A contribution to the theory of non-parametric statistical inference - the asymptotic equivalence of distributions based on normal theory and on randomisation theory, and an additional paper on Fisher's logarithmic distribution.

By

Samuel D. Silvey, M.A.

PREFACE

A central problem of non-parametric inference is that of testing, against specified alternatives, the hypothesis that the distribution functions of several different populations are identical. A common procedure when faced with a problem of this nature is to assume that all populations from which samples have been taken are normal and to establish a test of the hypothesis within the domain of normal populations. This assumption, of course, is introduced primarily because normal populations are particularly amenable to mathematical analysis. But in justification of its apparently rather limiting nature it is argued (i) that departure from normality of the populations will not affect to any great extent (at least for large samples) the sampling distribution of the statistic used to establish a critical region and (ii) that alternative approaches to the problem involving no assumptions regarding the populations, will (again at least for large samples) lead to similar inferences [5].

One such alternative approach that is often available is to use the same statistic as is used in 'normal theory' but to obtain a critical region from the distribution of the statistic over equally likely permutations of sample values. Leaving aside questions of the power of the test used justification of the normal theory approach on the basis (ii) rests on the equivalence for large samples of the randomisation distribution and the normal theory sampling distribution of the statistic used.

Discussion of randomisation distributions has generally been limited to finding the first four moments of the distributions [10], [13]. The first proof of the limiting form of a randomisation distribution was given by Hotelling and Pabst [4], in connection with rank correlation. More recently Wald and Wolfowitz [11], [12], Noether [8], Daniels [3], Madow [7] have discussed the limiting forms of the randomisation distributions of various statistics.

In the present thesis we discuss conditions under which the randomisation distributions and the normal theory distributions of statistics belonging to a certain class are asymptotically equivalent.

This thesis, which arose from a question put to me by Dr. C. E. V. Leser of Glasgow University, might be regarded as a development of a discussion initiated by, particularly, Wald and Wolfowitz [11]. The present viewpoint, however, is different. The main points of overlap with previous work on this subject are theorems 2·6 and 3·5. Theorem 2·6 is proved here under more general conditions than hitherto, while theorem 3·5 is proved by a new and more direct method.

The additional paper on the logarithmic distribution arose from a question concerning Professor Fisher's original paper on the subject [14] put to me by Mr. M.V. Brian of Glasgow University. Since it was written. Mr. F. J. Anscombe has drawn my attention to the fact that it is very similar to part of his own work on this subject [15].

I wish to thank Dr. R. A. Robb of Glasgow University whose lectures first stimulated my interest in mathematical statistics, who has kept this interest alive in subsequent discussions, and who supervised the major part of my research.

Finally I also wish to express my thanks, for a series of most interesting and instructive lectures, to the staff of the mathematical statistics department in Cambridge University - in particular to Dr. H. E. Daniels, who supervised my study there.

<u>Introduction</u> : The following notation will be used throughout.

An ordered set of variables (or fixed values) $\eta_1, \eta_2, \ldots\ldots \eta_n$, will be denoted by $\underset{\sim}{\eta_n}$

$$\bar{\eta}_n \equiv \frac{1}{n}(\eta_1 + \eta_2 + \cdots + \eta_n)$$

$$m_j(\underset{\sim}{\eta_n}) \equiv \frac{1}{n}\sum_{i=1}^{n}(\eta_i - \bar{\eta}_n)^j \qquad j = 2, 3, \ldots\ldots$$

$$m'_j(\underset{\sim}{\eta_n}) \equiv \frac{1}{n}\sum_{i=1}^{n}\eta_i^j \qquad j = 1, 2, 3, \ldots\ldots$$

$P\{R \mid I_1, I_2, \ldots\} \equiv$ the probability of the relation R given information $I_1, I_2, \ldots\ldots\ldots$

$$f(n) = O(n^\alpha) \equiv \frac{f(n)}{n^\alpha} \longrightarrow \text{a non-zero constant as } n \to \infty$$

$$f(n) = o(n^\alpha) \equiv \frac{f(n)}{n^\alpha} \longrightarrow 0 \text{ as } n \to \infty$$

$$f(n) \leq O(n^\alpha) \equiv f(n) \text{ either } O(n^\alpha) \text{ or } o(n^\alpha).$$

Let (ξ_i) $i = 1, 2, \ldots.$ be a sequence of random variables.

Let $H_0(\mu, \sigma)$ denote the hypothesis that the distribution function of ξ_i is $N(\mu, \sigma^2)$ for all i, i.e., $P\{\xi_i \leq a \mid H_0(\mu, \sigma)\} = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$.

$\underset{\sim}{\xi_n}$ can be represented by a point in an n-dimensional Euclidean space Wn(the sample space) with mutually $\perp^r$ axes $O\xi_1, O\xi_2, \cdots, O\xi_n$.

$H_0(\mu, \sigma)$ defines a probability measure in this space.

Let $t(\underset{\sim}{\xi_n})$ be a statistic which is a Borel measurable function and has the property that $P\{t(\underset{\sim}{\xi_n}) \leq a \mid H_0(\mu, \sigma)\}$ is independent of μ and σ. Then if ~~distribution~~ H_0 denotes the hypothesis that the distribution function of each ξ_i has the same normal form without specifying μ, σ, $P\{t(\underset{\sim}{\xi_n}) \leq a \mid H_0\}$ is meaningful and this is the normal theory distribution function of t.

Let x_i be a fixed value assumed by ξ_i $i = 1, 2, 3, \ldots.$ and let the sequence (x_i) be such that $m_2(x_n) \neq 0$ $n = 2, 3, \ldots.$

Let $\mathcal{X}_n$ denote the set of n! points in Wn obtained by permuting the coordinates $(x_1, x_2, \ldots x_n)$ in all possible ways. (In the case where some of the x's are equal a distinguishing label is supposed attached to each, so that there may be n! points in $\mathcal{X}_n$ - some of which may be coincident but not all since $m_2(x_n) \neq 0$.)

Let $\underset{\sim}{X}_n$ denote the general point of $\mathcal{X}_n$.

Let H_0' denote the hypothesis that each ξ_i has the same distribution function.

H_0' defines a probability measure in the set $\mathcal{X}_n$, the joint distribution of $X_1, X_2, \ldots X_n$ when H_0' is true being given by

$$P\{X_1 = x_{p_1}, X_2 = x_{p_2}, \ldots, X_n = x_{p_n} \mid H_0'\} = \frac{1}{n!}$$

for all permutations $(p_1, p_2, \ldots, p_n)$ of $(1, 2, \ldots, n)$

Then $P\{t(\underset{\sim}{X}_n) \leq a \mid H_0'\} = \frac{1}{n!} \nu\{t(\underset{\sim}{X}_n) \leq a\}$ where $\nu\{t(\underset{\sim}{X}_n) \leq a\}$ denotes the number of points of $\mathcal{X}_n$ for which $t(\underset{\sim}{X}_n) \leq a$, is the randomisation distribution of t.

We are interested in the equivalence

$$P\{t(\underset{\sim}{X}_n) \leq a \mid H_0'\} \equiv P\{t(\underset{\sim}{\xi}_n) \leq a \mid H_0\} \quad \text{(A)}$$

for large values of n.

There is a class of statistics for which the equivalence (A) has a certain geometrical significance and it is this class $\mathcal{Y}_n$ with which we shall be concerned.

$\mathcal{Y}_n$ is the class of statistics t with the properties

(i) If a is any constant

$$t(\xi_1 + a, \xi_2 + a, \ldots, \xi_n + a) = t(\xi_1, \xi_2, \ldots, \xi_n).$$

(ii) If $b > 0$ is a constant

$$t(b\xi_1, b\xi_2, \ldots, b\xi_n) = t(\xi_1, \xi_2, \ldots \xi_n).$$

i.e. t is homogeneous of degree 0.

(iii) $t(\underset{\sim}{X}_n^{(1)}) \neq t(\underset{\sim}{X}_n^{(2)})$ for some distinct pair of points $X_n^{(1)}, X_n^{(2)} \in \mathcal{X}_n$.

$\mathcal{J}_n$ includes many commonly occurring statistics for if $\xi_i' = \dfrac{\xi_i - \bar{\xi}_n}{\sqrt{m_2(\xi_n)}}$ then any function $t(\underset{\sim}{\xi}_n')$ has the properties (i) and (ii). Many statistics in common use are of the form $t(\underset{\sim}{\xi}_n')$. The property (iii) of the class $\mathcal{J}_n$ is introduced to exclude statistics t where randomisation distribution is trivial.

CHAPTER I The Geometrical Interpretation of the Problem.

1·1 Lemma : If $t \in \mathcal{J}_n$, $\mathcal{B}_{n-2}$ denotes the set of points ($\underset{\sim}{\xi}_n$) for which $\bar{\xi}_n = \bar{x}_n$ and $m_2'(\underset{\sim}{\xi}_n) = m_2'(\underset{\sim}{x}_n)$ and $C_{n-2}(t \le a)$ denotes the subset of $\mathcal{B}_{n-2}$ in which $t \le a$, then $P\{t(\underset{\sim}{\xi}_n) \le a \mid H_0\}$ is meaningful and is equal to $\dfrac{L_{n-2}\{C_{n-2}(t \le a)\}}{L_{n-2}\{\mathcal{B}_{n-2}\}}$ where L_r denotes r-dimensional Lebesgue measure.

Proof:

The truth of $H_0(\mu, \sigma)$ associates a probability weight $\frac{1}{\sigma^n(2\pi)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\xi_i-\mu)^2}$ with the point $\underset{\sim}{\xi}_n$ of Wn.

When $H_0(\mu, \sigma)$ is true all points of Wn for which $\sum_{i=1}^{n}(\xi_i-\mu)^2$ is constant are equally likely, i.e. all points on the surface of an n-sphere centred on $(\mu, \mu, \ldots\ldots, \mu)$ are equally likely. Hence if C_{n-r} is a region of $(n-r)$ dimensions contained in the surface of such a hypersphere while c_{n-r} is a measurable subset of C_{n-r},

$$P\{\underset{\sim}{\xi}_n \in c_{n-r} \mid H_0(\mu,\sigma), \underset{\sim}{\xi}_n \in C_{n-r}\} = \frac{L_{n-r}(c_{n-r})}{L_{n-r}(C_{n-r})}.$$

In particular the set C_{n-2} is contained on the surface of the hypersphere with equation

$$(\xi_1-\mu)^2+(\xi_2-\mu)^2+\cdots+(\xi_n-\mu)^2 = n\left[m_2(\underset{\sim}{x}_n)+(\mu-\bar{x}_n)^2\right]$$

for the hyperplane $\xi_1+\xi_2 \ldots\ldots+\xi_n = n\bar{x}_n$ meets this where it meets

$$\xi_1^2+\xi_2^2+\cdots+\xi_n^2 - 2n\mu\bar{x}_n + n\mu^2 = n\left[m_2(\underset{\sim}{x}_n)+(\mu-\bar{x}_n)^2\right]$$

$$\text{i.e.}\quad \xi_1^2+\xi_2^2+\cdots+\xi_n^2 = n\,m_2'(\underset{\sim}{x}_n)$$

Hence all points of C_{n-2} are equally likely and

$$P\{t(\underset{\sim}{\xi}_n) \le a \mid H_0(\mu,\sigma), \underset{\sim}{\xi}_n \in \mathcal{B}_{n-2}\} = \frac{L_{n-2}\{C_{n-2}(t \le a)\}}{L_{n-2}\{\mathcal{B}_{n-2}\}} \quad \ldots\ldots (1.1.1)$$

We have now to show that the R.S. of (1, 1, 1) is independent of μ, σ, $\bar{x}_n$ and $m_2'(x_n)$. It is clearly independent of μ, and σ for C_{n-2} depends on $\bar{x}_n$ and $m_2'(x_n)$ only.

Varying $\overline{x}_n$ is equivalent to translating C_{n-2} in a direction parallel to the line $\xi_1 = \xi_2 = \ldots = \xi_n$ and such translation leaves $C_{n-2}(t \leq a)$ relatively unaltered by the property (i) of t. Hence R.S. of (1.1.1) is independent of $\overline{x}_n$ and we can take $\overline{x}_n = 0$ so that the centre of C_{n-2} is the origin. Varying $m'_2(x_n)$ is equivalent to altering the radius of C_{n-2} and since by (ii) t is constant on any line of fixed direction through the origin the R.S. of (1.1.1) is independent of $m'_2(x_n)$.

Hence $P\{t \leq a \mid H_0(\mu, \sigma), \xi_n \in \mathcal{C}_{n-2}\}$ is independent of μ, σ, $\overline{x}_n$ and $m'_2(x_n)$ and it follows that $P\{t \leq a \mid H_0\}$ is meaningful and is equal to $\dfrac{L_{n-2}\{C_{n-2}(t \leq a)\}}{L_{n-2}\{\mathcal{C}_{n-2}\}}$.

<u>Note</u> : We have excluded the case $m'_2(x_n) = 0$ tacitly throughout the proof. Since given $H_0(\mu, \sigma)$ $m'_2(x_n) = 0$ with probability zero, the exclusion of this case does not affect the result.

<u>1·2 Lemma</u> : $C_{n-2} \supset \mathcal{X}_n$

This is immediately obvious for if $\underset{\sim}{X}_n$ is any point of $\mathcal{X}_n$ we have

$$X_1 + X_2 + \cdots + X_n = x_{\rho_1} + x_{\rho_2} \cdots + x_{\rho_n} = n\,\overline{x}_n$$

and

$$X_1^2 + X_2^2 \cdots + X_n^2 = x_{\rho_1}^2 + x_{\rho_2}^2 \cdots + x_{\rho_n}^2 = n\, m_2'(\underset{\sim}{x}_n).$$

and so $\underset{\sim}{X}_n \in \mathcal{C}_{n-2}$

Hence $\mathcal{X}_n \subseteq C_{n-2}$ and the equality sign can clearly be dropped.

<u>1·3</u> : If $V\{C_{n-2}(t \leq a)\}$ is the number of points of $\mathcal{X}_n$ occurring in $C_{n-2}(t \leq a)$

$$P\{t(\underset{\sim}{X}_n) \leq a \mid H_0'\} = \frac{V\{C_{n-2}(t \leq a)\}}{n!}$$

Hence the equivalence (A) occurs if and only if

$$\frac{V\{C_{n-2}(t \leq a)\}}{n!} = \frac{L_{n-2}\{C_{n-2}(t \leq a)\}}{L_{n-2}\{\mathcal{C}_{n-2}\}}$$

i.e. if and only if the number of points of $\mathcal{X}_n$ occurring in $C_{n-2}(t \leq a)$

is proportional to the measure of this set.

1.4 : Asymptotically the randomisation distribution and the distribution under normal theory of any statistics $t \in \mathcal{J}_n$ will be equivalent if the sequence (x_n) is such that the set of points $\mathcal{X}_n$ tends to be uniformly distributed throughout C_{n-2}: i.e. if the number of points occurring in any measurable subset of C_{n-2} tends to be proportional to its measure.

This is a very restrictive condition which, as will be seen and is intuitively to be expected, necessitates that the set of measures x_1, x_2, ...x_n should tend to be normally distributed as n increases.

For any particular t, however we require for (A) only that the points $\mathcal{X}_n$ should tend to be distributed uniformly relative to a particular class of subsets of C_{n-2} viz., the class of subsets of the form $C_{n-2}(t \leq a)$.

As is usually the case, this geometrical interpretation of the problem is more useful for its suggestiveness than for any other purpose. We continue the discussion on a narrower basis by considering subsets of $\mathcal{J}_n$ with the geometrical picture as a guide.

CHAPTER II The Class of Linear Combinations.

Let the sequence (R_n) where

$$R_n(\underset{\sim}{\xi_n}) = \frac{\sqrt{(n-1)}\sum_{i=1}^{n} y_{in}(\xi_i - \bar{\xi}_n)}{\sqrt{\left\{\sum_{i=1}^{n}(y_{in}-\bar{y}_{nn})^2 \sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2\right\}}}$$

where $\underset{\sim}{y_{nn}} = (y_{1n}, y_{2n}, \ldots\ldots, y_{nn})$ are assigned sets of constants for $n = 2, 3, 4, \ldots$ such that $m_2(\underset{\sim}{y_{nn}}) \neq 0$, define for each n, the statistic R, and let $\mathcal{R}$ be the class of statistics of the form R. From the point of view of randomisation theory $\mathcal{R}$ may be regarded as the set of standardised linear combinations. From the point of view of normal theory $\mathcal{R}$ is the set of standardised product-moment correlation coefficients.

Clearly $R_n \in \mathcal{Y}_n$

2·1 Theorem : A necessary and sufficient condition that

$$\mathop{\mathcal{L}}_{n\to\infty} P\{R_n(\underset{\sim}{x_n}) \leq a \mid H_0'\} = \mathop{\mathcal{L}}_{n\to\infty} P\{R_n(\underset{\sim}{\xi_n}) \leq a \mid H_0\} \qquad (B)$$

for all a and all $R \in \mathcal{R}$ is that the set of measures $x_1, x_2, \ldots\ldots x_n$ should tend to be normally distributed.

The proof is divided into several parts.

2·2 : Necessity. Consider the statistic R^0 defined by

$$R_n^0(\underset{\sim}{\xi_n}) = \frac{\sqrt{n}\,(\xi_1 - \bar{\xi}_n)}{\sqrt{\left\{\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2\right\}}}$$

i.e.

$$\underset{\sim}{y_{nn}} = (1, 0, 0, \ldots\ldots 0)$$

Without loss of generality we can take $\bar{x}_n = 0$ and $m_2(\underset{\sim}{x_n}) = 1$ so that the hypersphere C_{n-2} has equations

$$\xi_1 + \xi_2 + \cdots \quad + \xi_n = 0$$
$$\xi_1^2 + \xi_2^2 + \cdots \quad + \xi_n^2 = n .$$

R_n^0 is constant (and equal to C) on the intersection of C_{n-2} with the hypersphere $\xi_1 = C$.

Hence if a_1, a_2 are constants such that $-\sqrt{(n-1)} \le a_1 < a_2 \le \sqrt{(n-1)}$, $C_{n-2}(a_1 \le R_n^0 \le a_2)$ is that part of C_{n-2} contained between the hyperplanes $\xi_1 = a_1$, $\xi_1 = a_2$.

