DISTRIBUTION-FREE METHODS IN EXPERIMENTAL DESIGN

by

STEPHANUS GUSTAV REINACH

submitted to the Faculty of Agriculture

University of Pretoria

in partial fulfilment of the requirements for the

degree of

D.Sc.(Agric.)

PRETORIA

February, 1966

PROMOTORS: PROF. D.G. HAYLETT

PROF. D.J. STOKER

AAN MY OUERS

AAN ELNA

A C K N O W L E D G E M E N T S

At the completion of this thesis I want to thank my parents who made a university education possible for me and also for their keen interest in my progress during my years of studies.

To Prof. D.J. Stoker my thanks are due for guidance in the mathematical development of the statistical techniques presented in this thesis, for the insistance of sufficiently practical investigation of the applicability of the asymptotically valid results to finite samples, also for willingness to discuss problems arising, and for time spent on checking and correcting the theoretical derivations of this thesis.

A special word of thanks is due to Prof. D.G. Haylett for advice given and time spent on the agricultural orientation of this thesis.

Acknowledgement is also due to Mr. C.M. Havenga for his interest and guidance in connection with applications of mathematical statistics in the biological sciences; and to the Canadian Government for a scholarship awarded for post-graduate study in Canada.

The Department of Agricultural Technical Services is thanked for permission to use material and facilities in connection with the preparation of the thesis and the University of Pretoria for permission to use the 1 M 1620 computer.

C O N T E N T S

Chapter page

P A R T II

A P P L I C A T I O N

Practical application of the theoretical formulae derived in Part I are presented in Part II. Examples of the completely random, the balanced incomplete block, the randomized block and the latin square designs appear in Chapters VIII to XI. Applications of the distribution-free multiple comparison procedures for the completely random and randomized block designs are given in Chapter XII.

The original theory for the various methods of analysis dealt with was derived for arbitrary functions of ranks (provided these functions complied with specified conditions).

The distribution-free test statistics, in the special case of ranks, for a two-way classification of treatments are given in the beginning of each chapter in Part II. The method of subdivision of the sum of squares of the rank totals for treatments is fully illustrated in order to show how the theory can be generalized to any number of components.

The formulae for the special case of ranks instead of functions of ranks are given because:

1. The loss in efficiency due to the use of ranks instead of functions of ranks is usually small;
2. all the sampling studies on the applicability of the test statistics in Part I were for the special case of ranks; and
3. in practice the tendency will be to use ranks instead of functions of ranks since the calculations involved are easier.

An evaluation of the new methods appears at the end of each chapter in Part II.

Part II is written in such a way that a person with only an elementary knowledge of biometry should be able to understand and apply the tests derived without necessarily referring to all the details in Part I.

Research workers applying these methods are advised to read the Appendix p. 264 before applying the methods discussed in the following chapters.

C H A P T E R VIII

COMPLETELY RANDOM DESIGN

8.1 Introduction

For the requirements of the completely random design and its field of application the reader is referred to Cochran & Cox (1957), Federer (1955) and other standard statistical text books on experimental design.

In this chapter (§8.2) the formulae for a general distribution-free method of analysis based on ranks, for a two-way classification of treatments, where the design is a completely random one, are given. The method of subdivision of the total sum of squares of the rank totals for treatments, when the treatment combinations consist of two components, is fully illustrated in order to show how the theory can be generalized to any number of components.

The original theory for this method of analysis, for functions of ranks, was derived by Lemmer (1964) and in Chapter III of this thesis.

The calculations involved, for the special case of ranks are illustrated in Example 8.1. The calculations involved when the ranks are transformed to "normal values" by means of Van der Waerden's (1957) transformation are illustrated in Example 8.2.

In §8.4 a short summary of the method of analysis is given. The efficiency and an evaluation of the distribution-free methods are given in §8.5 and §8.6.

8.2 Test statistics

Suppose there are k treatments or treatment combinations which can be subdivided into two components A and B at m and p levels each respectively. Suppose further that all the $k=mp$ treatment combinations are replicated n times (every treatment combination must be replicated the same number of times for the theory to be valid).

Indicate the rank of X_{ijh} in the joint ranking (in increasing order) of the observations by r_{ijh}, $1 \leq r_{ijh} \leq N=kn$. If ties are present among the observations, the mean of the ranks they would have had, had they been different, is allocated to the tied observations in each tie (cf. Chapter III). X_{ijh} is the h-th observation in the (i,j)-th treatment combination ($i=1,\ldots,m$; $j=1,\ldots,p$; $h=1,\ldots,n$)*

Let now:

$$R_{ij} = \sum_h r_{ijh}$$

$$R_{i.} = \sum_j R_{ij}$$

$$\bar{R}_{i.} = p^{-1}R_{i.}$$

$$R_{.j} = \sum_i R_{ij}$$

$$\bar{R}_{.j} = m^{-1}R_{.j}$$

$$N = nmp$$

$$R_{..} = \sum_i\sum_j R_{ij} = \tfrac{1}{2}N(N+1)$$

$$\bar{R}_{..} = (mp)^{-1}R_{..} = \tfrac{1}{2}n(N+1)$$

Kruskal & Wallis (1952) defined the following test statistic (T_{tot}) to test the null hypothesis H_0 that the k-treatment combinations have the same effect. The test statistic rewritten in our notation becomes:

$$T_{tot} = 12[N(N+1)n]^{-1}\sum_i\sum_j(R_{ij} - \bar{R}_{..})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(8.1)$$

$$= 12[N(N+1)n]^{-1}\sum_i\sum_j[R_{ij} - \tfrac{1}{2}n(N+1)]^2$$

$$= 12[N(N+1)n]^{-1}\sum_i\sum_j R_{ij}^2 - 3(N+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots(8.2)$$

---

* From now on it will for the rest of this chapter be assumed that i takes the values from 1 to m; j the values from 1 to p and h the values from 1 to n. Both $m \& p \geq 2$.

It has been shown (by Kruskal & Wallis, 1952) that under H_o, T_{tot} defined in (8.1) is distributed approximately as X^2 with (k-1) degrees of freedom provided N=nk is not too small (i.e. N=nk >15; cf. Siegel, 1956, p. 185). When k=3 and $n \leq 5$ the X^2-approximation to the exact sampling distribution of T_{tot} is not sufficiently close. For such cases exact probabilities are presented in Table O of Siegel (1956).

The test statistics for components A, B and the interaction AB may now be obtained as follows, from equation (8.1):

$$T_{tot} = K\sum_{i}\sum_{j}(R_{ij} - \bar{R}_{..})^2$$

$$\text{where } K = 12[N(N+1)n]^{-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8.3)$$

$$T_{tot} = K\sum_{i}\sum_{j}[R_{ij}-\bar{R}_{i.}-\bar{R}_{.j}+\bar{R}_{..}+(\bar{R}_{i.}-\bar{R}_{..})+(\bar{R}_{.j}-\bar{R}_{..})]^2$$

$$= K\sum_{i}\sum_{j}[R_{ij}-\bar{R}_{i.}-\bar{R}_{.j}+\bar{R}_{..}]^2 + Kp\sum_{i}(\bar{R}_{i.}-\bar{R}_{..})^2+Km\sum_{j}(\bar{R}_{.j}-\bar{R}_{..})^2 \quad \ldots\ldots\ldots\ldots(8.4)$$

All the cross-terms are zero.

From equation (8.4) we now obtain:

$$T_A = Kp\sum_{i}(\bar{R}_{i.} - \bar{R}_{..})^2$$

$$= 12[N(N+1)np]^{-1}\sum_{i}[R_{i.} - \tfrac{1}{2}np(N+1)]^2$$

$$\text{since } \bar{R}_{..} = \tfrac{1}{2}n(N+1)$$

$$=12[N(N+1)np]^{-1}\sum_{i}R_{i.}^2 - 3(N+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots(8.5)$$

$$T_B = Km\sum_{j}(\bar{R}_{.j} - \bar{R}_{..})^2$$

$$= 12[N(N+1)nm]^{-1}\sum_{j}[R_{.j} - \tfrac{1}{2}nm(N+1)]^2$$

$$= 12[N(N+1)nm]^{-1}\sum_{j}R^2_{.j} - 3(N+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots(8.6)$$

$$T_{AB} = K\sum_{i}\sum_{j}[R_{ij} - \bar{R}_{i.} - \bar{R}_{.j} + \bar{R}_{..}]^2$$

$$= T_{tot} - T_A - T_B \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8.7)$$

The formulae in equations (8.2) and (8.5) to (8.7) are exactly the same as those in (3.18). The calculating formulae where a function of the ranks are considered are given in §3.5.

In Chapter III it was proved (for functions of ranks) that T_A, T_B and T_{AB} are, asymptotically for $N \to \infty$, under the hypothesis H_0 of no effects, mutually independently distributed as X^2 with (m-1), (p-1) and (m-1)(p-1) degrees of freedom respectively, provided certain conditions regarding specific function of ranks considered are satisfied. In the case of ranks themselves and Van der Waerdens transformation to "normal values" these conditions are satisfied. Under an alternative hypothesis H_a these components are not distributed asymptotically independent of each other. By means of sampling studies it was indeed found that the test for, say, component A becomes more conservative when the effects of the levels of component B are increased.

In the case of small samples the following approximations may be used:

1. If a component consists of two levels only and $N \leq 40$ the U-test of Mann-Whitney (1947) must be used for that component. Exact probabilities for the U-test are tabulated by Siegel (1956) up to $N \leq 40$ in Tables J and K; and
2. if a component consists of three levels and $N \leq 15$, the exact probabilities tabulated by Siegel (1956) in Table O (given for the Kruskal-Wallis test) must be used.

In all other cases the X^2-approximation of the test statistic for that component can be used.

If ties are present among the observations the method of allocating the mean rank, i.e. the mean of the ranks they would have had were they different, to the tied observations in each tie, keeps the total of the ranks constant but changes the sum of squares of the ranks.

Let t denote the length of the ties and define now:

$$T = \sum_{(t)} (t^3 - t) \qquad \ldots\ldots(8.8)$$

where $\sum_{(t)}$ denotes the summation over all ties.

Define further:

$$CT = 1 - T/(N^3-N) \qquad \ldots\ldots(8.9)$$

The only necessary alteration is that the different test criteria have to be divided by this correction term (CT) to give T_{tot}(corr), T_A(corr), T_B(corr) and T_{AB}(corr). It should be noted that if no ties are present, all t=1 and thus T=0. The correction term then becomes 1. According to Siegel (1956) the X^2-approximation of T_{tot} is applicable if the length of any tie does not exceed 20% of the number of observations. This will probably also be sufficient for the X^2-approximations of T_A, T_B and T_{AB} to be applicable.

