(v)dv = \sum_{j=1}^{m} b_j f_j(v)\, dv \,, \tag{3.23}$$

where $2m = p$,

$$f_j(v) = \begin{cases} \frac{1}{\Gamma(j)} v^{j-1} e^{-j} & v \geq 0 \\ 0 & \text{otherwise} \end{cases} \,, \tag{3.24}$$

and where the b_j's are suitable constants depending on m.

Proof: Let $p = 2m$.

$$\tilde{\nu}_k = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+2m}{2})}{\Gamma(\frac{1}{2})\Gamma(m)} \quad 2^k \,. \tag{3.26}$$

On expanding the right hand side of (3.26), we get

$$(3.27) \qquad \tilde{\nu}_k = \frac{(k+2m-2)(k+2m-4) \dots (k+4)(k+2)}{2^{m-1}} \cdot \frac{k!}{\Gamma(m)} \quad .$$

The moment generating function for the corresponding distribution is given by

$$\tilde{\phi}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \tilde{\nu}_k \quad ,$$

or

$$(3.28) \qquad \tilde{\phi}(t) = \Sigma \frac{t^k}{\Gamma(m)} \frac{(k+2m-2)(k+2m-4) \dots (k+4)(k+2)}{2^{m-1}} \quad .$$

This can be rewritten as

$$(3.29) \quad \tilde{\phi}(t) = c_{m-1} \frac{d^{m-1}}{dt^{m-1}} \Sigma t^{k+m-1} + c_{m-2} \Sigma t^{k+m-2} + \dots + c_1 \frac{d}{dt} \Sigma t^{k+1} + c_0 \Sigma t^k \quad ,$$

where $c_0 c_1 \dots c_{m-1}$ are constants depending on m and these are obtained by comparing the coefficients of like powers of k in (3.28) and (3.29). The uniqueness of the solution for $c_0 c_1 \dots c_{m-1}$ follows from the fact that each of the expressions (3.28) and (3.29) consists of a factor t^k multiplied by a polynomial of degree $m-1$ in k.

We can write (3.29) as

$$(3.30) \qquad \tilde{\phi}(t) = \sum_{i=1}^{m} c_{m-i} \sum_{k=0}^{\infty} \frac{d^{m-i}}{dt^{m-i}} (t^{k+m-i}) \quad .$$

For further simplification, we write $\sum_{k=0}^{\infty} t^{k+m-i}$ as $t^{m-i}(\sum_{k=0}^{\infty} t^k)$, which can be expressed as $\frac{t^{m-i}}{1-t}$, and the operations of summation and differentiation can be interchanged in the region of convergence of the series, namely $|t| < 1$.

Also, since

$$(3.31) \quad \frac{d^{m-i}}{dt^{m-i}}\left(\frac{t^{m-i}}{1-t}\right) = \frac{d^{m-i}}{dt^{m-i}} \left[\frac{-(1-t^{m-i})+1}{1-t} \right],$$

and further

$$= \frac{d^{m-i}}{dt^{m-i}} \left[-\sum_{\lambda=1}^{m-i-1} t^{\lambda} + \frac{1}{1-t} \right] = \frac{d^{m-i}}{dt^{m-i}} \left(\frac{1}{1-t}\right)$$

(3.30) becomes

$$\tilde{\phi}(t) = \sum_{i=1}^{m} c_{m-i} \frac{d^{m-i}}{dt^{m-i}} \left(\frac{1}{1-t}\right) .$$

This can be rewritten as

$$(3.32) \qquad \tilde{\phi}(t) = \sum_{i=1}^{m} c_{m-i} \frac{(m-i)!}{(1-t)^{m-i+1}} = \sum_{i=1}^{m} \frac{c^{*}_{m-i}}{(1-t)^{m-i+1}} .$$

It is well known that $\frac{1}{(1-t)^a}$ is the moment generating function of

$$(3.33) \qquad f_a(v) = \begin{cases} \frac{1}{\Gamma(a)} v^{a-1} e^{-v} & v \geq 0 \\ 0 & v < 0 \end{cases}$$

Hence we can write the distribution whose moment generating function is (3.32) as

(3.34) $$f(v)dv = \sum_{i=1}^{m} c^*_{m-i} \, f_{m-i}(v) \, dv \quad .$$

This can be expressed in a slightly better notation by writing j for m-i. This completes the proof of theorem 6.

Special Cases.

(i) p = 2.

The kth moment is given by

(3.35) $$\tilde{\nu}_k = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+2}{2})}{\sqrt{\pi}} \; 2^k \quad ,$$

which on simplification gives

(3.36) $$\tilde{\nu}_k = k!$$

The corresponding moment generating function is

(3.37) $$\tilde{\phi}(t) = \sum_{k=0}^{\infty} t^k \quad ,$$

which can be written as

(3.38) $$\tilde{\phi}(t) = \frac{1}{1-t} \quad .$$

From (3.38) we conclude that

(3.39) $$f(v)dv = \begin{cases} e^{-v}dv & \text{if } v \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad .$$

(ii) p = 4.

For p = 4 we have

(3.40) $$\tilde{\nu}_k = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+4}{2})}{\sqrt{\pi}} 2^k = \frac{k+2}{2} \quad k! \quad .$$

The moment generating function for this, namely

(3.41) $$\tilde{\phi}(t) = \sum_{k=0}^{\infty} \frac{t^k}{2}(k+2) \quad ,$$

can be rewritten as

(3.42) $$\tilde{\phi}(t) = \sum_k \frac{d}{dt}\left(\frac{t^{k+1}}{2}\right) + \sum_k \left(\frac{t^k}{2}\right) \quad .$$

This on simplification becomes

(3.43) $$\tilde{\phi}(t) = \frac{1}{2}\left[\frac{1}{(1-t)^2} + \frac{1}{(1-t)} \right] \quad .$$

The distribution of v is therefore given by

(3.44) $$f(v)dv = \frac{1}{2}(ve^{-v} + e^{-v})\, dv \quad .$$

(iii) p = 6.

The moments in this case are given by

(3.45) $$\tilde{\nu}_k = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+6}{2})}{\sqrt{\pi}\,\Gamma(3)} 2^k \quad ,$$

which simplifies to

(3.46) $$\tilde{\nu}_k = \frac{(k+4)(k+2)}{8} \quad k! \quad .$$

The corresponding moment generating function is

$$(3.47) \qquad \tilde{\phi}(t) = \sum_k t^k \frac{(k+4)(k+2)}{8} \quad ,$$

which can also be written as

$$(3.48) \qquad \tilde{\phi}(t) = \frac{1}{8} \sum_k \frac{d^2}{dt^2}(t^{k+2}) + \frac{3}{8} \sum_k \frac{d}{dt}(t^{k+1}) + \frac{3}{8} \sum_k (t^k) \quad .$$

Following the argument used in Theorem 6, this simplifies to

$$(3.49) \qquad \tilde{\phi}(t) = \frac{1}{8} \frac{d^2}{dt^2}\left(\frac{1}{1-t}\right) + \frac{3}{8} \frac{d}{dt}\left(\frac{1}{1-t}\right) + \frac{3}{8} \left(\frac{1}{1-t}\right) \quad ,$$

which gives

$$(3.50) \qquad \tilde{\phi}(t) = \frac{1}{4(1-t)^3} + \frac{3}{8(1-t)^2} + \frac{3}{8(1-t)} \quad .$$

The distribution to which this refers is obviously

$$(3.51) \qquad f(v)dv = \left(\frac{1}{8}v^2e^{-v} + \frac{3}{8}ve^{-v} + \frac{3}{8}e^{-v}\right)dv \quad .$$

4. An integral equation due to Wilks.

S. S. Wilks [55] considers the moments and distributions of some statistical coefficients related to samples from a multivariate normal population, and exhibits a new method of attack. He considers two integral equations which he calls Type A and Type B, and uses their solutions in deriving some now well known distributions. The first result adapted for the present use can be written as follows:

If

$$(4.1) \qquad \int_0^{\infty} v^k f(v)dv = B^k \frac{\Gamma(a_1+k)\Gamma(a_2+k)}{\Gamma(a_1)\Gamma(a_2)} ,$$

where k's and a's are real and positive and B and f(v) are independent of k, then

$$(4.2) \qquad f(v) = \frac{B^{-a_2} v^{a_2-1}}{\Gamma(a_1)\Gamma(a_2)} \int_0^{\infty} x^{a_1-a_2-1} e^{-x-\frac{v}{Bx}} dx .$$

The integral in (4.2) can be expressed in elementary functions when a_1-a_2 is half of an odd integer; and this case, as we shall see later, corresponds to the distribution of v defined in (2.3) for even values of p. If, however, a_1-a_2 is an integer, the integral is a Bessel function and this situation arises if p is odd. Before using (4.2) in finding the distribution of v, we shall, for the sake of completeness, add a note on Bessel functions.

5. A note on Bessel functions.

The equation

$$(5.1) \qquad z^2 \frac{d^2w}{dz^2} + z\frac{dw}{dz} + (z^2-n^2) = 0$$

is called Bessel's differential equation of order n, and Bessel functions are defined with reference to this equation. Its only singularities are at $z = 0$ and $z = \infty$.

A solution in series of (5.1) near the origin can be obtained by supposing that $w = z^c \sum_{i=0}^{\infty} a_i z^i$ is a solution. It is found that the discussion can be divided into four cases.

(a) $n \neq i$, $n \neq \frac{2i+1}{2}$ where i stands for an integer.

In this case there are two independent solutions:

$$J_n(z) \quad \text{and} \quad J_{-n}(z) \quad , \tag{5.2}$$

where

$$J_n(z) = \sum_{r=0}^{\infty} \frac{(-1)^r}{\Gamma(r+1)\Gamma(n+r+1)} \cdot \left(\frac{z}{2}\right)^{n+2r} \quad , \tag{5.3}$$

and is analytic for all values of z except possibly $z = 0$. It is called Bessel's function of the first kind.

(b) If $n = i$ an integer,

$J_n(z)$ and $J_{-n}(z)$ are two linearly dependent integrals satisfying the relation

$$J_{-n}(z) = (-1)^n J_n(z) \quad .$$

In this case the solutions are

$$J_n(z) \quad \text{and} \quad Y_n(z) \tag{5.4}$$

where

$$Y_n(z) = J_n(z) \log z - \frac{1}{2} \sum_{r=0}^{n-1} \frac{(n-r-1)!}{r!} \left(\frac{z}{2}\right)^{-n+2r} \tag{5.5}$$

$$- \frac{1}{2} \sum_{r=0}^{\infty} \frac{(-1)^r}{\Gamma(r+1)\Gamma(n+r+1)} \left(\frac{z}{2}\right)^{n+2r} [\phi(r) + \phi(n+r)] \ ,$$

where

(5.6) $\phi(r) = \frac{1}{1} + \frac{1}{2} + \ldots \frac{1}{r}$, $r = 1, 2, 3, \ldots$ and $\phi(0) = 0$.

$Y_n(z)$ is called Bessel's function of the second kind.

(c) If $n = \frac{2i+1}{2}$

(5.7) $J_n(z)$ and $J_{-n}(z)$ are two linearly independent integrals.

(d) If $n = 0$

(5.8) $$J_0(z) = \sum_{i=0}^{\infty} \frac{(-1)^r}{[\Gamma(r+1)]^2} \left(\frac{z}{2}\right)^{2r}$$

and

(5.9) $$Y_0(z) = J_0(z) \log z + \frac{z^2}{2} - \frac{z^4}{2^2 \cdot 4^2}\left(1 + \frac{1}{2}\right) + \frac{z^6}{2^2 \cdot 4^2 \cdot 6^2}\left(1 + \frac{1}{2} + \frac{1}{3}\right) - \ldots$$

are the two solutions.

$Y_0(z)$ is Bessel's function of the second kind of order zero.

Sometimes a function $G_n(z)$ is used instead of $J_{-n}(z)$ or $Y_n(z)$ asthe second solution of the Bessel's differential equation. It is defined by

(5.10) $$G_n(z) = \frac{\pi}{2 \sin n\pi} \left[J_{-n}(z) - e^{-in\pi} J_n(z) \right],$$

where n is not an integer; and

(5.11) $$G_n(z) = \frac{\partial}{\partial n} \left[\frac{J_{-n}(z) - e^{in\pi} J_n(z)}{2 \cos n\pi} \right],$$

when n is an integer.

If we put $z = iv$ in (5.1), the result is

$$v^2 \frac{d^2w}{dv^2} + v \frac{dw}{dv} - (n^2 + v^2) w = 0 \quad , \tag{5.12}$$

which is known as Bessel's transformed equation. Two solutions of (5.12), namely

$$I_n(v) = i^{-n} J_n(iv) = \sum_{r=0}^{\infty} \frac{1}{\Gamma(r+1)\Gamma(n+r+1)} \left(\frac{v}{2}\right)^{n+2r} \tag{5.13}$$

and

$$K_n(v) = i^n G_n(iv) = \frac{\pi}{2 \sin n\pi} \left[I_{-n}(v) - I_n(v) \right] , \tag{5.14}$$

are called respectively the modified Bessel functions of the first and second kinds of order n.

If n is a positive integer,

$$I_{-n}(v) = I_n(v) \quad , \tag{5.15}$$

and

$$K_n(v) = \lim_{\varepsilon \to 0} K_{n+\varepsilon}(v) \quad . \tag{5.16}$$

6. Distribution of v for odd values of p.

In (2.22) we proved that

$$\nu_k \sim \tilde{\nu}_k = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{k+p}{2})}{\Gamma(\frac{1}{2})\,\Gamma(\frac{p}{2})} \, 2^k \quad , \tag{6.1}$$

which gives only the principal term in the value of ν_k. Since we are not using the exact value but only an asymptotic approximation for the value of $\int_0^\infty v^k f(v)dv$, the results obtained by the use of (4.1) and (4.2) can not be presented as being final. Moreover, since the paper of Wilks referred to in Section 4 depends heavily on Stekloff's paper on the theory of closure as applied to the problem of moments [47], which is not easily available, the distribution for odd values of p is here presented on a heuristic basis. It may turn out to be the correct distribution, but it has to be left for further discussion and rigorization.

Consider again the equation (6.1). If

$$u = v^2 \quad , \tag{6.2}$$

then

$$E(u^k) = E(v^{2k}) = \frac{\Gamma(k+\frac{1}{2})\Gamma(k+\frac{p}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})} 2^{2k} \quad . \tag{6.3}$$

Comparing (5.3) with (4.1), we have

$$B = 4,\ a_1 = \frac{1}{2} \text{ and } a_2 = \frac{p}{2} \ . \tag{6.4}$$

In this case (4.2) gives

$$f(u)du = \frac{4^{-\frac{p}{2}} u^{\frac{p-2}{2}}}{\Gamma(\frac{1}{2})\ \Gamma(\frac{p}{2})} du \int_0^\infty x^{-\frac{p+1}{2}} e^{-x-\frac{u}{4x}} dx \quad . \tag{6.5}$$

Putting

(6.6) $$u = v^2 \quad \text{and} \quad p = 2m + 1 \quad ,$$

we get

(6.7) $$f(v)dv = \frac{v^{2m}dv}{2^{2m}\Gamma(\frac{1}{2})\Gamma(m)} \cdot \int_0^{\infty} x^{-m-1} e^{-(x+\frac{v^2}{4x})} dx \quad .$$

According to Watson [52, p. 183] the integral $\int_0^{\infty} x^{-m-1} e^{-x-\frac{v^2}{4x}} dx$, has been studied by Poisson, Glaisher, Kapteyn and others. The result stated in Watson is

(6.8) $$K_m(v) = \frac{1}{2}\left(\frac{v}{2}\right)^m \int_0^{\infty} x^{-m-1} e^{-(x+\frac{v^2}{4x})} dx \quad .$$

This reduces the distribution of v to the form

(6.9) $$f(v)dv = \frac{v^m}{2^{m-1}\sqrt{\pi}\,\Gamma(m)} K_m(v) \quad .$$

Putting $m = 0, 1, 2, \ldots$ in this, we get the distribution of v for $p = 1, 3, 5, \ldots$.