Now $\xi_1 = |c| \le \sqrt{(n-1)}$ meets C_{n-2} in the "surface" of an $(n-2)$ sphere of radius $\sqrt{n}\cos\theta$ where $\sin\theta = \frac{|c|}{\sqrt{(n-1)}}$

The 'surface' of an n-sphere of radius $r = 2\cdot\frac{\pi^{n/2}}{\Gamma(\frac{n}{2})} r^{n-1}$.

Hence

$$L_{n-2}\left\{C_{n-2}(a_1 \le R_n^0 \le a_2)\right\} = \int_{\sin^{-1}\frac{a_1}{\sqrt{(n-1)}}}^{\sin^{-1}\frac{a_2}{\sqrt{(n-1)}}} 2\cdot\frac{\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})}\left(\sqrt{n}\cos\theta\right)^{n-3}\sqrt{n}\,d\theta,$$

while

$$L_{n-2}\left\{\mathcal{B}_{n-2}\right\} = 2\frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\left(\sqrt{n}\right)^{n-2}$$

After reduction this gives

$$\frac{L_{n-2}\left\{C_{n-2}(a_1 \le R_n^0 \le a_2)\right\}}{L_{n-2}(\mathcal{B}_{n-2})} = \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})\sqrt{\pi}}\int_{\frac{a_1}{\sqrt{(n-1)}}}^{\frac{a_2}{\sqrt{(n-1)}}}\left(1-y^2\right)^{\frac{n-4}{2}}dy \quad \cdots\text{-}(2.2.1)$$

$$= \frac{\sqrt{n}}{\sqrt{(n-1)}}\frac{\Gamma(\frac{n-1}{2})}{\sqrt{\frac{n}{2}}\,\Gamma(\frac{n-2}{2})}\frac{1}{\sqrt{2\pi}}\int_{a_1}^{a_2}\left(1-\frac{y^2}{n-1}\right)^{\frac{n-4}{2}}dy.$$

Now as $n \to \infty$

$$\sqrt{\frac{n}{n-1}} \to 1 \qquad \frac{\Gamma(\frac{n-1}{2})}{\sqrt{(\frac{n}{2})}\,\Gamma(\frac{n-2}{2})} \to 1 \quad \text{and} \quad \left(1-\frac{y^2}{n-1}\right)^{\frac{n-4}{2}} \to e^{-\frac{y^2}{2}}$$

uniformly in the interval $a_1 \le y \le a_2$.

Hence

$$\mathop{\mathcal{L}}_{n\to\infty} P\left\{a_1 \le R_n^0(\underset{\sim}{\xi_n}) \le a_2 \,\middle|\, H_0\right\} = \frac{1}{\sqrt{2\pi}}\int_{a_1}^{a_2} e^{-\frac{y^2}{2}}dy.$$

Let $n\,\phi_n(a_1 \le x \le a_2)$ be the number of the values of the set $x_1, x_2, \ldots\ldots x_n$ for which $a_1 \le x_i \le a_2$.

Then $V\left\{C_{n-2}(a_1 \le R_n^0 \le a_2)\right\} = n!\,\phi_n(a_1 \le x \le a_2)$ for there are $(n-1)!$ points of $\mathcal{X}_n$ with each fixed value for ξ_1.

Hence $P\left\{a_1 \le R_n^0(\underset{\sim}{x_n}) \le a_2 \,\middle|\, H_0'\right\} = \phi_n(a_1 \le x \le a_2)$.

So in order that (B) be satisfied for R^0 it is necessary that

$$\mathop{\mathcal{L}}_{n\to\infty} \Phi_n(a_1 \le x \le a_2) = \frac{1}{\sqrt{2\pi}}\int_{a_1}^{a_2} e^{-\frac{y^2}{2}}\,dy$$

i.e. that the x's should tend to be normally distributed.

2·3 : Sufficiency :

2·3·1 Lemma : $\mathop{\mathcal{L}}_{n\to\infty} P\left\{ a_1 \le R_n(\underset{\sim}{\xi}_n) \le a_2 \mid H_0 \right\} = \frac{1}{\sqrt{2\pi}}\int_{a_1}^{a_2} e^{-\frac{x^2}{2}}\,dx$

for all $R \in \mathcal{R}$

Proof : 2·2·1 gives the distribution of $\frac{R_n^0}{\sqrt{(n-1)}}$ under normal theory, for writing $b_1 = \frac{a_1}{\sqrt{(n-1)}}$, $b_2 = \frac{a_2}{\sqrt{(n-1)}}$ we have, for $-1 \le b_1 < b_2 \le 1$

$$P\left\{ b_1 \le \frac{R_n^0(\underset{\sim}{\xi}_n)}{\sqrt{(n-1)}} \le b_2 \mid H_0 \right\} = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)\sqrt{\pi}}\int_{b_1}^{b_2}(1-y^2)^{\frac{n-4}{2}}\,dy.$$

Thus $\frac{R_n^0(\underset{\sim}{\xi}_n)}{\sqrt{(n-1)}}$ has the same distribution as has the product-moment correlation coefficient in samples of n from a bivariate normal population with zero correlation.

$\frac{R_n^0(\underset{\sim}{\xi}_n)}{\sqrt{(n-1)}}$ is in fact the product-moment correlation coefficient for the n sets of measures $(\xi_1, 1)$, $(\xi_2, 0)$, $(\xi_3, 0)$ ..... $(\xi_n, 0)$.

We show that every $R_n(\underset{\sim}{\xi}_n)$ has the same distribution when H_0 is true.

Let $R_n = \sqrt{(n-1)}\dfrac{\sum_{i=1}^{n} y_i(\xi_i - \bar{\xi}_n)}{\sqrt{\left\{\sum_{i=1}^{n}(y_i-\bar{y}_n)^2 \sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2\right\}}}$ be any member of $\mathcal{R}_n$

Then

$$\frac{R_n}{\sqrt{(n-1)}} = \frac{\sum_{i=1}^{n} y_i' \xi_i}{\sqrt{\sum_{i=1}^{n}(\xi_i-\bar{\xi}_n)^2}} \quad \text{where} \quad y_i' = \frac{y_i - \bar{y}_n}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y}_n)^2}}$$

Now let $\underset{\sim}{\eta}_n = P\underset{\sim}{\xi}_n$ be an orthogonal transformation such that

$$\eta_1 = \frac{1}{\sqrt{n}}(\xi_1 + \xi_2 + \cdots + \xi_n).$$

$$\eta_2 = \sum_{i=1}^{n} y_i'\xi_i$$

η_1, η_2 are orthogonal since $\sum_{i=1}^{n} y_i' = 0.$

If each ξ_i is $N(0, 1)$ and the ξ_s are independent, then each η_i is $N(0, 1)$ and the η's are independent.

Also we have $$\frac{R_n^2}{n-1} = \frac{y_2^2}{y_2^2+y_3^2+\cdots+y_n^2}$$

Hence when $H_0(0, 1)$ is true every R_n can be expressed in exactly the same form in terms of independent $N(0, 1)$ random variables.

So the distribution of every R_n when $H_0(0, 1)$ is true (and so when H_0 is true) is the same.

Hence under normal theory every R has the same limiting distribution and this is the limiting distribution of R^0.

This proves the lemma.

2·3·2 Lemma : Let $\alpha(h) = (\alpha_1, \alpha_2, \cdots \alpha_h)$ be a partition of an integer S.

Let $S_{\alpha(h)}(\underset{\sim}{x_n})$ be the symmetric polynomial,

$$S_{\alpha(h)}(\underset{\sim}{x_n}) = \sum x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_h}^{\alpha_h}$$

where summation extends over all ordered sets $(i_1, i_2, \ldots, i_h)$ of h integers from 1, 2,.....,n.

Then $$S_{\alpha(h)}(\underset{\sim}{x_n}) = n^h \prod_{i=1}^{h} m'_{\alpha_i}(\underset{\sim}{x_n}) + \tau(m')$$

where $\tau(m')$ is a sum of terms of the form $C(\alpha, \beta)\, n^{h'} \prod_{i=1}^{h'} m'_{\beta_i}(\underset{\sim}{x_n})$

and $\beta(h')$ is a partition of S such that each βi is the sum of one or more $\alpha_j s$, $h' < h$ and $C(\alpha, \beta)$ is a function of the partitions $\alpha(h)$, $\beta(h')$ independent of n .

This lemma states an identity in symmetric polynomials, and the proof is omitted as it occurs repeatedly (in slightly different form) e.g [7].

2·3·3 Lemma : Let $E(\emptyset | H)$ denote the expectation of the random variable $\emptyset$ whose distribution is defined by a hypothesis H.

Then $$E\left(X_{i_1}^{\alpha_1} X_{i_2}^{\alpha_2} \cdots X_{i_h}^{\alpha_h} \middle| H'_0 \right) = \frac{1}{n^{[h]}} S_{\alpha(h)}(\underset{\sim}{x_n}).$$

2·3·4 Lemma : Let r be a fixed integer > 2, and let $n > r$.

If $\bar{x}_n = 0$, $\bar{y}_n = 0$, $m_2(\underset{\sim}{x}_n) = m_2(\underset{\sim}{y}_n) = 1$ and

$$m_j(\underset{\sim}{x}_n) = \left.\begin{matrix} \dfrac{j!}{2^{j/2}(j/2)!} + o(1), & j \text{ even} \\ o(1) & , j \text{ odd} \end{matrix}\right\} \quad j = 3, 4, \cdots r.$$

then

$$E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i X_i\right)^r \Big| H_0'\right\} = \begin{cases} \dfrac{r!}{2^{r/2}\left(\frac{r}{2}\right)!} + o(1), & r \text{ even} \\ o(1) & , r \text{ odd} \end{cases}$$

Proof : Expanding $\left(\sum_{i=1}^{n} y_i X_i\right)^r$ by the multinomial expansion, taking expectations term by term and collecting terms gives

$$E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i X_i\right)^r \Big| H_0'\right\} = \frac{1}{n^{r/2}}\sum \frac{1}{n^{[h]}}\, C\{\alpha(h)\}\, S_{\alpha(h)}(\underset{\sim}{y}_n)\, S_{\alpha(h)}(\underset{\sim}{x}_n) \qquad \text{(C)}$$

where summation extends over all partitions of r and

$$C(\alpha(h)) = \frac{r!}{\alpha_1!\,\alpha_2!\cdots\alpha_h!}\;\frac{1}{\pi_1!\,\pi_2!\cdots\pi_p!}$$

where, when the α's are all chosen from p different positive integers $j_1, j_2, \ldots, j_p$, π_i of the α's are equal to j_i.

Now by lemma (2·3·2) since the number of terms in $\gamma(m')$ is independent of n and since $m_{\beta_i}(\underset{\sim}{x}_n)$ is at most O(1), $\beta_i \le r$

$$S_{\alpha(h)}(\underset{\sim}{x}_n) = n^h \prod_{i=1}^{h} m_{\alpha_i}(\underset{\sim}{x}_n) + Q_{\alpha(h)}(\underset{\sim}{x}_n).$$

where $Q_{\alpha(h)}(\underset{\sim}{x}_n)$ is at most $O(n^{h-1})$.

Hence

$$\frac{1}{n^{[h]}} S_{\alpha(h)}(\underset{\sim}{x}_n) = \prod_{i=1}^{h} m_{\alpha_i}(\underset{\sim}{x}_n) + o(1).$$

If any α_i is odd $\quad \prod_{i=1}^{h} m_{\alpha_i}(\underset{\sim}{x}_n) = o(1).$

If r is odd at least one α_i in each partition $\alpha(h)$ is odd and so since the number of terms on R.S. of (C) is independent of n,

$$E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i X_i\right)^r \Big| H_0'\right\} = o(1), \quad r \text{ odd}.$$

[$S_{\alpha(h)}(\underset{\sim}{y}_n) \leq O(n^{r/2})$. This can be seen by applying Lemma 2·3·2 to $S_{\alpha(h)}(\underset{\sim}{y}_n)$ and noting that , for $j \geq 2$

$$|m_j(\underset{\sim}{y}_n)| = \frac{1}{n}|y_1^j + y_2^j + \cdots + y_n^j| \leq \frac{1}{n}(y_1^2 + y_2^2 \cdots + y_n^2)^{j/2} = n^{\frac{j-2}{2}}]$$

If r is even , say r = 2q and if every α_i is even say $\alpha_i = 2\beta_i$. $\beta(h)$ is a partition of q and

$$\prod_{i=1}^{h} m_{\alpha_i}(\underset{\sim}{x}_n) = \prod_{i=1}^{h} \frac{(2\beta_i)!}{2^{\beta_i}\beta_i!}$$

$$= \frac{1}{2^q}\prod_{i=1}^{h}\frac{(2\beta_i)!}{\beta_i!}$$

So we have

$$E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i X_i\right)^{2q} \middle| H_0'\right\} = \frac{1}{n^q}\left\{\frac{1}{2^q}\sum\left[\frac{(2q)!}{\prod_{i=1}^{h}(2\beta_i)!}\prod_{i=1}^{h}\frac{(2\beta_i)!}{\beta_i!}\frac{1}{\pi_1!\pi_2!\cdots\pi_p!} S_{\alpha(h)}(\underset{\sim}{y}_n)\left(1+o(1)\right)\right] + o(1)\right.$$

summation now taking place over all partitions $\alpha(h)$ in which each α_i is even $\alpha_i = 2\beta_i$.

$$= \frac{(2q)!}{2^q q!}\frac{1}{n^q}\left\{\sum\frac{q!}{\beta_1!\cdots\beta_h!}\frac{1}{\pi_1!\pi_2!\cdots\pi_p}S_{\alpha(h)}(\underset{\sim}{y}_n)\left(1+o(1)\right)\right\} + o(1)$$

$$= \frac{(2q)!}{2^q q!}\left[\frac{y_1^2 + y_2^2 \cdots + y_n^2}{n}\right]^q + o(1)$$

$$= \frac{(2q)!}{2^q q!} + o(1).$$

and the lemma is proved.

2·3·5 Lemma : If, for sufficiently large n

$$\frac{m_j(\underset{\sim}{x}_n)}{[m_2(\underset{\sim}{x}_n)]^{j/2}} = \begin{cases} \dfrac{j!}{2^{j/2}(j/2)!} + o(1), & j \text{ even} \\ o(1), & j \text{ odd} \end{cases} \qquad j = 2, 3, 4, \cdots$$

then the randomisation distribution of any $R \in \mathcal{R}$ is asymptotically N(0, 1).

For writing $x_i' = \dfrac{x_i - \bar{x}_n}{\sqrt{m_2(\underset{\sim}{x}_n)}}$, $y_{in}' = \dfrac{y_{in} - \bar{y}_{nn}}{\sqrt{m_2(\underset{\sim}{y}_{nn})}}$.

any $R_n(X_n)$ can be written $\dfrac{\sqrt{(n-1)}}{n}\sum_{i=1}^{n} y_{in}' X_i'$ where the sequences ($\underset{\sim}{x}_n'$), ($\underset{\sim}{y}_{nn}'$) satisfy the conditions of lemma 2·3·4.

It follows that as $n \to \infty$ the rth moment of the distribution of $R_n(\underset{\sim}{X}_n)$

tends to that of a N(0, 1) distribution. Since the latter distribution is completely determined by its moments, the limiting distribution of $R_n(\underset{\sim}{X_n})$ is N(0, 1). This completes the proof of Theorem 2·1.

<u>2·4</u>: <u>Geometrical Interpretation</u> : The subset of C_{n-2} for which $a_1 \le R_n \le a_2$ may be considered as an "interval" of C_{n-2}

The geometrical interpretation of 2·1 is that the set of points $\mathcal{X}_n$ tends to be distributed uniformly relative to all "intervals" of C_{n-2} if and only if the measures $x_1, x_2, \ldots, x_n$ tend to be normally distributed.

However for uniform distribution of the set of points $\mathcal{X}_n$ relative only to various subsets of the class of all 'intervals' of C_{n-2}, much less restrictive conditions on the distributions of the measures $x_1, x_2, \ldots x_n$ is sufficient, as is shown by the following, theorem 2·6.

<u>2·5 Lemma</u> : Let <u>a</u> be a constant such that $0 \le a < \frac{1}{2}$ and let $b = \frac{1}{2} - a$.

If $\bar{\mathcal{X}}_n = \bar{y}_n = 0$, $m_2(\underset{\sim}{\mathcal{X}_n}) = m_2(\underset{\sim}{y_n}) = 1$

$$\left.\begin{aligned} m_j(\underset{\sim}{y_n}) &\le O\left[n^{a(j-2)}\right] \\ m_j(\underset{\sim}{x_n}) &= o\left[n^{b(j-2)}\right] \end{aligned}\right\} \quad j = 3, 4, \ldots r, \text{ r being a fixed integer } 3 \le r \le n,$$

then

$$E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i x_i\right)^r \Big| H_0'\right\} = \begin{cases} \dfrac{r!}{\left(\frac{r}{2}\right)!\,2^{r/2}} + o(1) & \text{r even} \\ o(1). & \text{r odd} \end{cases}$$

<u>Proof</u> : In 2·3·4 (C(we had

$$\text{2·5·1} \quad E\left\{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i x_i\right)^r \Big| H_0'\right\} = \frac{1}{n^{r/2}}\sum \frac{1}{n^{[h]}}\, C(\alpha(h))\, S_{\alpha(h)}(\underset{\sim}{y_n})\, S_{\alpha(h)}(\underset{\sim}{x_n})$$

summation extending over all partitions $\alpha(h)$ of r. $C[\alpha(h)]$ is independent of n and the number of terms on R.S. is independent of n .

Applying lemma 2·3·2 to $S_{\alpha(h)}(\underset{\sim}{y}_n)$ and $S_{\alpha(h)}(\underset{\sim}{x}_n)$ shows that the R.S. can be expressed as a sum of terms of the form

$$\underline{2\cdot5\cdot2}\quad C(\alpha(h))\, k\left[\alpha(h), \beta(h_1), \gamma(h_2)\right] \frac{1}{n^{r/2}} \frac{1}{n^{[R]}} \left\{ n^{h_1} \prod_{i=1}^{h_1} m_{\beta_i}(\underset{\sim}{y}_n) \right\} \left\{ n^{h_2} \prod_{j=1}^{h_2} m_{\gamma_j}(\underset{\sim}{x}_n) \right\}.$$

where

(i) the number of terms is independent of n .