In the case of a factorial design there is a second method of analysis called the single degree of freedom approach (cf. Anderson & Bancroft, 1952, p. 272 and Goulden, 1952, p. 88).

For k=mp treatments there exist (k-1) orthogonal components corresponding to the (k-1) orthogonal sets of constants $g_{ij}^{(r)}$ defined by:

$$\left.\begin{array}{lll} \text{i)} & \sum_{ij}\sum g_{ij}^{(r)} = 0 & r=1,\ldots,k-1 \\ \text{ii)} & \sum_{ij}\sum g_{ij}^{(r)} g_{ij}^{(r')} = 0 & r \neq r' \end{array}\right) \qquad \ldots\ldots(8.10)$$

Let for this method:

$$L_r = \sum_{ij}\sum g_{ij}^{(r)} R_{ij} \qquad r=1,\ldots,k-1 \qquad \ldots\ldots(8.11)$$

and then

$$\mathrm{var}(L_r) = (12)^{-1}N(N+1)n\sum_{ij}\sum[g_{ij}^{(r)}]^2 \qquad \ldots\ldots(8.12)$$

$$T_r = L_r^2/\mathrm{var}(L_r) \qquad r=1,\ldots,k-1 \qquad \ldots\ldots(8.13)$$

will under H_o be, asymptotically, for $N \to \infty$, independently distributed as X^2 with one degree of freedom each (see Chapter III).

By means of (8.10), (8.11) and (8.12) the linear and higher order effects of a component, say A, in a factorial design can for instance be tested if desired, provided the X^2-approximation of T_A is considered applicable.

In the case of tied observations T_r (8.13) is to be divided by CT (8.9) to yield T_r(corr).

The application of the formulae given above ~~are~~ is illustrated by means of an example in the next paragraph.

8.3 Application

Example 8.1 - To illustrate the application of the formulae given in §8.2 let us consider a hypothetical experiment consisting of two cultivars of maize C_1 and C_2 and three levels of nitrogen fertilizer N_1, N_2 and N_3 applied in all possible combinations. Each treatment combination was replicated six times and the experimental design was a completely random design consisting of 36 plots.

In Table 8.1 the yields (in bags per morgen) as well as the corresponding ranks of the individual plots are given.

The ranks are allocated from the smallest observation taking rank 1 to the largest observation taking rank N=36. The totals of the ranks within each treatment combination R_{ij} (i=1,2; j=1,2,3) are then calculated.

Table 8.2 gives the interaction of cultivars of maize x fertilizer. The rank totals of cultivars of maize and the fertilizer levels are calculated.

It should be noted that the total of the ranks must always add up to $\frac{1}{2}N(N+1) = \frac{1}{2} \times 36 \times 37 = 666$ in this case.

Table 8.1 - Yields and ranks (r) of maize cultivars (C) x fertilizer (N) experiment

C_1N_1 bpm	r_{11h} ranks	C_1N_2 bpm	r_{12h} ranks	C_1N_3 bpm	r_{13h} ranks	C_2N_1 bpm	r_{21h} ranks	C_2N_2 bpm	r_{22h} ranks	C_2N_3 bpm	r_{23h} ranks
15.2	9	13.8	3	15.9	16	13.4	2	13.0	1	15.7	14
15.3	10	15.0	7	16.9	25	13.9	4	15.1	8	16.7	24
14.9	6	15.6	13	17.6	28	16.1	18	16.0	17	18.0	29
14.5	5	15.8	15	18.2	30	17.0	26	16.2	19	18.8	33
15.4	11	16.6	23	18.3	31	15.5	12	16.3	20	18.9	34
16.4	21	17.5	27	19.0	35	16.5	22	18.4	32	19.2	36
R_{ij}	62		88		165		84		97		170

Table 8.2 - Interaction maize cultivars x fertilizer

	N_1	N_2	N_3	$R_{i.}$
C_1	62	88	165	315
C_2	84	97	170	351
$R_{.j}$	146	185	335	666

In the formulae of §8.2 cultivars of maize are now taken as component A and nitrogen fertilizer as component B, thus m=2, p=3, k=6 and n=6.

From equation (8.2) we have:

$$T_{tot} = 12[nN(N+1)]^{-1} \sum_{i=1}^{2} \sum_{j=1}^{3} R_{ij}^{2} - 3(N+1)$$

$$= 12[6 \times 36 \times 37]^{-1}[62^2+88^2+\ldots+97^2+170^2] - 3 \times 37$$

$$= 15.4$$

From equation (8.5) we have:

$$T_C = 12[npN(N+1)]^{-1} \sum_i R_{i.}^2 - 3(N+1)$$

$$= 12[6 \times 3 \times 36 \times 37]^{-1}[315^2 + 351^2] - 3 \times 37$$

$$= .33$$

From equation (8.6) we have:

$$T_N = 12[nmN(N+1)]^{-1} \sum_j R_{.j}^2 - 3(N+1)$$

$$= 12[6 \times 2 \times 36 \times 37]^{-1}[146^2+185^2+335^2] - 3 \times 37$$

$$= 14.95$$

From the table of the X^2-distribution we obtain $X^2_{.01}(2) = 9.21$ which means that T_N is highly significant. Note that, according to the rules given in the previous paragraph, the X^2-approximation is applicable in this case.

From equation (8.7) we have:

$$T_{NC} = T_{tot} - T_C - T_N$$

$$= 0.12$$

If the above was a one-way classification with six treatments the test statistic H defined by Kruskal & Wallis (cf. §2.3.4) would be the same as T_{tot} calculated in the example.

Suppose now the nitrogen fertilizer was chosen on a factorial basis, with the three levels being at equidistant intervals, then we could proceed to split T_N in two orthogonal components corresponding to the linear and quadratic effects of nitrogen fertilizer.

This step is only necessary if the value of T_N is larger than $X^2_{.05}(1) = 3.84$ because if the value of T_N is smaller than 3.84 no component of N can be significant at the 5% level even if it contains the whole value of T_N.

The values of the orthogonal polynomials corresponding to the linear and quadratic components of nitrogen fertilizer component are obtainable from the tables of Fisher & Yates (1938). They are

Linear (-1, 0, 1)

Quadratic (1, -2, 1)

where the first value in both cases corresponds to N_1 the second to N_2 and the third to N_3.

For the linear component the calculations proceed as follows:

$$\sum_i\sum_j[g_{ij}^{(r)}]^2 = m\sum_j[g_{.j}^{(r)}]^2$$

since the same set of values for the $g_{ij}^{(r)}$ are taken for all i.

$$= 2[(-1)^2 + 1^2] = 4$$

From equation (8.11) we have:

$$L_{N_1} = \sum_i\sum_j g_{ij}^{(r)} R_{ij} = \sum_i\sum_j g_{.j}^{(r)} R_{.j}$$

$$= -R_{.1} + R_{.3}$$

$$= -146 + 335 = 189$$

Equation (8.12) becomes:

$$\text{var}(L_{N_1}) = (12)^{-1} 36 \times 37 \times 6 \times 4$$

$$= 2664$$

Substituting the values above in (8.13) yields:

$$T_{N_1} = (189)^2/2664 = 13.41$$

Following the same procedure for the quadratic component we obtain:

It should be noted that in this case both the nonparametric and the parametric method of analysis lead to the same conclusion.

Example 8.2 - The use of the transformation on the ranks to "normal values" suggested by Van der Waerden (1957) will be illustrated with the aid of the data in the first example. The calculating formulae for functions of ranks have not been summarized in this chapter but **are directly obtain**able from §3.5.3 (p. 43). For the applicability of the X^2-approximation in this case see §3.8 (p. 62).

The ranks r_{ijh} of Table 8.1 are divided by the term $(N+1)=37$ to give fractions δ_{ijh}. These fractions are then transformed with the aid of the values in Table 2 of Van der Waerden (1957) page 334 $[Q(\delta_{ijh})]$.

The results obtained are presented in Table 8.5.

Table 8.6 is the interaction table for the transformed values.

With the aid of the formulae in §3.5.3 the different test statistics can now be calculated.

$$
\begin{aligned}
M_N &= \sum_{s=1}^{36} [Q(\delta_s)]^2 \\
&= (-0.70)^2 + (-0.61)^2 + \dots + 1.93^2 \\
&= 30.22
\end{aligned}
$$

Van der Waerden (1957) tabulated in Table 12 the value $N^{-1}M_N$ (if no ties are present). The value tabulated by him for N=36 is 0.836. If the above value is divided by N=36 we obtain 0.839 which is nearly the same except for rounding errors.

$$
\begin{aligned}
T_{tot} &= (N-1)(nM_N)^{-1}\sum_{i}\sum_{j} z_{ij.}^2 \\
&= 35(6 \times 30.22)^{-1}[(-3.76)^2+\dots+5.43^2] \\
&= 14.799
\end{aligned}
$$

Table 8.5 - Transformed values of the ranks in Table 8.1

δ_{11h}	$Q(\delta_{11h})$	δ_{12h}	$Q(\delta_{12h})$	δ_{13h}	$Q(\delta_{13h})$	δ_{21h}	$Q(\delta_{21h})$	δ_{22h}	$Q(\delta_{22h})$	δ_{23h}	$Q(\delta_{23h})$
.243	-0.70	.081	-1.40	.432	-0.17	.054	-1.61	.027	-1.93	.378	-0.31
.270	-0.61	.189	-0.88	.676	0.46	.108	-1.24	.216	-0.79	.649	-0.38
.162	-0.99	.351	-0.38	.757	0.70	.486	-0.04	.459	-0.10	.784	0.79
.135	-1.10	.405	-0.24	.811	0.88	.703	0.53	.514	0.04	.892	1.24
.297	-0.53	.622	0.31	.838	0.99	.324	-0.46	.541	0.10	.920	1.40
.568	0.17	.730	0.61	.946	1.61	.595	0.24	.865	1.10	.973	1.93
$z_{11.}$	-3.76	$z_{12.}$	-1.98	$z_{13.}$	4.47	$z_{21.}$	-2.58	$z_{22.}$	-1.58	$z_{23.}$	5.43

Table 8.6 - Interaction cultivars of maize x fertilizer

	N_1	N_2	N_3	$z_{i..}$
C_1	-3.76	-1.98	4.47	-1.27
C_2	-2.58	-1.58	5.43	1.27
$z_{.j.}$	-6.34	-3.56	9.90	0

$$T_N = (N-1)(nmM_N)^{-1}\sum_j z^2_{.j.}$$
$$= 35(6 \times 2 \times 30.22)^{-1}[(-6.34)^2+(-3.56)^2+9.90^2]$$
$$= 14.560$$

$$T_C = (N-1)(npM_N)^{-1}\sum_i z^2_{i..}$$
$$= 35(6 \times 3 \times 30.22)^{-1}[(-1.27)^2 + 1.27^2]$$
$$= 0.208$$

$$T_{NC} = T_{tot} - T_N - T_C = 0.031$$

To calculate the linear component of fertilizer we have from equation (3.25):

$$L_{N_1} = n^{-\frac{1}{2}}\sum_i\sum_j g^{(r)}_{ij}(z_{ij.} - n\,\bar{z}_{...})$$
$$= 6^{-\frac{1}{2}}\sum_i g^{(r)}_{.j}\,z_{.j.} \qquad \text{since } \bar{z}_{...} = 0$$
$$= 6^{-\frac{1}{2}}(-z_{.1.} + z_{.3.})$$
$$= 6^{-\frac{1}{2}}(6.34 + 9.90) = 6^{-\frac{1}{2}} \times 16.24$$

From equation (3.30) we have:

$$\mathrm{var}(L_{N_1}) = (N-1)^{-1}M_N\sum_i\sum_j[g^{(r)}_{ij}]^2$$
$$=(30.22 \times 4)/35$$

Substituting the above values in equation (3.32) yields:

$$T_{N_1} = L^2_{N_1}/\mathrm{var}(L_{N_1}) = 12.73$$

In a similar way the quadratic component can be calculated, viz.

$$T_{N_q} = 1.83$$

Comparing these values with those obtained in Example 8.1 we see that **from** both analyses the same conclusions are drawn.