7. The use of a differential equation in the evaluation of an integral.

In Section 6 we found that

(7.1) $$f(v)dv = \frac{v^{p-1}dv}{2^{p-1}\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})} \int_0^{\infty} x^{-\frac{p+1}{2}} e^{-x-\frac{v^2}{4x}} dx \quad ,$$

where p is the number of variates in the underlying normal distributioms.

A known technique for evaluating

(7.2) $$\phi(v) = \int_0^{\infty} x^{-\frac{p+1}{2}} e^{-x-\frac{v^2}{4x}} dx$$

is as follows: Put $x = \frac{z^2}{2}$ in (7.2) to get

(7.3) $$\phi(v) = 2^{\frac{p+1}{2}} \int_0^{\infty} \frac{1}{z^p} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} dz \quad .$$

Now we define

(7.4) $$\psi(v) = \int_0^{\infty} z^{\frac{1}{p-2}} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} dz \quad ,$$

where

$$v \geq 0 \; .$$

Since the conditions for differentiation under the integral sign are satisfied, we differentiate $\psi(v)$ with respect to v and get

(7.5) $$\psi'(v) = -v \int_0^{\infty} \frac{1}{z^p} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} dz \quad .$$

Similarly

(7.6) $$\psi''(v) = -\int_0^{\infty} \frac{1}{z^p} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} + v^2 \int_0^{\infty} \frac{1}{z^{p+2}} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} dz \quad .$$

Now

$$(7.7)\quad \int_0^{\infty} \left(-\frac{1}{z^{p-2}} - \frac{p-1}{z^p} + \frac{v^2}{z^{p+2}}\right) e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} dz$$

$$= \int_0^{\infty} \frac{d}{dz} \left[\frac{1}{z^{p-1}} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})}\right] dz = \frac{1}{z^{p-1}} e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})} \Big|_0^{\infty}$$

which is equal to zero identically therefore, using this identity we obtain, from (7.4), (7.5) and (7.6), the following differential equation

$$(7.8)\qquad v\,\psi''(v) + (p-2)\,\psi'(v) - v\,\psi(v) = 0 \quad .$$

The value of $\phi(v)$ can be found by using the solution of this and the fact that

$$(7.9)\qquad \phi(v) = -\frac{\psi'(v)}{v} \cdot 2^{\frac{p+1}{2}} .$$

8. The asymptotic distribution of v for even and odd values of p.

We shall, in this section, derive the distributions of v again by starting with the result,

$$(8.1)\qquad f(v)dv = \frac{v^{p-1}dv}{2^{p-1}\Gamma(\frac{1}{2})\Gamma(\frac{p}{2})} \int_0^{\infty} x^{-\frac{p+1}{2}} e^{-x-\frac{v^2}{4x}} dx \quad , \quad \text{and}$$

by evaluating the integral involved by the help of (7.8) and (7.9). We divide the discussion into two cases.

Case A. $p = 2, 4, \ldots$

(A_1) Let $p = 2$

The differential equation (7.8), in this case, reduces to

(8.2) $$(D^2-1)\,\Psi(v) = 0$$

where the symbol D stands for the operation of differentiation.

(8.3) $$\Psi(v) = Ae^{v} + Be^{-v}$$

is the solution of (8.2). Also for $p = 2$

(8.4) $$\Psi(v) = \int_0^\infty e^{-\frac{1}{2}(z^2+\frac{v^2}{z^2})}\,dz$$

by definition.

This gives

(8.5) $$\Psi(0) = \sqrt{\frac{\pi}{2}} \quad \text{and} \quad \Psi(\infty) = 0$$

Thus

(8.6) $$\Psi(v) = \sqrt{\frac{\pi}{2}}\, e^{-v}$$

and hence

(8.7) $$-\frac{\Psi(v)}{v}\,2^{3/2} = 2\sqrt{\pi}\,\frac{e^{-v}}{v} = \phi(v)$$

where $\phi(v)$ stands for the integral occuring in (8.1).

Hence we have the result

(8.9) $$f(v)dv = e^{-v}\,dv \quad .$$

(A_2) Let $p = 4$.

Here

$$(8.9) \qquad \Psi(v) = \int_0^\infty \frac{1}{z^2} e^{-\frac{1}{2}(z^2 + \frac{v^2}{z^2})} dz ,$$

but from (8.4) and (8.6)

$$(8.10) \qquad \int_0^\infty e^{-\frac{1}{2}(z^2 + \frac{v^2}{z^2})} dv = \sqrt{\frac{\pi}{2}}\, e^{-v} .$$

Differentiating both sides of (8.10) with respect to v and dividing by $-v$, we get

$$(8.11) \qquad \Psi(v) = \sqrt{\frac{\pi}{2}}\, \frac{e^{-v}}{v} .$$

Thus

$$(8.12) \qquad \phi(v) = \frac{-\Psi'(v)}{v}\, 2^{5/2} = 4\sqrt{\pi}\, \frac{ve^{-v}+e^{-v}}{v^3} .$$

Hence from (8.1)

$$(8.13) \qquad f(v)dv = \frac{1}{2}(e^{-v} + ve^{-v})\, dv$$

(A_3) <u>p = 6</u>.

Here

$$(8.14) \qquad \Psi(v) = \int_0^\infty \frac{1}{z^4} e^{-\frac{1}{2}(z^2 + \frac{v^2}{z^2})} dz .$$

Using the reasoning of example 2, we get

$$(8.15) \qquad \Psi(v) = \sqrt{\frac{\pi}{2}}\, \frac{ve^{-v}+e^{-v}}{v^3} .$$

Therefore

$$\phi(v) = 2^3 \sqrt{\pi}\ \frac{v^2e^{-v}+3ve^{-v}+3e^{-v}}{v^5} \qquad (8.16)$$

This value substituted in (8.9) gives the following distribution for $p = 6$.

$$f(v)\,dv = \frac{3e^{-v}+3ve^{-v}+v^2e^{-v}}{8}\,dv \quad .$$

The process can obviously be carried on to get the distribution of v for all even values of p.

Case B. $p = 3, 5, \ldots$

(B_1) Let $p = 3$.

The differential equation satisfied by $\psi(v)$ in this case reduces to

$$v\,\psi''(v) + \psi'(v) - v\,\psi(v) = 0 \quad , \qquad (8.17)$$

which is the modified Bessel equation of order zero, and is satisfied by

$$\psi(v) = K_0(v) \quad .$$

Therefore

$$\phi(v) = -\frac{\psi'(v)}{v}\,2^2 = -4\,\frac{K_0'(v)}{v} \quad . \qquad (8.18)$$

But

$$K_0'(v) = -K_1(v) \qquad (8.19)$$

(see for instance, Watson [52, p. 79])

Therefore

$$\phi(v) = \frac{4K_1(v)}{v} \tag{8.20}$$

Substitution of this in (8.1) gives

$$f(v)dv = \frac{2vdv}{\pi} K_1(v) \tag{8.21}$$

(B_2) Let $p = 5$.

Here

$$\Psi(v) = \int_0^\infty \frac{1}{z^3} e^{-\frac{1}{2}(z^2 + \frac{v^2}{z^2})} dz \quad . \tag{8.22}$$

But from (B_1)

$$\int_0^\infty \frac{1}{z} e^{-\frac{1}{2}(z^2 + \frac{v^2}{z^2})} dz = K_0(v) \quad . \tag{8.23}$$

Hence, on differentiating with respect to v and transposing suitable factors, we get

$$\Psi(v) = \frac{-K_0'(v)}{v} = \frac{K_1(v)}{v} \quad . \tag{8.24}$$

Thus

$$\phi(v) = \frac{-vK_1'(v) - K_1(v)}{v^3} \quad , \tag{8.25}$$

which, by using the formula

$$vK_n'(v) - nK_n(v) = -vK_{n+1}(v) \tag{8.26}$$

gives

(8.27) $$\phi(v) = \frac{K_2(v)}{v^2} ,$$

and consequently

(8.28) $$f(v)dv = \frac{v^2}{12\pi} K_2(v)dv$$

as the distribution for $p = 5$.

This process can obviously be continued to obtain the asymptotic distribution of v for all odd values of p.

This section also shows that we get the same distribution of v for $p = 2m$ by the two methods, namely

(1) The use of the moment generating function ,

(2) The application of the integral equation given in Section 4.

9. Note on the construction of tables.

Case A (When $p = 2m$)

The distribution of v in this case is

(9.1) $$f(v)dv = \sum_{j=1}^{m} b_j f_j(v)dv ,$$

where

$$f_j(v) = \begin{cases} \frac{1}{\Gamma(j)} v^{j-1} e^{-j} & v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and where the b_j's are constants which can be found for any given integral value of m.

The evaluation of the integrals of the type $\int_x^{\infty} f(v)dv$ can obviously be made to depend on tables of χ^2 distribution with even degrees of freedom. For illustration it will be enough to consider the cases when $p = 2$ and $p = 4$.

(9.2) (A_1) <u>For $p = 2$</u>,

$$f(v)dv = e^{-v}dv$$

and the substitution $v = \frac{\chi^2}{2}$ shows that

$$P(v > \tfrac{a}{2}) = P(\chi^2 > a) \tag{9.3}$$

which gives the method of tabulating areas for the distribution of v.

In this particular case it may be more convenient to use the tables of exponential function.

(A_2) <u>$p = 4$</u>.

Here

$$f(v)dv = \tfrac{1}{2}(e^{-v} + ve^{-v})dv \quad . \tag{9.4}$$

Putting $v = \frac{\chi^2}{2}$, this becomes

$$f(\chi^2)d\chi^2 = \frac{1}{2}\left[\frac{1}{2}e^{-\frac{\chi^2}{2}}d(\chi^2) + \frac{1}{4}\chi^2 e^{-\frac{\chi^2}{2}}d(\chi^2)\right] \quad . \tag{9.5}$$

The two frequency functions inside the square brackets are χ^2 frequency functions for two and four degrees of freedom. Consider the following table giving tail areas for these distributions.

(9.6)

ν	4.00	4.2	4.4	4.6	4.8	5.0	5.2
2	.13534	.12246	.11080	.10026	.09072	.08209	.07427
4	.40601	.37962	.35457	.33085	.30844	.28730	.26739

from table 7, Pearson and Hartley [35]. Averaging these as suggested by (9.5), we have the following table for $P(\chi' \geq x) = P(v > \frac{x}{2})$,

x	4	4.2	4.4	4.6	4.8	5.0	5.2
p	.27068	.25104	.23269	.21556	.19958	.18470	.17083

From this table it is possible by linear interpolation or by using the formulae for interpolation when the arguments are not equally spaced, to find the values of x corresponding to $P = .25$, $P = .20$ etc.

Similar remarks apply to the construction of tables for $p = 6$, 8, ...

Case B. $p = 2m + 1$.

For this case we proved in Section 6 that

$$f(v)dv = C\, v^m K_m(v)dv$$

Tables for these distributions can be constructed by using the series for $K_m(v)$ and integrating term by term.

10. Summary of Chapter III.

In this chapter we have discussed the distribution of $v = |nm_3|$ for large values of n. The kth moment $E(v^k)$ is found in Section 3 by following the methods of integration developed in Chapter II. These moments have been used in finding the asymptotic distribution of v for even values of p by the help of the corresponding moment generating function. For obtaining the large sample distribution of v for odd values of p, use has been made of an integral equation due to S. S. Wilks.

CHAPTER IV

AN ASYMPTOTIC SERIES EXPANSION FOR THE DISTRIBUTION OF $w = m_3$ IN THE NULL CASE

1. Introduction.

Harter [18] has obtained the distribution of m_3 as a double series by starting with the joint distribution of m_1, m_2 and m_3 of Wald in the special case when $\rho_i = 0 = \xi_i$, which we call the null case, and which has been the subject of our discussion in the preceeding chapters. The series obtained by Harter would present difficulties in practical applications, since in any practical situation the number n, which is determined by the sizes of the two samples, will not be very small. For large n the investigator wishes to use that distribution of m_3 in which the ratio of each term after the first to the preceeding term is of order n^{-1}. It is also obvious that the main point in getting such a series is to obtain terms beyond the first. Of these, however, the second and third approximations are of chief interest and are doubtless easier to calculate than any of those of higher order. Because of these considerations in this chapter we shall obtain the first three terms in the distribution of $w = |m_3|$ as an asymptotic series. For the first approximation the constant of integration will be found, and the method of finding the tail areas for the construction of tables will also be discussed.

It might be noted that the statistic w is $\frac{1}{n}$ times the statistic v defined in Chapter II. Towards the end we shall also compare the result of this chapter with that of Chapter III.

2. <u>An asymptotic series for the distribution.</u>

We consider the joint probability distribution of m_1, m_2 and w, which is the same as the probability distribution of m_1, m_2 and m_3 except for the constant of integration because

$$f(m_1, m_2, m_3) = f(m_1, m_2, -m_3) \quad .$$

Let C denote the constant of integration. Then

(2.1) $f(m_1, m_2, w)dm_1dm_2dw =$

$$C(m_1m_2-w^2)^{\frac{p-3}{2}}\left[(1-m_1)(1-m_2)-w^2\right]^{n-p-1}dm_1dm_2dw \quad .$$

The region of integration is determined by

$$\text{(2.2)} \qquad D = \begin{array}{ll} m_1m_2-w^2 & \geq 0 \\ (1-m_1)(1-m_2)-w^2 & \geq 0 \\ 1 \geq m_1 & \geq 0 \\ 1 \geq m_2 & \geq 0 \end{array} ,$$

which also determines the range $0 \leq w \leq \frac{1}{2}$ for the variate $w = |m_3|$.

To integrate with respect to m_1 and m_2, we shall keep w fixed, and put $m_1 = x + y$ and $m_2 = x - y$. This gives

(2.3) $f(x,y,w)dxdydw = 2C(x^2-y^2-w^2)^{\frac{p-3}{2}}\left[(x-1)^2-y^2-w^2\right]^{\frac{n-p-1}{2}} dxdydw \quad .$

For fixed w, the two expressions in the brackets in $f(x,y,w)$ are zero on hyperbolas in the (x,y)-plane. Moreover $x + y = 0$ and $x - y = 0$ are the asymptotes of $x^2-y^2=w^2$, and $x + y - 1 = 0$ and $x - y - 1 = 0$ are the asymptotes of the hyperbola $(x - 1)^2 - y^2 = w^2$. The region of integration for x and y is thus the area enclosed by the two hyperbolas and is shown in the figure on the adjoining page.