(ii) $\left.\begin{matrix}\gamma(h_2)\\ \beta(h_1)\end{matrix}\right\}$ is a partition of r in which $\left\{\begin{matrix}h_2 \le h\\ h_1 \le h\end{matrix}\right\}$ and each $\left\{\begin{matrix}\gamma\\ \beta\end{matrix}\right\}$ is the sum of one or more α's.

(iii) $k\left[\alpha(h), \beta(h_1), \gamma(h_2)\right]$ is independent of n and is equal to 1 if $h_1 = h_2 = h$ so that $\alpha(h)$, $\beta(h_1)$ and $\gamma(h_2)$ are the same partition of r.

We consider the order of the term 2·5·2.

If any $\gamma_j > 2$ and every $\beta_i \geqslant 2$ the order of the term 2·5·2 is

$$o\left[n^{h_1+h_2-h-r/2+a(r-2h_1)+b(r-2h_2)}\right]$$
$$= o\left[n^{2bh_1+2ah_2-h}\right] \quad \text{since } a+b=\tfrac{1}{2}.$$
$$= o\left[n^{2bh+2ah-h}\right] = o(1) \quad \text{since } h_1 \le h,\ h_2 \le h,\ b>0$$

If any $\gamma_i = 1$ the term 2·5·2 is zero.

Also if any $\beta_i = 1$ the term 2·5·2 is zero.

Hence if r is odd, the R.S. of 2·5·1 is O(1), for if r is odd at least one γ_j in each term is odd i.e. is either 1 or is greater than 2.

The only terms which can be greater than O(1) are those in which each $\gamma_j = 2$. This necessitates r even.

<u>If r is even</u> and $h_1 = \frac{r}{2}$ while $\gamma_1 = \gamma_2 = \ldots\ldots = \gamma_{[r/2]} = 2$ then

(i) the term 2·5·2 is 0 if any $\beta_i = 1$

(ii) if each $\beta_i \geqslant 2$ the term 2·5·2 is not greater than

$$O\left(n^{2bh_1+2ah_2-h}\right)$$

and this is O(1) unless $h_1 = h_2 = h$ for $2bh_1 + 2ah_2 \le h$, the equality sign occurring only when $h_1 = h_2 \neq h$.

Hence the only term of the form 2·5·2 which may not be $o(1)$ is the term in which $h_1 = h_2 = h$ and $\beta(h_1) = \gamma(h_2) = \alpha(h) = (2, 2, \ldots\ldots 2)$

For this term $k = 1$, $C\left[\alpha(h)\right] = \dfrac{r!}{\left(\frac{r}{2}\right)! \, 2^{r/2}}$ by 2·3·4 and the term itself is

$$\frac{r!}{\left(\frac{r}{2}\right)! \, 2^{r/2}} + o(1).$$

This completes the proof.

2·6 Theorem : If $0 \leq a < \frac{1}{2}$ and $b = \frac{1}{2} - a$ and $\mathcal{R}_a$ is the subset of the class $\mathcal{R}$ for which $\dfrac{m_j(y_n)}{\left[m_2(y_n)\right]^{j/2}} \leq O\left[n^{a(j-2)}\right]$ $\quad j = 3, 4, \ldots\ldots n$

then the randomisation distribution of any $R \in \mathcal{R}_a$ is $N(0, 1)$ asymptotically provided $\dfrac{m_j(x_n)}{\left[m_2(x_n)\right]^{j/2}} = o\left[n^{b(j-2)}\right]$ $\quad j = 3, 4, \ldots\ldots\ldots n.$

Proof : This follows from 2·5 in the same ways as does 2·3·5 from 2·3·4.

Theorem 2·6 is a more general form of a theorem by Wald and Wolfowitz. [11]

2·7: General Discussion of Chapter 2.

Theorem 2·1 goes part of the way towards answering the question - is the asymptotic randomisation distribution of any statistic equivalent to its asymptotic normal theory distribution, when in fact the population from which the sample has been drawn is normal? Intuition would indicate that this must be so. Theorem 2·1 answers the question for a particular class of statistics - the class $\mathcal{R}$ of correlation coefficients.

Theorem 2·6 in the general form proved here, is also mainly of theoretical interest in that it demonstrates how mild are the conditions sufficient to ensure that the asymptotic normal theory distribution and the asymptotic randomisation distribution of a statistic $R \in \mathcal{R}$ be equivalent.

From this point of view it strengthens the conviction that the normal theory approach for large samples leads to more or less the same results as the randomisation approach.

From a slightly different viewpoint this theorem emerges as a particular case of the Central Limit Theorem when the random variables in the sequence concerned are not independent. This aspect of the Central Limit Theorem has not been fully treated at all and the main results in this field which deal with cases in which certain subsequences are 'nearly independent' do not seem to be applicable in the case of the sequence (X_i) of random variables which are 'far from independent'.

The main interest of this paper, however, lies in investigating the asymptotic equivalence of the two approaches with reference to statistics commonly used in tests of significance. While the results of the next chapter are more useful in this respect, some commonly occurring statistics are members of the class $\mathcal{R}$, and the corresponding cases of Theorem 2·6 are of some interest per se.

2·7·1 : Asymptotic normality of the product-moment rank correlation coefficient. This result was first proved by Hotelling and Pabst. It is derived from Theorem 2·6 setting

$$y_{in} = i \qquad i = 1, 2, \ldots\ldots n$$

$$x_i = i \qquad i = 1, 2, \ldots\ldots n.$$

The corresponding $R(\underset{\sim}{X_n})$ is clearly the standardised product moment rank correlation coefficient.

Also for these values

$$\frac{m_j(\underset{\sim}{y_{nn}})}{[m_2(\underset{\sim}{y_{nn}})]^{j/2}} = \frac{m_j(\underset{\sim}{x_n})}{[m_2(\underset{\sim}{x_n})]^{j/2}} \leq O(1) \qquad j = 3, 4, \cdots$$

for

$$\bar{x}_n = \tfrac{1}{2}(n+1)$$

$$m_2(\underset{\sim}{x}_n) = \frac{1}{12}(n^2-1)$$

$$m_j(\underset{\sim}{x}_n) = \sum_{\alpha=1}^{j} {}^jC_\alpha \left[\tfrac{1}{2}(n+1)\right]^{j-\alpha}(-1)^\alpha m'_\alpha(\underset{\sim}{x}_n)$$

where $$m'_\alpha(\underset{\sim}{x}_n) = \frac{1}{n}\sum_{i=1}^{n} i^\alpha = O[n^\alpha]$$

Hence $$m_j(\underset{\sim}{x}_n) \not> O[n^j]$$

and $$\frac{m_j(\underset{\sim}{x}_n)}{[m_2(\underset{\sim}{x}_n)]^{j/2}} \not> O(1)$$

Hence the conditions of Theorem 2·6 are satisfied a fortiori.

This provides an interesting illustration of the weakness of the normal theory assumption in the case of large samples. For in this case the assumption that the set of measures 1, 2,......n is a random sample from a normal population would lead to more or less the same result as (here) the more conventional randomisation approach.

2·7·2 : Sampling without Replacement from a Finite Population.

Let m be the arithmetic mean of a random sample of fn (where f, the sampling fraction, is a constant between 0 and 1) drawn without replacement from a finite population of n members with mean μ and variance σ^2. Then the sampling distribution of m is asymptotically $N\left[\mu, \frac{\sigma^2}{n}(1-f)\right]$

This result was proved by Madow [7].

It is derived from Theorem 2·6 by setting

$$\begin{aligned} y_{in} &= \frac{1}{fn} - \frac{1}{n} \qquad i = 1, 2, \cdots fn; \\ &= -\frac{1}{n} \qquad i = fn+1, \cdots, n. \end{aligned}$$

and considering the x's as the measures of the population.

If now we regard the first fn measures of a permutation of the population values as a random sample of fn members drawn without replacement from the population, the statistic $R_m(\underset{\sim}{X}_n)$ corresponding to the above values of the y's is given by

$$R_m(\underset{\sim}{X}_n) = \frac{\sqrt{fn}\,(m-\mu)}{\sigma\sqrt{(1-f)}} .$$

Since $\bar{y}_n = 0$ and $\sum_{i=1}^{n} y_{in}^2 = fn\,\frac{1}{n^2}\left(\frac{1}{f}-1\right)^2 + n(1-f)\cdot\frac{1}{n^2}$

$$\doteq \frac{1}{n}\left(\frac{1}{f}-1\right);$$

also $\bar{x}_n = \mu$ and $\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2$.

For the statistic R_m we have

$$m_2(\underset{\sim}{y}_{nn}) = \frac{1}{n^2}\left(\frac{1}{f}-1\right)$$

and

$$\begin{aligned} m_j(\underset{\sim}{y}_{nn}) &= \frac{1}{n}\left\{fn\left(\frac{1}{fn}-\frac{1}{n}\right)^j + n(1-f)\left(-\frac{1}{n}\right)^j\right\} \\ &= \frac{1}{n^j}\left\{f\left(\frac{1}{f}-1\right)^j + (-1)^j(1-f)\right\}. \end{aligned}$$

$$\therefore \quad \frac{m_j(\underset{\sim}{y}_{nn})}{\left[m_2(\underset{\sim}{y}_{nn})\right]^{j/2}} = O(1) \qquad j = 3, 4, \cdots\cdots$$

Hence by Theorem 2·6 R_m is asymptotically distributed in the N(0, 1) form provided the population moments are such that

$$\frac{m_j(\underset{\sim}{x_n})}{[m_2(\underset{\sim}{x_n})]^{j/2}} = o\left(n^{\frac{j-2}{2}}\right) \qquad j = 3, 4, \cdots$$

This condition is satisfied by almost all populations - effectively the only exception being the case where almost all individuals in the population have the same measure ['almost all' in the sense that the number of individuals with this measure is O(n)].

2·7·3 Studentisation of $R_m(\underset{\sim}{X_n})$:

If in $R_m(\underset{\sim}{X_n})$ we replace σ by S, its sample estimate given by $S^2 = \frac{1}{fn-1}\sum_{i=1}^{fn}(X_i - m)^2$ and let the resulting statistic be denoted by $t(\underset{\sim}{X_n})$, then $t(\underset{\sim}{X_n})$ is the analogue for a sample drawn without replacement from a finite population, of Student's t for normal theory.

Theorem: If $\sigma^2 = O(1)$ and if $\frac{m_j(\underset{\sim}{x_n})}{[m_2(\underset{\sim}{x_n})]^{j/2}} = o\left(n^{\frac{j-2}{2}}\right)$, $j > 2$ [i.e. if $m_j(\underset{\sim}{x_n}) = o\left[n^{\frac{j-2}{2}}\right]$, $j > 2$, in this case], then the randomisation distribution of $t(\underset{\sim}{X_n}) \sim N(0, 1)$.

Proof : We show firstly that S^2 converges in probability to σ^2.
In this proof $m_j \equiv m_j(\underset{\sim}{x_n})$ - the population moments

$\mathcal{S}_{\alpha(k)} \equiv \mathcal{S}_{\alpha(k)}(\underset{\sim}{x_n})$ - the population symmetric polynomials

where as in 2·7·2. x_i, i=1,2,.....n are the population measures.
Without loss of generality we may take $\mathcal{S}_1 = 0$.

We have

$$E\{S^2 \mid H_0'\} = \frac{fn}{fn-1}m_2 - \frac{1}{fn-1}m_2 - \frac{\mathcal{S}_{11}}{n(n-1)}$$

$$= m_2 + \frac{1}{n(n-1)}\mathcal{S}_2' \quad \text{since } \mathcal{S}_1'^2 = \mathcal{S}_2' + \mathcal{S}_{11}' \text{ and } \mathcal{S}_1' = 0$$

$$= m_2\left[1 + \frac{1}{n-1}\right]$$

$$= \sigma^2.$$

Also $$E\left\{s^4/H_0'\right\} = \frac{1}{(fn-1)^2} E\left\{\left[\sum_{i=1}^{fn} X_i^2 - \frac{1}{fn}\left(\sum_{i=1}^{fn} X_i\right)^2\right]^2 \Big/ H_0'\right\}.$$

Now $$E\left\{\left(\sum_{i=1}^{fn} X_i^2\right)^2 \Big/ H_0'\right\} = \frac{fn}{n} S_4' + \frac{fn(fn-1)}{n(n-1)} S_{22}'$$
$$= fn\, m_4 + \frac{fn(fn-1)}{n(n-1)}\left[S_2'^2 - S_4'\right]$$
$$= fn\left[1 - \frac{fn-1}{n-1}\right] m_4 + \frac{fn(fn-1)}{n-1}\, n\, m_2^2 .$$

$$\therefore \frac{1}{(fn-1)^2} E\left\{\left(\sum_{i=1}^{fn} X_i^2\right)^2 \Big/ H_0'\right\} = m_2^2\left\{1 + O\left(\frac{1}{n}\right) + \frac{fn^2(1-f)}{(fn-1)^2(n-1)}\,\frac{m_4}{m_2^2}\right\}.$$
$$= m_2^2\{1 + o(1)\} \quad \text{since } \frac{m_4}{m_2^2} = o(n)$$
$$= m_2^2 + o(1) \quad \text{since } m_2 = O(1)$$
$$= \sigma^4 + o(1).$$

Similarly the remaining terms in $E\{S^4/H_0'\}$ can be shown to be $O(1)$.

Hence $$var(s^2) = E\{s^4/H_0'\} - \left[E(s^2/H_0')\right]^2$$
$$= \sigma^4 + o(1) - \left(\sigma^4 + o(1)\right)$$
$$= o(1)$$

It follows by applying Tchebycheff's Inequality that S^2 converges in probability to σ^2.

Then, since the randomisation distribution $R_m \sim N(0, 1)$ it follows by a theorem of Cramer [2] that the randomisation distribution of $t(\underset{\sim}{X_n}) \sim N(0, 1)$.

We thus arrive at the analogue of the t-test in the case of finite populations, and this test holds for nearly all large populations.

Of course this result can also be used to establish confidence limits for the mean of a finite population.

If m, S^2 be respectively the mean and variance of a sample of n individuals drawn at random without replacement from a finite population of N individuals where $\frac{n}{N} = f$ and N is large, then provided the population is not a very unusual one, confidence limits (5%) for the population mean are given by $m \pm 1{\cdot}96 \frac{s}{\sqrt{n}} \sqrt{(1-f)}$

These are the main pmactical implications of Theorem 2·6 - indeed of chapter II.

CHAPTER III

3·1 : It has been stated that the results of Chapter II are mainly of theoretical interest. Many more results 'justifying' the normal assumption in commonly occurring tests of significance can be derived from consideration of the asymptotic joint randomisation distribution of more than one statistic belonging to the class $\mathcal{R}$.

Of course it does not follow that, if $R_1 \in \mathcal{R}$ and $R_2 \in \mathcal{R}$ are separately asymptotically normally distributed under H_0' , they are jointly asymptotically normal under H_0'. From the geometrical standpoint we are apparently asking more for asymptotic joint normality.

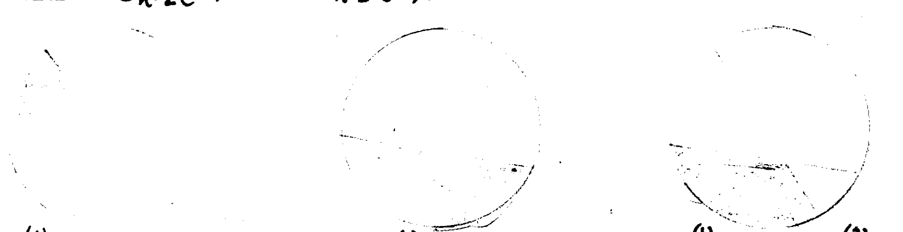
In the notation of section 1, under normal theory the joint distribution of $R_1^{(n)}$, $R_2^{(n)}$ where $(R_1^{(n)})$, $(R_2^{(n)})$ define respectively the statistics R_1 , R_2 , is given by

$$P\left\{ R_1^{(n)}(\xi_n) \leq a,\ R_2^{(n)}(\xi_n) \leq b, \middle| H_0 \right\} = \frac{L_{n-2}\left[C^{(1)}_{n-2}(a) \cap C^{(2)}_{n-2}(b) \right]}{L_{n-2}(\mathcal{C}_{n-2})}.$$

where $C^{(1)}_{n-2}(a)$ is the subset of C_{n-2} for which $R_1^{(n)} \leq a$

$C^{(2)}_{n-2}(b)$ " " " " " " " $R_2^{(n)} \leq b$.

Hence as in section 1, for asymptotic equivalence of the joint distributions of R_1 and R_2 under normal theory and under randomisation theory we require 'uniform distribution of the set $\mathcal{X}_n$ relative to the class of subsets of the form $C^{(1)}_{n-2}(a) \cap C^{(2)}_{n-2}(b)$'. This is apparently asking more than 'uniform distribution of the set $\mathcal{X}_n$ relative to the class of subsets of the form $C^{(1)}_{n-2}(a)$ or $C^{(2)}_{n-2}(b)$.'



$C^{(1)}_{n-2}(a)$ $C^{(2)}_{n-2}(b)$ $C^{(1)}_{n-2}(a) \cap C^{(2)}_{n-2}(b)$.

From the geometrical standpoint, however, it is very difficult to appreciate just how much more we ask for asymptotic equivalence of the joint distribution of R and R under normal and randomisation theories - a three dimensional picture does not give much help in appreciating this situation - though it is quite clear that it is necessary to consider separately the question of the asymptotic joint distribution of more than one $R \in \mathcal{R}$.

Sufficient conditions for asymptotic normality of the joint randomisation distribution of a finite number of statistics belonging to the class $\mathcal{R}$ can be obtained by a straightforward generalisation of the methods of Chapter II. Wald and Wolfowitz discuss this question by a different method, their argument being much more subtle than the present straightforward generalisation.