8.4 Summary of the procedure

These are the steps in the distribution-free analysis of a two-way classification of treatments in a completely random design using ranks:

1. Rank all the observations for the k treatment combinations in a single series, allocating ranks from 1 to N, and mean ranks to tied observations;

2. determine the value of R_{ij} (the sum of the ranks) for each of the k treatment combinations;

3. set up an interaction table of the R_{ij}'s and obtain the total of the ranks $R_{i.}$ and $R_{.j}$ for the two components A and B;

4. calculate the test statistics T_{tot}, T_A, T_B and T_{AB} with the aid of the formulae (8.2), (8.5), (8.6) and (8.7);

5. if applicable, components with one degree of freedom may now be calculated with the aid of formula T_r in equation (8.13);

6. if a large proportion of the observations are tied (i.e. if the lengths of the ties are substantial), compute the CT defined in (8.9) and divide the test statistics calculated in 4. and 5. above by this term;

7. the method of assessing the significance of any component, say A, at m levels depends on the value of m and N.

a) If $m=2$ and $N \leq 40$ the U-test of Mann- Whitney (1947) must be used. Exact probabilities for the U-test are tabulated in Siegel (1956) in Tables J and K.

b) If $m=3$ and $N \leq 15$, Table O in Siegel (1956) must be used to determine the associated probability under H_0 of an T_A value as large as that observed.

c) In all other cases, the significance of a value as large as the observed value T may be assessed by reference to the X^2-table.

8.5 Efficiency of the distribution-free tests given in §8.2

Lemmer (1964) in his Chapter V proved that the asymptotic relative efficiency (a.r.e.) (for shift alternatives) for any component say A, when compared with the F-test, in the case of normal distributed variates, is $3/\pi$, provided the effects of the remaining components are zero.

From sampling studies done (cf. §3.7) it is seen that the value of T_A is depressed when the effects of the levels of the remaining components are not zero. This means that the a.r.e of the test based on T_A (compared with the F-test for normally distributed variates) in general will be less than $3/\pi$.

To conclude it can be stated that the a.r.e. for alternatives of shift, of the distribution-free test statistics given in §8.2, when compared with the F-test, in the case of normal distributed variates, is unknown but in general smaller than $3/\pi$.

From theoretical consideration (cf. Lermer, 1964) it can be expected that the Van der Waerden type of tests (cf. Example 8.2) will generally be more efficient than the tests discussed in this chapter (based on ranks), although they are more difficult to apply.

8.6 Evaluation of the new distribution-free methods

Some of the advantages of the distribution-free methods when compared with the F-test can be summarised as follows:

1. The calculations involved may be less than those for the F-test;

2. the distribution of the test statistics ~~are~~ is independent of the distribution of the population from which the original observations come and the corresponding tests can be applied as long as the observations can be ranked. (This is not strictly true in the case where ties exist among the observations);

3. a 2^n factorial design which is not replicated can be

analysed by means of the distribution-free methods without any problem. In the parametric case it is necessary to use the higher order interactions as estimates of the error variance before the F-test can be applied, which is not a very sound practice.

The most important disadvantage of the distribution-free test statistics given in § 8.2 is that their a.r.e., compared to the F-test, is unknown but in general less than $3/\pi$ if the requirements of the F-test are met in the case of shifting alternatives.

It is thus recommended that the distribution-free tests should only be used if the F-test is not applicable or if doubt exists as to the applicability of the F-test.

C H A P T E R IX

BALANCED INCOMPLETE BLOCK DESIGNS

9.1 Introduction

The balanced incomplete block design is used in practice when a large number of treatments must be tested and the number of plots available in a homogeneous block is smaller than the number of treatments.

For a full discussion on requirements, fields of application, arangement of experimental material, randomization and statistical analysis in the parametric case, the reader is referred to Cochran & Cox (1957).

The distribution-free test statistics, in the special case of ranks, for a two-way classification of treatments are given in §9.2. The method of subdivision of the sum of squares of the rank totals for treatments, is fully illustrated in order to show how the theory can be generalized to any number of components.

The theory for this method of analysis is given in Chapter V, Part I.

The application of the formulae is illustrated with an example in §9.3. A summary of the steps in the analysis appears in §9.4 and an evaluation of this method of analysis in §9.5.

9.2 Test statistics

Suppose there are two factors or components A and B at m and p levels each respectively. (There can be any number of components and levels). Thus there are $k=mp$ treatment combinations in total. Suppose further:

t = number of treatment combinations (plots) in each block ($k \geqslant t$), each treatment combination appears in a block only once;

n = number of times any treatment combination is replicated, and is a constant for all treatments;

b = number of blocks = kn/t;

λ = number of times any two treatment combinations appears in the same block

= n(t-1)/(k-1) and is a constant for all possible pairs of treatments.

Within each block the observations are ranked from 1 for the smallest value to t for the largest. If tied observations are present within a block the tied observations in each tie are given the mean rank of the ranks they would have had if they had been different (cf. Chapter IV). The observations in each block are ranked independently of the observations in any other block.

Let now r_{ijh} be the rank of X_{ijh} where X_{ijh} is the h-th replication of the (i,j)-th treatment combination (i=1,...,m; j=1,...,p; h=1,...,n).* It should be noted that these indexes have no connection with the blocks.

Let now:

$$R_{ij} = \sum_h r_{ijh}$$

$$R_{i.} = \sum_j R_{ij}$$

$$R_{.j} = \sum_i R_{ij}$$

$$R_{..} = \sum_i\sum_j R_{ij} = \tfrac{1}{2}bt(t+1)$$

$$\bar{R}_{i.} = p^{-1}R_{i.}$$

$$\bar{R}_{.j} = m^{-1}R_{.j}$$

$$\bar{R}_{..} = (mp)^{-1}R_{..} = \tfrac{1}{2}n(t+1) = \tfrac{1}{2}bt(t+1)/k$$

$$N = nmp = bt$$

Durbin (1951) defined the following test statistic

---

* For the rest of this chapter it is assumed that i takes the values 1,...,m; j the values 1,...,p and h the values 1,...,n.

(T_{tot}) to test the null hypothesis H_o that the k treatment combinations have the same effect. The test statistic rewritten in our notation becomes:

$$T_{tot} = 12[\lambda k(t+1)]^{-1}\sum_{ij}\sum[R_{ij} - \bar{R}_{..}]^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(9.1)$$

$$= 12[\lambda k(t+1)]^{-1}\sum_{ij}\sum[R_{ij} - \tfrac{1}{2}n(t+1)]^2$$

$$= 12[\lambda k(t+1)]^{-1}\sum_{ij}\sum R_{ij}^2 - 3n^2(t+1)\lambda^{-1} \quad \ldots\ldots\ldots\ldots(9.2)$$

It was shown by Durbin (1951) that under H_o, T_{tot} defined in (9.1) is asymptotically for $n \rightarrow \infty$, distributed as X^2 with (k-1) degrees of freedom.

No exact probabilities are available in the case of small samples. In §4.10 an estimate was obtained as to when the X^2-approximation may be applicable. It was estimated that if the X^2-approximation was applicable for a given k (number of treatments) and n' (number of replications) in the randomized block design it may also be applicable for the same k and

$$n = n't(k-1)/k(t-1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9.3)$$

replications in the balanced incomplete block design.

The test statistics for the components A, B and the interaction AB may be obtained from equation (9.1) in a similar way as in equation (8.4). They are given by:

$$T_A = 12[\lambda pk(t+1)]^{-1}\sum_{i}[R_{i.} - \tfrac{1}{2}np(t+1)]^2$$

$$= 12[\lambda pk(t+1)]^{-1}\sum_{i}R_{i.}^2 - 3n^2(t+1)\lambda^{-1} \quad \ldots\ldots\ldots\ldots\ldots(9.4)$$

$$T_B = 12[\lambda mk(t+1)]^{-1}\sum_{j}[R_{.j} - \tfrac{1}{2}nm(t+1)]^2$$

$$= 12[\lambda mk(t+1)]^{-1}\sum_{j}R_{.j}^2 - 3n^2(t+1)\lambda^{-1} \quad \ldots\ldots\ldots\ldots\ldots(9.5)$$

$$T_{AB} = T_{tot} - T_A - T_B \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9.6)$$

The formulae given in equations (9.2), (9.4), (9.5) and (9.6) are exactly the same as those in equation (4.94).

No exact probabilities are available for the distribution of the test statistics of the components in the case of small samples for the following reasons:

1. For a given number of treatments there exist several designs which can be used, e.g. when k=5, designs with block sizes of 2, 3 and 4 may be used; and

2. this design is very seldom used if the treatments form a two- or higher-way layout since it is then usually possible to confound some of the higher order interaction(s) with blocks, thus reducing the block sizes (cf. Cochran & Cox, 1957).

It may reasonably be assumed that if the χ^2-approximation is applicable for T_{tot} that it will also be applicable for the different components.

In the case where tied observations are present among the observations within any block the procedure is as follows:

Let $d_{s\alpha_s}$ denote the length of the α_s-th tie in the s-th block and define

$$C_s = \sum_{\alpha_s} (d^3_{s\alpha_s} - d_{s\alpha_s}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9.7)$$

where $\sum_{\alpha_s}$ denotes the summation over all different values in the s-th block.

Define further:

$$CT = 1 - \sum_{s=1}^{b} C_s/bt(t^2-1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9.8)$$

The only necessary alteration is that the different test criteria have to be divided by this correction term (CT) to give T_{tot}(corr), T_A(corr), T_B(corr) and T_{AB}(corr).