The coordinates of the points of intersection A and B of the two hyperbolas are

$$A = \left[\frac{1}{2}, \left(\frac{1}{4} - w^2\right)^{\frac{1}{2}} \right] .$$

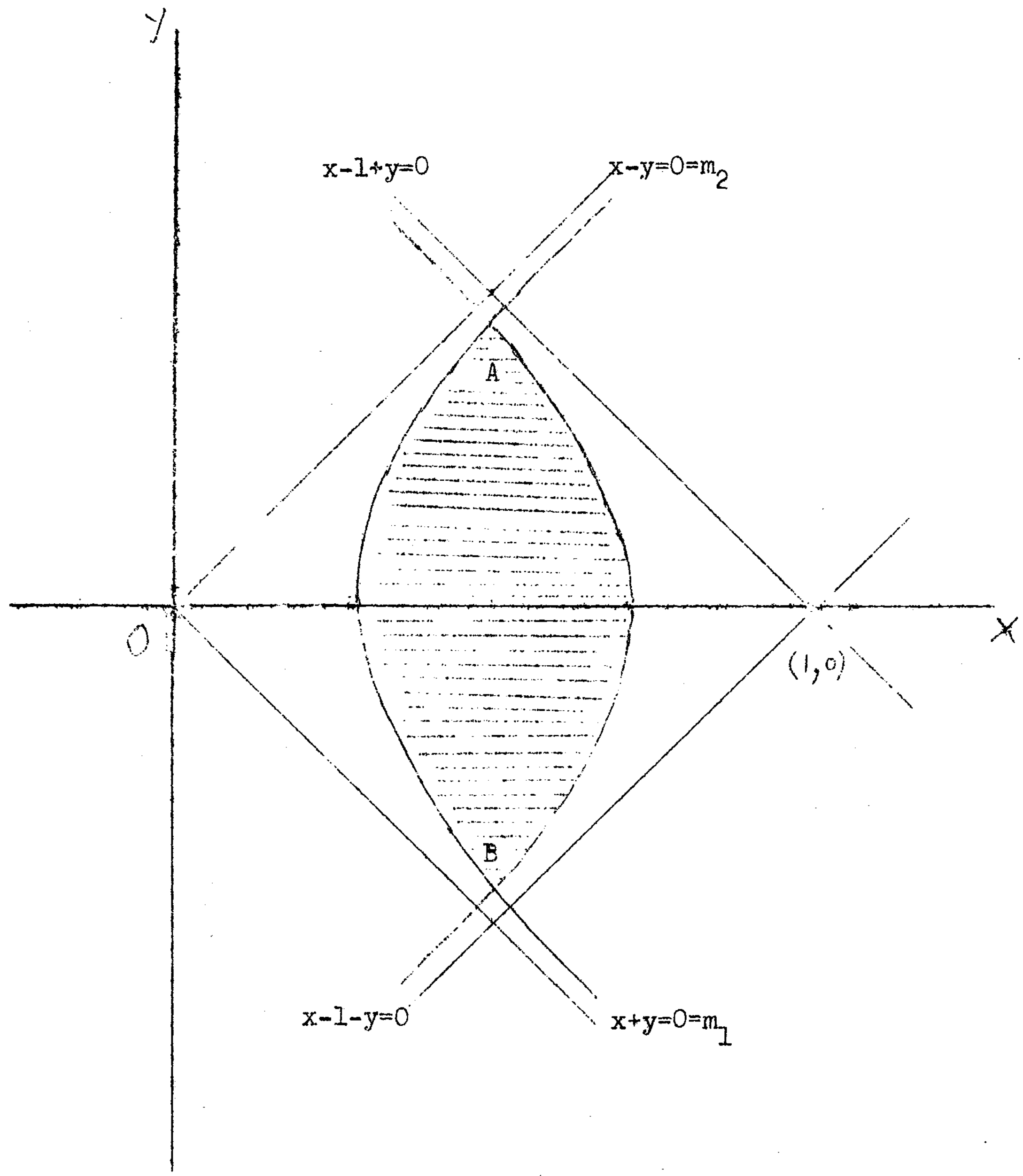
$$B = \left[\frac{1}{2}, -\left(\frac{1}{4} - w^2\right)^{\frac{1}{2}} \right] .$$

The probability distribution of w will be given by the following double integral:

$$(2.4) \quad f(w)dw = 2C \int_{y=-\sqrt{\frac{1}{4}-w^2}}^{\sqrt{\frac{1}{4}-w^2}} \int_{x=\sqrt{y^2+w^2}}^{1-\sqrt{y^2+w^2}} (x^2-y^2-w^2)^{\frac{p-3}{2}} \left[(x - 1)^2-y^2 - w^2\right]^{\frac{n-p-1}{2}} dx\, dy\, dw .$$

Put

$$(2.5) \qquad \frac{p-3}{2} = r$$


Y
x-1+y=0
x-y=0=m2
A
O
(1,0)
X
B
x-1-y=0
x+y=0=m1

and

$$\frac{n-p-1}{2} = q \quad .$$

Also replace the positive root $\sqrt{y^2+w^2}$ by a .

To perform the integration with respect to x, we shall suppose y to be constant. Using (2.5) and by noticing the symmetry of the integrand in y, we can write (2.4) as $f(w)dw$, where

$$(2.6) \quad f(m) = 4c \int_{y=0}^{\sqrt{\frac{1}{4}-w^2}} \int_{x=a}^{1-a} (x^2-a^2)^r \left[(r-1)^2-a^2 \right]^q dx\, dy \; .$$

Let

$$(2.7) \qquad (1-v)^2 = \frac{(x-1)^2-a^2}{1-2a} \quad .$$

This transformation sets up a one-to-one correspondence between the values of x and the values of v. Furthermore, as x increases from a to $1-a$, v increases monotonically from zero to one.

From (2.7) we have the following:

$$(2.8) \qquad (x-1)^2-a^2 = (1-2a)(1-v)^2$$

$$(2.9) \qquad x = 1 - \sqrt{a^2+(1-2a)(1-v)^2}$$

$$(2.10) \quad x^2-a^2 = \frac{2a(1-2a)}{1-a} v \left[1 + \frac{1-3a+a^2}{2a(1-a)} v - \frac{a(1-2a)}{2(1-a)^2} v^2 + \dots \right]$$

$$(2.11) \quad dx = \frac{1-2a}{1-a}(1-v)\left[1- \frac{2(1-2a)}{(1-a)^2}\, v + \frac{1-2a}{(1-a)^2}\, v^2 \right]^{-\frac{1}{2}} dv \, .$$

To examine the convergence of the series in v which will be obtained as a result of this transformation, we regard v as a complex variable and equate to zero the quantity under the radical sign in (2.9). Thus, if v_0 denotes a singularity, then $(1-v_0)^2 = \frac{-a^2}{1-2a}$, which gives

$$(2.12) \qquad v_0 = 1 \pm \frac{ia}{\sqrt{1-2a}} \, .$$

This shows that the two singularities are situated on the line parallel to the imaginary axis at unit distance and are equidistant from the point 1. Also

$$(2.13) \qquad v_0{}^2 = \frac{(1-a)^2}{1-2a}$$

or

$$v_0 = \frac{1-a}{\sqrt{1-2a}} \, .$$

Both the singularities lie outside the unit circle around the origin, since $\frac{(1-a)^2}{1-2a} > 1$ because $a^2 > 0$.

Using the transformation from x to v, we get

$$(2.14)\quad f(w) = 2^{r+2}C \int_{y=0}^{\sqrt{\frac{1}{2} - w^2}} \frac{(1-2a)}{(1-a)^{r+1}}\, a^r \int_{v=0}^{1} v^r (1-v)^{2q+1}$$

$$\left[1 + \frac{1-3a+a^2}{2a(1-a)}\, v - \frac{a(1-2a)}{2(1-a)^2}\, v^2 + \ldots \right]^r .$$

$$\left[1 - \frac{2(1-2a)}{(1-a)^2}\, v + \frac{1-2a}{(1-a)^2}\, v^2 \right]^{-\frac{1}{2}} dv\, dy \quad .$$

We write (2.14) after expanding the last two binomials, but omitting terms involving cubes and higher powers of v since they will not affect the first three terms of the desired asymptotic series.

This gives

$$(2.15)\quad f(w) = 2^{r+2}C \int_{y=0}^{\sqrt{\frac{1}{4} - w^2}} \frac{(1-2a)^{q+r+1}}{(1-a)^{r+1}}\, a^r$$

$$\int_{v=0}^{1} v^r (1-v)^{2q+1} \left\{ 1 + r\, \frac{(1-3a+a^2)}{2a(1-a)}\, v + \right.$$

$$\left. \frac{r}{2(1-a)^2} \left[\frac{(r-1)(1-3a+a^2)^2}{4a^2} - a(1-2a) \right] v^2 + \ldots \right\}$$

$$\left[1 + \frac{1-2a}{(1-a)^2}\, v - \frac{(a^2+4a-2)(1-2a)}{2(1-a)^4}\, v^2 \ldots \right] dv\, dy \quad .$$

This can further be reduced to

$$(2.16)\qquad f(w) = 2^{r+2}C \int_{y=0}^{\sqrt{\frac{1}{4}-w^2}} \frac{(1-2a)^{2+r+1}}{(1-a)^{r+1}}\, a^r$$

$$\int_{v=0}^{1} v^r(1-v)^{2q+1} \Big\{ 1+v\Big[\frac{1-2a}{(1-a)^2} + \frac{r(1-3a+a^2)}{2a(1-a)}\Big]$$

$$+v^2 \Big[\frac{r(r-1)(1-3a+a^2)^2}{a^2(1-a)^2} - \frac{ra(1-2a)}{2(1-a)^2} -$$

$$\frac{(a^2+4a-2)(1-2a)}{2(1-a)^4} + \frac{r(1-3a+a^2)(1-2a)}{2a(1-a)^3}\Big] + \cdots \Big\}\, dv\, dy \;.$$

Integration with respect to v after replacing q and r by their values in terms of n and p, gives

$$(2.17)\quad f(w)=2^{\frac{p+1}{2}}C \int_y \frac{(1-2a)^{\frac{n-2}{2}}a^{\frac{p-3}{2}}}{(1-a)^{\frac{p-1}{2}}} \; \frac{\Gamma(n-p+1)\Gamma(\frac{p-1}{2})}{\Gamma(\frac{2n-p+1}{2})}$$

$$\Big\{ 1+ \Big[\frac{1-2a}{(1-a)^2} + \frac{(p-3)(1-3a+a^2)}{4a(1-a)}\Big]\frac{p-1}{2n-p+1} + \Big[\frac{(p-3)(p-5)(1-3a+a^2)^2}{32a^2(1-a)^2}$$

$$- \frac{(p-3)a(1-2a)}{4(1-a)^2} - \frac{(a^2+4a-2)(1-2a)}{2(1-a)^4} + \frac{(p-3)(1-3a+a^2)(1-2a)}{2a(1-a)^3}\Big].$$

$$\frac{(p^2-1)}{(2n-p+1)(2n-p+3)} + \quad \ldots . \_7 \Big\} \, dy \, ,$$

in which a is an abbreviation for $\sqrt{y^2+w^2}$. We shall write

$$C_1 = 2^{\frac{p+1}{2}} \, C \frac{\Gamma(n-p+1)\Gamma(\frac{p-1}{2})}{\Gamma(\frac{2n-p+1}{2})} \quad . \tag{2.18}$$

To integrate with respect to y we make the transformation

$$z = \frac{2w - 2\sqrt{y^2+w^2}}{2w-1} \quad . \tag{2.19}$$

The limits of integration for z will be zero and one, since those of y are zero and $\sqrt{\frac{1}{4}-w^2}$. Also z is a monotonically increasing function of y. This transformation will change the integrand essentially into the product of two factors, one of which is a high power of 1-z and the other a series of ascending powers of z. Thus (2.19) will change the integral into a sum of beta functions suitable for giving an asymptotic series for the distribution of w.

To effect this transformation we have to find the values of various factors involved, and we have

$$(y^2+w^2)^{\frac{1}{2}} = w \, \_/ \, 1 + \frac{1-2w}{2w} z \, \_7 \quad , \tag{2.20}$$

(2.21) $$y^2 + w^2 = w^2 \left[1 + \frac{1-2w}{w} z + \frac{(1-2w)^2}{4w^2} z^2 \right] ,$$

(2.22) $$1 - (y^2+w^2)^{\frac{1}{2}} = (1 - w)\left[1 - \frac{1-2w}{2(1-w)} z \right] ,$$

(2.23) $$1-2(y^2 + w^2)^{\frac{1}{2}} = (1 - 2w)(1 - z) ,$$

and

(2.24) $$\frac{dy}{dz} = \frac{(1-2w)^{\frac{1}{2}}\left[2w + (1-2w)z \right]}{2z^{\frac{1}{2}}\left[4w+(1-2w)z \right]^{\frac{1}{2}}}$$

$$= \frac{w^{\frac{1}{2}}(1-2w)^{\frac{1}{2}}}{2} z^{-\frac{1}{2}}\left[1+ \frac{1-2w}{2w} z \right]\left[1 + \frac{1-2w}{4w} z \right]^{-\frac{1}{2}} .$$

The singularities z_1, z_2 and z_3 of the resulting series in z are determined by the equations

(2.25) $$1 + \frac{1-2w}{2w} z = 0$$

(2.26) $$1 - \frac{1-2w}{2(1-w)} z = 0$$

and

(2.27) $$1 + \frac{1-2w}{4w} z = 0$$

respectively.

From these we have

(2.28) $$z_1 = \frac{2w}{2w-1}$$

(2.29) $$z_2 = \frac{2(1-w)}{1-2w}$$

and

(2.30) $$z_3 = \frac{4w}{2w-1} \quad .$$

Since the range of w is from 0 to $\frac{1}{2}$, we find from the above three equations

(2.31) $$-\infty < z_1 < 0$$

(2.32) $$2 < z_2 < \infty$$

and

(2.33) $$-\infty < z_3 < 0 \quad .$$

In otherwords, two of the singularities lie on the negative half of the real axis and one on the positive side in the z plane. To be able to get a convergent series in z we have to make sure that these singularities do not lie in the unit circle around the origin. To examine this, we proceed to find the range of values of w for which $z > 1$.

(i) $|z_1| > 1$ if $\frac{2w}{1-2w} > 1$ or if $w > \frac{1}{4}$

(ii) $|z_2| > 1$ if $\frac{2(1-w)}{1-2w} > 1$ or if $2 > 1$ which is true

(iii) $|z_3| > 1$ if $\frac{4w}{1-2w} > 1$ or if $w > \frac{1}{6}$.

These investigations indicate that the resulting series in z will converge for $w > \frac{1}{4}$, which does not cover the whole range of values of w, which is zero to $\frac{1}{2}$. We shall, however, proceed to make this transformation and subsequently find the probability distribution of w as a series of powers of $\frac{1}{n}$. Even the first approximation of the resulting series will be shown to give close results, especially for finding the right hand tail areas.

Making the transformation (2.19) in (2.17), we get:

$$(2.34)\ f(w)=C_1 \frac{(1-2w)^{\frac{n-1}{2}} w^{\frac{p-2}{2}}}{2(1-w)^{\frac{p-1}{2}}}$$

$$\int_{z=0}^{1} z^{-\frac{1}{2}}(1-z)^{\frac{n-2}{2}} \left[1+\frac{1-2w}{2w} z\right]^{\frac{p-1}{2}} \left[1-\frac{1-2w}{2(1-w)} z\right]^{-\frac{p-1}{2}}$$

$$\left[1+\frac{1-2w}{4w} z\right]^{-\frac{1}{2}} \left\{1+\frac{p-1}{2n-p+1}\ \phi_1(z,w)\right.$$

$$\left. + \frac{p^2-1}{(2n-p+1)(2n-p+3)}\ \phi_2(z,w)+ \ \cdots \right\} dz \quad ,$$

where $\phi_1(z,w)$ and $\phi_2(z,w)$ can be written down after making the transformation in the relevant factors in (2.17). To get three terms in the probability distribution of w we need only retain the term independent of z and the term containing z from $\phi_1(z,w)$ and the term independent of z from $\phi_2(z,w)$.