Like Wald and Wolfowitz we discuss the case of the joint randomisation distribution of only two R's, $R_1, R_2 \in \mathcal{R}$. Clearly a similar argument introducing only additional algebraic complexity can be applied to the case of more than two R's. No new principle would be involved.

3·2 : Joint -partitions : Let u, v be fixed +ve integers with $u<n$ $v<n$ where n is some given integer.

Let q, r be +ve integers and p a +ve integer or zero such that

$p \leq \min(q, r)$

$q \leq u$

$r \leq v$

and let $s = q + r - p$

Let $\{ \alpha_1, \beta_1 ; \alpha_2, \beta_2 ; \cdots\cdots ; \alpha_s, \beta_s \} \equiv (\alpha, \beta)_s$

be a set of S pairs of numbers (+ve integers or zero) (α_i, β_i) where

$$0 < \alpha_1 \leq \alpha_2 \leq \ldots\ldots \leq \alpha_q$$

and $(\alpha_1, \alpha_2, \ldots\ldots\ldots \alpha_q)$ is a partition of u.

and $\alpha_i = 0 \quad i = q+1, \ldots\ldots S.$

Also $\beta_i \neq 0 \quad i = 1, \ldots\ldots\ldots p, q+1, \ldots\ldots S$

$= 0 \quad i = p+1, \ldots\ldots q.$

and $(\beta_1, \beta_2, \ldots\ldots\ldots \beta_p, \beta_{q+1}, \beta_{q+2}, \ldots\ldots \beta_s)$ is a partition of v

with $\beta_{q+1} \leq \beta_{q+2} \leq \ldots\ldots\ldots \leq \beta_s.$

If some of the α's from $\alpha_1, \ldots\ldots, \alpha_p$ are equal then the β's corresponding to equal α's are to be placed in ascending order of magnitude.

$(\alpha, \beta)s$ will be called a joint partition of u and v, of order S. Two joint partitions $(\alpha, \beta)s$ and $(\alpha', \beta')s'$ are the same if and only if $S = S'$ and $\alpha_i = \alpha'_i$, $\beta_i = \beta'_i$ $i = 1, 2, \ldots, s.$

3·3 : Let ξ_i, and η_i $i = 1, 2, \ldots, n$ be two sets of n variables.

Let $(i_1, i_2, \ldots\ldots, i_s)$ be an ordered set of S distinct integers from 1, 2, ..., n.

Let $S_{(\alpha,\beta)_s}[\underset{\sim}{\xi_n}, \underset{\sim}{\eta_n}]$ denote the expression

$$\sum_{I} \left(\xi_{i_1}^{\alpha_1} \eta_{i_1}^{\beta_1}\right)\left(\xi_{i_2}^{\alpha_2} \eta_{i_2}^{\beta_2}\right) \ldots\ldots \left(\xi_{i_s}^{\alpha_s} \eta_{i_s}^{\beta_s}\right),$$

where $(\alpha, \beta)_s$ is a joint partition of u and v and summation extends over all sets $(i_1, i_2, \ldots\ldots i_s)$.

Let

$$m_{\gamma_j \delta_j}(\underset{\sim}{\xi_n}, \underset{\sim}{\eta_n}) = \frac{1}{n} \sum_{i=1}^{n} \xi_i^{\gamma_j} \eta_i^{\delta_j}$$

Then

3.3.1 Lemma : $S_{(\alpha,\beta)_s}[\underset{\sim}{\xi_n}, \underset{\sim}{\eta_n}]$ can be expressed as a sum of terms of the form $\quad C_{(\gamma,\delta)_{s'}}\, n^{s'} \prod_{j=1}^{s'} m_{\gamma_j\delta_j}(\underset{\sim}{\xi}_n, \underset{\sim}{\eta}_n)$, where for each such term

(i) $S' \leq S$ and $S' = S$ only when $(\gamma, \delta)_{s'} = (\alpha, \beta)_s$.

(ii) each γ_j is the sum of one or more α's and each δ_j the sum of the corresponding β's.

(iii) $C_{(\gamma,\delta)_{s'}}$ does not depend on n and the number of terms is independent of n.

This is a generalisation of Lemma 2·3·2.

Proof: Assume the result true for every pair partition of order S - 1, of any pair of integers u', v', with $u' \leq u$ and $v' \leq v$. Now

$$\left\{ \sum_{i_1, i_2, \ldots i_{s-1}} \left(\xi_{i_1}^{\alpha_1} \eta_{i_1}^{\beta_1}\right)\left(\xi_{i_2}^{\alpha_2} \eta_{i_2}^{\beta_2}\right) \cdots \left(\xi_{i_{s-1}}^{\alpha_{s-1}} \eta_{i_{s-1}}^{\beta_{s-1}}\right)\right\}\left\{\sum_{i=1}^{n} \xi_i^{\alpha_s} \eta_i^{\beta_s}\right\}$$

$$\equiv \sum_{i_1, i_2, \ldots, i_s} \left(\xi_{i_1}^{\alpha_1} \eta_{i_1}^{\beta_1}\right) \cdots\cdots \left(\xi_{i_{s-1}}^{\alpha_{s-1}} \eta_{i_{s-1}}^{\beta_{s-1}}\right)\left(\xi_{i_s}^{\alpha_s} \eta_{i_s}^{\beta_s}\right).$$

$$+ \sum_{j=1}^{s-1} \left\{ \sum_{i_1, i_2, \ldots, i_{s-1}} \left(\xi_{i_1}^{\alpha_1} \eta_{i_1}^{\beta_1}\right) \cdots \left(\xi_{i_{j-1}}^{\alpha_{j-1}} \eta_{i_{j-1}}^{\beta_{j-1}}\right)\left(\xi_{i_j}^{\alpha_j+\alpha_s} \eta_{i_j}^{\beta_j+\beta_s}\right)\left(\xi_{i_{j+1}}^{\alpha_{j+1}} \eta_{i_{j+1}}^{\beta_{j+1}}\right) \cdots\cdots \left(\xi_{i_{s-1}}^{\alpha_{s-1}} \eta_{i_{s-1}}^{\beta_{s-1}}\right)\right\},$$

by actual multiplication of the L.H.S.

Now let $\quad u' = u - \alpha_s$

$\qquad\qquad v' = v - \beta_s$.

Then $\{\alpha_1, \beta_1 ; \alpha_2, \beta_2 ; \cdots\cdots ; \alpha_{s-1}, \beta_{s-1}\} = (\alpha, \beta)_{s-1}$, say, is a joint partition of order S - 1 of u' and v' .

Also $\{\alpha_1, \beta_1 ; \alpha_2, \beta_2 ; \cdots\cdots ; \alpha_{j-1}, \beta_{j-1} ; \alpha_j + \alpha_s ; \beta_j + \beta_s ; \alpha_{j+1}, \beta_{j+1} ; \cdots\cdots ; \alpha_{s-1}, \beta_{s-1}\}$ is a joint partition of u and v of order S - 1 and of the form $(\gamma, \delta)_{s-1}$

Let this be written as $(\gamma, \delta)_{s-1}^{j}$

Then the above identity may be written thus - omitting the ξs and ηs for typographical brevity

$$S_{(\alpha,\beta)_{s-1}}\, n . m_{\alpha_s\beta_s} \equiv S_{(\alpha,\beta)_s} + \sum_{j=1}^{s-1} S_{(\gamma,\delta)_{s-1}^{j}}.$$

i.e. $$S_{(\alpha,\beta)_s} \equiv n\, m_{\alpha_s \beta_s}\, S_{(\alpha,\beta)_{s-1}} - \sum_{j=1}^{s-1} S^{j}_{(\gamma,\delta)_{s-1}} .$$

On the basis of the initial assumptions each S' on R.S. can be expressed in the required form and since the number of S_s in R.S. is independent of **n**, then $S_{(\alpha,\beta)_s}$ can be expressed in required form.

The result is true trivially for every joint partition of order 1 of any pair of integers u', v'.

It is therefore true for partitions of higher order of any pair of integers.

Lemma 2·3·2 follows as a special case by putting $v = 0$.

3·4 : Orthogonal Members of the Class $\mathcal{R}$.

Let R_1, R_2 be two statistics of the class $\mathcal{R}$ [section 2] defined respectively by

$$\underset{\sim}{y}^{(1)}_{nn} = \left(y^{(1)}_{1n}, y^{(1)}_{2n}, \cdots , y^{(1)}_{nn} \right)$$

and

$$\underset{\sim}{y}^{(2)}_{nn} = \left(y^{(2)}_{1n}, y^{(2)}_{2n}, \cdots , y^{(2)}_{nn} \right).$$

R_1, R_2 will be said to be orthogonal statistics if for each **n**,

$$\sum_{i=1}^{n} \left(y^{(1)}_{in} - \bar{y}^{(1)}_{nn} \right) \left(y^{(2)}_{in} - \bar{y}^{(2)}_{nn} \right) = 0$$

In what follows, without loss of generality, we will take

$$\bar{y}^{(1)}_{nn} = \bar{y}^{(2)}_{nn} = 0 \quad , \quad \sum_{i=1}^{n} \left(y^{(1)}_{in} \right)^2 = \sum_{i=1}^{n} \left(y^{(2)}_{in} \right)^2 = n$$

Also it is convenient to write $\underset{\sim}{y}^{(i)}_{nn}$ as $\underset{\sim}{y}_{in}$ where $\underset{\sim}{y}_{in} = (y_{i1}, y_{i2}, \cdots ; y_{in})$ though this does not imply that for $n_1 < n_2$ $\underset{\sim}{y}_{in_1}$ is a subset of $\underset{\sim}{y}_{in_2}$.

3·5 : Theorem: Let R_1, R_2 be orthogonal statistics belonging to the class $\mathcal{R}$. Then the joint randomisation distribution of $R^{(n)}_1(\underset{\sim}{X}_n)$ and $R^{(n)}_2(\underset{\sim}{X}_n)$ is asymptotically normal if, when,

$$m_{j_1 j_2}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n}) \leq O\left[n^{a(j-2)} \right] \text{ where } j_1 + j_2 = j > 2 \text{ and } 0 \leq a < \tfrac{1}{2},$$

then

$$\frac{m_j(x_n)}{[m_2(x_n)]^{j/2}} = o\left[n^{b(j-2)} \right] \text{ where } b = \tfrac{1}{2} - a, \quad j = 3, 4, \cdots .$$

<u>Proof</u> : Let M_{uv}, where u and v are fixed +ve integers, denote the (u, v)th moment of the joint randomisation distribution of $R_{(1)}^{(n)}(\underset{\sim}{X}_n)$ and $R_2^{(n)}(\underset{\sim}{X}_n)$.

Without loss of generality we take $m_2(\underset{\sim}{x}_n) = 1$ and $\bar{x}_n = 0$. Then

$$M_{uv} = E\left\{\left[R_1^{(n)}(\underset{\sim}{X}_n)\right]^u \left[R_2^{(n)}(\underset{\sim}{X}_n)\right]^v \,\middle|\, H_0'\right\}$$

$$= \frac{1}{n^{\frac{u+v}{2}}} E\left\{(y_{11}X_1 + y_{12}X_2 \cdots + y_{1n}X_n)^u (y_{21}X_1 + y_{22}X_2 + \cdots + y_{2n}X_n)^v \,\middle|\, H_0'\right\}$$

By multiplying out each bracket and taking expectations term by term it is clear that M_{uv} consists of a linear sum of terms of the form

$$\frac{1}{n^{\frac{u+v}{2}}} C_{(\alpha,\beta)_s} S_{(\alpha,\beta)_s}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n}) \frac{1}{n^{[s]}} S_{\epsilon(s)}(\underset{\sim}{x}_n) \qquad \text{(A)}$$

where $\epsilon_i = \alpha_i + \beta_i$.

$C(\alpha, \beta)_s$ is a constant independent of <u>n</u>.

The number of terms in the sum is also independent of <u>n</u>, depending solely on the number of joint partitions of u and v. Now applying Lemma 3·3·1 to $S_{(\alpha,\beta)_s}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n})$ and lemma 2·3·2 to $S_{\epsilon(s)}(\underset{\sim}{x}_n)$ it is clear that the term (A) contributes to M_{uv} terms (whose number is independent of <u>n</u>) of the form

$$\frac{1}{n^{\frac{u+v}{2}}} C_{(\alpha,\beta)_s}^{(\gamma,\delta)_{s_1}} \frac{1}{n^{[s]}} n^{s_1} \prod_{j=1}^{s_1} m_{\gamma_j\delta_j}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n})\, C_{\epsilon(s)}^{K(s_2)} n^{s_2} \prod_{j=1}^{s_2} m_{K_j}(\underset{\sim}{x}_n). \qquad \text{(B)}$$

where

(i) $(\gamma_1, \gamma_2, \ldots\ldots, \gamma_{s_1})$ is a partition of u with each γ the sum of one or more α_s [In this partition we allow of zero γs]

(ii) $(\delta_1, \delta_2, \ldots\ldots\ldots, \delta_{s_1})$ is a partition of v with each δ the sum of the corresponding β_s [Again we allow of zero δs] .

(iii) $(K_1, K_2, \ldots, K_{S_2})$ is a partition of $(u+v)$ with each K the sum of one or more ϵs.

(iv) the C's are constants which do not depend on n.

We consider the order of the term (B).

If any $K_j = 1$, or any pair (γ_j, δ_j) is (1,0), (0, 1) or (1, 1) the term B is 0.

If any $K_j > 2$, then as in theorem 2.5 the term (B) is either $o(1)$ or 0.

Again as in theorem 2.5, if every $K_j = 2$ the term (B) is $O(1)$ unless $S_1 = S_2 = S$.

If $S_1 = S_2 = S$, then $K_j = \gamma_j + \delta_j$, and $\gamma_j = \alpha_j$, $\delta_j = \beta_j$

Collecting these results the term B is either 0 or $o(1)$ unless possibly we have u and v both even and

$$(\alpha, \beta)_s = \Big[\underbrace{2,0\ ;\ 2,0\ ;\ \cdots\ ;2,0}_{\frac{u}{2}\text{ terms}}\ ;\ \underbrace{0,2\ ;\ 0,2\ ;\ \cdots\ ;\ 0,2}_{\frac{v}{2}\text{ terms}}\Big]$$

and the term B is that term in which $S_1 = S_2 = \frac{u+v}{2}$ derived from the term of the form A corresponding to this particular joint partition of u and v.

Evaluating the constants concerned shows that

$$M_{uv} = \frac{u!}{(\frac{u}{2})!\,2^{u/2}}\ \frac{v!}{(\frac{v}{2})!\,2^{v/2}} + o(1) \quad \text{if } u \ \& \ v \text{ are both even,}$$
$$= o(1) \quad \text{if } u \ \& \ v \text{ are not both even.}$$

Since the bivariate normal distribution is completely determined by its moments if follows that the asymptotic joint randomisation distribution of R_1, R_2 is subject, to the conditions of the theorem, normal and that R_1, R_2 are asymptotically independently distributed. The extension to the case of more than two R's is obvious - the proof will clearly go through in exactly the same manner. We thus get the following theorem.

<u>3·6 : Theorem</u> : Let R_1 , R_2 ,....., R_m be m mutually orthogonal statistics belonging to the class $\mathcal{R}$; if

$$m_{j_1 j_2 \cdots j_m}\left(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n}, \cdots, \underset{\sim}{y}_{mn}\right) \leq O\left[n^{a(j-2)}\right]$$

where $\qquad j_1 + j_2 + \cdots + j_m = j, \quad$ for $j = 3, 4, \cdots, \qquad 0 \leq a < \frac{1}{2}$

and $\qquad \dfrac{m_j(\underset{\sim}{x}_n)}{[m_2(\underset{\sim}{x}_n)]^{j/2}} = o\left[n^{b(j-2)}\right] \quad$ where $b = \frac{1}{2} - a$,

the joint randomisation distribution of R_1 , R_2 ,....R_m is asymptotically multinormal and R_1 , R_2 ...... R_m are asymptotically mutually independent.

Theorem 3·6 does not of course establish asymptotic equivalence of the joint distributions of R_1 , R_2, ... R_m under normal theory and under randomisation theory. Its interest lies mainly in its subsequent use to establish asymptotic equivalence of the distributions of certain other statistics in the two cases.

CHAPTER IV : Quadratic Forms :

Many statistics in common use are based on quadratic forms in random variables - especially those statistics used in analysis of variance. It therefore seems natural to proceed from a discussion of linear forms to a discussion of quadratic forms. It is in this connection that Theorem 3·6 becomes important.

4·1 : Quadratic Forms under Normal Theory : Let $\underset{\sim}{\Xi}_n$ denote the column vector $\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$ in the random variables $\xi_1, \xi_2 \ldots\ldots \xi_n$; and

$\underset{\sim}{\Xi}_n'$ denote the transpose of $\underset{\sim}{\Xi}_n$

Let $\underset{\sim}{A}n$ be a symmetric $n \times n$ matrix ($a_{ij}^{(n)}$) such that $\sum_{j=1}^{n} a_{ij}^{(n)} = 0$

i = 1, 2,......n.

With this condition $\underset{\sim}{A}n$ has a latent vector $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ corresponding to the latent root zero.

Let the remaining latent roots of $\underset{\sim}{A}n$ be $\lambda_1^{(n)}, \lambda_2^{(n)} \ldots\ldots, \lambda_{n-1}^{(n)}$.

$\underset{\sim}{\Xi}_n' \underset{\sim}{A}n \underset{\sim}{\Xi}_n$ represents a quadratic form Qn, say, in the random variables ξ_i .

Let the sequence of matrices (An) define for each n the statistic Q.

Q_n can be reduced by an orthogonal transformation

$$\underset{\sim}{H}_n = \underset{\sim}{P}_n \underset{\sim}{\Xi}_n \quad \text{where} \quad \underset{\sim}{H}_n = \begin{bmatrix} \eta_{1n} \\ \eta_{2n} \\ \vdots \\ \eta_{nn} \end{bmatrix} \quad \text{and} \quad \eta_{nn} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i$$

to the form $Q_n = \lambda_1^{(n)} \eta_{1n}^2 + \lambda_2^{(n)} \eta_{2n}^2 + \ldots\ldots\ldots + \lambda_{n-1}^{(n)} \eta_{n-1,n}^2$

where if $H_0(0, 1)$ is true the η_{in}s are independent N(0, 1) random variables, so that η_{in}^2 is distributed as χ^2 with one degree of freedom (written $\chi^2_{[1]}$).