It should be noted that if no ties are present within any block, all the $d_{s\alpha_s}=1$ and thus $C_s=0$. The correction term then has the value 1.

In the case of a factorial design there exists a second method of analysis namely the single degree of freedom approach (cf. Anderson & Bancroft, 1952, p. 272 and Goulden, 1952, p. 88).

For k=mp treatments there exist (k-1) orthogonal components corresponding to the (k-1) orthogonal sets of constants $g_{ij}^{(r)}$ defined by:

$$\left.\begin{array}{lll} \text{i)} & \sum\sum_{ij} g_{ij}^{(r)} = 0 & r=1,\ldots,k-1 \\ \text{ii)} & \sum\sum_{ij} g_{ij}^{(r)} g_{ij}^{(r')} = 0 & r\neq r' \end{array}\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9.9)$$

Let for this method (cf. equation (4.95)):

$$L_r = \sum\sum_{ij} g_{ij}^{(r)} R_{ij} \qquad r=1,\ldots,k-1 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(9.10)$$

then

$$\text{var}(L_r) = [12(k-1)]^{-1} n(t+1)(t-1)k \sum\sum_{ij} [g_{ij}^{(r)}]^2$$

$$= (12)^{-1} \lambda k(t+1) \sum\sum_{ij} [g_{ij}^{(r)}]^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(9.11)$$

since $\lambda = n(t-1)/(k-1)$

$$T_r = L_r^2/\text{var}(L_r) \qquad r=1,\ldots,k-1 \qquad \ldots\ldots\ldots\ldots\ldots(9.12)$$

will under H_0 be, asymptotically for $n \rightarrow \infty$, independantly distributed as X^2 with one degree of freedom (see Chapter IV).

By means of equation (9.12) the linear and higher order effects of a component, say T_A, in a factorial design, can for instance be tested if desired.

In the case of tied observations in any of the blocks T_r (9.12) is to be divided by CT (9.8) to yield T_r(corr).

The application of the formulae given above is illustrated in the next paragraph.

9.3 Application

This example is for illustration purposes only as it has not been proved that the asymptotic approximation is applicable.

No suitable example was available and it was decided to take the data from Table XIII-I in Federer (1955) p. 420.

Table 9.1 - Yields and ranks (r) of a 2 x 5 factorial experiment

Blocks	Treatment combinations																			
	P_1N_5		P_1N_4		P_1N_3		P_1N_2		P_1N_1		P_2N_5		P_2N_4		P_2N_3		P_2N_2		P_2N_1	
	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r
1	9.7	4	8.7	3			5.4	2	5.0	1										
2			9.6	4	8.8	3					5.6	2							3.6	1
3			9.0	4			7.3	3			3.8	1	4.3	2						
4	9.3	4			8.7	3			6.8	2			3.8	1						
5	9.0	4					7.5	3							4.2	2			2.8	1
6			9.6	4									5.1	3	4.6	2	3.6	1		
7			9.8	4					7.4	3					4.4	2			3.8	1
8									9.4	4			6.3	3			5.1	2	2.0	1
9	9.3	3.5	9.3	3.5	8.2	2											3.3	1		
10							8.7	3	9.0	4	6.0	2					3.3	1		
11	9.7	4									6.7	3	6.6	2					2.8	1
12					9.3	4	8.1	3									3.7	2	2.6	1
13	9.8	4									7.3	3			5.4	2	4.0	1		
14					9.0	4	8.3	3					4.8	2	3.8	1				
15					9.3	4			8.3	3	6.3	2			3.8	1				
R_{ij}		23.5		22.5		20		17		17		13		13		10		8		6

The ten treatments of Federer's example were artificially divided into two components which will be called methods of planting (P_1 & P_2) and nitrogen fertilizer at five levels (N_1, N_2, N_3, N_4 & N_5). The design was a balanced incomplete block design with 15 blocks of four plots each. Each treatment combination appears six times in the experiment as a whole.

Suppose the data given in Table 9.1 were bags per morgen of maize harvested on the plots. The design used as well as the yields with their corresponding ranks appear in Table 9.1.

Table 9.2 - Interaction table of ranks for a 2 x 5 factorial experiment

	N_1	N_2	N_3	N_4	N_5	$R_{i.}$
P_1	17	17	20	22.5	23.5	100
P_2	6	8	10	13	13	50
$R_{.j}$	23	25	30	35.5	36.5	150

The different constants defined in §9.2 are as follows:
$\lambda = 2$, $k = 10$, $t = 4$, $n = 6$.

If k=10, n'=2 replications are sufficient in the case of the randomized block design for the X^2-approximation to be applicable (cf. Chapter X). With the aid of equation (9.3) we can now estimate the number of replications needed in the balanced incomplete block design above for the X^2-approximation to be valid (cf. §4.10).

$n = n't(k-1)/k(t-1) = (2 \times 4 \times 9)/(10 \times 3) = 2.4$

i.e. 3 replications will be sufficient and we have 6. Thus it seems safe to use the X^2-approximation for the distributions of the test statistics.

From Table 9.2 the following values can be calculated:

$$\sum_i\sum_j R_{ij}^2 = 2574.50$$

$$\sum_i R_{i.}^2 = 12500.00$$

$$\sum_j R_{.j}^2 = 4646.50$$

Substituting the above values in equations (9.2), (9.4), (9.5) and (9.6) yields:

$$T_{tot} = 12[\lambda k(t+1)]^{-1}\sum_i\sum_j R_{ij}^2 - 3n^2(t+1)/\lambda$$
$$= 12(2 \times 10 \times 5)^{-1} \times 2574.00 - 3 \times 6^2 \times 5/2$$
$$= 38.94$$

$$T_P = 12[\lambda kp(t+1)]^{-1}\sum_i R_{i.}^2 - 3n^2(t+1)/\lambda$$
$$= 12(2 \times 10 \times 5 \times 5)^{-1} \times 12500 - 3 \times 6^2 \times 5/2$$
$$= 30.00$$

$$T_N = 12[\lambda nk(t+1)]^{-1}\sum_j R_{.j}^2 - 3n^2(t+1)/\lambda$$
$$= 12(2 \times 2 \times 10 \times 5)^{-1} \times 4646.50 - 3 \times 6^2 \times 5/2$$
$$= 8.79$$

$$T_{NP} = T_{tot} - T_N - T_P$$
$$= 0.15$$

If the nitrogen fertiliser was originally applied at equidistant levels, the linear and higher order effects may be calculated with the aid of equation (9.12). The values of the orthogonal polynomials are obtainable from the tables of Fisher & Yates (1938). The values corresponding to the levels N_1, N_2, N_3, N_4 and N_5 for the linear component are (-2, -1, 0, 1, 2). The following may now be calculated.

$$\sum_i\sum_j[g_{ij}^{(r)}]^2 = m\sum_j[g_{.j}^{(r)}]^2 = 2[(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2]$$
$$= 20$$

since the same set of values for the $g_{ij}^{(r)}$ are taken for all i, we write $g_{ij}^{(r)} = mg_{.j}^{(r)}$.

From (9.10) we have:

$$L_r = \sum_i \sum_j g_{ij}^{(r)} R_{ij} = \sum_j g_{.j}^{(r)} R_{.j}$$

$$= -2R_{.1} - R_{.2} + R_{.4} + 2R_{.5}$$

$$= -2(23) - 25 + 35.5 + 2(36.5)$$

$$= 37.5$$

From (9.11) we have:

$$\operatorname{var}(L_r) = (12)^{-1} \lambda k(t+1) \sum_i \sum_j [g_{ij}^{(r)}]^2$$

$$= (12)^{-1} \times 2 \times 10 \times 5 \times 20 = 166.67$$

Substituting these values in (9.12) yields:

$$T_{N_1} = (37.5)^2 / 166.67 = 8.44$$

where N_1 means the linear effect of N.

The rest of the fertiliser component is obtained by subtraction.

$$T_{N_r} = T_N - T_{N_1} = 8.79 - 8.44 = 0.35$$

The results obtained can be summarised as in Table 9.3

Table 9.3 - Distribution-free analysis of variance

Source of variation		DF	T	
Ploughing		1	30.00**	
Fertiliser	Linear	1		8.44**
Fertiliser		4	8.79	
Fertiliser	Rest	3		0.35
Interaction		4	0.15	
Total		9	38.94	

** significant at the 1% level

9.4 Summary of the procedure

The following are the steps in the distribution-free analysis of a two-way classification of treatments in a balanced incomplete block design by means of ranks:

1. Rank the observations within each block from 1 for the smallest to t for the largest value, allocating mean ranks to tied observations;

2. determine the value of R_{ij} (the sum of the ranks) for each of the k treatment combinations;

3. set up an interaction table for the R_{ij}'s and obtain the total of the ranks $R_{i.}$ & $R_{.j}$ corresponding to the two components A and B;

4. calculate the values of the test statistics T_{tot}, T_A, T_B and T_{AB} with the aid of the formulae (9.2), (9.4), (9.5) and (9.6);

5. if applicable, components with one degree of freedom may now be calculated with the aid of equation (9.12);

6. if a large proportion of the observations within a block are tied (i.e. if the lengths of the ties are substantial), compute CT defined in (9.8) and divide the test statistics calculated in 4. and 5. above by this term;

7. the significance of any component is assessed by making use of the X^2-tables. It should be noted that for a given number of treatment combinations (k), an estimate of the number of replications needed for the X^2-approximation to be applicable is obtainable from equation (9.3), (cf. also §4.10). No exact probabilities are available for small samples.

9.5 Evaluation of the new distribution-free methods

Some of the advantages of the new distribution-free method are as follows:

1. The calculations involved may be less than the corresponding parametric analysis;

2. the tests are independent of the distribution of the observations. (This is not strictly true in the case where ties exist among the observations.)

Some of the disadvantages are the following:

1. No exact probabilities are available in the case of small samples;

2. The behaviour of the test statistics under an alternative hypothesis is unknown. However, it seems reasonable to expect that the tests will generally tend to be conservative under an alternative hypothesis;

3. The asymptotic relative efficiency of the method given in §9.2 when compared with the F-test is unknown, but in general less than $3t/\pi(t+1)$ if the requirements of the F-tests are met in the case of shifting alternatives (cf. §5.14 of this thesis where the randomized block design is considered; and Van Elteren & Noether, 1959).

To summarize it seems as if this method of analysis does not contribute very much to the practical analysis of experiments. From the theoretical viewpoint however, the derivation of the test statistics ~~are~~ is of great importance since the formulae for a randomized block design are readily obtainable as a special case from those for a balanced incomplete block design.