If we retain only those terms in the various expansions involved in (2.34) which contribute to the first three terms of the desired series, we get

$$(2.35)\quad f(w) = C_1 \frac{(1-2w)^{\frac{n-1}{2}} w^{\frac{p-2}{2}}}{2(1-w)^{\frac{p-1}{2}}}$$

$$\int_{z=0}^{1} z^{-\frac{1}{2}}(1-z)^{\frac{n-2}{2}} \left[1+ \frac{(p-1)(1-2w)}{4w} z + \frac{(p-1)(p-3)(1-2w)^2}{32w^2} z^2 + \dots\right]$$

$$\left[1+ \frac{(p-1)(1-2w)}{4(1-w)} z + \frac{(p^2-1)(1-2w)^2}{32(1-w)^2} z^2 + \dots\right]$$

$$\left[1- \frac{1-2w}{8w} z - \frac{3(1-2w)^2}{128w^2} z^2 + \dots\right]$$

$$\Bigg\{ 1+ \frac{p-1}{2n-p+1} \Big(\left[\frac{1-2w}{(1-w)^2} + \frac{(p-3)(1-3w+w^2)}{4w(1-w)} \right]$$

$$+ \left[\frac{(1-2w)^2}{(1-w)^3} - \frac{1-2w}{(1-w)^2} + \frac{(p-3)(4w^4-11w^2+7w-1)}{8w^2(1-w)^2} \right] z + \dots \Big)$$

$$+ \frac{p^2-1}{(2n-p+1)(2n-p+3)} \cdot \Big[\frac{(p-3)(p-5)(1-3w+w^2)^2}{32w^2(1-w)^2} -$$

$$\frac{(p-3)w(1-2w)}{4(1-w)^2} - \frac{(w^2+4w-2)(1-2w)}{2(1-w)^4} + \frac{(p-3)(1-3w+w^2)(1-2w)}{2w(1-w)^3} + \dots \Big] \Bigg\}\, dz\ .$$

Further simplification gives

$$(2.36)\quad f(w) = C_1 \frac{(1-2w)^{\frac{n-1}{2}} w^{\frac{p-2}{2}}}{2(1-w)^{\frac{p-1}{2}}}$$

$$\int_{z=0}^{1} z^{-\frac{1}{2}}(1-z)^{\frac{n-2}{2}} \Big\{ 1 + \Big[-\frac{(p-1)(1-2w)}{4w} + \frac{(p-1)(1-2w)}{4(1-w)}$$

$$- \frac{1-2w}{8w} \Big] z + \Big[\frac{-(p-1)(p-3)(1-2w)^2}{32w^2} + \frac{(p^2-1)(1-2w)^2}{32(1-w)^2}$$

$$- \frac{3(1-2w)^2}{128w^2} + \frac{(p-1)^2(1-2w)^2}{16w(1-w)} - \frac{(p-1)(1-2w)^2}{32w(1-w)}$$

$$- \frac{(p-1)(1-2w)^2}{32w^2} \Big] z^2 + \ldots \Big\}.$$

$$\Big\{ 1 + \frac{p-1}{2n-p+1} \Big[A + Bz + \ldots \Big] + \frac{p^2-1}{(2n-p+1)(2n-p+3)} \Big[E + \ldots \Big] + \ldots \Big\} \, dz ,$$

where A, B and E are functions of w, and are known explicitly from (2.35). For the sake of brevity we write this as

$$(2.37)\quad f(w) = C_1 \frac{(1-2w)^{\frac{n-1}{2}} w^{\frac{p-2}{2}}}{2(1-w)^{\frac{p-1}{2}}} \int_{z=0}^{1} z^{-\frac{1}{2}}(1-z)^{\frac{n-2}{2}} \Big[1 + A_1 z + B_1 z^2 + \ldots \Big].$$

$$\Big\{ 1 + \frac{p-1}{2n-p+1} \Big[A + Bz + \ldots \Big] + \frac{p^2-1}{(2n-p+1)(2n-p+3)} \Big[E + \ldots \Big] + \ldots \Big\} \, dz .$$

The integral involved in this can be written as

$$(2.38)\quad \int_{z=0}^{1} z^{-\frac{1}{2}}(1-z)^{\frac{n-2}{2}} \left\{ 1 + \left[\frac{p-1}{2n-p+1} A + A_1 z \right] + \right.$$

$$\left. \left[\frac{(p^2-1)E}{(2n-p+1)(2n-p+3)} + \frac{p-1}{2n-p+1}(B+AA_1)z + B_1 z^2 \right] + \cdots \right\} dz \quad ,$$

where the terms in curly brackets are arranged in three blocks according as they contribute to first, second and third approximations.

Integrating with respect to z, we have from these

$$(2.39)\quad f(w) = C_1 \frac{(1-2w)^{\frac{n-1}{2}} w^{\frac{p-2}{2}}}{2(1-w)^{\frac{p-1}{2}}} \cdot$$

$$\frac{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} \left\{ 1 + \left[\frac{p-1}{2n-p+1} A + \frac{1}{n+1} A_1 \right] \right.$$

$$+ \left[\frac{(p^2-1)E}{(2n-p+1)(2n-p+3)} + \frac{p-1}{(2n-p+1)(n+1)} (B+AA_1) \right.$$

$$\left. \left. + \frac{3}{(n+3)(n+1)} B_1 \right] + \cdots \right\} \quad ,$$

where

$$A = \frac{1-2w}{(1-w)^2} + \frac{(p-3)(1-3w+w^2)}{4w(1-w)} \quad ,$$

$$B = \frac{(1-2w)^2}{(1-w)^3} - \frac{1-2w}{(1-w)^2} + \frac{(p-3)(4w^4-11w^2+7w-1)}{8w^2(1-w)^2} ,$$

$$A_1 = \frac{(p-1)(1-2w)}{4w} + \frac{(p-1)(1-2w)}{4(1-w)} - \frac{1-2w}{8w} ,$$

$$B_1 = \frac{(p-1)(p-3)(1-2w)^2}{32w^2} + \frac{(p^2-1)(1-2w)^2}{32(1-w)^2} - \frac{3(1-2w)^2}{128w^2}$$

$$+ \frac{(p-1)^2(1-2w)^2}{16w(1-w)} - \frac{(p-1)(1-2w)^2}{32w(1-w)} - \frac{)o-1)(1-2w)^2}{32w^2} ,$$

and

$$E = \frac{(p-3)(p-5)(1-3w+w^2)^2}{32w^2(1-w)^2} - \frac{(p-3)w(1-2w)}{4(1-w)^2} - \frac{(w^2+4w-2)(1-2w)}{2(1-w)^4}$$

$$+ \frac{(p-3)(1-3w+w^2)(1-2w)}{2w(1-w)^3} .$$

Furthermore,

$$\frac{1}{2n-p+1} = \frac{1}{2n}\left(1-\frac{p-1}{2n}\right)^{-1} = \frac{1}{2n} + \frac{p-1}{4n^2} + O\left(\frac{1}{n^3}\right) ,$$

$$\frac{1}{n+1} = \frac{1}{n} - \frac{1}{n^2} + O\left(\frac{1}{n^3}\right) ,$$

$$\frac{1}{(2n-p+1)(2n-p+3)} = \frac{1}{4n^2} + O\left(\frac{1}{n^3}\right) ,$$

$$\frac{1}{(2n-p+1)(n+1)} = \frac{1}{2n^2} + O\left(\frac{1}{n^3}\right) ,$$

$$\frac{1}{(n+3)(n+1)} = \frac{1}{n^2} + O(\frac{1}{n^3}) ,$$

Using these in (2.39), the first three terms of the series can be written as follows:

$$(2.40)\quad f(w)dw = K\,(1-2w)^{\frac{n-1}{2}}\, w^{\frac{p-2}{2}}\,(1-w)^{-\frac{p-1}{2}}\Big\{1+\frac{1}{n}\Big[\frac{p-1}{2}A + A_1\Big] +$$

$$\frac{1}{n^2}\Big[\frac{(p-1)^2}{4}A - A_1 + \frac{p^2-1}{4}E + \frac{p-1}{2}(B+AA_1)+3B_1\Big] + O(n^{-3})\Big\} .$$

Explicit expressions for the first, second and third approximations can be written by substituting the values of A, B, A_1, A_2 and E from (2.39). This series can be written as

$$(2.41)\quad f(w) = Kf_1(w)\Big[1+\frac{1}{n}\lambda(w)+\frac{1}{n^2}\mu(w)+O(\frac{1}{n^3})\Big] ,$$

and is such that the ratio of each term to the preceeding term is $\frac{1}{n}$, and is the desired asymptotic series.

3. <u>The constant of integration for the first approximation</u>.

The first approximation to the distribution of w is

$$(3.1)\quad f(w) = K_1 w^{\frac{p-2}{2}}(1-w)^{-\frac{p-1}{2}}(1-2w)^{\frac{n-1}{2}} + O(n^{-1})$$

where $0 \leq w \leq \frac{1}{2}$. The constant of integration can be found by starting with the constant of integration of the triple integral found in

Chapter II or directly by integration from (3.1) as follows:

$$(3.2) \qquad \frac{1}{K_1} = \int_0^{\frac{1}{2}} w^{\frac{p-2}{2}} (1-w)^{-\frac{p-1}{2}} (1-2w)^{\frac{n-1}{2}} dw \quad .$$

Let $1-2w = y$, then

$$(3.3) \qquad \frac{1}{K_1} = \frac{1}{\sqrt{2}} \int_0^1 y^{\frac{n-1}{2}} (1-y)^{\frac{p-2}{2}} (1+y)^{-\frac{p-1}{2}} dy \quad .$$

By comparing with the hypergeometric integral, we can write

$$(3.4) \qquad \frac{1}{K_1} = \frac{1}{\sqrt{2}} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{p}{2})}{\Gamma(\frac{n+p+1}{2})} F(\frac{p-1}{2}, \frac{n+1}{2}, \frac{n+p+1}{2}, -1) \quad .$$

Using the formula

$$(3.5) \qquad F(a,b,c,x) = (1-x)^{-a} F(a, c-b, c, \frac{x}{x-1}) \quad ,$$

we get

$$(3.6) \qquad \frac{1}{K_1} = \frac{1}{2^{\frac{p}{2}}} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{p}{2})}{\Gamma(\frac{n+p+1}{2})} F(\frac{p-1}{2}, \frac{p}{2}, \frac{n+p+1}{2}, \frac{1}{2}) \quad .$$

Since the third coefficient in the hypergeometric series involved is large, the value of the series can be approximated by 1 for large n, that is

$$(3.7) \qquad \lim_{n \to \infty} F(\frac{p-1}{2}, \frac{p}{2}, \frac{n+p+1}{2}, \frac{1}{2}) = 1 \quad ,$$

and we can write

(3.8) $$\frac{1}{K_1} \sim \frac{1}{2^{\frac{p}{2}}} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{p}{2})}{\Gamma(\frac{n+p+1}{2})} .$$

Using Stirling's approximation, it can be further simplified to

(3.9) $$\frac{1}{K_1} \sim \frac{\Gamma(\frac{p}{2})}{n^{\frac{p}{2}}} ,$$

giving approximately the constant of integration in the first approximation.

4. The tail areas for the first approximation.

Let

(4.1) $$\phi(x) = \frac{n^{\frac{p}{2}}}{\Gamma(\frac{p}{2})} \int_{\frac{x}{2}}^{\frac{1}{2}} w^{\frac{p-2}{2}} (1-w)^{-\frac{p-1}{2}} (1-2w)^{\frac{n-1}{2}} \, dw .$$

The transformation $y = 1-2w$ reduces it to

(4.2) $$\phi(x) = \frac{n^{\frac{p}{2}}}{2^{1/2}\Gamma(\frac{p}{2})} \int_0^{1-x} y^{\frac{n-1}{2}} (1-y)^{\frac{p-2}{2}} (1+y)^{-\frac{p-1}{2}} \, dy .$$

This is the kind of integral we will have to evaluate for finding probabilities of the type $P(w > \frac{x}{2})$.

We write the integral involved in (4.2) as:

(4.3) $$\mu(t) = \int_0^t y^{\frac{n-1}{2}} \lambda(y) \, dy ,$$

where $\lambda(y) = (1-y)^{\frac{p-2}{2}}(1+y)^{-\frac{p-1}{2}}$. Integration by parts gives

$$(4.4) \qquad \mu(t) = \frac{2}{n+1}\, y^{\frac{n+1}{2}}\, \lambda(y) \Big|_0^t - \frac{2}{n+1}\int_0^t y^{\frac{n-1}{2}} \lambda'(y)dy \ .$$

This gives

$$(4.5) \qquad \mu(t) = \frac{2}{n+1}\, t^{\frac{n+1}{2}}\lambda(t) - \frac{2}{n+1}\int_0^t y^{\frac{n+1}{2}} \lambda'(y)\, dy \ .$$

Performing another integration, we get

$$(4.6) \quad \mu(t) = \frac{2}{n+1}\, t^{\frac{n+1}{2}}\lambda(t) - \frac{4}{(n+1)(n+3)}\, t^{\frac{n+3}{2}}\lambda'(t) +$$

$$\frac{4}{(n+1)(n+3)}\int_0^t y^{\frac{n+3}{2}}\lambda''(y)\, dy \ .$$

From (4.5) we can write

$$(4.7) \qquad \left|\mu(t) - \frac{2}{n+1}\, t^{\frac{n+1}{2}}\lambda(t)\right| = \frac{2}{n+1}\int_0^t y^{\frac{n+1}{2}}\lambda'(y)\, dy \ ,$$

where $\lambda(y) = (1-y)^{\frac{p-2}{2}}(1+y)^{-\frac{p-1}{2}}$. $0 \leq y \leq 1$

$$(4.8) \qquad \left|\lambda'(y)\right| = +\, \frac{p-2}{2}(1-y)^{\frac{p-4}{2}} + \frac{p-1}{2}(1+y)^{-\frac{p+1}{2}} \ ,$$

and this shows that the maximum value of $\lambda'(y)$ corresponds to the minimum value of y, which is zero, and

$$(4.9) \qquad \left|\lambda'(y)\right| \le B = \frac{2p-3}{2} .$$

By taking account of this bound we can write (4.7) as

$$(4.10) \qquad \left|\mu(t) - \frac{2}{n+1} t^{\frac{n+1}{2}} \lambda(t)\right| \le \frac{2p-3}{n+1} \int_0^t y^{\frac{n+1}{2}} dy ,$$

which is the same as

$$(4.11) \qquad \mu(t) - \frac{2}{n+1} t^{\frac{n+1}{2}} \lambda(t) \le \frac{(2p-3)2t^{\frac{n+3}{2}}}{(n+1)(n+3)} .$$

The range of t is zero to one, and thus equation (4.11) asserts that given ε, we can find N such that

$$(4.12) \qquad \left|\mu(t) - \frac{2}{n+1} t^{\frac{n+1}{2}} \lambda(t)\right| < \varepsilon \text{ for } n > N .$$

$\frac{2}{n+1} t^{\frac{n+1}{2}} \lambda(t)$ can therefore be taken as an asymptotic approximation of $\mu(t)$. Using this in (4.2) we get

$$(4.13) \qquad \phi(x) = \frac{2^{\frac{1}{2}} n^{\frac{p}{2}} (1-x)^{\frac{n+1}{2}} x^{\frac{p-2}{2}} (2-x)^{-\frac{p-1}{2}}}{\Gamma(\frac{p}{2})(n+1)} \left\{1 + O(n^{-1})\right\} .$$

We can write this as

$$(4.14) \quad P(w > \frac{x}{2}) = \frac{2^{\frac{1}{2}} n^{\frac{p-2}{2}}}{\Gamma(\frac{p}{2})} (1-x)^{\frac{n+1}{2}} x^{\frac{p-2}{2}} (2-x)^{-\frac{p-1}{2}} \left[1 + O(n^{-1})\right]$$

which provides a formula suitable for finding the tail areas of the type (4.1).

5. Comparison with the results of Chapter III.

We have shown in this chapter that if we omit terms of order $\frac{1}{n}$

$$(5.1)\qquad f(w)dw = C\ w^{\frac{p-2}{2}}(1-w)^{-\frac{p-1}{2}}(1-2w)^{\frac{n-1}{2}}dw \quad .$$

To get the corresponding first approximation for the distribution of the statistic $v = nm_3$, we put $w = \frac{v}{n}$ in this, and get

$$(5.2)\qquad f(v)dv = \frac{C}{n^p}\ v^{\frac{p-2}{2}}(1-\frac{v}{n})^{-\frac{p-1}{2}}(1-\frac{2v}{n})^{\frac{n-1}{2}}\ dv \quad .$$

Since

$$(i)\qquad \frac{\lim_{n\to\infty}(1-\frac{2v}{n})^{\frac{n-1}{2}}}{e^{-v}} = 1 \quad ,$$

and

$$(ii)\qquad \frac{\lim_{n\to\infty}(1-\frac{v}{n})^{-\frac{p-1}{2}}}{1} = 1 \quad ,$$

we can write (5.2) as

$$(5.3)\qquad f(v)\ dv = \text{Const.}\ v^{\frac{p-2}{2}}\ e^{-v}\ dv\ [1 + O(\frac{1}{n})] \quad ,$$

which states that for large n, v is approximately distributed as

$\chi^2/2$ where the χ^2 has p degrees of freedom.