Furthermore this transformation shows that Qn is distributed independently of $\bar{\xi}_n = \frac{1}{n}(\xi_1 + \xi_2 + \cdots + \xi_n) = \sqrt{n}\, y_{nn}$ when $H_0(0, 1)$ is true. In general the distribution function of a linear combination of independent $\chi^2_{[1]}$ random variables is very complicated. However under the wide conditions of the Central Limit Theorem, when the number of non zero $\lambda^{(n)}_s$ tends to infinity as $n \to \infty$ the distribution of Qn will be asymptotically normal under $H_0(0, 1)$. If the number of non-zero $\lambda^{(n)}_s$ remains finite as $n \to \infty$ the distribution of $\sum_{i=1}^{n-1} \lambda_i^{(n)} y_{in}^2$ will not necessarily tend to normality but may still tend to a limiting form which is completely determined by its moments.

We will suppose that this is so, i.e., that the limiting form of the distribution function of Q exists under $H_0(0, 1)$ and that it is of a form completely determined by its moments.

We will denote $\mathcal{L}_{n\to\infty} P\{Q_n \leq a \mid H_0(0,1)\}$ by $F_Q(a)$.

4.2 : The statistic Q does not, of course, belong to the class $\mathcal{Y}$ with which we are concerned.

Let
$$\mathcal{U}_n(\underset{\sim}{\xi}_n) = \frac{Q_n(\underset{\sim}{\xi}_n)}{m_2(\underset{\sim}{\xi}_n)}.$$

Then the statistic $\mathcal{U}$ defined by the sequence ($\mathcal{U}_n$) does belong to the class $\mathcal{Y}$ since

(i) $\mathcal{U}_n$ is homogeneous of degree 0 in the ξs

(ii) $$(\underset{\sim}{\Xi}_n + \underset{\sim}{a}_n)' \underset{\sim}{A}_n (\underset{\sim}{\Xi}_n + \underset{\sim}{a}_n) \quad \text{where } \underset{\sim}{a}_n = \begin{bmatrix} a \\ a \\ \vdots \\ a \end{bmatrix}$$
$$= \underset{\sim}{\Xi}_n' \underset{\sim}{A}_n \underset{\sim}{\Xi}_n + \underset{\sim}{a}_n' \underset{\sim}{A}_n(\underset{\sim}{\Xi}_n + \underset{\sim}{a}_n) + \underset{\sim}{\Xi}_n' \underset{\sim}{A}_n \underset{\sim}{a}_n$$

Also
$$\underset{\sim}{a}_n' \underset{\sim}{A}_n = [0, 0, \cdots 0] \quad \text{and} \quad \underset{\sim}{A}_n \underset{\sim}{a}_n = (\underset{\sim}{a}_n' \underset{\sim}{A}_n)' = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

because $\sum_{j=1}^{n} a_{ij}^{(n)} = 0.$

(iii) We suppose that the third property of the class $\mathcal{Y}$ is satisfied. Now when $H_0(0, 1)$ is true $m_2(\underset{\sim}{\xi_n})$ converges in probability to 1.

Hence

$$P\{U_n \le a \mid H_0\} = P\{U_n \le a \mid H_0(0,1)\} \quad \text{since } U \in \mathcal{Y}$$

$\sim$ $F_q(a)$ by a theorem of Cramer since $m_2(\underset{\sim}{\xi_n})$ converges in probability to 1. [2]

$$= \mathcal{L}_{n\to\infty} P\{Q_n \le a \mid H_0(0,1)\}.$$

<u>4·3</u> : We consider now $U_n(\underset{\sim}{X_n})$.

Let $\underset{\sim}{X_n}$ denote the column vector $\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$

When dealing with the randomisation distribution of U_n we may, without loss of generality consider $m_2(\underset{\sim}{x_n}) = 1$.

The transformation $\underset{\sim}{\overline{Y}_n} = \underset{\sim}{P_n}\,\underset{\sim}{\overline{X}_n}$ reduces $\underset{\sim}{\overline{X}_n'}\,\underset{\sim}{A_n}\,\underset{\sim}{\overline{X}_n}$ to the form $\sum_{i=1}^{n} \lambda_i^{(n)} Y_{in}^2$.

$$Y_{in} = p_{i1}^{(n)} X_1 + p_{i2}^{(n)} X_2 + \cdots + p_{in}^{(n)} X_n \qquad i = 1, 2, \cdots n-1,$$

where

$$\sum_{j=1}^{n} p_{ij} = 0 \qquad i = 1, 2, \cdots, n-1,$$

$$\sum_{j=1}^{n} p_{ij}^2 = 1$$

The statistic Y_i defined by the sequence (Y_{in}) belongs to the class $\mathcal{R}$. Now suppose all sets of m different Y_is, $Y_{i_1}, Y_{i_2}, \cdots Y_{i_m}$ satisfy the conditions imposed on $R_1, R_2, \ldots\ldots, R_m$ in Theorem 3·6. Then, by Theorem 3·6 if $m_j(\underset{\sim}{x_n}) = o\left[n^{b(j-2)}\right]$ $j = 3, 4, \ldots\ldots n$.

<u>4·3·1</u> : $$E\left\{Y_{i_1}^{2\alpha_1} Y_{i_2}^{2\alpha_2} \cdots Y_{i_m}^{2\alpha_m} \mid H_0'\right\} = E\left\{\xi_{i_1}^{2\alpha_1} \xi_{i_2}^{2\alpha_2} \cdots \xi_{i_m}^{2\alpha_m} \mid H_0(0,1)\right\} + \tau.$$

where $\tau = o(1)$.

and $(\alpha_1, \alpha_2, \ldots\ldots \alpha_m)$ is a partition of a fixed integer $r < n$.

We consider now

$$\left| E\left\{ [U_n(\underset{\sim}{X}_n)]^r / H_0' \right\} - E\left\{ [U_n(\underset{\sim}{\xi}_n)]^r / H_0 \right\} \right|$$

$$= \left| E\left\{ [U_n(\underset{\sim}{X}_n)]^r / H_0' \right\} - E\left\{ [Q_n(\underset{\sim}{\delta}_n)]^r / H_0(0,1) \right\} \right| + \epsilon$$

where $\epsilon = o(1)$ with respect to $\underline{n}$ by 4·2.

$$= \left| E\left\{ \sum_{par} \sum_{i_1, i_2, \dots i_m} Y_{i_1 n}^{2\alpha_1} \cdots Y_{i_m n}^{2\alpha_m} (\lambda_{i_1}^{(n)})^{\alpha_1} \cdots (\lambda_{i_m}^{(n)})^{\alpha_m} / H_0' \right\} - \right.$$

$$\left. - E\left\{ \sum_{par} \sum_{i_1, i_2 \dots i_m} \eta_{i_1 n}^{2\alpha_1} \cdots \eta_{i_m n}^{2\alpha_m} (\lambda_{i_1}^{(n)})^{\alpha_1} \cdots (\lambda_{i_m}^{(n)})^{\alpha_m} / H_0(0,1) \right\} \right| + \epsilon$$

where $\sum_{par}$ denotes summation over partitions $(\alpha_1, \alpha_2, \dots \dots \alpha_m)$ of r,

$\sum_{i_1, i_2, \dots i_m}$ denotes summation over ordered sets $i_1, i_2, \dots \dots i_m$.

$$= \left| \sum_{par} \sum_{i_1 \dots i_m} (\lambda_{i_1}^{(n)})^{\alpha_1} \cdots (\lambda_{i_m}^{(n)})^{\alpha_m} \, \delta\binom{\alpha_1 \cdots \alpha_m}{i_1 \cdots i_m} \right| + \epsilon \quad \text{------} \quad (4.3.2)$$

where $\delta\binom{\alpha_1 \cdots \alpha_m}{i_1 \cdots i_m}$ is a function of the partition and the integers $i_1, \dots \dots \dots i_m$ which is $O(1)$ by 4·3·1.

It follows immediately that if $\sum_{i=1}^{n-1} |\lambda_i^{(n)}| \leq O(1)$ the r^{th} moments of the distribution of U_n under normal theory and randomisation theory are asymptotically equal. Then since we are supposing that the limiting distribution of U_n under normal theory exists and is determined by its moments, it follows that the asymptotic distributions of U for the two cases are equivalent.

The geometrical interpretation of this result, as considered in Cahpter I, still applies but by this stage has outlived its usefulness.

Furthermore the conditions set up in this section while sufficient for asymptotic equivalence are by no means necessary. They are designed more for the case of statistics U whose numeratorQ is defined by a sequence of quadratic forms(Q_n) in which the rank of Q_n remains finite as $n \to \infty$, than for those in which the rank of $Q_n \to \infty$ as $n \to \infty$. This is illustrated by the following two examples.

4·4 : Analysis of Variance - One way Classification : In this we consider the set $\xi_1, \xi_2, \dots, \xi_n$ of random variables to be divided into k classes, the i^{th} class consisting of n_i variables where $\sum_{i=1}^{k} n_i = n$ and each n_i is $O(n)$. It is therefore convenient to write the set $\xi_1, \xi_2, \dots \xi_n$ in the form $\xi_{11}, \xi_{12}, \dots \xi_{1n_1}, \xi_{21}, \xi_{22}, \dots \xi_{2n_2}, \dots, \xi_{k1}, \xi_{k2}, \dots \xi_{kn_k}$.

Let
$$\bar{\xi}_{in} = \frac{1}{n_i}\sum_{j=1}^{n_i} \xi_{ij}$$
and
$$Q_n = \sum_{i=1}^{k} n_i \left(\bar{\xi}_{in} - \bar{\xi}_n\right)^2$$
where $\bar{\xi}_n = \frac{1}{n}\sum_{j=1}^{n} \xi_j = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} \xi_{ij}$, in present notation.

When $H_0(0, 1)$ is true Q_n is distributed as $\chi^2_{[k-1]}$

Under an orthogonal transformation $\underset{\sim}{H}_n = \underset{\sim}{P}_n \underset{\sim}{\Xi}_n$ in which
$$y_{in} = c_{in}\left\{\sum_{j=1}^{i} \bar{\xi}_{jn} - i\,\bar{\xi}_{i+1,n}\right\} \quad i = 1, 2, \dots, k-1,$$
where C_{in} is a normalising constant, and in which
$$y_{nn} = \sqrt{n}\,\bar{\xi}_n$$
-the well-known Helmert transformation -

Q_n reduces to the form $y_{1n}^2 + y_{2n}^2 + \dots + y_{k-1,n}^2$.

Thus the limiting form of the distribution of Q_n under normal theory is $\chi^2_{[k-1]}$, a form completely determined by its moments.

Also $Y_{1n}, Y_{2n}, \dots, Y_{k-1,n}$ satisfy the condition imposed on the R's in Theorem 3·6 with a = θ.

For we have, for $1 \le i \le k-1$
$$\underset{\sim}{p}_{in} \propto \Big(\underbrace{\tfrac{1}{n_1}, \tfrac{1}{n_1}, \dots \tfrac{1}{n_1}}_{n_1 \text{ terms}}, \underbrace{\tfrac{1}{n_2}, \tfrac{1}{n_2}, \dots \tfrac{1}{n_2}}_{n_2 \text{ terms}}, \dots, \underbrace{\tfrac{1}{n_i}, \dots \tfrac{1}{n_i}}_{n_i \text{ terms}}, \underbrace{-\tfrac{i}{n_{i+1}}, \dots -\tfrac{i}{n_{i+1}}}_{n_{i+1} \text{ terms}}, 0, 0, \dots 0\Big)$$

Write
$$c_{in} = \sqrt{\left\{\frac{1}{n}\left(\frac{1}{n_1} + \frac{1}{n_2} \dots + \frac{1}{n_i} + \frac{i^2}{n_{i+1}}\right)\right\}}.$$

Then $\underset{\sim}{p}_{in} \propto \underset{\sim}{y}_{in}$, where

$$\underset{\sim}{y}_{in} = \frac{1}{C_{in}}\left(\frac{1}{n_1}, \cdots \frac{1}{n_1}, \frac{1}{n_2}, \cdots \frac{1}{n_2}, \cdots, \frac{1}{n_i} \cdots \frac{1}{n_i}, -\frac{i}{n_{i+1}}, \cdots -\frac{i}{n_{i+1}}, 0, - 0\right).$$

Also $\quad m_2(\underset{\sim}{y}_{in}) = 1$

If $i_1 < i_2 < \ldots < i_h$ are $h(\geq 1)$ integers from $1, 2, \ldots, k-1$, and $j_1, j_2, \ldots j_h$ are h positive integers and such that $j_1 + j_2 + \ldots + j_h = j$, then

$$m_{j_1 j_2 \cdots j_h}(\underset{\sim}{y}_{i_1 n}, \underset{\sim}{y}_{i_2 n}, \cdots \underset{\sim}{y}_{i_h n}) = \frac{1}{n}\,\frac{1}{C_{i_1 n}^{j_1} \cdots C_{i_h n}^{j_h}}\left[\frac{1}{n_1^{j-1}} + \frac{1}{n_2^{j-1}} \cdots + \frac{1}{n_{i_1}^{j-1}} + \frac{(-i_1)^j}{n_{i_1+1}^{j-1}}\right]$$

$$= O(1)$$

since $C_{in} = O(\frac{1}{n})$ and $n_i = O(n)$ $i = 1, 2, \ldots, k-1$.

So $Y_{1n}, Y_{2n}, \ldots, Y_{k-1,n}$ satisfy the conditions of Theorem 3·6, with a = 0.

It follows that if $\quad \dfrac{m_j(\underset{\sim}{x}_n)}{[m_2(\underset{\sim}{x}_n)]^{j/2}} = o\left(n^{\frac{j-2}{2}}\right) \qquad j = 3, 4, \ldots\ldots$

then the asymptotic randomisation distribution of $U_n = Q_n / m_2(\underset{\sim}{x}_n)$ has the $\chi^2_{[k-1]}$ form.

This forms the basis of the randomisation 'justification' of the normal theory assumption in the case of a one-way classification analysis of variance (see later).

The exact expression for the first three moments of the randomisation distribution of Q_n in the case where $n_1 = n_2 = \ldots = n_k$ is discussed in appendix 1.

<u>4·5</u> : <u>Serial Correlation</u> : While 4·3·2 is O(1) if $\sum_{i=1}^{n-1}|\lambda_i^{(n)}| \leq O(1)$ it is quite apparent that the δs may be such that 4·3·2 is O(1) although this condition does not hold. So this condition may not be necessary for asymptotic equivalence of the normal theory and randomisation distributions of statistids of the form U_n. The fact that indeed it is not necessary is shown by the following.

Let $Q_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\xi_i - \bar{\xi}_n)(\xi_{i+1} - \bar{\xi}_n)$ where $\xi_{n+1} \equiv \xi_1$

Then $U_n = Q_n(\xi_n) / m_2(\xi_n)$ is the standardised serial correlation coefficient of lag 1.

The distribution of U_n under normal theory has been given by Anderson [1] Asymptotically the normal theory distribution of U_n is N(0,1). The randomisation distribution of $U_n(\underset{\sim}{X}_n)$ has been shown by Noether [9], in widening conditions given by Wald and Wolfowitz to be N(0,1) provided

$$m_j(x_n) / [m_2(x_n)]^{j/2} = o\left(n^{\frac{j-2}{4}}\right) \quad j = 3, 4, \ldots\ldots,$$

So subject to this condition the asymptotic distributions of this statistic obtained by the two approaches are equivalent. It will now be shown that all the conditions of 4·3 are satisfied by U_n except the condition that $\sum_{i=1}^{n-1} |\lambda_i^{(n)}| \leq O(1)$.

If $Q_n(\underset{\sim}{\xi}_n) = \underset{\sim}{\xi}'_n \underset{\sim}{A}_n \underset{\sim}{\xi}_n$ and $\underset{\sim}{I}_n$ is the unit $n \times n$ matrix then $\det(\underset{\sim}{A}_n - \lambda \underset{\sim}{I}_n)$ is a circulant and the roots of $\det(\underset{\sim}{A}_n - \lambda \underset{\sim}{I}_n) = 0$ are easily shown to be 0 and $\frac{1}{\sqrt{n}} \cos \frac{2k\pi}{n}$ $k = 1, 2, \ldots, n-1$,

Hence here $\lambda_i^{(n)} = \frac{1}{\sqrt{n}} \cos \frac{2i\pi}{n}$ $\quad i = 1, 2, \ldots\ldots\ldots n - 1$

If $\underline{n}$ is odd, say $n = 2n' + 1$, and $\underset{\sim}{P}_n$ is an orthogonal matrix in which

$$\underset{\sim}{p}_{in} \propto \left[\sqrt{2}\cos\left(\frac{2i\pi}{n}\right), \sqrt{2}\cos\left(2 \cdot \frac{2i\pi}{n}\right), \ldots\ldots\ldots, \sqrt{2}\cos\left(n \cdot \frac{2i\pi}{n}\right)\right] = \underset{\sim}{y}_{in}$$

for $i = 0, 1, 2, \ldots, n'$.

and

$$\underset{\sim}{p}_{in} \propto \left[\sqrt{2}\sin\left(\frac{2(i-n')\pi}{n}\right), \sqrt{2}\sin\left(2 \cdot \frac{2(i-n')}{n}\pi\right), \ldots, \sqrt{2}\sin\left(n \cdot \frac{2(i-n')\pi}{n}\right)\right] = \underset{\sim}{y}_{in}$$

for $i = n'+1, \ldots, n$.

then the substitution $\underset{\sim}{I}_n = \underset{\sim}{P}_n \underset{\sim}{\xi}_n$ reduces $Q_n(\underset{\sim}{\xi}_n)$ to the form $\frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \cos \frac{2i\pi}{n} \cdot y_{in}^2$.