CHAPTER X

RANDOMIZED BLOCK DESIGN

10.1 Introduction

For the requirements and fields of application of the randomized block design the reader is referred to any of the standard statistical text books on experimental design (e.g. Cochran & Cox, 1957).

The distribution-free test statistics, in the special case of ranks, for a two-way classification of treatments, are given in §10.2. The method of subdividing the sum of squares of the rank totals for treatments, is fully illustrated in order to show how the theory can be generalized to any number of components.

The theory for this method of analysis is given in Chapter V, Part I.

The applications of the formulae are illustrated with examples in §10.3. A short summary of the steps in the analysis is given in §10.4. A discussion of the efficiency, as well as an evaluation of the new methods, is given in §10.5 and §10.6 respectively.

10.2 Test statistics

Suppose there are two factors or components A and B each at m and p levels respectively. (There can be any number of levels and/or components). Thus there are in all k=mp treatment combinations. Suppose these treatment combinations are replicated n times and the design is a randomized block design.

Let r_{ijh} be the rank of X_{ijh} when the observations in the h-th replication or block are ranked, independently, giving rank 1 to the smallest and k to the largest observation in the block, i.e. $1 \leq r_{ijh} \leq k$ (cf. Chapter V

& Friedman, 1937). X_{ijh} is the observation in the h-th replication of the (i,j)-th treatment combination (i=1,.. ..,m; j=1,...,p; h=1,...,n).*

If ties are present within a replication the tied observations in each tie are given the mean rank of the ranks they would have had if they had been different (cf. Chapter V). Let now:

$$R_{ij} = \sum_h r_{ijh}; \qquad N = nmp;$$

$$R_{i.} = \sum_j R_{ij}; \qquad \bar{R}_{i.} = p^{-1}R_{i.};$$

$$R_{.j} = \sum_i R_{ij}; \qquad \bar{R}_{.j} = m^{-1}R_{.j};$$

$$R_{..} = \sum_{ij}\sum R_{ij} = \tfrac{1}{2}nk(k+1); \qquad \bar{R}_{..} = (mp)^{-1}R_{..} = \tfrac{1}{2}n(k+1).$$

Friedman (1937) defined the following test statistic (T_{tot}) to test the null hypothesis H_o that the k treatment combinations have the same effect. The test statistic rewritten in our notation becomes:

$$T_{tot} = 12[nk(k+1)]^{-1}\sum_{ij}\sum(R_{ij} - \bar{R}_{..})^2 \qquad \text{................(10.1)}$$

$$= 12[nk(k+1)]^{-1}\sum_{ij}\sum[R_{ij} - \tfrac{1}{2}n(k+1)]^2$$

$$= 12[nk(k+1)]^{-1}\sum_{ij}\sum R_{ij}^2 - 3n(k+1) \qquad \text{...............(10.2)}$$

It can be shown that under H_o, T_{tot} defined in (10.1) is, asymptotically for $n \to \infty$, distributed as X^2 with (k-1) degrees of freedom. Siegel (1956) in Table N gives exact probabilities associated with values as large as an observed T_{tot} for k=3, n=2 to 9 and k=4, n=2 to 4. For n and/or k larger than these values, the associated probability may be determined by reference to the X^2-distribution with (k-1) degrees of freedom.

The test statistics for components A, B and the

---

* For the rest of this chapter it will be assumed that the index i takes the values from 1 to m; j the values from 1 to p and h the values from 1 to n. Both m & p $\geq$ 2.

interaction AB may be obtained from equation (10.1) as follows:

$$T_{tot} = K\sum_i\sum_j(R_{ij} - \bar{R}_{..})^2$$

where

$$K = 12[nk(k+1)]^{-1} \qquad \ldots\ldots(10.3)$$

$$T_{tot} = K\sum_i\sum_j[R_{ij}-\bar{R}_{i.}-\bar{R}_{.j}+\bar{R}_{..}+(\bar{R}_{i.}-\bar{R}_{..})+(\bar{R}_{.j}-\bar{R}_{..})]^2$$

$$= K\sum_i\sum_j(R_{ij}-\bar{R}_{i.}-\bar{R}_{.j}+\bar{R}_{..})^2 + K\sum_i\sum_j(\bar{R}_{i.}-\bar{R}_{..})^2 + K\sum_i\sum_j(\bar{R}_{.j}-\bar{R}_{..})^2$$

all cross-products are zero.

$$= K\sum_i\sum_j(R_{ij}-\bar{R}_{i.}-\bar{R}_{.j}+\bar{R}_{..})^2 + Kp\sum_i(\bar{R}_{i.}-\bar{R}_{..})^2 + Km\sum_j(\bar{R}_{.j}-\bar{R}_{..})^2 \qquad \ldots\ldots(10.4)$$

$$= T_{AB} + T_A + T_B$$

From (10.4) we now obtain:

$$T_A = Kp\sum_i(\bar{R}_{i.} - \bar{R}_{..})^2$$

$$= 12[npk(k+1)]^{-1}\sum_i[R_{i.} - \tfrac{1}{2}np(k+1)]^2$$

since $\bar{R}_{..} = \frac{1}{2}n(k+1)$

$$= 12[npk(k+1)]^{-1}\sum_i R_{i.}^2 - 3n(k+1) \qquad \ldots\ldots(10.5)$$

$$T_B = Km\sum_j(\bar{R}_{.j} - \bar{R}_{..})^2$$

$$= 12[nmk(k+1)]^{-1}\sum_j[R_{.j} - \tfrac{1}{2}nm(k+1)]^2$$

$$= 12[nmk(k+1)]^{-1}\sum_j R_{.j}^2 - 3n(k+1) \qquad \ldots\ldots(10.6)$$

$$T_{AB} = K\sum_i\sum_j(R_{ij} - \bar{R}_{i.} - \bar{R}_{.j} + \bar{R}_{..})^2$$

$$= T_{tot} - T_A - T_B \qquad \ldots\ldots(10.7)$$

The formulae in equations (10.2) and (10.5) to (10.7) correspond to those in (5.66). The calculating formulae for the case where a function of the ranks is considered, are given in §5.5, p. 120.

The test statistics T_A, T_B and T_{AB} are under H_o, asymptotically for $n \to \infty$, mutually independently distributed as X^2 with (m-1), (p-1) and (m-1)(p-1) degrees of freedom respectively (cf. Chapters IV & V). Under an alternative hypothesis H_a these components are not asymptotically independantly distributed. By means of sampling studies it was found (cf. §5.10) that the test for, say, component A becomes more conservative when the effects of the levels of component B are increased.

To obtain the probabilities of obtaining a value as large as or larger than the calculated value of a test statistic (say T_A) in the case of small samples, one of the following alternatives should be followed depending on the value of k, m and n.

1. Exact probabilities for T_A are given in Table 10.1 for the following values of k, m and n:

a) k=4, m=2, n=2 to 4; and

b) k=6, m=2 and 3, n=2.

For T_B the values of m and p should be interchanged.

2. For all other values of k, m and n the X^2-distribution may be used with (m-1) degrees of freedom. If, however, any of the components consists of two levels only, a continuity correction of a $\frac{1}{2}$ should be applied (i.e. a $\frac{1}{2}$ subtracted from the highest rank total and a $\frac{1}{2}$ added to the smallest rank total for that component) before the value of the corresponding test statistic is calculated.

In the case of ties the following correction term is used (cf. Kendall, 1955, p. 100):

$$CT = 1 - \sum_h \tau_h / nk(k^2-1) \qquad \text{.............................(10.8)}$$

where

$$\tau_h = \sum_{(t_h)} (t_h^3 - t_h);$$

t_h = length of any tie within replicate h;

$\sum_{(t_h)}$ = summation over all different values in the h-th replication.

Table 10.1 - Exceeding probabilities of obtaining a value as large as or larger than the calculated value T_A

	T_A	Exact Prob.		T_A	Exact Prob.
	4.8	0.056		9.143	0.0007
m=2	2.7	0.167		8.143	0.0037
k=4	1.2	0.445		7.429	0.0111
n=2	0.3	0.779		7.000	0.0230
	0.0	1.000		6.857	0.0289
	7.2	0.009		6.143	0.0407
	5.0	0.037		5.571	0.0556
m=2	3.2	0.121		5.286	0.0733
k=4	1.8	0.284		5.143	0.0859
n=3	0.8	0.518	m=3	4.429	0.1156
	0.2	0.825	k=6	4.000	0.1556
	0.0	1.000	n=2	3.857	0.1748
	9.60	0.002		3.571	0.1956
	7.35	0.008		3.000	0.2519
	5.40	0.029		2.714	0.3141
m=2	3.72	0.079		2.286	0.3585
k=4	2.40	0.179		1.857	0.4563
n=4	1.35	0.340		1.714	0.5170
	0.60	0.571		1.286	0.5778
	0.15	0.849		1.000	0.7111
	0.00	1.000		0.571	0.7985
	7.714	0.005		0.429	0.8844
	6.095	0.015		0.143	0.9793
	4.667	0.040		0.000	1.0000
m=2	3.429	0.090			
k=6	2.381	0.170			
n=2	1.524	0.290			
	0.857	0.455			
	0.381	0.655			
	0.095	0.880			
	0.000	1.000			

The only necessary alteration is that the different test statistics have to be divided by this correction term (CT) to give T_{tot}(corr), T_A(corr), T_B(corr) and T_{AB}(corr).

It should be noted that if no ties are present all $t_h=1$ and thus all $\mathbf{C}_h=0$. The correction term then becomes 1.

In the case of a factorial design there is a second method of analysis called the single degree of freedom approach (cf. Anderson & Bancroft, 1952, p. 272 and Goulden, 1952, p. 38).

For k=mp treatments there exist (k-1) orthogonal components corresponding to the (k-1) orthogonal sets of constants $g_{ij}^{(r)}$ defined by:

$$\left.\begin{array}{lll} \text{i)} & \sum_{ij}\sum g_{ij}^{(r)} = 0 & r=1,\ldots,k-1 \\ \text{ii)} & \sum_{ij}\sum g_{ij}^{(r)} g_{ij}^{(r')} = 0 & r \neq r' \end{array}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10.9)$$

Let for this method:

$$L_r = \sum_{ij}\sum g_{ij}^{(r)} R_{ij} \qquad r=1,\ldots,k-1 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(10.10)$$

then

$$\mathrm{var}(L_r) = (12)^{-1} nk(k+1) \sum_{ij}\sum [g_{ij}^{(r)}]^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(10.11)$$

(cf. (5.66) and Chapters IV & V).

$$T_r = L_r^2/\mathrm{var}(L_r), \qquad r=1,\ldots,k-1 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(10.12)$$

will under H_0, be asymptotically for $n \rightarrow \infty$, independently distributed as X^2 with one degree of freedom each.