This is not in agreement with the probability distribution of v obtained in Chapter III, and this discrepancy can be easily explained by the fact that the convergence of the series in w in (2.38) was not guaranteed for the whole region of w. It was seen on pages 101-103, that convergence of the series in z from which we obtained f(w) by integration would be obtained only if $w > \frac{1}{4}$. However, it appears that the two distributions would give close results if we are interested in the tail areas. As an illustration we shall find approximately the 5 % point for p = 4 by the results of the two chapters.

Example. To find x such that $P(v > x) = .05$ in the two cases

(1) $$f_1(v)\,dv = \frac{1}{2}(ve^{-v} + e^{-v})\,dv$$

(2) $$f_2(v)\,dv = ve^{-v}\,dv$$

From table 7, [35] we have the following values for the probability integral of the χ^2-distribution.

Giving $P(\chi^2 \geq \chi_0^2)$

χ_0^2 / d.f	7.8	8.0	8.2	8.4	8.6	8.8
2	.02024	.01832	.01657	.01500	.01357	.01228
4	.09919	.09158	.08452	.07798	.07191	.06630
Sum	.11943	.10990	.10109	.09298	.08548	.07858
$\frac{1}{2}$ Sum	.05972	.05495	.05055	.04649	.04274	.03929

The last row gives probabilities of the type $P(2v > \chi_0^2)$ or $P(v > v_0)$ say, and shows that approximately $P(v > 4.1) = .05$ if we use (1). Also from the same table we find that $P(v > 3.9) = .05$ if we use (2).

6. Summary of Chapter IV.

In this chapter we have considered the problem of obtaining an asymptotic series for the distribution of $w = m_3$ by starting with the joint distribution of m_1, m_2 and m_3 in the null case. The desired distribution is obtained by integrating out m_1 and m_2 over a lens shaped region enclosed by the two hyperboles $m_1 m_2 = w^2$(a constant) $= (1-m_1)(1-m_2)$. To perform the two integrations, one with respect to $x = m_1 + m_2$ and the other with respect to $y = m_1 - m_2$ we have, at each stage, regarded the other variables as constants and found a transformation which changes to integrand into a function of the form $\sum_i c_i(1 - z)^b z^i$, where b is a large number, and where z varies from zero to one. This leads us to a result of the type

$$f(w) = K \left[f_1(w) + \frac{1}{n} f_2(w) + \frac{1}{n^2} f_3(w) + \ldots \right] .$$

First three terms of the asymptotic series have been obtained in this manner. For the first approximation we have also found the constant of integration, and discussed the method of finding tail areas.

In Section 5 we have compared the results of this chapter with those of Chapter III.

CHAPTER V

THE APPLICATION OF TCHEBYCHEFF-MARKOFF INEQUALITIES TO A SPECIAL CASE

1. Introduction.

This chapter will be confined to the discussion of the special case $p = 3$ and $N_1 + N_2 = 20$. In this case, starting from the joint distribution of m_1, m_2 and m_3, we shall find the moments of the exact statistic V. These moments will then be used in setting up bounds to probabilities of the type $P(V \leq \xi)$ by the use of some investigations due to Tchebycheff and Markoff [48], [49]. This will provide, on the one hand, some exact results of some importance for this case, and on the other hand illustrate what can be done when the first few moments of an unknown distribution are known.

2. The integral over D_1.

As before, we shall denote by I_1 and I_2 the integrals over D_1 and D_2. Thus

$$(2.1) \qquad I_1 = \iiint_{D_1} \left[(1-m_1)(1-m_2)-m_3^2\right]^7 \, dm_1 \, dm_2 \, dm_3 .$$

Expanding the integral by the use of the binomial theorem, we get

$$(2.2) \qquad \sum_{j=0}^{7} \binom{7}{j} \iiint_{D_1} (1-m_1-m_2)^{7-j} \, (m_1 m_2 - m_3^2)^j \, dm_1 \, dm_2 \, dm_3 \; .$$

Each of the integrals in this sum can be calculated by first putting $m_3 = (m_1 m_2 t)^{1/2}$ to integrate with respect to t, then following the procedure of Section 6, Chapter II. The result is

$$(2.3) \qquad I_1 = \pi \Big[\frac{1}{90 \cdot 16} + \frac{1}{32 \cdot 30 \cdot 11} + \frac{}{48 \cdot 88 \cdot 12} + \frac{1}{16 \cdot 98 \cdot 12 \cdot 11 \cdot 13} + \frac{1}{32 \cdot 26 \cdot 88 \cdot 8} + \frac{1}{32 \cdot 130 \cdot 64 \cdot 6} + \frac{1}{128 \cdot 32 \cdot 120 \cdot 8} + \frac{1}{256 \cdot 128 \cdot 16 \cdot 17} \Big],$$

which is therefore the integral over D_1.

3. The integral over D_2.

This has been found exactly in Chapter II, but an independent derivation based on geometrical considerations could be given here.

$$(3.1) \qquad I_2 = \iiint_{D_2} \Big[(1-m_1)(1-m_2) - m_3^2 \Big]^{\frac{n-4}{2}} \, dm_1 \, dm_2 \, dm_3 \; .$$

The value of I_2 we want will be obtained by putting $n = 18$.

(3.2) In (3.1) put $u = (m_1 + m_2)/\sqrt{2}$

and $v = (m_1 - m_2)/\sqrt{2}$.

Then let $v = r \cos \Theta$

and $\sqrt{2}m_3 = z = r \sin \Theta$.

Then

$$(3.3) \quad I_2 = \frac{1}{\sqrt{2}} \int_{u=\frac{1}{\sqrt{2}}}^{\sqrt{2}} \int_{r=0}^{\sqrt{u^2 + 2(1-\sqrt{2}\,u)}} \int_{\Theta=0}^{2\pi} \left(1 - \sqrt{2}u + \frac{u^2 - r^2}{2}\right)^{\frac{n-4}{2}} 2\,d\Theta\,dr\,du .$$

Integration with respect to Θ and r is immediate, and yields

$$(3.4) \quad I_2 = \frac{2\sqrt{2}\,\pi}{n-2} \int_{\frac{1}{\sqrt{2}}}^{\sqrt{2}} \left(1 - \sqrt{2}\,u + \frac{u^2}{2}\right)^{\frac{n-2}{2}} du.$$

Putting $u = \sqrt{2}\,x$ and integrating we obtain

$$(3.5) \quad I_2 = \frac{\pi}{(n-1)(n-2)2^{n-3}} \quad ;$$

and for $n = 18$, it reduces to

(3.6) $$I_2 = \frac{\pi}{17 \cdot 16 \cdot 2^{15}} \quad .$$

4. The integral over D.

The value of I, the integral over D, is

(4.1) $$I_1 + I_2 = \frac{\pi}{18 \cdot 17 \cdot 4} \quad .$$

5. Moments of V.

The kth moment ν_k is given by :

(5.1) $$E(V^k) = \frac{1}{I_1+I_2} \Big[\iiint_{D_1 + D_2} (18m_3)^k (1-m_1-m_2+m_1m_2-m_3^2)^{7-k} dm_1dm_2dm_3 \Big] .$$

Due to the symmetry of $f(m_1m_2m_3)$ the joint density function of m_1, m_2 and m_3 in m_3, $E(V^k) = 0$ for odd values of k.

Putting k = 2, 4 and 6 in(5.1) and integrating as in Section 2, we get the following moments:

(5.2) $$\mu_2 = 6.5571637176,$$
$$\mu_4 = 459.6304942728,$$
$$\mu_6 = 25661.8465464 \quad .$$

We know that $V \geq n\, m_3$ since V equals $n\, m_3$ divided by a quantity which cannot exceed one, and in fact remains less than one inside D. Thus, since the range of $18m_3$ is from -9 to 9, the range of V in the case under consideration is larger, say from $-a$ to a where $a > 9$.

We shall now use the moments obtained above in setting up bounds for $P(V < \xi)$ to get an idea of the exact simpling distribution of V.

6. Some results due to Tchebycheff and Markoff.

We shall, in this section state without proof the results which lead to the historic inequalities which were announced by Tchebycheff and proved by Markoff, and which we shall use with the moments obtained in the preceeding section.

Theorem I. Any three consecutive polynomials in an arbitrary sequence $\{p_i(x)\}$ of orthogonal polynomials satisfy the relation

$$(6.1) \qquad p_m(x) = (a_m x + b_m)\, p_{m-1}(x) - c_m p_{m-2}(x) \; ,$$

where $p_m(x)$ stands for the mth orthogonal polynomial. Here a_m, b_m and c_m are constants, $a_m > 0$ and $c_m > 0$. If the highest coefficient of $p_m(x)$ is denoted by k_m , we have

$$(6.2) \qquad a_m = \frac{k_m}{k_{m-1}} \text{ and } c_m = \frac{a_m}{a_{m-1}}$$

The recurrence formula (6.1) is also true for $m = 1$ if we define $p_{-1}(x) = 0$.

Theorem II. The roots of the equation $p_m(x) = 0$, where $p_m(x)$ is the orthogonal polynomial of degree m associated with the weight function $\alpha(x)$ on the interval (a,b) , are all real and distinct; and all of them lie in the range of definition of the polynomials.

Theorem III. If

$$(6.3) \qquad \psi_m(z) = \int_a^b \frac{p_m(z) - p_m(x)}{z-x} \, d\alpha(x) ,$$

where $\alpha(x)$ is the weight function of the system of polynomials, then ψ's satisfy the same relation (6.1) as the p's, though with different initial conditions.

Definition. Let

$$(6.4) \qquad \frac{\psi_m(z)}{p_m(z)} = \sum_{i=1}^{m} \frac{\rho_i}{z-c_i} ,$$

where $\psi_m(z)$ is of degree m-1 in z, whereas $p_m(z)$ is of degree m. The c_i are the roots of $p_m(z) = 0$. Suppose $a < c_1 < c_2 \ldots < c_m < b$, where (a,b) is the range of the basic function $\alpha(x)$. Then ρ_i are called the Christoffel numbers.

Theorem IV. The Christoffel numbers are positive, and

$$\sum_{i=1}^{m} \rho_i = \int_a^b d\alpha(x) = \alpha(b) - \alpha(a).$$

Note: Because of theorem IV there exist numbers $d_1 < d_2 < \ldots < d_{m-1}$ lying between a and b such that

$$\rho_i = \alpha(d_i) - \alpha(d_{i-1}) , \; i = 1,2,\ldots m; \; d_0 = a, \; d_m = b.$$

Theorem V. $c_1, c_2 \ldots c_m$ alternate with $d_1, d_2 \ldots d_{m-1}$; that is $c_i < d_i < c_{i+1}$;

more precisely

$$\alpha(c_i + 0) - \alpha(a) < \alpha(d_i) - \alpha(a) = \sum_{j=1}^{i} \rho_j$$

$$< \alpha(c_{i+1} - 0) - \alpha(a), \; i = 1,2,\ldots m - 1 .$$

That is if F(x) is the class of cummulative distributions having the given moments, then

$$\int_a^{c_i} dF(x) < \int_a^{d_i} dF(x) = \sum_{j=1}^{i} \rho_j < \int_a^{c_{i+1}} dF(x) .$$

7. Application of Tchebycheff-Markoff Theorems to this Example.

We have the following matrix of the moments of the distribution studied in this chapter:

$$\begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \mu_1 & \mu_2 & \mu_3 & \mu_4 & \mu_5 \\ \mu_2 & \mu_3 & \mu_4 & \mu_5 & \mu_6 \\ \mu_3 & \mu_4 & \mu_5 & \mu_6 & \mu_7 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 6.557163716 & 0 & 459.630494 \\ 0 & 6.557163716 & 0 & 459.630494 & 0 \\ 6.557163716 & 0 & 459.630494 & 0 & 256661.8465 \\ 0 & 459.630494 & 0 & 256661.8465 & 0 \end{bmatrix} ,$$

in which all four principal diagonel matrices are positive definite.

Let

(7.1) $$p_4(x) = a_0 + a_2x^2 + a_4x^4$$

be one of the orthogonal polynomials in the sequence corresponding to the frequency function of V which gives rise to the moments found in Section 5. Then, by the definition of orthogonal polynomials,

$$(7.2) \qquad E[p_4(x) \cdot \theta_k(x)] = 0 \quad \text{if } k < 4.$$

Taking $\theta_k(x) = x^k$ for k = 0, and 2 and noting that odd moments are zero, we get the following equations for finding the coefficients in (7.1):

$$a_0 + 6.5571637176a_2 + 459.6304942728a_4 = 0$$

(7.3) and

$$6.5571637176a_0 + 459.6304942728a_2 + 256661.8465464a_4 + 0$$

Taking $a_4 = 1$ in (7.3), we obtain

(7.4) and

$$a_0 = 3532.38864$$
$$a_2 = -608.802724$$

as the solution of (7.3).

Thus

$$p_4(x) = x_4 - 608.802724x^2 + 3532.38864 \tag{7.5}$$

Solving $p_4(x) = 0$ we get the following four values of x, arranged in increasing order of magnitude

$$-34.726,\ -3.423,\ 3.423,\ \text{and}\ 34.726 \tag{7.6}$$

Now the function $\psi_4(z)$ can be found using its definition given in (6.20), which gives

$$\psi_4(z) = \mathop{E}_{x} \frac{(z^4 - x^4) + a_2(z^2 - x^2)}{z - x} \;; \tag{7.7}$$

and, using (5.2) and (7.4), this becomes

$$\psi_4(z) = z^3 - 602.24556z \;. \tag{7.8}$$

<u>The Christoffel Numbers</u>.

The Christoffel numbers as defined in Section 6 are the numbers ρ_i given by

$$(7.9) \qquad \frac{\psi_m(z)}{p_m(z)} = \sum_{i=1}^{m} \frac{\rho_i}{z - c_i} .$$

So we have to split

$$(7.10) \qquad \frac{z^3 - 602.24556z}{z^4 - 608.802724z^2 + 3532.38864}$$

into partial fractions. Write (7.10) as

$$(7.11) \qquad \frac{\rho_1}{z + 34.\ 26} + \frac{\rho_2}{z + 3.423} + \frac{\rho_3}{z - 3.423} + \frac{\rho_4}{z - 34.726}$$

Comparing the coefficients of like powers of z in (7.10) and (7.11) we get

$$\rho_1 = \rho_4 = .251$$
$$\rho_2 = \rho_3 = .248$$

approximately, where ρ_i corresponds to c_i , the ith root of the equation (7.5).

Thus we get the following table giving bounds to probabilities by using the formula, (6.5).

Table 7.12

Limits for ξ	Bounds for $P(V < \xi) = P$
$\xi < -34.726$	$P \leq .251$
$-34.726 < \xi \leq -3.423$	$0 < P \leq .499$
$-3.423 < \xi \leq 3.423$	$.251 < P \leq .747$
$3.423 < \xi \leq 34.726$	$.499 < P \leq 1$
$34.726 < \xi$	$.747 < P \leq 1$

It can be seen that the bounds given above are far from being close. For obtaining bounds which are sufficiently close and therefore useful we would have to calculate a large number of moments. The labor involved in finding enough moments, and proceeding with subsequent investigation based on those, however, is prohibitive of any such investigation in these pages.

CHAPTER VI

NON-NULL CASE

1. Introduction.

This chapter will be devoted to the study of the non-null case. In these first few sections we will consider the joint probability distribution of m_1, m_2 and m_3 given by Sitgreaves [45]. This distribution corresponds to the statistic

$$W = \sum_i \sum_j s^{ij}(z_i - \frac{\bar{y}_j+\bar{x}_j}{2})(\bar{y}_j - \bar{x}_j)$$

and has been obtained under the restriction that the mean vectors of the two populations are proportional to each other. For large n we shall convert this into a different form. Some of the difficulties in proceeding beyond that point will also be discussed.