Now for $1 \leq i \leq n'$,

$$m_2(\underset{\sim}{y}_{in}) = \frac{1}{n} \cdot 2 \sum_{k=1}^{n} \cos^2\left(k \cdot \frac{2i\pi}{n}\right)$$

$$= \frac{1}{n} \sum_{k=1}^{n} \left[1 + \cos\left(k \cdot \frac{4i\pi}{n}\right)\right] = 1.$$

while for $n' < i \leq n - 1$

$$m_2(\underset{\sim}{y}_{in}) = 1, \text{ Similarly.}$$

Hence if $i_1, i_2, \dots i_h$ are h distinct integers from $1, 2, \dots, n-1$ and $j_1, j_2, \dots j_h$, are h positive integers such that $j_1 + j_2 + \dots j_h = j$ then $m_{j_1 j_2 \dots j_h}(y_{i_1 n}, y_{i_2 n}, \dots y_{i_h n}) = \frac{2^{h/2}}{n}$ times a sum of terms of multiples of cosines and sines, the modulus of each term certainly being less than or equal to 1.

It follows that $m_{j_1 j_2 \dots j_h}(y_{i_1 n}, y_{i_2 n}, \dots y_{i_h n}) \leq O(1)$.
Thus all sets of $\underline{h}$ different Y_{in}s corresponding to non-zero $\lambda_i^{(n)}$s satisfy the conditions of Theorem 3·6 with a = 0.
Also the limiting distribution function of u_n being N(0, 1) is completely determined by its moments.

But
$$\sum_{i=1}^{n-1} |\lambda_i^{(n)}| = \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \left|\cos\frac{2i\pi}{n}\right|$$
$$> \frac{1}{\sqrt{n}} \sum_{i=1}^{n''} \cos\frac{2i\pi}{n}$$

where n'' is the greatest integer less than $n/4$.

$$= \frac{1}{\sqrt{n}} \frac{\cos\left(\frac{\pi}{n} + \frac{n''}{n}\pi\right) \sin\frac{n''\pi}{n}}{\sin \pi/n}$$
$$= \frac{\sqrt{n}}{\pi} \frac{\cos\left[(1+n'')\frac{\pi}{n}\right] \sin\frac{n''\pi}{n}}{\sin\frac{\pi}{n} / \frac{\pi}{n}} = O(n^{1/2})$$

This emphasises the weakness in the conditions set out in 4·3. In fact these conditions are more or less useless for studying quadratic forms whose rank tends to infinity with $\underline{n}$ because in almost all these cases, this final condition is not satisfied.

I have tried to weaken the condition $\sum_{i=1}^{n-1} \lambda_i^{(n)} \leq O(1)$ by studying the remainders δ in 4·3·2. But an approach along these lines would seem to lead to conditions so complicated as to be useless.

I have also tried to establish conditions which would not necessitate reduction of the quadratic form to canonical form but this approach (naturally enough) was even less promising than the former.

CHAPTER V :

In this chapter we deal with some practical implications of results obtained in previous chapters.

5·1 : Analysis of Variance.

5·1·1 : One-way Classification : Here we have k samples, the i^{th} sample containing n_i measures x_{i1}, x_{i2} ......x_{in_i}, i = 1, 2, ... k.

$$\sum_{i=1}^{k} n_i = n \quad , \quad x_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij} \quad , \quad x_{..} = \frac{1}{n}\sum_i\sum_j x_{ij} \; .$$

Under normal theory we assume that x_{ij} can be expressed in the form

$x_{ij} = \alpha_i + \epsilon_{ij}$ where α_1, α_2, ... α_k are constants and ϵ_{ij} is a value of a normal random variable whose distribution does not depend on i or j.

Let H_{0A} denote the hypothesis that var $\alpha = 0$.

We test H_{0A} by means of the statistic

$$F = \frac{\frac{1}{k-1}\sum_{i=1}^{k} n_i (x_{i.} - x_{..})^2}{\frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{i.})^2} \; .$$

large values being significant.

When H_{0A} (0, 1) is true, as $n \to \infty$, the denominator of F converges in probability to unity. Hence the distribution of F $\sim$ the distribution of the numerator and the numerator is distributed as $\frac{1}{k-1}\chi^2_{[k-1]}$, under $H_{0A}(0, 1)$ i.e., asymptotically, under normal theory, F is distributed as $\frac{1}{k-1}\chi^2_{[k-1]}$.

Under randomisation theory we assume that x_{ij} can be expressed in the form $x_{ij} = \alpha_i + \epsilon'_{ij}$ where α_1, α_2 ...α_k are constants and ϵ'_{ij} is a value of a random variable (not necessarily normal), whose distribution does not depend on i or j.

H'_{0A} denotes the hypothesis that var $\alpha = 0$.

distribution of $\dfrac{\frac{1}{k-1}\sum_{i=1}^{k} n_i (x_{i.} - x_{..})^2}{\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{..})^2}$ is $\frac{1}{k-1}\chi^2_{[k-1]}$

asymptotically, subject to a very mild restriction.

Now $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{..})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{i.})^2 + \sum_{i=1}^{k} n_i (x_{i.} - x_{..})^2$

$$\therefore \quad 1 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{i.})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{..})^2} + \frac{\sum_{i=1}^{k} n_i (x_{i.} - x_{..})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{..})^2}$$

The second term on the R.S. converges in probability to 0 as $n \to \infty$

$$\sum_{i=1}\sum_{j=1}(x_{ij} - x_{..})^2 \qquad \sum_{i=1}\sum_{j=1}(x_{ij} - x_{..})^2$$

The second term on the R.S. converges in probability to 0 as $n \to \infty$

Therefore the first .... to 1 as $n \to \infty$

Therefore $\dfrac{\frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{i.})^2}{\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - x_{..})^2}$ converges in probability to 1 as $n \to \infty$

Hence the randomisation distribution of F $\sim \frac{1}{k-1}\chi^2_{[k-1]}$

i.e., a randomisation test using the same statistic as the usual normal theory test is asymptotically equivalent to the latter and so for large values of n the normal assumption is in almost all cases unimportant.

Again the condition $\dfrac{m_j(\underset{\sim}{x_n})}{[m_2(\underset{\sim}{x_n})]^{j/2}} = o\left[n^{\frac{j-2}{2}}\right]$ is satisfied in particular, for the set $\underset{\sim}{x_n} = (1, 2, 3, \cdots, n-1, n)$.

So if the original set of measures are ranks, the analysis of variance technique can be applied, without qualification.

5·1·2 : Two-way Classification : The set up in which we compare the randomisation and normal theory approaches in this paragraph is the following.

pq measures are arranged in p A-classes and q B-classes, the measure in the i^{th} A-class and j^{th} B-class being x_{ij}.

$$x_{i.} = \frac{1}{q}\sum_{j=1}^{q} x_{ij}, \qquad x_{.j} = \frac{1}{p}\sum_{i=1}^{p} x_{ij}, \qquad x_{..} = \frac{1}{pq}\sum_{i}\sum_{j} x_{ij}$$

$$S_A^2 = q\sum_{i=1}^{p}(x_{i.} - x_{..})^2, \qquad S_B^2 = p\sum_{j=1}^{q}(x_{.j} - x_{..})^2$$

$$S_R^2 = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(x_{ij} - x_{i.} - x_{.j} + x_{..}\right)^2$$

Under normal theory we assume that x_{ij} can be expressed in the form $x_{ij} = \alpha_i + \beta_j + \epsilon_{ij}$ where α_i $i = 1, 2, \ldots p$, β_j $j = 1, \ldots q$ are constants and ϵ_{ij} is a value of a normal random variable ϵ whose distribution does not depend on i or j.

H_{oA} denotes the hypothesis that var $\alpha = 0$

H_{oB} .... var $\beta = 0$

We test H_{oA} by means of the statistic $F_A = \frac{1}{p-1} S_A^2 \Big/ \frac{1}{(p-1)(q-1)} S_R^2$.

We test H_{oB} .... $F_B = \frac{1}{q-1} S_B^2 \Big/ \frac{1}{(p-1)(q-1)} S_R^2$.

These tests are independent of each other.

Under randomisation theory we make a similar assumption omitting only normality of the distribution of ϵ .

In this case H'_{oA} and H'_{oB} denote the corresponding hypotheses.

As in 5·1·1 we can use the statistics F_A and F_B in the randomisation approach and if p, q are large, the randomisation distribution of F_A and F_B are equivalent to their normal distributions. For with the initial assumption, whether or not H'_{oB} is true, if H'_{oA} is true then all permutations of the measures y_{ij} where $y_{ij} = x_{ij} - x_{.j}$, are equally likely and this leads as in the one-way classification case to the same asymptotic (with regard to p) distribution of F_A as under normal theory subject to the mild condition on the y's that has been applied previously. Similarly for F_B .

It is clear that the whole analysis of variance technique can be built up in this way, and that provided we are dealing with large numbers, in almost all cases the normal assumption can be dropped without altering to any extent the result of the test.

5·2 : The Problem of m-rankings and a Generalisation of it: The problem of m-rankings is the following.

m individuals each place n objects in order according to their valuation of them. The question is whether there is any association between the ways in which the different individuals order the objects. We discuss the following more general problem which arises in connection with analysis of variance.

In the notation of 5·1·2 let p = n and q = m.

In 5·1·2 the assumption that x_{ij} can be expressed in the form

$$x_{ij} = \alpha_i + \beta_j + \epsilon_{ij} \quad ....(1)$$

is not always justifiable. There may be no reason to suppose that the only effect of the B-classification say, may be a 'change of mean' as this assumption implies. In this case the weaker assumption that x_{ij} can be expressed in the form

$$x_{ij} = \alpha_i + \epsilon_i^{(j)} \quad .... (2)$$

where $\alpha_1, \alpha_2, ..., \alpha_n$ are constants and $\epsilon_i^{(j)}$ is a value of the random variable $\epsilon^{(j)}$ whose distribution depends on j (not necessarily only with regard to the mean), may be more appropriate. With the assumption (2) arises the question of testing the hypothesis H''_{oA} that var $\alpha = 0$. If H''_{oA} is true, then on the assumption (2), for fixed j all permutations of the values $x_{1j}, x_{2j}, ..., x_{nj}$ are equally likely. This is so for $j = 1, 2, ..., m$.

Pitman and Welsh have proposed a test of H'_{oA} on this basis though neither has discussed in detail the asymptotic (for $n \to \infty$) properties of the test. [10], [13].

While this problem arises from weakening assumptions in the analysis of variance approach, as it tacitly recognised by Pitman, it is most conveniently treated in a similar fashion to the problem of m-rankings. For if H''_{0A} is true there is no tendency for measures of relatively similar magnitude to be grouped together in the same A-class, while if var $\alpha \neq 0$, there is such a tendency.

Accordingly we let $\quad x'_{ij} = \dfrac{x_{ij} - x_{\cdot j}}{\sqrt{\left\{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - x_{\cdot j})^2\right\}}}$.

Let $\quad S_A'^2 = \sum_{i=1}^{n} x_{i.}'^2$

Then as in analysis of variance $S_A'^2$ is used as the basis of a statistic for testing H'_{0A}.

We standardise $S_A'^2$ by setting

$$W = m\sqrt{\left\{\frac{m}{2n(m-1)}\right\}}\left(S_A'^2 - \frac{n}{m}\right)$$

and show that the distribution of W over equally likely permutations within each B-class is asymptotically N(0, 1).

We have
$$S_A'^2 = \frac{1}{m^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{m} x'_{ij}\right)^2$$
$$= \frac{1}{m^2}\left(\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}'^2 + \sum_{\substack{j,k=1 \\ j\neq k}}^{m} n\, r_{jk}\right)$$

where $\quad r_{jk} = \frac{1}{n}\sum_{i=1}^{n} x'_{ij}\, x'_{ik} =$ product moment correlation coefficient between the measures in the j^{th} and k^{th} B classes.

Now $\quad \frac{1}{m^2}\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}'^2 = \frac{n}{m}$

Hence $\quad W = \dfrac{\sqrt{2}}{\sqrt{m(m-1)}}\sum_{\substack{j,k=1 \\ j\neq k}}^{m}\sqrt{n}\, r_{jk}$

Now the distribution of $\sqrt{n}\, r_{jk}$ over equally likely permutations within each of the j^{th} and k^{th} B classes is exactly the same as its distribution over equally likely permutations of the j^{th} class only for a given permutation of the k^{th} class.

Hence by section 2 the distribution of each $\sqrt{n}\,r_{jk} \sim N(0, 1)$ provided the u^{th} moment of the set $(x_{1j}, x_{2j}, \ldots\ldots, x_{nj})$ is $O\left[n^{\frac{u-2}{4}}\right]$, $u = 3, 4, \ldots$ for each j........... 5·2·1.

Furthermore it can be shown that subject to this restriction the set of ${}^mC_2\, r_{jk}$s are asymptotically independently distributed over equally likely permutations within each B-class........5·2·2.

Thus $\sqrt{\left\{\frac{m(m-1)}{2}\right\}}\, W$ is asymptotically distributed over equally likely permutations within each B-class as the sum of $\frac{m(m-1)}{2}$ independent $N(0, 1)$ random variables.

Hence subject to these restrictions W is asymptotically distributed in the $N(0, 1)$ form.

The statement 5·2·2 requires some justification. For suppose we consider $m = 3$.

Then we have the three correlation coefficients r_{12}, r_{13}, r_{23}.

It is immediately obvious that these are pair-wise independently distributed e.g. r_{12} and r_{13} are independent, since given any value of r_{12} the distribution of r_{13} is obtained from all permutations of the third B-class relative to the first. But it is by no means obvious that the three variables are independent, asymptotically. In fact this would seem to be untrue since given $r_{12} = 1$ and $r_{13} = 1$, r_{23} is of necessity 1.

However, the fact that they are asymptotically independently distribution can be proved if the condition 5·2·1 is satisfied. The method of proof is similar to that used in Theorem 3·5 and only a sketch is given here as the details are almost identical with these of Theorem 3·5.

We consider $E\left[(\sqrt{n}\,r_{12})^{u}(\sqrt{n}\,r_{23})^{v}(\sqrt{n}\,r_{31})^{w} \,/\, H''_{0A}\right]$

We then define a joint partition of the three integers u, v, w, by an obvious generalisation of 4·3.

Denote such a joint partition by $(\alpha, \beta, \gamma)_s$.

By considering the expansion of $(\sqrt{n}\,r_{12})^{u}(\sqrt{n}\,r_{23})^{v}(\sqrt{n}\,r_{31})^{w}$ we find that $E\left[(\sqrt{n}\,r_{12})^{u}(\sqrt{n}\,r_{23})^{v}(\sqrt{n}\,r_{31})^{w} \,/\, H''_{0A}\right]$ can be expressed as the sum of a number (independent of n) of terms of the form

$$5\cdot2\cdot3. \quad C_{(\alpha,\beta,\gamma)_s} \frac{n^{[s]}}{n^{\frac{u+v+w}{2}}} \frac{1}{n^{[h_1]}} S'_{\epsilon_1(h_1)}(\underset{\sim}{x}'_{n1}) \frac{1}{n^{[h_2]}} S'_{\epsilon_2(h_2)}(\underset{\sim}{x}'_{n2}) \frac{1}{n^{[h_3]}} S'_{\epsilon_3(h_3)}(\underset{\sim}{x}'_{n3})$$

where

$C_{(\alpha,\beta,\gamma)_s}$ is a constant independent of n.

$\epsilon_1(h_1)$ is a partition of $(u+w)$ formed by the h_1 non-zero sums $\epsilon_{1i} = (\alpha_i + \gamma_i)$ of the joint-partition $(\alpha, \beta, \gamma)_s$;

$\epsilon_2(h_2)$ " " " " $(u+v)$ " " " h_2 " " " $\epsilon_{2j} = (\alpha_j + \beta_j)$ of the joint-partition $(\alpha, \beta, \gamma)_s$;

$\epsilon_3(h_3)$ " " " $(v+w)$ " " " h_3 " " $\epsilon_{3k} = (\alpha_k + \beta_k)$ of the joint-partition $(\alpha, \beta, \gamma)_s$;

and $S'_{\epsilon_i(h_i)}(\underset{\sim}{x}'_{ni})$ denotes a symmetric polynomial (as in 2·3·2) in the elements $(x'_{1i}, x'_{2i}, \ldots\ldots, x'_{ni})$ of the i^{th} B-class.

Noting that $h_1 + h_2 + h_3 \geqslant 2S$ we find that if each $\epsilon \geqslant 2$ and some $\epsilon > 2$ the term 5·2·3 is $o(1)$, if $\mu_j(\underset{\sim}{x}'_{ni}) = o(n^{\frac{j-2}{4}})$ $j = 3, 4, \cdots$ $i = 1, 2, 3.$

Also if one or more $\epsilon = 1$ the term 5·2·3 is easily shown to be $o(1)$.

Subject to this condition.

The remainder of the proof is similar to that of Theorem 3·5, The facts that r_{12}, r_{13}, and r_{23} have, asymptotically, independent randomisation distributions and that in particular $r_{12} = r_{13} = 1$ implies $r_{23} = 1$ for <u>all</u> n, seem to be mutually contradictory.

That they are not in fact so, is explained as follows.

Let (x_i, y_i, z_i) i = 1, 2,......,n be n sets of measures such that

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} z_i = 0,$$

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} z_i^2 = 1,$$

$$r_{12} = \sum x_i y_i = a \quad , \quad r_{13} = \sum x_i z_i = b\ .$$

r_{12} and r_{13} are respectively the product moment correlation coefficients between x and y and x and z.

We ask what are the limits of variation of $r_{23} = \sum_{i=1}^{n} y_i z_i$ subject to these conditions.

Introducing Lagrangian constant λ_1, λ_2, $\frac{1}{2}\lambda_3$, λ_4, we have stationary values of r_{23} given by the solutions of the above equations and the set of n equations,

$$\lambda_1 z_i + \lambda_2 + \lambda_3 y_i + \lambda_4 x_i = 0 \quad , \quad i = 1, 2, \cdots, n.$$

$\lambda_2 = 0$ by addition.

From this set of n equations we find easily that if r_0 is a stationary value of r_{23} then

$$b\lambda_1 + a\lambda_3 + \lambda_4 = 0$$

$$r_0\lambda_1 + \lambda_3 + a\lambda_4 = 0$$

$$\lambda_1 + r_0\lambda_3 + b\lambda_4 = 0$$

and hence $r_0 = ab \pm \sqrt{\{(1 - a^2)(1 - b^2)\}}$.

Clearly r_{23} has a maximum and a minimum value and these stationary values must be these.