By means of (10.10), (10.11) and (10.12) the linear and higher order effects of a component, say T_A, in a factorial design, can for instance be tested if desired provided the X^2-approximation of T_A is considered applicable.

In the case of tied observations, T_r (10.12) is to be divided by CT (10.8) to yield T_r(corr).

10.3 Application

Example 10.1 - To illustrate the application of the formulae a hypothetical factorial design with three cultivars of maize C_1, C_2 and C_3 and four levels of nitrogen fertilizer N_1, N_2, N_3 and N_4 is considered. The design used was a randomized block design, with six replications.

The yields in bags/morgen for the different plots as well as the corresponding ranks are presented in Table 10.2.

Within each replication the observations are ranked from the smallest to the largest value and the totals of the ranks for each treatment combination are obtained. These totals are set out in Table 10.2.

Table 10.2a is the interaction table of cultivars of maize x fertilizer. Cultivars of maize are now taken as component A and nitrogen fertilizer as component B in

Table 10.1 - Yields and ranks (r) of cultivars of maize (C) x fertilizer (N) experiment

Replicates	Treatment combinations											
	C_1N_1	r_{11h}	C_1N_2	r_{12h}	C_1N_3	r_{13h}	C_1N_4	r_{14h}	C_2N_1	r_{21h}	C_2N_2	r_{22h}
	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r
1	15.2	4	17.4	8	14.8	3	15.8	5	13.3	1	17.1	6
2	13.1	2	16.0	9	12.6	1	14.0	4	13.2	3	15.9	8
3	16.4	3	16.0	2	16.9	4	17.1	5	17.2	6	20.4	11
4	14.9	4	15.9	7	18.3	11	14.8	3	13.2	1	15.0	5
5	14.6	2	15.2	3	17.2	9	13.3	1	15.3	4	16.0	6
6	15.4	6	17.8	9	17.4	8	18.6	10	14.8	5	12.3	1
	R_{11}	21	R_{12}	38	R_{13}	36	R_{14}	28	R_{21}	20	R_{22}	37

Replicates	Treatment combinations											
	C_2N_3	r_{23h}	C_2N_4	r_{24h}	C_3N_1	r_{31h}	C_3N_2	r_{32h}	C_3N_3	r_{33h}	C_3N_4	r_{34h}
	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r
1	18.6	10	17.2	7	14.4	2	17.5	9	19.7	12	19.3	11
2	15.1	6	18.0	12	16.3	11	16.2	10	15.5	7	14.6	5
3	20.3	10	19.0	7	15.3	1	19.9	9	21.6	12	19.5	8
4	17.4	9	18.0	10	14.2	2	15.8	6	19.2	12	16.3	8
5	17.0	8	17.7	10	15.4	5	16.5	7	20.5	12	18.2	11
6	16.5	7	18.8	11	13.7	2	14.1	3	14.5	4	20.3	12
	R_{23}	50	R_{24}	57	R_{31}	23	R_{32}	44	R_{33}	59	R_{34}	55

Table 10.2a - Interaction cultivars (C) x fertilizer (N)

	N_1	N_2	N_3	N_4	$R_{i.}$
C_1	21	38	36	28	123
C_2	20	37	50	57	164
C_3	23	44	59	55	181
$R_{.j}$	64	119	145	140	468

formulae (10.5) to (10.7), thus m=3, p=4, k=12 and n=6.

To calculate the test statistics for the different components we proceed as follows:

From equation (10.2) we have:

$$T_{tot} = 12[nk(k+1)]^{-1} \sum_{i=1}^{3} \sum_{j=1}^{4} R_{ij}^2 - 3n(k+1)$$
$$= 12(6 \times 12 \times 13)^{-1}(21^2+\ldots\ldots+55^2) - 3 \times 6 \times 13$$
$$= 28.23$$

From equations (10.5) to (10.7) we obtain:

$$T_C = 12[npk(k+1)]^{-1} \sum_i R_{i.}^2 - 3n(k+1)$$
$$= 12(6 \times 4 \times 12 \times 13)^{-1}(123^2 + 164^2 + 181^2) - 3 \times 6 \times 13$$
$$= 5.70$$

$$T_N = 12[nmk(k+1)]^{-1} \sum_j R_{.j}^2 - 3n(k+1)$$
$$= 12(6 \times 3 \times 12 \times 13)^{-1}(64^2+119^2+145^2+140^2) - 3 \times 6 \times 13$$
$$= 17.63$$

$$T_{CN} = T_{tot} - T_C - T_N$$
$$= 28.23 - 5.70 - 17.63$$
$$= 4.90$$

There are no ties among the observations within any replication and thus it is not necessary to calculate CT in equation (10.8).

If the calculated value of any of the test statistics is larger than the value needed to be significant by means of the X^2 distribution at the 5% level with one degree of freedom (i.e. 3.841), we may proceed to subdivide the test statistic into components with one degree of freedom, if it makes any sense, i.e. if there is linear contrast(s) among the levels of the component, orthogonal to all other contrasts, we would like to make (e.g. calculating the linear and higher order effects of a fertilizer being applied at more than two levels). The reason for this is that if the test statistic for one of the components with one degree of freedom is nearly as large as the test statistic for the original component, the test for this component with one degree of freedom will be significant.

In the above example cultivars cannot be subdivided into single degree of freedom components, and arethus ignored. If, however, the fertilizer were applied on a factorial basis, where the different levels of the fertilizer are equidistan from each other, the linear, quadratic and cubic components can be calculated with the aid of orthogonal polynomials (cf. the tables of Fisher & Yates, 1938). The constants obtained from the tables for the linear component which correspond to the four fertilizer levels are (-3, -1, 1, 3). Since these constants are the same for all three cultivars of corn we denote them by $g_{.j}$ (j=1,2,3,4).

With the aid of equations (10.10) to (10.12) the linear effect of nitrogen fertilizer can be calculated.

$$\sum_i\sum_j g_{ij}^2 = m\sum_j g_{.j}^2 = 3[(-3)^2 + (-1)^2 + 1^2 + 3^2] = 60$$

$$L_{N_1} = \sum_i\sum_j g_{ij}R_{ij} = \sum_j g_{.j}R_{.j}$$

$$= [(-3)64 + (-1)119 + 145 + 3(140)]$$

$$= 254$$

where L_{N_1} indicates the linear effect of nitrogen fertilizer.

$$\mathrm{var}(L_{N_1}) = (12)^{-1}nk(k+1)\sum_i\sum_j g_{ij}^2$$

$$= (12)^{-1} \times 6 \times 12 \times 13 \times 60$$

$$= 4680$$

$$T_{N_1} = L_{N_1}^2/\mathrm{var}(L_{N_1})$$

$$= (254)^2/4680 = 13.79$$

The test statistic for the rest of the fertilizer effect is obtained by subtracting this last value from T_N. The results obtained above can be summarized as in Table 10.3. (Note that, according to the rules given in the previous paragraph, the X^2-approximation is applicable in this case.)

The observations were also analysed according to the parametric method of analysis and the results obtained are tabulated in Table 10.4.

Table 10.3 - Distribution-free analysis of variance

Source of variation	DF	T
Cultivars	2	5.70
Fertilizer	3	17.63**
Linear	1	13.78**
Rest	2	3.85
Interaction	6	4.90
Total	11	28.23

** significant at the 1% level

Table 10.4 - Parametric analysis of variance

Source of variation	DF	F
Replicates	5	
Cultivars	2	3.74*
Fertilizer	3	10.35**
Linear	1	25.40**
Rest	2	2.82
Interaction	6	1.53
Error	55	
Total	71	

* significant at the 5% level
** significant at the 1% level

Both the parametric and nonparametric methods showed the linear effect of nitrogen fertiliser significant at the 1% level. The parametric method further showed cultivars significant at the 5% level whilst the nonparametric method did not declare it significant.

Example 10.2 - To illustrate the method of analysis for the special case where use is made of the transformation of the ranks to "normal values" suggested by Van der Waerden (1957) (see §5.5.3), the same observations as in Example 10.1 is used. The applicability of the X^2-approximation in this case, can be based on the same argument as given in the last sentence of §3.8.

Divide the ranks in Table 10.2 by (k+1)=13 and these values (δ_{ijh}) are then transformed with the aid of the values in Table 2 of Van der Waerden (1957) p. 334 [$Q(\delta_{ijh})$].

The results obtained are given in Table 10.5 together with the totals $z_{ij.}$ for each treatment combination.

Table 10.5 - Transformed values (T_r) of the ranks in Table 10.2

Replicates	Treatment combinations					
	$Q(\delta_{11h})$	$Q(\delta_{12h})$	$Q(\delta_{13h})$	$Q(\delta_{14h})$	$Q(\delta_{21h})$	$Q(\delta_{22h})$
	T_r	T_r	T_r	T_r	T_r	T_r
1	-0.50	0.29	-0.74	-0.29	-1.43	-0.10
2	-1.02	0.50	-1.43	-0.50	-0.74	0.29
3	-0.74	-1.02	-0.50	-0.29	-0.10	1.02
4	-0.50	0.10	1.02	-0.74	-1.43	-0.29
5	-1.02	-0.74	0.50	-1.43	-0.50	-0.10
6	-0.10	0.50	0.29	0.74	-0.29	-1.43
$z_{ij.}$	-3.88	-0.37	-0.86	-2.51	-4.49	-0.61

Replicates	Treatment combinations					
	$Q(\delta_{23h})$	$Q(\delta_{24h})$	$Q(\delta_{31h})$	$Q(\delta_{32h})$	$Q(\delta_{33h})$	$Q(\delta_{34h})$
	T_r	T_r	T_r	T_r	T_r	T_r
1	0.74	0.10	-1.02	0.50	1.43	1.02
2	-0.10	1.43	1.02	0.74	0.10	-0.29
3	0.74	0.10	-1.43	0.50	1.43	0.29
4	0.50	0.74	-1.02	-0.10	1.43	0.29
5	0.29	0.74	-0.29	0.10	1.43	1.02
6	0.10	1.02	-1.02	-0.74	-0.50	1.43
$z_{ij.}$	2.27	4.13	-3.76	1.00	5.32	3.76

Table 10.6 is the interaction table for the transformed values, as well as the totals for cultivars of maize $z_{i..}$ (i=1,2,3) and nitrogen fertilizer $z_{.j.}$ (j=1,2,3,4).