The next section will deal with the distribution of $U = \sum_i \sum_j s^{ij} z_i(\bar{y}_j - \bar{x}_j)$ for the case $p = 1$, and on the assumption that n is large so that s^2 can be replaced by σ^2. This assumption reduces U to the product of two normal variates whose distribution is known; see for instance [2], [8] and [27]. It has not been possible to extend this to the case $p > 1$.

In Section 7 of this chapter we have exemplified the differential method which was quite popular with statisticians a few years ago. The illustration deals with the finding of approximations to the mean and variance of the statistic U for large samples by taking into account the sampling fluctuations of the sample means and covariances. Higher moments can also be found but the algebra involved is very heavy.

The concluding section of the chapter deals with a practical suggestion for modifying the variance of the discriminant function of R. A. Fisher by taking into account the sampling fluctuations of the means. The sample covariances can be taken as the population covariances when n is large. Thus the statistic U in this case becomes $U^* = \sum_i \sum_j \sigma^{ij} z_i (\bar{y}_j - \bar{x}_j)$.

2. The joint distribution.

The joint distribution of m_1, m_2 and m_3 given by Sitgreaves is

$$(2.1)\ f(m_1,m_2,m_3)\ dm_1 dm_2 dm_3 = \frac{\Gamma(\frac{n+1}{2}) e^{-\frac{\lambda^2}{2}(k_1^2+k_2^2)}}{\Gamma(\frac{n-p+2}{2})\Gamma(\frac{n-p+1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{1}{2})} |M|^{\frac{p-3}{2}} |I-M|^{\frac{n-p-1}{2}}$$

$$\sum_{j=0}^{\infty} \frac{\Gamma(\frac{n+2}{2}+j)}{\Gamma(\frac{p}{2}+j)j!} \left(\frac{\lambda^2}{2}\right)^j (k_1^2 m_1 + 2k_1k_2m_3 + k_2^2 m_2)^j \, dm_1 dm_2 dm_3$$

where

$$M = \begin{bmatrix} m_1 & m_3 \\ m_3 & m_2 \end{bmatrix} , \quad \lambda = \delta' \Sigma^{-1} \delta ,$$

and $k_1\delta$ and $k_2\delta$ are the mean vectors.

Using the notation of the confluent hypergeometric series, we can write (2.1) as

$$(2.3)\quad f(m_1 m_2 m_3) = \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{n+2}{2})\, e^{-\frac{\lambda^2}{2}(k_1^2+k_2^2)}}{\Gamma(\frac{n-p+2}{2})\Gamma(\frac{n-p+1}{2})\Gamma(\frac{p-1}{2})\Gamma(\frac{1}{2})}$$

$$|M|^{\frac{p-3}{3}}\ |I-M|^{\frac{n-p-1}{2}}\quad F(\tfrac{n+2}{2}, \tfrac{p}{2}, x) \quad ,$$

where

$$x = \frac{\lambda^2}{2}(k_1^2 m_1 + 2k_1 k_2 m_3 + k_2^2 m_2) \quad .$$

The function $F(a,c,x)$ is also written as $\phi(a,c,x)$ or as ${}_1F_1(a,c,x)$, and is known as the confluent hypergeometric function.

3. Notes for reference on confluent hypergeometric functions.

Consider the hypergeometric series

$$(3.1)\qquad F(a,b,c,z) = 1 + \frac{a.b}{c}\, z + \frac{a(a+1)(b)(b+1)}{c(c+1)}\, \frac{z^2}{2!} + \ldots \quad ,$$

in which we suppose that both a and c are positive. $F(a,b,c,\frac{x}{b})$ gives a power series with b as the radius of convergence. It defines an analytic function with singularities at 0, b, and ∞. The limiting case of this series as $b \rightarrow \infty$ defines an entire function whose singularity at ∞ is the confluence of two singularities of $F(a,b,c,\frac{x}{b})$ and which can be written as

$$(3.2)\quad F(a,c,x) = 1 + \frac{a}{c}\frac{x}{1!} + \frac{a(a+1)}{c(c+1)}\frac{x^2}{2!} + \cdots ,$$

It satisfies the confluent hypergeometric equation:

$$(3.3)\qquad x\frac{d^2y}{dx^2} + (c - x)\frac{dy}{dn} - ay = 0 \quad .$$

According to Bateman [4], the asymptotic behavior of $\phi(a, c, x)$ as $a \to \infty$ has been discussed by Perron , Tricomi and Taylor. An asymptotic form of F uniformly valid in the neighborhood of $x=0$ given by Taylor is

$$(3.4)\quad F(a,c,x) = \Gamma(c)(Kx)^{\frac{1}{2} - \frac{c}{2}} e^{\frac{x}{2}} J_{c-1}[2(Kx)^{\frac{1}{2}}] + O(K^{-1}) ,$$

where c and Kx are bounded, and $K = c/2 - a$, and J is the notation for Bessel functions.

If x is bounded and bounded away from zero and $\arg x - \arg K \leq \pi$, then

$$(3.5)\quad F(a,c,x) = 2^{\frac{1}{2}}\Gamma(c)e^{\frac{x}{2}} K^{\frac{3}{4} - \frac{c}{2}} x^{\frac{3}{4} - \frac{c}{2}}$$
$$c_1 e^{2i(Kx)^{\frac{1}{2}}} + c_2 e^{-2i(Kx)^{\frac{1}{2}}} + (Kx)^{-\frac{1}{2}} \Theta[\exp \operatorname{Im}(2K^{\frac{1}{2}} x^{\frac{1}{2}})] ,$$

where with s an integer, we have

$$c_1 = (2\pi)^{-\frac{1}{2}} e^{i\pi(s - \frac{1}{4})(2c-1)} \quad x$$

$$(2s-2)\pi + \varepsilon \leq \arg(Kx)^{\frac{1}{2}} \leq (2s+1)\pi - \varepsilon$$

and

$$c_2=(2\pi)^{-\frac{1}{2}} e^{i\pi(s+\frac{1}{4})(2c-1)x}$$

$$(2s-1)\pi+\varepsilon \leq \arg(Kx)^{\frac{1}{2}} \leq (2s+2)\pi-\varepsilon ,$$

and where $Im(y)$ denotes the imaginary part of y. The first of these results will be used in simplifying the distribution given in (2.3).

4. The asymptotic form of $f(m_1,m_2,m_3)$.

For large n, we have, by using (3.6),

$$(4.1)\quad F(\tfrac{n+2}{2},\tfrac{p}{2},x)=\Gamma(\tfrac{p}{2})(-\tfrac{2n+4-p}{4}x)^{\frac{2-p}{4}}e^{\frac{x}{2}}J_{\frac{p-2}{2}}[i\sqrt{(2n+4-p)x}]+O(\tfrac{1}{2n+4-p}).$$

Let

$$(4.2)\quad p=2+4q \quad ,$$

where q is an integer. Then, for large n ,

$$(4.3)\quad F(\tfrac{n+2}{2},\tfrac{p}{2},x)\quad \Gamma(2q+1)(2q-n-1)^{-q}e^{\frac{x}{2}}J_{2q}[i\sqrt{(2n+2-4q)x}] .$$

Using the relation

$$(4.4)\quad I_n(z)=i^{-n}J_n(iz) \quad ,$$

where $I_n(z)$ stands for Bessel functions with purely imaginary argument, (4.3) becomes

$$(4.5)\quad F(\tfrac{n+2}{2},\tfrac{p}{2},x)\quad (-1)^q\Gamma(2q+1)(2q-n-1)^{-q}e^{\frac{x}{2}}I_{2q}[\sqrt{(2n-4q+2)x}] .$$

Using this, we can write the joint density function of m_1, m_2 and m_3 as

$$(4.6)\ f(m_1 m_2 m_3)\, dm_1\, dm_2\, dm_3 = Ce^{-\frac{\lambda^2}{2}(k_1^2+k_2^2)} |M|^{\frac{p-3}{2}} |I-M|^{\frac{n-p-1}{2}}$$

$$e^x I_{\frac{p-2}{2}} \left[\sqrt{(2n+4-p)x} \right] dm_1\, dm_2\, dm_3 \quad .$$

for all p satisfying (4.2); where

$$x = \frac{\lambda^2}{2}(k_1^2 m_1 + 2k_1 k_2 m_3 + k_2^2 m_2) \quad .$$

4A. The difficulties in proceeding further.

Various methods have been tried to proceed beyond this point, but none seems to work w ell. The main difficulty, even at this stage, is, that the coefficients in the expansion of the Bessel function involved are increasing. As a consequence of this one would not be justified in omitting terms in the expansion of $I_{\frac{p-2}{2}} \left[\sqrt{2n+4-px} \right]$ beyond the first few, and discuss the distribution of nm_3. The difficulty would probably be removed if we consider small values of n, and try to integrate over the lens-shaped region of Chapter IV to find the distribution of m_3, but the objection to that would be that m_3 or nm_3 is not a suitable statistic for small values of n. This discussion, therefore, had to be left at this point.

5. The distribution of U for large n and p=1, an independent approach.

The statistic U reduces to $z(\bar{y}-\bar{x})/\sigma^2$ for large n, since s^2 is then found from a large sample, and can therefore be replaced by σ^2 to which it approximates. This does not imply that the sample means can also be replaced by their population values since for n to be large it is enough that one of the sample sizes is large. Moreover, none of the means has as many degrees of freedom as the variance.

The distribution of $z(\bar{y}-\bar{x})/\sigma^2$ can be found under both the hypotheses

(1) $z \in \pi_1$ which is $N(\mu,\sigma^2)$

(2) $z \in \pi_2$ which is $N(\nu,\sigma^2)$

as follows.

Let $z \in \pi_1$. The statistic U can be taken as the product of two normal variates z which is $N(\mu,\sigma^2)$, and $z^* = \frac{\bar{y}-\bar{x}}{\sigma^2}$ which is

$$N\left(\frac{\nu-\mu}{\sigma^2}, \frac{N_1+N_2}{N_1N_2\sigma^2}\right) .$$

We can, instead of z and z^*, consider the variables u and v, where u is $N(m,\sigma^2)$ and v is $N(0,\sigma'^2)$, where

$\sigma'^2 = \frac{N_1+N_2}{N_1N_2\sigma^2}$ and $m = \mu - \frac{\nu-\mu}{\sigma^2}$ or $\nu - \frac{\nu-\mu}{\sigma^2}$, according as $z \in \pi_1$ or $z \in \pi_2$.

The distribution of the product of two independent normal variates is known from the work of C. C. Craig [8], Aroian [2] and others, but for the sake of completeness we shall include a derivation.

Definition. x is said to be a Bessel variate if

$$(5.1)\quad f(x)dx = C\, x^{\frac{p-1}{2}}\, e^{-ax}\, I_{p-1}(b\, x^{\frac{1}{2}})dx \qquad 0 \le x < \infty \quad ,$$

where $I_{p-1}(bx^{\frac{1}{2}})$ is the modified Bessel function of the first kind as defined by Watson [52].

We shall now state without proof two lemmas.

Lemma 1. If x is $N(m,\sigma^2)$, then $\chi'^2 = \frac{x^2}{\sigma^2}$ is a Bessel variate.

In fact, if $\lambda^2 = m^2/\sigma^2$, then

$$(5.2)\, f(\chi'^2, d\chi'^2 = \frac{1}{\lambda^2}\, e^{-\frac{1}{2}\lambda^2}\, (\chi'^2)^{-\frac{1}{4}}\, e^{-\frac{1}{2}\chi'^2}\, I_{-\frac{1}{2}}(\lambda^2\chi'^2)^{\frac{1}{2}} d\chi'^2 \quad ,$$

which shows that χ'^2 is a Bessel variate with $a = \frac{1}{2}$, $b = \lambda$.

Lemma 2. If x_1 and x_2 are two independent Bessel variates with respective distributions

$$f(x_j)dx_j = Cx_j^{\frac{p-1}{2}}\, e^{-x_j}\, I_{p-1}(bx_j^{\frac{1}{2}})dx_j, \qquad (j = 1, 2) \quad ,$$

then the distribution of $\xi = x_1 - x_2$ is given by

$$(5.3)\quad f(\xi)d\xi = \frac{e^{-\frac{1}{2}b^2}}{\sqrt{\pi}}\left(\frac{\xi}{2}\right)^{\frac{2p-1}{2}} \sum_{r=0}^{\infty}\left(\frac{b}{2}\right)^{2r}\frac{\xi^2}{r!\Gamma(p+r)} K_{p+r-\frac{1}{2}}(\xi)\, d\xi ,$$

where $K_m(\xi)$ is the modified Bessel function of the second kind as defined by Watson, or in Whittaker and Watson [54, p. 373].

The distribution of $U = uv$ where u is $N(m,\sigma^2)$ and v is $N(0,\sigma'^2)$.

Let $\eta = \frac{u}{\sigma} + \frac{v}{\sigma}$, and $\zeta = \frac{u}{\sigma} - \frac{v}{\sigma}$. Then $uv = \text{const}\ (\eta^2 - \zeta^2)$,

where η is $N(\frac{m}{\sigma}, \sqrt{2})$ and ζ is $N(\frac{m}{\sigma}, \sqrt{2})$. Thus by lemma 1, both η^2 and ζ^2 are Bessel variates with $a = 1$, $b = \frac{m^2}{\sigma^2}$ and $p = \frac{1}{2}$ as parameters; and by lemma 2, therefore, the distribution of the product is given by

$$(5.4)\quad f(U)dU = \frac{e^{-\frac{1}{2}\frac{m^2}{\sigma^2}}}{\sqrt{\pi}} \sum_{r=0}^{\infty} \frac{b^{2r}\, U^r}{r!\Gamma(r+\frac{1}{2})2^{2r}}\, K_2(U)dU$$

and by noticing that

$$(2r)! = 2^{2r}\, r!\, \Gamma(r+\tfrac{1}{2})/\sqrt{\pi} \quad .$$

We can rewrite this as

$$(5.5)\quad f(U)dU = \frac{e^{-\frac{1}{2}\frac{m^2}{\sigma^2}}}{\pi} \sum_{r=0}^{\infty}\left(\frac{m}{\sigma}\right)^{2r} \frac{U^r}{(2r)!}\, K_r(U)dU \quad .$$

Replacing m by $\mu - \frac{\nu-\mu}{\sigma^2}$ and by $\nu - \frac{\nu-\mu}{\sigma^2}$, we get the distributions under the two hypotheses.

6. The asymptotic mean and variance of the statistic $U = \sum_i \sum_j s^{ij} z_i(\bar{y}_j - \bar{x}_j)$ by the differential method.

We shall, in this section, find the mean and variance of U, approximately for large samples, by a method which was formerly quite popular and still is sometimes used. The object of the section is mainly to exemplify this method, which can sometimes be applied in getting moments of an unknown distribution. Some of the sets of conditions under which the method is applicable are discussed by Cramer [9] in Chapter 27, but we shall, like statisticians in the past, apply it without stopping to verify the validity of the application. Because of the heavy algebra involved it will be enough if we confine ourselves to the discussion of the first two moments.