If a = b = 1 then $r_{23} = 1$, as we had above.

But if $a = b = 0$ then r_{23} can lie between ± 1 ; and if a and b are both very small the limits of variations of r_{23} are very nearly ± 1.
Now the randomisation distributions of r_{12} and r_{13} discussed above are independent of each other and each distribution is asymptotically $N(0, \frac{1}{n})$.

Hence when $n \to \infty$ with probability $\to 1$, a and b can be taken to be $O(\frac{1}{\sqrt{n}})$. And so it is quite possible that asymptotically the randomisation distribution of the variables r_{12} , r_{13} , r_{23} are in fact independent. This resolves the apparent contradiction above.

Returning to the asymptotic distribution of W considered above, if $x_{ij} = i$, $i = 1,2,\ldots,n$, $j = 1,2,\ldots\ldots, m$, then as has been seen previously the $x_i{}'_{js}$ satisfy the condition that $m_\alpha(\underset{\sim}{x}'_{\lambda j}) = O(n^{\frac{\alpha-2}{4}})$ $\alpha = 3,4,\ldots\ldots$.
It follows that if the values x_{ij} from which we start are ranks, then the distribution of $W \sim N(0, 1)$.
Hence the distribution of $S_A'^2 \sim N\left(\frac{n}{m}, \frac{2n(m-1)}{m^3}\right)$.

For the case of ranks Kendall [6] has introduced a coefficient of concordance which we will denote by W_1 where

$$W_1 = \frac{1}{n} S_A'^2$$

According to the above the distribution of $W_1 \sim N\left(\frac{1}{m}, \frac{2(m-1)}{nm^3}\right)$.
Kendall has suggested [6] that the distribution function $F(W_1)$ of W_1 is given asymptotically by

$$dF = \frac{1}{B(p,q)} W_1^{p-1}(1-W_1)^{q-1}\, dW_1$$

where

$$p = \tfrac{1}{2}(n-1) - \tfrac{1}{m}$$

$$\nu = (m-1)\left\{\tfrac{1}{2}(n-1) - \tfrac{1}{m}\right\}.$$

The distribution suggested by Kendall tends as $n \to \infty$ to the $N\left(\frac{1}{m}, \frac{2(m-1)}{nm^3}\right)$ form, as is easily shown.

This proves that Kendall's suggestion is, in this sense, correct.

5·3: <u>The Multiple Correlation Coefficient</u>: The case of the multiple correlation coefficient provides an interesting application of an aspect of Theorem 3·6 that has not yet been discussed. This aspect is the following.

Theorem 3·6 stated that the joint randomisation distribution of a set $R_1, R_2, \ldots R_m$ of mutually orthogonal members of the class $\mathcal{R}$ is asymptotically multinormal subject to certain conditions.
It follows that subject to these conditions the joint randomisation distribution of m linearly independent linear combinations of $R_1, R_2, \ldots\ldots, R_m$ is asymptotically multinormal.
The question then arises - under what conditions is the joint distribution of a set $R_1, R_2, \ldots\ldots, R_m$ which are not necessarily orthogonal, asymptotically multinormal. This leads to the following corollary to Theorem 3·6.

5·3·1 Lemma : Let R_i, $i = 1,2,\ldots\ldots m$, be a set of m members of the class $\mathcal{R}$,

where $\quad R_i^{(n)}(\underset{\sim}{X_n}) = \frac{1}{\sqrt{n}}\left(y_{i1}X_1 + y_{i2}X_2 + \cdots\cdots + y_{in}X_n\right)$

and $\quad \sum_{\alpha=1}^{n} y_{i\alpha} = 0\,, \quad m_2(\underset{\sim}{y_{in}}) = 1\,, \quad i = 1,2,\cdots, m$

Also we suppose that $\sum_{i=1}^{n} x_i = 0$ and $m_2(\underset{\sim}{x_n}) = 1$

Let $\quad r_{ij}^{(n)} = \frac{1}{n}\sum_{\alpha=1}^{n} y_{i\alpha}\,y_{j\alpha}$

If (a) $\quad m_{j_1 j_2 \cdots j_m}(\underset{\sim}{y_{1n}}, \underset{\sim}{y_{2n}}, \cdots\cdots, \underset{\sim}{y_{mn}}) \leq O\left[n^{a(j-2)}\right]$ where $0 \leq a < \frac{1}{2}$

$j_1 + j_2 \ldots\ldots + j_m = j$, and $j = 3,4,\ldots\ldots\ldots$;

(b) $\quad r_{ij}^{(n)} \rightarrow r_{ij}$ as $n \rightarrow \infty$ and the matrix

$$\underset{\sim}{V} = \begin{bmatrix} 1 & r_{12} & \cdots\cdots & r_{1m} \\ r_{21} & 1 & \cdots\cdots & r_{2m} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots\cdots & 1 \end{bmatrix}$$

is non-singular and so positive definite ;

then if $m_j(\underset{\sim}{x}_n) = 0\left[n^{b(j-2)}\right]$ where $b = \frac{1}{2} - a$ for $j = 3, 4, \cdots$

the joint randomisation distribution of $R_1, R_2, \ldots\ldots R_m$ is asymptotically multinormal.

<u>Proof</u> : Let $\underset{\sim}{V}^{(n)}$ be the matrix $(r_{ij}^{(n)})$ with $r_{ii}^{(n)} = 1$ and let n be chosen so large that this matrix is non-singular.

Let $\underset{\sim}{C}^{(n)} = (c_{ij}^{(n)})$ be an orthogonal $m \times m$ matrix such that

$$\underset{\sim}{C}^{(n)\prime}\,\underset{\sim}{V}^{(n)}\,\underset{\sim}{C}^{(n)} = \begin{bmatrix} \lambda_1^{(n)} & 0 \cdots & 0 \\ 0 & \lambda_2^{(n)} \cdots & 0 \\ \vdots & & \\ 0 & \cdots\cdots & \lambda_m^{(n)} \end{bmatrix}$$

where $\lambda_1^{(n)}, \lambda_2^{(n)}, \cdots, \lambda_m^{(n)}$ are the (non-zero) latent roots of $\underset{\sim}{V}^{(n)}$

Let
$$\grave{R}_i^{(n)} = \frac{1}{\sqrt{\lambda_i^{(n)}}}\left(c_{1i}^{(n)} R_1^{(n)} + c_{2i}^{(n)} R_2^{(n)} + \cdots + c_{mi}^{(n)} R_m^{(n)}\right) \qquad i = 1, 2, \cdots, m$$

The set $\grave{R}_1, \grave{R}_2, \ldots\ldots, \grave{R}_m$ is an orthogonal set of members of the class $\mathcal{R}$ for if

$$\grave{R}_i^{(n)} = \frac{1}{\sqrt{n}}\left(\grave{y}_{i1} X_1 + \grave{y}_{i2} X_2 + \cdots + \grave{y}_{in} X_n\right).$$

then
$$\grave{y}_{i\alpha} = \frac{1}{\sqrt{\lambda_i^{(n)}}}\left(c_{1i}^{(n)} y_{1\alpha} + c_{2i}^{(n)} y_{2\alpha} + \cdots + c_{mi}^{(n)} y_{m\alpha}\right)$$

Also
$$\sum_{\alpha=1}^{n} \grave{y}_{i\alpha}\grave{y}_{j\alpha} = \frac{1}{\sqrt{(\lambda_i^{(n)}\lambda_j^{(n)})}} \sum_{k=1}^{m} c_{ki}^{(n)}\left[c_{1j}^{(n)} r_{k1}^{(n)} + c_{2j}^{(n)} r_{k2}^{(n)} + \cdots + c_{mj}^{(n)} r_{km}^{(n)}\right]$$

$$= \frac{n}{\sqrt{(\lambda_i^{(n)}\lambda_j^{(n)})}} \sum_{k=1}^{m} \lambda_j^{(n)} c_{ki}^{(n)} c_{kj}^{(n)}$$

$$= n\sqrt{\left(\frac{\lambda_j^{(n)}}{\lambda_i^{(n)}}\right)}\,\delta_{ij} \quad \text{where} \quad \delta_{ij} = \begin{matrix} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{matrix}.$$

So we have
$$\sum_{\alpha=1}^{n} \grave{y}_{i\alpha}\grave{y}_{j\alpha} = 0, \quad i \neq j$$

and
$$m_2(\underset{\sim}{\grave{y}}_{in}) = 1$$

Also since $m_{j_1 j_2 \cdots j_m}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n}, \cdots, \underset{\sim}{y}_{mn}) \leq O\left[n^{a(j-2)}\right] \quad j = 3, 4, \cdots,$

since $\lambda_i^{(n)} \to \lambda_i$ as $n \to \infty$ where λ_i is non-zero.

and since $|c_{ij}| < 1$ it is clear that

$$m_{j_1 j_2 \cdots j_m}(\grave{y}_{1n}, \grave{y}_{2n}, \cdots, \grave{y}_{mn}) \leq O[n^{a(j-2)}], \quad j = 3, 4, \cdots$$

It follows by Theorem 3·6 that the joint randomisation distribution of $\grave{R}_1$, $\grave{R}_2$, ... R_m is asymptotically multinormal.

Then since R_1, R_2 ..... R_m can be expressed as linear combinations of $\grave{R}_1$, $\grave{R}_2$, ..., R_m, their joint randomisation distribution is asymptotically multinormal.

This completes the proof of the lemma.

5·3·2 : We now apply this lemma to the case of the multiple correlation coefficient.

Let ξ, η_1, η_2, ..., η_m be $(m + 1)$ random variables.

Let H_o' denote the hypothesis that the distribution of ξ is independent of the distribution of the ηs

Let $(x_\alpha, y_{1\alpha}, y_{2\alpha}, \ldots\ldots, y_{m\alpha})$ $\alpha = 1, 2, \ldots\ldots n$, be n sets of values assumed by these random variables $n = 1, 2, 3, \ldots\ldots$

Without loss of generality in what follows we will assume that, in previous notation $\bar{x}_n = 0$, $m_2(\underset{\sim}{x}_n) = 1$,

$$\bar{y}_{in} = 0 \quad \text{and} \quad m_2(\underset{\sim}{y}_{in}) = 1 \qquad i = 1, 2, \cdots, m.$$

for all values of n.

Let $$v_{ij}^{(n)} = \frac{1}{n} \sum_{\alpha=1}^{n} y_{i\alpha} y_{j\alpha}$$

Let $\underset{\sim}{X}_n$ in previous notation be a permutation of $\underset{\sim}{x}_n$

Let $$S_i^{(n)} = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} X_\alpha y_{i\alpha}$$

Let $\underset{\sim}{V}^{(n)}$ be the $(m \times m)$ matrix $(v_{ij}^{(n)})$ where $v_{ii}^{(n)} = 1$

We assume that as $n \to \infty$ $\underset{\sim}{V}^{(n)} \to \underset{\sim}{V}$ a non-singular $m \times m$ matrix.

It is easily shown that $R^{(n)}$, the multiple correlation coefficient of X on y_1, y_2, ..., y_m is given by

$$n R^{(n)2} = \underset{\sim}{S}^{(n)'} \left(\underset{\sim}{V}^{(n)}\right)^{-1} \underset{\sim}{S}^{(n)}$$

where
$$\underset{\sim}{S}^{(n)} = \begin{bmatrix} S_1^{(n)} \\ S_2^{(n)} \\ \vdots \\ S_m^{(n)} \end{bmatrix}$$

Thus $n R^{(n)2}$ is a quadratic form in the $S_i^{(n)}$'s.

When H_0' is true all permutations $\underset{\sim}{X}_n$ are equally likely and the joint distribution of $X_1, X_2, \ldots, X_n$ is thus defined.

If
$$m_{j_1 j_2 \ldots j_m}(\underset{\sim}{y}_{1n}, \underset{\sim}{y}_{2n}, \ldots, \underset{\sim}{y}_{mn}) \leq O\left[n^{a(j-2)}\right]$$

$j_1 + j_2 \ldots\ldots + j_m = j \quad j = 3, 4, \ldots$ then by lemma 5·3·1 the joint randomisation distribution of the $S_i^{(n)}$s is asymptotically multinormal.

Also
$$E\left\{ S_i^{(n)} S_j^{(n)} / H_0' \right\} = \frac{1}{n} E\left\{ \left(\sum_{\alpha=1}^{n} X_\alpha y_{i\alpha}\right)\left(\sum_{\alpha=1}^{n} X_\alpha y_{j\alpha}\right) \Big| H_0' \right\}$$

$$= \frac{1}{n}\left[m_2(\underset{\sim}{x}_n) \sum_{\alpha=1}^{n} y_{i\alpha} y_{j\alpha} + \frac{1}{n(n-1)} S_{11}(\underset{\sim}{x}_n) \sum_{\substack{\alpha,\beta=1 \\ \alpha \neq \beta}}^{n} y_{i\alpha} y_{j\beta} \right]$$

where $S_{11}(\underset{\sim}{x}_n)$ is the symmetric polynomial in the x's as before.

$$= r_{ij}^{(n)} - \frac{m_2(x_n)}{n(n-1)} \left(-\sum_{\alpha=1}^{n} y_{i\alpha} y_{j\alpha} \right)$$

$$= r_{ij}^{(n)} + \frac{r_{ij}^{(n)}}{n-1} \longrightarrow r_{ij} \quad \text{as } n \longrightarrow \infty.$$

Hence the asymptotic correlation matrix of the $S_i^{(n)}$s is $\underset{\sim}{V}$.

Then since $n R^{(n)2} = \underset{\sim}{S}^{(n)'}\left(\underset{\sim}{V}^{(n)}\right)^{-1} \underset{\sim}{S}^{(n)}$ it follows that asymptotically the randomisation distribution of nR^2 is of the χ^2 form with m degrees of freedom.

<u>5·3·3</u> Once again this result is the same as the normal theory result. For it is well-known that if the joint distribution of $\xi, \eta_1, \ldots\ldots, \eta_m$ is multinormal and the distribution of ξ is independent of that of $\eta_1, \eta_2, \eta_3, \ldots\ldots, \eta_m$, then the sampling distribution function $F(R^2)$ of R^2 is given by

$$dF = \frac{1}{B\left(\frac{n-m-1}{2}, \frac{m}{2}\right)} \left(1-R^2\right)^{\frac{n-m-1}{2}-1} \left(R^2\right)^{\frac{m}{2}-1} dR^2$$

From this it is easily shown that the sampling distribution of $nR^2 \sim \chi^2$ with m degrees of freedom.

Again, subject to very mild restrictions, the randomisation approach and the normal theory approach are asymptotically equivalent.

Conclusion

The applications given in Chapter V cover a fairly wide field and there are probably many more statistics, less widely used, to which the theory established in Chapters 1 - 4 is applicable.

Of course any results of this paper are purely of mathematical interest. In practice it is necessary only to find the first few moments of a randomisation distribution in order to get a good enough approximation to it. This is commonly done, though it does often require a fair amount of algebraic manipulation (e.g. appendix 1).

The main mathematical interest is the extreme weakness of conditions sufficient to ensure asymptotic equivalence of the various randomisation and normal-theory distributions discussed. And this is just another facet of the theorem on which so much statistical theory is based - the Central Limit Theorem, though this thesis does little more than touch the fringe of this aspect of it.

APPENDIX I

In this we find exact expressions for the first two moments of the randomisation distribution of the sum of squares between classes for a one-way classification analysis of variance.

Let $x_{11}, x_{12}, \dots, x_{1n}, x_{21} \dots x_{2n}, \dots x_{pn}$ be a given set of numbers

Let $X_{11}, X_{12}, \dots, X_{1n}, X_{21}, \dots X_{2n}, \dots X_{pn}$ be a permutation of this set.

Let $X_{i.} = \frac{1}{n}\sum_{\alpha=1}^{n} X_{i\alpha}$

and $v_b = n \sum_{i=1}^{p} X_{i.}^2$ where without loss of generality we suppose $\sum_i \sum_j x_{ij} = 0$.

Suppose all permutations of the given set are equally likely

We have

$$v_b = \frac{1}{n}\left\{ X_{11}^2 + X_{12}^2 \cdots + X_{pn}^2 + 2\left[p \cdot {}^nC_2 \text{ terms of the form } X_{ij}X_{ik}\right]\right\}.$$

$$= \frac{1}{n} S_2 + \frac{2}{n} u$$

, say where u denotes the expression in square brackets and $S_{(\lambda)}$ denotes a symmetric polynomial in the pn values of x as previously.

We have

$$E(v_b) = \frac{1}{n} S_2 + \frac{2}{n} \frac{p \cdot {}^nC_2}{pn(pn-1)} S_{11}$$

Now $S_1^2 = S_2 + S_{11}$ and since $S_1 = 0$, $S_{11} = -S_2$.

$$\therefore E(v_b) = \frac{1}{n} S_2 \left[1 - \frac{n-1}{pn-1}\right] = \frac{p-1}{pn-1} S_2$$

$$= (p-1) k_2$$

where k_j is the j^{th} k-statistic of the set of x's.

Again $n^2 E(v_b^2) = E(S_2 + 2u)^2$

$$= S_2^2 + 4 S_2 E(u) + 4E(u^2).$$

From the above $E(u) = -\frac{n-1}{2(pn-1)} S_2$

Also $E(u^2) = p\,{}^nC_2 \frac{1}{(pn)^{[2]}} S'_{22} + 2p\,{}^nC_2 (n-2) \frac{1}{(pn)^{[3]}} S'_{211} +$

$$+ \, p \cdot {}^nC_2 (p \cdot {}^nC_2 - 2n+3) \frac{1}{(pn)^{[4]}} S'_{1111} .$$

We have

$$S_2'^2 = S'_4 + S'_{22}$$
$$S'_2 S_1'^2 = S'_4 + 2S'_{31} + S'_{22} + S'_{211}$$
$$S'_3 S'_1 = S'_4 + S'_{31}$$
$$S_1'^4 = S'_4 + 4S'_{31} + 3S'_{22} + 6S'_{211} + S'_{1111} .$$

and so since $S'_1 = 0$,

$$S'_{22} = S_2'^2 - S'_4$$
$$S'_{211} = 2S'_4 - S_2'^2 .$$
$$S'_{1111} = -6S'_4 + 3S_2'^2 .$$

Hence collecting up terms we get after some algebraic reduction

$$n^2 E(v_b^2) = \left\{ 1 - \frac{n-1}{pn-1} + \frac{n^2(n-1)p(p-1)}{(pn-1)(pn-2)(pn-3)} \right\} S_2'^2 - \frac{2n^2(n-1)p(p-1)}{(pn-1)(pn-2)(pn-3)} S'_4$$

or in terms of the k's

$$E(v_b^2) = \frac{np-1}{np+1}(p^2-1)k_2^2 - \frac{2(n-1)(p-1)}{n(np+1)} k_4 .$$

So if n is large and k_4 small, $E(v_b^2) \doteqdot (p^2-1)k_2^2$.