Table 10.6 - Interaction (transformed ranks) of cultivars maize x fertilizer

	N_1	N_2	N_3	N_4	$z_{i..}$
C_1	-3.88	-0.37	-0.86	-2.51	-7.62
C_2	-4.49	-0.61	2.27	4.13	1.30
C_3	-3.76	1.00	5.32	3.76	6.32
$z_{.j.}$	-12.13	0.02	6.73	5.38	0

From equations (4.5) and (4.6) we now calculate:

$$\sigma^2 = (nk)^{-1}\sum_{ijh}\sum\sum z_{ijh}^2 \qquad \text{since } \bar{z}_{...} = 0$$

$$= (6 \text{ x } 12)^{-1}[(-0.50)^2+(-1.02)^2+\ldots+(1.02)^2+(1.43)^2]$$

$$= 0.663$$

Van der Waerden (1957) tabulated in Table 12 the values for σ^2 (if no ties are present). The value tabulated by him for k=12 values is 0.661, which is nearly the same as the one above. The difference is probably due to rounding errors.

With the aid of the formulae in equations (4.86) to (4.89) (cf. §5.5.3) we now proceed to calculate the values of the different test statistics.

$$T_{tot} = (k-1)(nk\sigma^2)^{-1}\sum_{ij}\sum z_{ij.}^2$$

$$= 11(6 \text{ x } 12 \text{ x } 0.663)^{-1}[(-3.88)^2 + \ldots + 3.76^2]$$

$$= 28.25$$

$$T_N = (k-1)(nmk\sigma^2)^{-1}\sum_{j=1}^{4} z_{.j.}^2$$

$$= 11(6 \text{ x } 3 \text{ x } 12 \text{ x } 0.663)^{-1}[(-12.13)^2+0.02^2+6.73^2+5.38^2]$$

$$= 17.78$$

$$T_C = (k-1)(npk\sigma^2)^{-1}\sum_{i=1}^{3} z_{i..}^2$$

$$= 11(6 \text{ x } 4 \text{ x } 12 \text{ x } 0.663)^{-1}[(-7.62)^2+1.30^2+6.32^2]$$

$$= 5.74$$

$$T_{NC} = T_{tot} - T_N - T_C$$
$$= 28.25 - 17.78 - 5.74$$
$$= 4.73$$

For similar reasons as in Example 10.1 we now proceed to calculate the test statisticscorresponding to the linear effect of nitrogen fertilizer with the aid of equations (4.90) to (4.92).

$$L_{N_1} = n^{-\frac{1}{2}}\sum_{i}\sum_{j} g_{.j} z_{ij.} = n^{-\frac{1}{2}}\sum_{j} g_{.j} z_{.j.}$$
$$= 6^{-\frac{1}{2}}[(-3)(-12.13) + (-1)(0.02) + 6.73 + 3(5.38)]$$
$$= 6^{-\frac{1}{2}} \times 59.24$$
$$\mathrm{var}(L_{N_1}) = (k-1)^{-1} k\sigma^2 \sum_{i}\sum_{j} g^2_{.j}$$
$$= (11)^{-1} \times 12 \times 0.663 \times 3[(-3)^2 + (-1)^2 + 1^2 + 3^2]$$
$$= (11)^{-1} \times 477.36$$

Thus

$$T_{N_1} = L^2_{N_1}/\mathrm{var}(L_{N_1}) = (11 \times 59.24^2)/6 \times 477.36 = 13.48$$

The test statistic for the rest of the fertilizer effect is obtained by subtracting this last value from T_N. The results obtained above can be presented as in Table 10.7.

Table 10.7 - Distribution-free analysis of variance

Source of variation		DF		T	
Cultivars		2		5.745	
Fertilizer	Linear	3	1	17.78**	13.47**
	Rest		2		4.30
Interaction		6		4.73	
Total		11		28.25	

** significant at the 1% level

With this analysis the same conclusions are drawn as in Example 10.1 where ranks themselves were considered (i.e. the linear effect of nitrogen fertilizer significant at the 1% level and cultivars not significant).

Example 10.3 - This example illustrates the analysis of a 3^3 factorial design with 4 replicates. The special case of ranks will be considered (i.e. an extension of the formulae in §10.2.

To illustrate the application of the formulae for more than two components the example in Cochran & Cox (1957) p. 196 was used. In this example the effects of three levels of nitrogen, three of phosphorus and three of potash on the germination of lettuce seedlings, conducted at the seed laboratary, Iowa State College, were tested. The seed was thoroughly mixed, and divided into 108 samples of about 60 seeds each and germinated. At the end of 5 to 7 days the normal seedlings were counted. The data in Table 10.8 show the number of normal lettuce plants as well as their corresponding ranks.

The computations proceed as follows:

1. Rank the observations within each replication from the smallest to the largest. For tied observations give the mean rank of the ranks they would have had if they had been different. In this example k=27=number of treatment combinations. The indexes i,j,h can take on three values 1, 2, 3 each and they sum over the levels of the components N, P and K respectively;
2. calculate $R_{ijh} = \sum_{s=1}^{4} r_{ijhs}$ as in Table 10.8;
3. obtain the three interaction tables for N, P and K;
4. calculate the different test statistics as follows:

$$T_N = 12[n \times 3^2 \times k(k+1)]^{-1} \sum_{i=1}^{3} R_{i..}^2 - 3n(k+1)$$

$$= 12(4 \times 9 \times 27 \times 28)^{-1}(605^2 + 507^2 + 400^2) - 3 \times 4 \times 28$$

$$= 9.27$$

It should be noted that $n \times 3^2 = 36$ is the number of plots (ranks) added together to obtain the total $R_{i..}$ (i=1,2,3) which is squared. This can be applied right

Table 10.8 - Number and rank (r) of normal lettuce plants emerging in a 3^3 experiment (based on data from Cochran & Cox, 1957)

Treatments	Rep. 1		Rep. 2		Rep. 3		Rep. 4		Total ranks
n p k	No.	r	No.	r	No.	r	No.	r	R_{ijh}
0 1 2	11	2	37	23.5	30	22	41	22	69.5
1 2 2	11	2	22	11	19	8.5	29	9.5	31.0
2 2 0	13	5	34	20.5	19	8.5	26	5	39.0
2 0 2	12	4	27	14	22	13.5	35	16	47.5
1 0 1	11	2	14	6	26	17	31	12.5	37.5
0 2 1	30	11	13	5	27	18.5	29	9.5	44.0
0 0 0	41	17.5	40	27	38	25	52	27	96.5
1 1 0	21	7.5	38	25	27	18.5	32	14	65.0
2 1 1	21	7.5	12	3	25	15.5	28	7.5	33.5
2 0 1	42	19	34	20.5	11	1	42	23	63.5
1 2 1	20	6	39	26	18	5.5	31	12.5	50.0
2 1 0	24	9	37	23.5	17	4	24	3	39.5
0 1 1	38	14	35	22	19	8.5	38	19.5	64.0
1 1 2	39	15	19	9	22	13.5	37	17.5	55.0
1 0 0	61	27	32	18	42	27	39	21	93.0
0 0 2	40	16	16	7.5	29	21	33	15	59.5
2 2 2	46	21.5	12	3	20	11	15	1.5	37.0
0 2 0	44	20	21	10	19	8.5	47	25	63.5
1 1 1	46	21.5	12	3	37	24	28	7.5	56.0
1 0 2	48	23	31	16.5	28	20	50	26	85.5
2 2 1	58	26	33	19	21	12	15	1.5	58.5
0 0 1	53	24	30	15	40	26	37	17.5	82.5
0 1 0	54	25	26	12.5	36	23	44	24	84.5
2 1 2	37	13	26	12.5	13	2	27	6	33.5
0 2 2	41	17.5	31	16.5	14	3	25	4	41.0
1 2 0	25	10	16	7.5	18	5.5	30	11	34.0
2 0 0	32	12	6	1	25	15.5	38	19.5	48.0

Table 10.9 - N x P interaction table (ranks)

	P_o	P_1	P_2	$R_{i..}$
N_o	238.5	218.0	148.5	605.0
N_1	216.0	176.0	115.0	507.0
N_2	159.0	106.5	134.5	400.0
$R_{.j.}$	613.5	500.5	398.0	

Table 10.10 - N x K interaction table (ranks)

	K_o	K_1	K_2
N_o	244.5	190.5	170.0
N_1	192.0	143.5	171.5
N_2	126.5	155.5	118.0
$R_{..l}$	563.0	489.5	459.5

Table 10.11 - P x K interaction table (ranks)

	K_o	K_1	K_2
P_o	237.5	183.5	192.5
P_1	189.0	153.5	158.0
P_2	136.5	152.5	109.0

through as a general rule namely (except for the constant $12[k(k+1)]^{-1}$) the rank total which is squared should be divided by the number of observations (ranks) from which it is obtained.

$$T_P = 12[36 \times k(k+1)]^{-1} \sum_{j=1}^{3} R^2_{.j.} - 3n(k+1)$$

$$= 12(36 \times 27 \times 28)^{-1}(613.5^2+500.5^2+398^2) - 3 \times 4 \times 28$$

$$= 10.25$$

From Table 10.9

$$T_{NP} = 12[12 \times k(k+1)]^{-1} \sum_{i=1}^{3} \sum_{j=1}^{3} R^2_{ij.} - 3n(k+1) - T_N - T_P$$

$$= 12(12 \times 27 \times 28)^{-1}(238.5^2+\ldots+134.5^2) - 3 \times 4 \times 28 - 9.27 - 10.25$$

$$= 4.31$$

$$T_K = 12[36k(k+1)]^{-1} \sum_{h=1}^{3} R^2_{..h} - 3n(k+1)$$

$$= 12(36 \times 27 \times 28)^{-1}(563^2+489.5^2+459.5^2) - 3 \times 4 \times 28$$

$$= 2.50$$

From Table 10.10

$$T_{NK} = 12(12 \times 27 \times 28)^{-1}(244.5^2+\ldots+118^2) - 3 \times 4 \times 28 - 9.27 - 2.50$$

$$= 4.01$$

From Table 10.11

$$T_{PK} = 12(12 \times 27 \times 28)^{-1}(237.5^2+\ldots+152.5^2) - 3 \times 4 \times 28 - 10.25 - 2.50$$

$$= 1.98$$

The value of T_{NPK} **was** not calculated since it is partly confounded with blocks in this example.

Calculate the correction term as defined in equation (10.8).

In replication 1 there are three ties of two values and one tie of three values.

$$C_1 = [3(2^3 - 2) + (3^3 - 3)] = 42$$

In replication 2 there are five ties of two and one of three.

$$C_2 = [5(2^3 - 2) + (3^3 - 3)] = 54$$

In replication 3 there are four ties of two values and one tie of four values.

$$C_3 = [4(2^3 - 2) + (4^3 - 4)] = 84$$

In replication 4 there are six ties of two values.

$$C_4 = 6(2^3 - 2) = 36$$

Thus $\sum_{s=1}^{4} C_s = 216$

Substituting this value in (10.8) yields:

$$CT = [1 - 216/4 \times 27(27^2-1)] = 0.9973$$

All the test statistics calculated above should be divided by this correction term. The results obtained are summarized in Table 10.12.