Let

$$U = \sum_i \sum_j s^{ij} z_i(\bar{y}_j - \bar{x}_j) \tag{6.1}$$

be written as

$$\sum_{i=1}^{p} b_i z_i \quad , \tag{6.2}$$

where

$$b_i = \sum_j s^{ij}(\bar{y}_j - \bar{x}_j) \quad . \tag{6.3}$$

We define

$$ds_{ij} = s_{ij} - \sigma_{ij} \quad , \tag{6.4}$$

Then

(6.5) $$E(ds_{ij}) = 0 \quad .$$

Following this definition, we let

(6.6) $$ds^{ij} = s^{ij} - \sigma^{ij} \quad ,$$

where

(6.7) $$\left[s^{ij} \right] = \left[s_{ij} \right]^{-1} \quad .$$

We note that

(6.8) $$E(ds^{ij}) \neq 0 \quad .$$

<u>To find</u> $E(ds^{ij})$:

Let s^{ij} be expanded in Taylor's series. We have

(6.9) $$s^{ij} = \sigma^{ij} + \sum_k \sum_l \frac{\partial \sigma^{ij}}{\partial \sigma_{kl}} ds_{kl} + \frac{1}{2} \sum_k \sum_l \sum_r \sum_t \frac{\partial^2 \sigma^{ij}}{\partial \sigma_{kl} \partial \sigma_{rt}} ds_{kl} ds_{rt} + \cdots \; .$$

Therefore

(6.9) $$E(ds^{ij}) = E\left[\sum_k \sum_l \frac{\sigma^{ij}}{\sigma_{kl}} ds_{kl} + \frac{1}{2} \sum_k \sum_l \sum_r \sum_t \frac{\partial^2 \sigma^{ij}}{\partial \sigma_{kl} \partial \sigma_{rt}} ds_{kl} ds_{rt} + \cdots \right] .$$

Since $E(ds_{kl}) = 0$, this reduces to

(6.10) $$E(ds^{ij}) = \frac{1}{2} E\left[\sum_k \sum_l \sum_r \sum_t \frac{\partial^2 \sigma^{ij}}{\partial \sigma_{kl} \partial \sigma_{rt}} ds_{kl} ds_{rt} + \cdots \right] .$$

To evaluate this, we have to find

(6.11) $$\frac{\partial^2 \sigma^{ij}}{\partial \sigma_{kl} \partial \sigma_{rt}} \quad \text{and} \quad E(ds_{kl} ds_{rt}) \quad .$$

The latter of these is known from Hotelling [23] as

$$(6.12) \qquad E(ds_{kl}ds_{rt}) = \frac{\sigma_{kr}\sigma_{lt} + \sigma_{kt}\sigma_{lr}}{n} \quad .$$

To find the second order derivatives involved we proceed as follows:

Consider the identities

$$(6.13) \qquad \sum_i \sigma^{ij}\sigma_{ik} = \delta_k^j \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases} \quad .$$

Differentiating (6.13) partially with respect to $\sigma_{\alpha\beta}$, we have

$$(6.14) \qquad \sum_i \left[\sigma_{ik} \frac{\partial \sigma^{ij}}{\partial \sigma_{\alpha\beta}} + \sigma^{ij}(\delta_i^\alpha \delta_k^\beta + \delta_k^\alpha \delta_i^\beta - \delta_{ik}^{\alpha\beta}) \right] = 0 \quad ,$$

where

$$(6.15) \qquad \delta_{ik}^{\alpha\beta} = \begin{cases} 1 & \text{if } \quad i = k = \alpha = \beta \\ 0 & \quad \text{otherwise} \end{cases} \quad .$$

Using

$$(6.16) \qquad \sum_k \sigma^{km}\sigma_{ik} = \delta_i^m$$

in (6.14) and simplifying, we get

$$(6.17) \qquad \frac{\partial \sigma^{mj}}{\partial \sigma_{\alpha\beta}} = - \sigma^{\alpha j}\sigma^{\beta m} - \sigma^{\beta j}\sigma^{\alpha m} + (\sigma^{jj})^2 \delta_{\alpha\beta}^{mj} \quad ,$$

which provides a formula for the first derivatives.

If the covariance matrix Σ is the identity matrix, as can be supposed for the statistic U, which is known to be invariant under non-singular linear transformations, then the σ's can be replaced by Kronecker deltas with the same suffices. Thus (6.17) simplifies to

$$(6.18)\quad \frac{\partial\sigma^{mj}}{\partial\sigma_{\alpha\beta}} = -\delta_{\alpha}^{j}\,\delta_{\beta}^{m} - \delta_{\beta}^{j}\,\delta_{\alpha}^{m} + \delta_{\alpha\beta}^{mj} \quad ,$$

which states that

$$(6.19)\quad \frac{\sigma^{ii}}{\sigma_{ii}} = -1 \quad ,$$

$$(6.20)\quad \frac{\sigma^{ij}}{\sigma_{ij}} = -1 \quad ,$$

and that the derivative with respect to any other element is zero.

To obtain $\frac{\partial^2\sigma^{ij}}{\partial\sigma_{\alpha\beta}\partial\sigma_{r\delta}}$ we differentiate partially with respect to $\sigma_{\gamma\delta}$ the equation:

$$(6.21)\quad \frac{\partial\sigma^{ij}}{\partial\sigma_{\alpha\beta}} = -\sigma^{\alpha j}\sigma^{\beta i} - \sigma^{\beta j}\sigma^{\alpha i} + (\sigma^{jj})^2\delta_{\alpha\beta}^{ij} \quad .$$

This gives:

$$(6.22)\quad \frac{\partial^2\sigma^{ij}}{\partial\sigma_{\alpha\beta}\partial\sigma_{r\delta}} = -\sigma^{\alpha j}\frac{\partial\sigma^{\beta i}}{\partial\sigma_{\gamma\delta}} - \sigma^{\beta i}\frac{\partial\sigma^{\alpha j}}{\partial\sigma_{\gamma\delta}} - \sigma^{\beta j}\frac{\partial\sigma^{\alpha i}}{\partial\sigma_{\gamma\delta}}$$

$$- \sigma^{\alpha i}\frac{\partial\sigma^{\beta j}}{\partial\sigma_{\gamma\delta}} + 2\sigma^{jj}\frac{\partial\sigma^{jj}}{\partial\sigma_{\gamma\delta}}\delta_{\alpha\beta}^{ij} \ ,$$

Using (6.21) in (6.22) and replacing the σ's by δ's as before, we obtain

$$(6.23)\quad \frac{\partial^2\sigma^{ij}}{\partial\sigma_{\alpha\beta}\partial\sigma_{\gamma\delta}} = \delta_{\alpha}^{j}(\delta_{\delta}^{\beta}\delta_{\gamma}^{i} + \delta_{\gamma}^{\beta}\delta_{\delta}^{i} - \delta_{\gamma\delta}^{\beta i}) +$$

$$\delta_{\beta}^{i}(\delta_{\delta}^{\alpha}\delta_{\gamma}^{j} + \delta_{\delta}^{j}\delta_{\gamma}^{\alpha} - \delta_{\gamma\delta}^{\alpha j}) +$$

$$\delta^{j}_{\beta}(\delta^{\alpha}_{\delta}\delta^{i}_{\gamma} + \delta^{i}_{\delta}\delta^{\alpha}_{\gamma} - \delta^{\alpha i}_{\gamma\delta}) +$$

$$\delta^{i}_{\alpha}(\delta^{\beta}_{\delta}\delta^{j}_{\gamma} + \delta^{j}_{\delta}\delta^{\beta}_{\gamma} - \delta^{\beta j}_{\gamma\delta}) +$$

$$2 \left[2\delta^{j}_{\delta}\delta^{j}_{\gamma} - \delta^{jj}_{\gamma\delta} \right] \delta^{ij}_{\alpha\beta} \quad .$$

This gives

(6.24) $$\frac{\partial^2 \sigma^{ii}}{\partial \sigma^2_{ii}} = 6$$

(6.25) $$\frac{\partial^2 \sigma^{ij}}{\partial \sigma^2_{ij}} = 0, \quad (i \neq j) ;$$

and all other derivatives of the second order are also zero.

Using these results in (6.10), we have

(6.26) $$E(ds^{ij}) = \frac{6p}{n} + O(n^{-2}) \quad .$$

To find $E(ds^{ij}ds^{gh})$

(6.27) $$E(ds^{ij}ds^{gh}) = \sum_k \sum_l \sum_r \sum_t \frac{\partial\sigma^{ij}}{\partial\sigma_{kl}} \cdot \frac{\partial\sigma^{gh}}{\partial\sigma_{rt}} \cdot E(ds_{kl}ds_{rt}) \quad .$$

Using (6.12), we can reduce this to

(6.28) $$E(ds^{ij}ds^{gh}) = \frac{1}{n} \sum_k \sum_l \sum_r \sum_t \frac{\partial\sigma^{ij}}{\partial\sigma_{kl}} \frac{\partial\sigma^{gh}}{\partial\sigma_{rt}} \left[\sigma_{kr}\sigma_{lt} + \sigma_{dt}\sigma_{lr} \right] + \ldots \; .$$

Substituting the values of the first order derivative in terms of δ's from (6.18) on the supposition of Σ being the identity matrix, we get

(6.29) $E(ds^{ij}ds^{gh}) \sim \frac{1}{n} \sum_k \sum_l \sum_r \sum_t [\delta_k^j \delta_l^i + \delta_k^i \delta_l^j - \delta_{kl}^{ij}]$

$$[\delta_r^h \delta_t^g + \delta_r^g \delta_t^h - \delta_{rt}^{gh}][\delta_k^r \delta_l^t + \delta_k^t \delta_l^r] = \frac{1}{n} \delta_{ijgh} \text{ say} .$$

In terms of σ's we can use the notation

(6.30) $$E(ds^{ij}ds^{gh}) \sim \frac{\sigma_{ijgh}}{n} .$$

These results will now be used in finding $E(U)$ and Var (U).

(6.31) <u>To find</u> $E(U) = E[\sum_i \sum_j s^{ij} z_i (\bar{y}_j - \bar{x}_j)]$.

Since s^{ij}, z_i, $\bar{y}_j - \bar{x}_j$ are allindependently distributed, the expectation of the product is equal to the product of the expectations.

(6.32) $$E(z_i) = \mu_i \quad \text{if} \quad z \in \pi_i$$

(6.33) $$E(\bar{y}_j - \bar{x}_j) = \nu_j - \mu_j = d_j \quad \text{say}$$

(6.34) $$E(s^{ij}) = \sigma^{ij} + \frac{6p}{n} \quad \text{from} \quad (6.26)$$

$$= \begin{cases} 1 + \frac{6p}{n} & \text{if } i = j \\ \frac{6p}{n} & \text{if } i \neq j \end{cases} .$$

Thus

(6.35) $$E(U) = \sum_i \sum_j \mu_i(\nu_j - \mu_j) [\delta_i^j + \frac{6p}{n}] .$$

This reduces to

(6.36) $$E(U) = (1 + \frac{6p}{n}) \sum_{i=1}^{p} \mu_i(d_i) + \frac{6p}{n} \sum_{\substack{i \\ i \neq j}} \sum_{j} \mu_i(d_j) \quad .$$

<u>To find</u> Var(U)

We can write

(6.37) $$U = \sum_{i=1}^{p} b_i z_i = \sum_{j=1}^{p} b_j z_j \quad ,$$

where

(6.38) $$b_i = \sum_k s^{ik} (\bar{y}_k - \bar{x}_k) \quad .$$

Define

(6.39) $$\beta_i = \sum_k \sigma^{ik} d_k ,$$

where

(6.40) $$d_k = \nu_k - \mu_k \quad .$$

Then in π_1,

(6.41) $$\sigma_U = \sum_i \sum_j [\beta_i \beta_j \sigma_{ij} + \mu_i \mu_j \sigma_{b_i b_j} + \cdots]$$

To find $\sigma_{b_i b_j}$, we start with

(6.42) $$\delta b_i = \sum_k [\delta s^{ik} d_k + \sigma^{ik} \delta(\bar{y}_k - \bar{x}_k) + \cdots]$$

and

(6.43) $$\delta b_j = \sum_m [\delta s^{jm} d_m + \delta^{jm} \delta(\bar{y}_m - \bar{x}_m) + \cdots] \quad .$$

From these two equations:

$$(6.44)\quad \sigma_{b_i b_j} = \sum_k \sum_m \left[d_k d_m E(\delta s^{ik} . \delta s^{jm}) + \sigma^{ik}\sigma^{jm} E\left[\delta(\bar{y}_k - \bar{x}_k) . \delta(\bar{y}_m - \bar{x}_m) \right] + \ldots \right.$$

Since

$$(6.45)\qquad E(\delta s^{ik} \delta s^{jm}) \sim \frac{6_{ijkm}}{n}$$

from (6.30) and

$$(6.46)\qquad E\left[\delta(\bar{y}_k - \bar{x}_k)\ \delta(\bar{y}_m - \bar{x}_m) \right] = \sigma_{\bar{y}_k - \bar{x}_k,\ \bar{y}_m - \bar{x}_m} = \sigma_{km}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) ,$$

(6.44) can be written as

$$(6.47)\qquad \sigma_{b_i b_j} \sim \sum_k \sum_m \frac{\sigma_{ijkm}}{n} d_k d_m + \sigma^{ik}\sigma^{jm}\sigma_{km}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) .$$

But

$$\sum_m \sigma^{jm}\sigma_{km} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases} ;$$

therefore we obtain

$$(6.48)\qquad \sigma_{b_i b_j} \sim \sum_k \sum_m \frac{\sigma_{ijkm}}{n} d_k d_m + \sigma^{ij}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) .$$

Using (6.48) in (6.41), we get

$$(6.49)\quad \sigma_U^2 \sim \sum_i \sum_j \left\{ \beta_i \beta_j \sigma_{ij} + \mu_i \mu_j \left[\sum_k \sum_m \frac{\sigma_{ijkm}}{n} d_k d_m + \sigma^{ij}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \right] \right\} .$$

Replacing β_i and β_j by their values from (6.39), we reduce this to

$$(6.50)\quad \sigma_U^2 \sim \sum_i \sum_j \sum_k \sum_m \left[\sigma^{ik}\sigma^{jm}\sigma_{ij} + \frac{\sigma_{ijkm}}{n} \mu_i \mu_j \right] d_k d_m + \sum_i \sum_j \sigma^{ij}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \mu_i \mu_j .$$

If we suppose that $\Sigma = I$, then (6.50) reduces to

(6.51) $$\sigma_U^2 \sim \Delta^2 + \frac{1}{n} \sum_i \sum_j \sum_k \sum_m \mu_i \mu_j \, d_k d_m \, \delta_{ijkm} + p\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \mu_i^2 \; .$$

where

$$\Delta = \sum_k \sum_m \sigma^{km} \, d_k d_m \quad .$$

7. <u>Correction term for the variance of the linear discriminant function.</u>

In this section we shall find the variance of the statistic

(7.1) $$U^* = \sum_i \sum_j \sigma_{z_i}^{ij} \; (\bar{y}_j - \bar{x}_j) \quad ,$$

which is the same as U with s^{ij} replaced by σ^{ij} because of the supposition that n is large. If N_1 and N_2 are both large then $\bar{y}_j - \bar{x}_j$ can be replaced by the corresponding difference in the populations, namely $\nu_j - \mu_j$, giving for U^* a linear function of normal variates. As an improvement we shall find the variance of U^* by taking into account the sampling flucutations due to the difference of sample means.