Again $n^3 E(v_b^3) = S_2'^3 + 6S_2'^2 E(u) + 12 S'_2 E(u^2) + 8E(u^3)$.

We already have $E(u)$, $E(u^2)$.

$$E(u^3) = p \cdot {}^nC_2 \frac{S'_{33}}{(pn)^{[2]}} + 3p.n^{[3]} \frac{S'_{321}}{(pn)^{[3]}} + p\,n^{[4]} \frac{S'_{3111}}{(pn)^{[4]}} + p\,n^{[3]} \frac{S'_{222}}{(pn)^{[3]}} +$$
$$+ \, p\,{}^nC_2 \left\{ p\,{}^nC_2 - 2n+3 + 4(n-2)(n-3) + 6({}^nC_2 - 2n+3) + 2(p-1)\,{}^nC_2 \right\} \frac{S'_{2211}}{(pn)^{[4]}} +$$
$$+ \, pn(n-1)\left\{ (pn(n-1) - 6n+12)\frac{n-2}{2} + [n(n-1) - 4n+6](n-4) + (p-1)n^{[3]} \right\} \frac{S'_{21111}}{(pn)^{[5]}}$$
$$+ \, \frac{1}{8} pn(n-1) \left\{ 3(p-1)n^{[4]} + (p-1)(p-2)\left(n^{[2]}\right)^2 + (n-2)^{[4]} \right\} \frac{S'_{111111}}{(pn)^{[6]}}$$

It is easily shown that

$$S'_{33} = -S'_6 + S_3'^2$$
$$S'_{321} = 2S'_6 - S'_4 S'_2 + 2S_3'^2$$
$$S'_{3111} = -6S'_6 + 3S'_4 S'_2 + S_3'^2$$

$$S'_{222} = 2S'_6 - 3S'_4S'_2 + S'^3_2$$
$$S'_{2211} = -6S'_6 + 5S'_4S'_2 + 2S'^2_3 - S'^3_2$$
$$S'_{21111} = 24S'_6 - 18S'_4S'_2 - 8S'^2_3 + 3S'^3_2$$
$$S'_{111111} = -120S'_6 + 90S'_4S'_2 + 40S'^2_3 - 15S'^3_2$$

The subsequent expression of $E(v_b^3)$ in terms of the k's involves a great deal of algebra and seems hardly worth while. On carrying this through and picking out terms of O(1) we get

$$E(v_b^3) = (p^2-1)(p+3)k_2^3 + o(1)$$

This verifies that the first three moments of the randomisation distribution of $\frac{v_b}{k_2}$ are, to the first order of approximation, those of a χ^2 variable with $(p-1)$ degrees of freedom.

1 Anderson R.L.: Distribution of the Serial Correlation Coefficient: Ann. Maths. Stats. XIII (1942).

2 Cramer H.: Mathematical Methods of Statistics (1946) pp 254-255.

3 Daniels H.E.: The Relation between Measures of Correlation in the Universe of Sample Permutations : Biometrika 33 (1944).

4 Hotelling H and Pabst M.R. : Rank Correlation and Tests of Significance Involving No Assumptions of Normality. Ann. Maths. Stats. 7 (1936).

5 Fisher R.A.: The Design of Experiments - 4th Edition (1947) § 21.

6 Kendall M.G.: The Advanced Theory of Statistics (1946) Vol 1. p 411.

7 Madow W.G.: On the Limiting Distribution of Estimates Based on Samples from Finite Universes: Ann. Math. Stats. XIX 4 (1948).

8. Noether G.E.: On a Theorem of Wald and Wolfowitz: Ann. Math. Stats. XX 3 (1949).

9 Noether G.E.: Asymptotic Properties of the Wald - Wolfowitz Test of Randomness : Ann. Math. Stats. XXI 2 (1950).

10 Pitman E.J.G.: Significance Tests which may be Applied to Samples from Any Population: I. Supp J.R.S.S. 4 (1937): II Supp. J.R.S.S. 4(1937) III Biometrika 29 (1938).

11 Wald A. and Wolfowitz J.: Statistical Tests Based on Permutations of the Observations: Ann. Math. Stats. XV 4 (1944).

12 Wald A. and Wolfowitz J.: An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation: Ann. Math.Stats. XIV 4(1943).

13 Welch B.L.: On the z-test in Randomised Blocks and Latin Squares Biometrika 29 (1937).

ADDITIONAL PAPER

A note on the logarithmic distribution.

In a series of N trials with constant probability π of success in each trial, the probability of S successes is a term of a binomial distribution viz., ${}^{N}C_{S}\, \pi^{S}(1-\pi)^{N-S}$

The most familiar method of deriving the Poisson distribution (giving the probability $e^{-s}\dfrac{m^{s}}{s!}$ of S successes in a large number of trials where the probability of success in each trial is small) is to let $N \to \infty$ and $\pi \to 0$ in the binomial distribution so that $N\pi \to m$.

Fisher's logarithmic distribution can be derived from the negative binomial distribution in a similar manner.

It is convenient to have in mind a particular set-up where the negative binomial distribution applies. The following set-up was chosen, despite its artificiality, for ease of interpretation.

If, from an effectively infinite volume of liquid containing organisms, a sample of given volume is drawn, the probability that the sample contains n organisms may be taken as the n^{th} term of a Poisson distribution. The mean m of this distribution depends on the volume of the sample. A sample of such volume that the corresponding Poisson distribution has mean m will be called a sample of type m.

If the volume of the sample is not fixed in advance but is chosen at random so that the probability that the sample type lies between m and m + dm is $\dfrac{1}{\Gamma(k)}\, p^{-k} m^{k-1} e^{-m/p}\, dm$ then the probability that the sample contains n organisms can easily be shown to be a term of a negative binomial distribution viz., $\dfrac{\Gamma(k+n)}{\Gamma(k)\, n!}\dfrac{p^{n}}{(1+p)^{k+n}}$

Suppose V samples are drawn in this way.

Then the expected number of samples containing n organisms is

$$V\,\frac{\Gamma(k+n)}{\Gamma(k)\, n!}\,\frac{p^{n}}{(1+p)^{k+n}}.$$

Now let $V \to \infty$ and $k \to 0$ so that $Vk \to \alpha$, a constant. Writing $x = \frac{p}{1+p}$ gives in the limit the expected number of samples containing n organisms to be $\frac{\alpha x^n}{n}$, a term of the logarithmic distribution.

Now let S denote the number of samples drawn containing at least one organism, and let N denote the total number of organisms obtained in drawing V samples, in the original case.

The probability $P(a_1, a_2, \ldots\ldots)$ of drawing

V - S samples containing no organisms

a_1 .... 1 organism

a_2 .... 2 organisms

and so on, where

$$a_1 + a_2 + \ldots\ldots = S$$

$$a_1 + 2a_2 + \ldots\ldots = N$$

is given by

$$P(a_1, a_2 \ldots) = \frac{V!}{(V-S)!\, a_1!\, a_2! \ldots}(1+p)^{-k(V-S)} \prod_n \left\{ \frac{\Gamma(k+n)}{\Gamma(k)\, n!} \frac{p^n}{(1+p)^{k+n}} \right\}^{a_n}.$$

Letting $V \to \infty$ and $k \to 0$, $Vk \to \alpha$, and writing $x = p/1+p$ gives

$$\mathcal{L}\, P(a_1, a_2 \ldots) = L(a_1, a_2, \ldots) = (1-x)^{\alpha} \prod_{n=1}^{\infty} \left(\frac{\alpha x^n}{n}\right)^{a_n} \frac{1}{a_n!}$$

L is then the likelihood function in the limiting case i.e. the logarithmic distribution case. This likelihood function leads to the maximum likelihood estimators $\hat{\alpha}$ and $\hat{x}$ of α and x respectively as the solutions of the equations

$$S = -\hat{\alpha} \log(1 - \hat{x})$$

$$N = \frac{\hat{x}}{1 - \hat{x}} \hat{\alpha}.$$

<u>The Limiting Form of the Joint Distribution of S and N.</u>

In order to find the joint sampling distribution of S and N for the logarithmic case, we have to sum L over all values of $a_1, a_2, \ldots\ldots$ such that

$$a_1 + a_2 + a_3 + \ldots\ldots = S$$

and $$a_1 + 2a_2 + 3a_3 + \ldots\ldots = N.$$

On performing this summation we get

$$P\{S = S, N = N\} = (1-x)^{\alpha} \frac{\alpha^S}{S!} \frac{x^N}{N!} f_S^N(0) \qquad \begin{matrix} 0 \le S \le \infty \\ S \le N \le \infty \end{matrix}$$

where
$$f_S(z) = [-\log(1-z)]^S$$
$$f_S^N(0) = \left[\frac{d^N}{dz^N} f_S(z)\right]_{z=0}$$

Hence the distribution of S is given by

$$P\{S = S\} = (1-x)^{\alpha} \frac{\alpha^S}{S!} \sum_{N=S}^{\infty} f_S^N(0) \frac{x^N}{N!}$$

$$= (1-x)^{\alpha} \frac{\alpha^S}{S!} \sum_{N=0}^{\infty} f_S^N(0) \frac{x^N}{N!} \quad \text{since } f_S^N(0) = 0, \text{ for } N = 0, 1, \cdots S-1.$$

$$= (1-x)^{\alpha} \frac{\alpha^S}{S!} f_S(x).$$

$$= (1-x)^{\alpha} \frac{\alpha^S}{S!} [-\log(1-x)]^S.$$

Hence in the logarithmic case S is distributed in the Poisson form with mean $-\alpha \log(1-x)$.

The moment generating function of the joint distribution of S and N is given by

$$E\left(e^{t_1 S + t_2 N}\right) = (1-x)^{\alpha} \sum_{S=0}^{\infty} \frac{(\alpha e^{t_1})^S}{S!} \sum_{N=S}^{\infty} \frac{(xe^{t_2})^N}{N!} f_S^N(0)$$

$$= (1-x)^{\alpha} \sum_{S=0}^{\infty} \frac{(\alpha e^{t_1})^S}{S!} \sum_{N=0}^{\infty} \frac{(xe^{t_2})^N}{N!} f_S^N(0)$$

$$= (1-x)^{\alpha} \sum_{S=0}^{\infty} \frac{(\alpha e^{t_1})^S}{S!} f_S(xe^{t_2})$$

$$= (1-x)^{\alpha} \sum_{S=0}^{\infty} \frac{[-\alpha e^{t_1} \log(1-xe^{t_2})]^S}{S!}$$

$$= \frac{(1-x)^{\alpha}}{(1-xe^{t_2})^{\alpha e^{t_1}}}$$

Putting $t_1 = 0$ we get the moment generating function of the distribution of N to be $\dfrac{(1-x)^{\alpha}}{(1-xe^{t_2})^{\alpha}}$.

It follows that N is distributed in the negative binomial form

and $$P\{N=N\} = (1-x)^{\alpha}\,\frac{\Gamma(\alpha+N)}{\Gamma(\alpha)}\,\frac{x^N}{N!}$$

Given S > 0 the distribution of N is of some interest. It is given by

$$P\{N=N/S\} = [-\log(1-x)]^{-S}\,\frac{x^N}{N!}\,f_S^N(0)$$

This distribution, the binomial distribution, the negative binomial distribution, and the Poisson all fall into the same class, as can be seen as follows :

Let a random variable ξ take the discrete set of values 0, 1,2,.....

and let

$$P\{\xi=n\} = \frac{1}{g(a)}\,\frac{a^n}{n!}\,g^n(0)$$ where g(z) is a 'suitable' function.

If $g(z) = (1+z)^s$ where S is a + ve integer then ξ has the binomial distribution with mean $p = a/1+a$.

For $g^n(0) = \frac{S!}{(S-n)!}$ and $P\{\xi=n\} = \frac{1}{(1+\frac{p}{1-p})^S}\,\frac{p^n}{(1-p)^n}\,{}^SC_n = {}^SC_n\,p^n(1-p)^{S-n}$.

If $g(z) = (1-Z)^{-\alpha}$ then $g^n(0) = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}$

and $P\{\xi=n\} = (1-a)^{\alpha}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}\,\frac{a^n}{n!}$ i.e., ξ has a negative binomial distribution.

If $g(z) = e^z$ then $g^n(0) = 1$

and $P\{\xi=n\} = e^{-a}\,\frac{a^n}{n!}$ i.e. ξ has a Poisson distribution.

Clearly other distributions of this type can be derived by giving g(z) different forms.

Returning to the original discussion we have

$$E(SN) = \frac{\partial^2}{\partial t_1\,\partial t_2}\left[\frac{(1-x)^{\alpha}}{(1-xe^{t_2})^{\alpha e^{t_1}}}\right]_{t_1=t_2=0}$$

$$= (1-x)^{\alpha}\left[\frac{\partial}{\partial t_1}\,\frac{\alpha x e^{t_1+t_2}}{(1-xe^{t_2})^{\alpha e^{t_1}+1}}\right]_{t_1=t_2=0}$$

$$= \alpha x(1-x)^{\alpha}\left[\frac{\partial}{\partial t_1}\frac{e^{t_1}}{(1-x)^{\alpha}e^{t_1}+1}\right]_{t_1=0}$$

$$= \alpha x(1-x)^{\alpha}\left[\frac{1}{(1-x)^{\alpha+1}} - \frac{\alpha\log(1-x)}{(1-x)^{\alpha+1}}\right]$$

$$= \frac{\alpha x}{1-x}\left\{1-\alpha\log(1-x)\right\}$$

Hence $\quad \operatorname{cov}(SN) = \dfrac{\alpha x}{1-x}$

Also we have var S = $-\alpha\log(1-x)$ and var N = $\dfrac{\alpha x}{(1-x)^2}$.
The results obtained here for var S, var N and cov(S, N) differ from these obtained by Professor Fisher in his discussion of the logarithmic distribution in connection with sampling a population of butterflies. This difference is accounted for as follows :-

Firstly we establish a correspondence between the butterfly problem and the negative binomial set-up above by making

Type of sample $\longrightarrow$ Species of butterfly.

By establishing this correspondence we automatically associate with a species a number m, which is the mean of the Poisson distribution arising from sampling the particular species (see below).
We suppose also that the probability of the number m associated with a species chosen at random lying between m and m + dm is $\frac{1}{\Gamma(k)}\, p^{-k} m^{k-1} e^{-m/p}\, dm$.
This maintains the correspondence.

We further assume that the result of our sampling activity is to fix attention on V species, chosen at random from the population and then to sample these species [This corresponds to drawing V samples].

With this correspondence we may replace 'type of sample containing n organisms' by 'species represented by n individuals throughout the above discussion.

Interest lies in the limiting case where $V \to \infty$ and $k \to 0$. Increasing V of course means increasing the number of species on which we fix attention but even as $V \to \infty$ we retain an element of randomness in the choice of species on which we fix attention.

In calculating var S and var N Professor Fisher considers a situation into which this element of randomness does not enter. Consequently the results here obtained for these quantities should be greater than his results. Mathematically this difference is brought out as follows :

If $f(m) = \frac{1}{\Gamma(k)} p^{-k} m^{k-1} e^{-m/p}$ and $P = \int_{m_1}^{m_2} f(m)\,dm$, where $0 < m_1 < m_2$, then in choosing V samples at random, probability that Δ of these have a value of m between m and m + dm is ${}^{V}C_{\Delta} P^{\Delta} (1-P)^{V-\Delta}$

As $V \to \infty$ and $k \to 0$, this probability $\to \dfrac{(\alpha P')^{\Delta} e^{-\alpha P'}}{\Delta !}$

where $P' = \int_{m_1}^{m_2} \frac{1}{m} e^{-m/p}\,dm$.

In other words in the situation discussed here, in the limiting case, the number of species on which we fix attention for sampling with a value of m between m_1 and m_2 does not converge in probability to its expected value $\alpha \int_{m_1}^{m_2} \frac{1}{m} e^{-m/p}\,dm$.

Professor Fisher evaluated var S in the limiting case on the basis that 'the distribution of species according to m' is given by $\frac{\alpha}{m} e^{-m/p}\,dm$. He was envisaging quite a different situation by considering the sampling technique such as to fix attention on species to be sampled so that there are, in the limiting case $\frac{\alpha}{m} e^{-m/p}\,dm$ such species whose associated number lies between m and m + dm.

Finally there are two less artificial situations in which the results here obtained might be applicable.

(i) A population of individuals consists of many different species. In each of various different 'districts' a very large number of species are present but not all species occur in any one district. The number m associated with a species is the expected number of this species observed as a result of a fixed sampling method in any 'district' in which the species is present. In the limiting case most species are very rare in the districts in which they occur. We then assume that sampling by a fixed method in a chosen district is equivalent to choosing a random selection of species from the population and then sampling this selection by the fixed sampling method.

(ii) A population of individuals consists of V different species. The medium in which the population exists is heterogeneous. So that in sampling a single species the probability of observing n members of the species may be taken to be $\frac{\Gamma(k+n)}{\Gamma(k)\,n!}\frac{p^n}{(1+p)^{k+n}}$ the same for every species. Every species in this situation has exactly the same status in the population. The limiting case now corresponds to sampling a population of individuals consisting of many species all of which are equally rare - quite a different situation from case I.

References :

(14) Fisher R.A. : The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. Journal of Anaimal Ecology 12.1943. (Also Author's Note in reprint 43 in "Contributions to Mathematical Statistics" published by Wiley 1950).

(15) Anscombe F.J: Sampling Theory of the Negative Binomial and Logarithmic Series Distributions : Biometrika 37 1950.