We have

(7.2) $$E(x_i) = \mu_i, \quad E(y_i) = \nu_i$$

(7.3) $$E(x_i - \mu_i)(x_j - \mu_j) = \sigma_{ij} = E(y_i - \nu_i)(y_j - \nu_j)$$
$$= E(z_i - \mu_i)(z_j - \mu_j) \quad .$$

Let

(7.4) $$d_i = \nu_i - \mu_i$$

and

$$[\sigma^{ij}] = [\sigma_{ij}]^{-1} \quad .$$

The correction term for the variance of U^* is the variance of U^* on the assumption that $z = (z_1, z_2, \ldots, z_p)$ is fixed.

$$(7.5) \qquad U^* = \sum_i \sum_j \sigma^{ij}(\bar{y}_j - \bar{x}_j) z_i$$

can be written as

$$(7.6) \qquad U^* = \sqrt{\frac{1}{N_1} + \frac{1}{N_2}} \sum_i \sum_j \sigma^{ij} w_j z_i \quad ,$$

where

$$(7.7) \qquad w_r = \frac{y_r - x_r}{\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \quad ,$$

so that

$$(7.8) \qquad \sigma_{w_r w_t} = \sigma_{rt} = \sigma_{z_r z_t} \quad .$$

$$(7.9) \qquad U^* + \delta U^* = \sqrt{\frac{1}{N_1} + \frac{1}{N_2}} \sum_i \sum_j \sigma^{ij}(w_j + \delta w_j) z_i \quad .$$

Hence

$$(7.10) \qquad \delta U^* = \sqrt{\frac{1}{N_1} + \frac{1}{N_2}} \sum_i \sum_j \sigma^{ij} \delta w_j z_i \quad ,$$

and thus

$$(7.11) \qquad \sigma'^2_{\delta U^*} = \left(\frac{1}{N_1} + \frac{1}{N_2}\right) \sum_i \sum_j \sigma^{ij} z_i z_j$$

which will give the correction term. Since

(7.12) $$E \sum_i \sum_j \sigma^{ij} z_i z_j = \Sigma_i \Sigma_j \sigma^{ij}(\sigma_{ij} + d_i d_j) \quad ,$$

but

(7.13) $$\sum_i \sigma_{ir}\sigma^{ij} = \delta_r^j \quad ,$$

and

(7.14) $$\sum_i \sum_j \sigma^{ij} d_i d_j = \Delta^2 \quad \text{say.}$$

(7.12) gives

(7.15) $$E \sum_i \sum_j \sigma^{ij} z_i z_j = p + \Delta^2 \quad .$$

Adding this to the variance of the linear discriminant function we have, for the corrected variance,

(7.16) $$\sigma^2_{U^*} = (1 + \frac{1}{N_1} + \frac{1}{N_2}) \Delta^2 + p(\frac{1}{N_1} + \frac{1}{N_2}) \quad .$$

This formula shows that the variance Δ^2 based on the assumption "N_1 and N_2 are large" is an underestimate of the correct variance of the discriminant function, but that the difference approaches zero as rapidly as N_1 and N_2 approach infinity.

CHAPTER VII

SOME RELATED UNSOLVED PROBLEMS

In this chapter we shall describe very briefly some unsolved problems related to the problem of classification.

1. <u>On classification statistics of Wald and Anderson</u>.

(a) The preceeding discussion deals mainly with the distribution of the approximate statistic $v = nm_3$, that is the statistic whose distribution approximates the distribution of V where

$$V = \frac{nm_3}{(1-m_1)(1-m_2)-m_3^2}$$

for large n. We have discussed mainly the null case, and much work needs to be done in getting its distribution in the non-null case for the two statistics,

(1) Discussed by Wald [50]

(2) Discussed by Sitgreaves [45].

(b) The exact treatment of the sampling distribution of V, both in the central and the non-central cases is still wanting.

2. <u>The quadratic discriminators</u>.

Let μ and ν denote the mean vectors of two p-variate normal populations, and Σ_1 and Σ_2 the two covariance matrices. There are three situations that may arise in discussing the problem of classification, namely

$$(2.1)\qquad \begin{array}{lll} & \text{(a)} & \mu \neq \nu \ \text{ and } \ \Sigma_1 = \Sigma_2 \quad , \\ & \text{(b)} & \mu = \nu \ \text{ and } \ \Sigma_1 \neq \Sigma_2 \quad , \\ \text{and} & \text{(c)} & \mu \neq \nu \ \text{ and } \ \Sigma_1 \neq \Sigma_2 \quad . \end{array}$$

If we suppose all the population parameters to be known, then in these three situations we get the following three statistics,

$$U_a = \sum_i \sum_j \alpha^{ij} z_i(\nu_j - \mu_j)$$

$$(2.2)\qquad U_b = \sum_i \sum_j (z_i - \mu_i)(z_j - \mu_j) \left[\sigma^{ij} - \sigma^{ij*} \right]$$

$$U_c = \sum_i \sum_j \left[(z_i-\mu_i)(z_j-\mu_j)\sigma^{ij} - (z_i-\nu_i)(z_j-\nu_j)\sigma^{ij*} \right]$$

where in U_a, $[\sigma^{ij}] = [\sigma_{ij}]^{-1}$, $[\sigma_{ij}]$ being the common co-variance matrix of the two populations, and σ^{ij}, σ^{ij*} in U_b and U_c refer to the two covariance matrices in the two populations. Thus the distribution problem underlying (2.1) (b) and (c) are those of a general indefinite quadratic form with zero expectations of the normal variates in (b) but not in (c). The importance of this problem has been stressed by Hotelling [22]. This, of course, is under the assumption that the population parameters are known which amounts to saying that $N_1 \longrightarrow \infty$ and $N_2 \longrightarrow \infty$, and would be a first step in discussing the distributions of the statistics

$$(2.3)\qquad W_b = \sum_i \sum_j (z_i - \bar{x}_i)(z_j - \bar{x}_j) \left[s^{ij} - s^{ij*} \right]$$

and

$$W_c = \sum_i \sum_j [(z_i - \bar{x}_i)(z_j - \bar{x}_j)s^{ij} - (z_i - \bar{y}_i)(z_j - \bar{y}_j)s^{ij*}] ,$$

which are obtained from U_b and U_c by replacing the population values in U_b and U_c by their estimates from the samples.

3. Possibility of a different approach.

(a) It may be desirable to discuss the distribution of $U = \sum_i \sum_j s^{ij} z_i(\bar{y}_j - \bar{x}_j)$ by some independent method. Very often it is a good start to examine in what form the non-centrality parameters would enter into the distribution. The answer to this sometimes provides a key to the solution of the distribution problem. Furthermore some questions related to the behavior of the test can be answered even without finding the actual distribution in the non-null case.

(b) It might be worth while to try some altogether different approach. It is possible that we run into some simpler distribution problems. Papers of Rao [32] and Roy [35] should be useful in this connection.

4. Efficiency.

(a) The idea of efficiency in problems on classification needs to be developed systematically. Kossack [26] took 1-P as the index of efficiency where P is the common probability of the two types of misclassification over variations of the parameters involved. He, however, considered only the univariate case. Pitman [37] defined it as the ratio of two sample sizes.

These and other ideas can be examined in this connection.

(b) If there are more statistics than one for the same situation, then some measure of relative efficiency is needed.

(c) The discriminant function of R.A. Fisher or the statistic $\sum_i \sum_j \sigma^{ij} z_i(\nu_j - \mu_j)$ are based on the assumption that $\Sigma_1 = \Sigma_2$. One important problem that calls for investigation is to examine how good is the linear discriminant function when actually $\Sigma_1 \neq \Sigma_2$.

5. The greater mean vector.

Bahadur and Robbins [3] and Robbins [42] have pointed out that even in the univariate case of sorting numerous objects known to belong to one or the other of two normal populations with the same known variance, the obvious rule of classifying an object to the population whose mean is closer to the measure of the object, may not be the best rule. Their objections apply to the corresponding multivariate situations and should be considered in problems of classification in multivariate analysis.

BIBLIOGRAPHY

[1] Anderson, T.W., "Classification by Multivariate Analysis", Psychometrika, Vol. XVI(1951), pp. 31-50.

[2] Arion, L. A., "The Probability Function of the Product of two Normally Distributed Variables", Annals of Mathematical Statistics, Vol. XVIII(1947), pp. 265-271.

[3] Bahadur R.R. and Robbins H. E., "The Problem of the Greater Mean", Annals of Mathematical Statistics, Vol. XXI (1950), pp. 469-487.

[4] Bateman, Harry, Higher Transcendental Functions. Vol. I and II, McGraw Hill Book Company, Inc., 1953.

[5] Bose, R. C., "On the Exact Distribution and Moment Coefficients of the D^2-Statistics", Sankhya, Vol. II (1935-1936), pp. 143-154.

[6] Chernoff, Herman, "Large Sample Theory", Annals of Mathematical Statistics, Vol. XXVII (1956), pp. 1-23.

[7] Cochran, W. G. and Bliss, C. I., "Discriminant Function with Covariance", Annals of Mathematical Statistics, Vol. XIX (1948), pp. 151-176.

[8] Craig, C. C. "On the Frequency Function of xy", Annals of Mathematical Statistics, Vol. VII(1936), pp. 1-15.

[9] Cramer, Harold, Mathematical Methods of Statistics. Princeton University Press, 1951.

[10] Fisher, R. A., "The Use of Multiple Measurements in Taxonomic Problems", Annals of Eugenics, Vol. VII(1936), pp. 179-188.

[11] Fisher, R. A., "The Statistical Utilization of Multiple Measurements", Annals of Eugenics, Vol. VIII(1938), pp. 376-386.

[12] Fix, Evelyn, and Hodges, J. L., "Discriminatory Analysis: Non-Parametric Discrimination: Consistency Problems", School of Aviation Medicine. Project number 21-49-004 (1951).

[13] Ford, W. B., Studies in Divergent Series and Summability. The Macmillan Co., New York, 1916.

[14] Goursat, Edourad, (Hedrick, E.R., translator), A Course in Mathematical Analysis. Vol. I, Ginn and Company, New York, 1904.

[15] Grad, Arthur, and Solomon, Herbert, "Distribution of Quadratic Forms and Some Applications", *Annals of Mathematical Statistics*, Vol. XXVI(1955), pp. 464-477.

[16] Gurland, John, "Distribution of Quadratic Forms and Ratios of Quadratic Forms", *Annals of Mathematical Statistics*, Vol. XXII (1953), pp. 416-427.

[17] Gurland, John, "Distribution of Definite and Indefinite Quadratic Forms", *Annals of Mathematical Statistics*, Vol. XXVI(1955), pp. 122-128.

[18] Harter, H. L. "On the Distribution of Wald's Classification Statistic", *Annals of Mathematical Statistics*, Vol. XXII (1951), pp. 58-67.

[19] Hotelling, Harold, "New Light on the Correlation Coefficient and its Transforms", *Journal of the Royal Statistical Society*, Series B, Vol. XV, No. 2 (1953), pp. 193-232.

[20] Hotelling, Harold, *Notes on Approximation Techniques*. (unpublished), 1955.

[21] Hotelling, Harold, "Some New Methods for the Distribution of Quadratic Forms", Abstract, *Annals of Mathematical Statistics*, Vol. XIX (1948), p. 119.

[22] Hotelling, Harold, "Multivariate Analysis", *Statistics and Mathematics in Biology*, Iowa State College Press, (1954), pp. 67-80.

[23] Hotelling, Harold, "Relation Between Two Sets of Variates", *Biometrika*, Vol. XXVIII (1936), pp. 321-377.

[23a] Hotelling, Harold, "The Generalization of Student's Ratio", *Annals of Mathematical Statistics*, Vol.II(1931), pp. 360-378.

[24] Hotelling, Harold, "A Generalized T-test and Measure of Multivariate Dispersion", *Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 23-41.

[24a] Isaacson, S. L., "Problems in Classifying Populations", *Statistics and Mathematics in Biology*. Iowa State College Press (1954), pp. 107-119.

[24b] Kendall, M. G., Notes on Multivariate Analyses, Institute of Statistics, mimeograph series No. 95, (1954).

[25] Kolmogoroff, A. N., Foundations of the Theory of Probability. Chelsea Publishing Company, New York, (1950).

[26] Kossack, C. F., "Some Techniques for Simple Classification", Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, (1945-46), pp. 345-352.

[27] Laha, R. G., "On Some Properties of the Bessel Function Distribution", Bulletin of the Calcutta Mathematical Society. Vol. XVVI, No. 1, (1954), pp. 59-72.

[28] MacRobert, T. M., Functions of a Complex Variable. Second Edition, Macmillan and Company, Limited, London, 1933.

[29] Mann, H. B. and Wald, Abraham, "On Stochastic Limit and Order Relationship", Annals of Mathematical Statistics, Vol. XIV, (1943), pp. 265-275.

[30] McCarthy, M.D., "On the Application of z-test to Randomized Blocks", Annuals of Mathematical Statistics, Vol. X(1939), pp. 337-359.

[31] Mises, R. V., "On the Classification of Observed Data into Distinct Groups", Annuals of Mathematical Statistics, Vol. 16, (1945), pp. 68-73.

[32] Neyman, Jerzy, and Pearson, E.S., "Contributions to the Theory of Testing Statistical Hypotheses", Statistical Research Memoirs, Vol. I(1936), pp. 1-161.

[33] Ogawa, Junjiro, "Remark on Wald's Paper On a Statistical Problem Arising in the Classification of an Individual into One of Two Groups", Institute of Statistics, Mimeograph Series No.

[34] Pachares, James, "Note on the Distribution of a Definite Quadratic Form", Annals of Mathematical Statistics, Vol. XXVI(1955), pp. 128-131.

[35] Pearson E.S. and Hartley, H. O., Biometrika Tables for Statisticians, Vol. I, Cambridge, at University Press, 1954.

[36] Pearson, Karl, Tables of Incomplete Beta Functions, Cambridge University Press, 1934.

[37] Pitman, E.J.G., Lecture Notes on Non-Parametric Statistical Inference. (unpublished).

[38] Rao, C. R., "The Utilization of Multiple Measurements in Problems of Biological Classification", Journal of the Royal Statistical Society, Series B, Vol. X (1948), pp. 159-193.

[39] Rao, C. R., Advanced Statistical Methods in Biometric Research. John Wiley and Sons (1952), New York.

[40] Rao, C. R., "A General Theory of Discrimination When the Information about Alternative Populations is based on Samples", Annals of Mathematical Statistics, Vol. XXV(1954), pp. 651-670.

[41] Robbins, H. E. and Pitman, E. J. G., "Application of the Method of Mixtures to Quadratic Forms in Normal Variates", Annals of Mathematical Statistics, Vol. XX (1949), pp. 552-560.

[42] Robbins, H. E., "Asymptotically Subminimax Solutions of Compound Statistical Decision Problems", Proceedings of the Second Berkeley Symposium in Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 131-148.

[43] Roy, S. N., "On a Heuristic Method of Test Construction, and its Use in Multivariate Analysis", Annals of Mathematical Statistics, Vol. XXIV, 1953, pp. 220-238.

[44] Roy, S. N., A Report on Some Aspects of Multivariate Analysis, North Carolina Institute of Statistics, Mimeograph Series, No. 121, 1954.

[45] Sitgreaves, Rosedith, "On the Distribution of Two Random Matrices used in Classification Procedures", Annals of Mathematical Statistics, Vol. XXIII(1952), pp. 263-270.

[46] Smith, C.A.B., "Some Examples of Discrimination", Annals of Eugenics, Vol. 13(1947), pp. 272-282.

[47] Stekloff, W., "Quelques Applications Nouvelles de la Theorie de Fermeture au Probleme de Representation Approchee de Moments", Memoire de l'Academie Imperiale des Sciences de St. Petersbourg. Vol. XXXII, No. 4, (1914), pp.

[48] Szego, Gâbor, Orthogonal Polynomials. American Mathematical Society Colloquium Publication, Vol. XXIII, 1939.

[49] Uspensky, J. V., Introduction to Mathematical Probability. McGraw Hill Book Company, Inc., 1937.

[50] Wald, Abraham, "On a Statistical Problem Arising in the Classification of an Individual into One of Two Groups", Annals of Mathematical Statistics, Vol. XV, 1944, pp. 145-162.

[51] Wald, Abraham, Selected Papers in Statistics and Probability. McGraw-Hill Book Company, Inc., New York, 1955.

[52] Watson, G. N., Theory of Bessel Functions. Second Edition, Cambridge University Press, 1945.

[53] Welch, B. L., "Note on Discriminant Functions", Biometrika, Vol. XXXI (1939), pp. 218-220.

[54] Whittaker, E. T. and Watson, G. N., A Course of Modern Analysis. Fourth Edition, Cambridge University Press, 1952.

[55] Wilks, S. S., "On Some Generalizations of the Analysis of Variance", Biometrika, Vol. XXIV. (1932), pp. 471-494.