It should be noted that the number of observations is large enough for the X^2-approximation to be applicable.

Table 10.12 - Distribution-free analysis of normal lettuce plants in a 3^3 experiment

Source of variation	DF	T	T(corr)
N	2	9.27	9.30**
P	2	10.25	10.28**
K	2	2.50	2.51
NP	4	4.31	4.32
NK	4	4.01	4.02
PK	4	1.98	1.99

** significant at the 1% level

Comparing these results with the results for the parametric analysis of variance given in Cochran & Cox (1957) p. 199, it is seen that both these methods showed N and highly significant.

Next the linear and quadratic components of the fertilizers (except K since $T_K < 3.84$) are calculated with the aid of orthogonal polinomials given in the tables of Fisher & Yates (1938). From the tables the values for the two components are:

Linear -1 0 1

Quadratic 1 -2 1

where the three values of the two components given above correspond to the 1, 2 and 3 levels of the different fertilizers respectively and since they are the same for the two remaining components we indicate them by $g_{i..}$, $g_{.j.}$ and $g_{..h}$ for N, P and K respectively.

or the linear component of the fertilizers we calculate:

$$\sum_{ijh}\sum\sum g^2_{i..} = \sum_{ijh}\sum\sum g^2_{.j.} = \sum_{ijh}\sum\sum g^2_{..h} = 9[(-1)^2 + 1^2] = 18$$

For N_1 we get from (10.10):

$$L_{N_1} = \sum_{ijh}\sum\sum g_{i..}R_{ijh} = \sum_{i} g_{i..}R_{i..} = -605 + 400 = -205$$

From (10.11) and (10.12) we have:

$$T_{N_1} = 12(-205)^2/(4 \times 27 \times 28 \times 18) = 9.26$$

where N_1 indicates the linear effect of nitrogen fertilizer. Similarily

$$T_{P_1} = 10.23$$

The value of T_{K_1} is not calculated since T_K is already smaller than 3.841, the value required for T_{K_1} to be significant at the 5% level.

The quadratic components may be calculated in the same way as the linear components or they can be obtained by subtraction. The interactions N_1P_1, N_1K_1 etc. can be calculated as follows:

$$N_1P_1 = (N_3 - N_1)(P_3 - P_1) = N_3P_3 - N_3P_1 - N_1P_3 + N_1P_1$$

$$= g_{33.} - g_{31.} - g_{13.} + g_{11.}$$

$$N_1K_1 = g_{3.3} - g_{3.1} - g_{1.3} + g_{1.1}$$

Thus

$$\sum_{ijh}\sum\sum g^2_{ij.} = \sum_{ijh}\sum\sum g^2_{i.h} = 3[1^2 + (-1) + (-1)^2 + 1^2] = 12$$

For N_1P_1 we get from Table 10.9

$$\sum_{ijh}\sum\sum g_{ij.}R_{ijh} = \sum_{ij}\sum g_{ij.}R_{ij.} = 134.5 - 159.0 - 148.5 + 238.5$$
$$= 65.5$$

$$T_{N_1P_1} = 12(65.5)^2/(4 \times 27 \times 28 \times 12) = 1.42$$

Similarily

$$T_{P_1K_1} = 1.44$$

The results obtained are summarized in Table 10.13.

Comparing these results with those obtained in Cochran & Cox (1957) p. 199 using parametric procedures, it is seen that both methods **showed** N_1 and P_1 significant at the 1% level. In addition the parametric method fur-

Table 10.13 - Distribution-free analysis of individual components

Source of variation	DF	T	T(corr)
N_1	1	9.26	9.29**
N_q	1	0.01	0.01
P_1	1	10.23	10.26**
P_q	1	0.01	0.01
N_1P_1	1	1.42	1.42
P_1K_1	1	1.44	1.45

** significant at the 1% level

ther showed K_1 and N_1K_1 significant at the 5% level.

10.4 Summary of the procedure

These are the steps in the distribution-free analysis of a two-way classification of treatments in a randomized block design by means of ranks:

1. Rank the observations within each replication from 1 for the smallest value to k for the largest value, allocating mean ranks to tied observations;

2. determine the value of R_{ij} (the sum of the ranks) for each of the k treatment combinations;

3. set up an interaction table of the R_{ij}'s and obtain the total of the ranks $R_{i.}$ and $R_{.j}$ corresponding to the two components A and B;

4. calculate the test statistics T_{tot}, T_A, T_B and T_{AB} with the aid of formulae (10.2), (10.5), (10.6) and (10.7);

5. if applicable, components with one degree of freedom may now be calculated with the aid of equation (10.12);

6. if a large proportion of the observations within a replicate are tied (i.e. if the lengths of the ties are substantial) compute CT defined in (10.8) and divide the test statistics calculated in 4. and 5. above by this term;

7. the method of assessing the significance of any component, say A, at m levels depends on the value of m, k and n.

a. Exact probabilities are tabulated in Table 10.1 for k=4, m=2, n=2 to 4 and k=6, m=2 and 3, n=2.

b. For all other values of k, m and n the X^2-approximation should be used at (m-1) degrees of freedom. It should be noted that a continuity correction of a $\frac{1}{2}$ should be applied to the rank totals, if the component contains two levels only, before the value of the corresponding test statistic is calculated. If the value of this component must be subtracted to obtain the interaction (cf. equation (10.7)) we first calculate the test statistic for the component at two levels without a continuity correction, use this value to calculate the interaction and then calculate the test statistic again with the continuity correction and use this value to test the specific component.

10.5 Efficiency of the distribution-free tests given in §10.2

In the special case of ranks the highest asymptotic relative efficiency (a.r.e.) obtainable for any of the test statistics say T_A, when compared with the F-test, for alternatives of shift, is $3k/\pi(k+1)$, in the case of normal distributed variates (cf. Lemmer, 1964, p. 168 and §5.14). This is true if the effects of the levels of the remaining components are zero.

From sampling studies (cf. §5.10) it was found that the value of say T_A was depressed when the effects of the levels of the remaining components were not zero. As a consequence of this the a.r.e. of T_A (compared ~~to~~ with the F-test when the requirements of the F-test are met) will generally be smaller than $3k/\pi(k+1)$.

From theoretical consideration (cf. Lemmer, 1964) it can be expected that the Van der Waerden type of tests (cf. Example 10.2) will generally be more efficient than the tests discussed in this chapter (based on ranks), although they are more difficult to apply.

10.6 Discussion and evaluation

Some of the advantages of the distribution-free test statistics presented in this chapter, when compared with the corresponding F-test for a randomized block design, are the following:

1. The distribution-free tests are in contrast to the F-test independent of assumptions such as additivity (i.e. no block x treatment interaction) and normality;
2. the calculations involved are probably a little easier than those for the F-test in the special case of ranks; and
3. in the case of a split plot design there are no problems in obtaining the correct error term since the distribution-free tests are independent of an estimate of the error variance.

The most important disadvantage of the distribution-free test statistics is the fact that the a.r.e. of the test statistics, when compared with the F-test, if the requirements of the latter are met, is unknown but less than $3k/\pi(k+1)$ for shift alternatives in the special case of ranks.

Thus we may conclude by saying that the F-test should always be used when its requirements are fully met. In all other cases preference should be given to the distribution-free tests.

CHAPTER XI

LATIN SQUARE DESIGNS

11.1 Introduction

In this chapter three distribution-free methods of analysis are presented for latin square designs. The first two are for a latin square design replicated once only and the third for a latin square design replicated several times. All three of these methods of analysis are given for the case where the treatments form a one-way classification. If the treatments form a multi-way classification the test statistics for the different components can be obtained in a similar manner ~~as~~ to that in equation (8.4).

For the randomization, requirements and analysis of variance in the parametric case of latin square designs the reader is referred to Cochran & Cox (1957).

The formulae for the three methods of analysis are given in §11.2 for the special case of ranks. The formulae for functions of ranks are obtainable from Chapter VI (p. 144).

An illustration of the application of the formulae is given in §11.3. A summary of the procedure is given in §11.4 and an evaluation of the three methods of analysis in §11.5.

11.2 Methods of analysis for a latin square design

Three methods of analysis are presented for the analysis of a latin square design. The formulae are given for the special case of ranks only. A full discussion of these methods as well as the formulae for functions of ranks are obtainable in §6.5 (p. 153).

11.2.1 Method I - The latin square treated as a completely random design

All the observations in the latin square are ranked giving rank 1 to the smallest observation and rank $N=k^2$ (k = number of rows, columns or treatments) to the largest observation. If tied observations are present the tied observations in each tie are given the mean rank of the ranks they would have had if they had been different.

Let $Y_{ij(h)}$ be the observation in the i-th row, j-th column where h corresponds to the treatment applied to the plot (i,j) and let $r_{ij(h)}$ be the rank allocated to $Y_{ij(h)}$ in the joint ranking. Let further R_h be the total of the ranks corresponding to the h-th treatment.

The test statistic to test the null hypothesis that the effect of the treatments is the same is obtainable from (6.7) (p. 154) as a special case and is exactly the same as the one defined by Kruskal & Wallis (1952)(cf. (8.2) with n=k). In the notation of this chapter it becomes:

$$T = 12k[N(N+1)]^{-1} \sum_{h=1}^{k} (\bar{R}_h - \bar{r}_{..})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (11.1)$$

$$= 12[N(N+1)k]^{-1} \sum_{h=1}^{k} R_h^2 - 3(N+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (11.2)$$

where

$\bar{R}_h = k^{-1}R_h$

$\bar{r}_{..} = k^{-2}\sum_i\sum_j r_{ij(h)}$

$N = k^2$, and k is the number of treatments, rows or columns.

If $k=3$ the exact probabilities tabulated for the Kruskal-Wallis (1952) test statistic may be used to determine whether the value of the calculated test statistic (T) is significant. For $k>3$ the tables of the X^2-distribution may be used at $(k-1)$ degrees of freedom.

If ties are present the correction term given in equation (8.9) must be calculated and T (equation (11.2)) must be divided by CT to yield T(corr).

The application of the method of analysis discussed above will be illustrated in Example 11.1.

11.2.2 Method II - The latin square treated as a randomized block design

The rows (or columns) are considered as blocks and ranks ($r_{ij(h)}$) are allocated within rows (or columns) to the observations ($Y_{ij(h)}$) as described in Chapter X. To tied observations the mean rank is allocated as usual. If it is known beforehand that say the effect of rows is larger than that of columns, ranks should be allocated within rows. If however no prior knowledge is available, either can be selected at random as blocks.

Let R_h again be the total of the ranks corresponding to the h-th treatment.

The test statistic to test the null hypothesis that the effect of all treatments is the same is obtainable from (6.10) as a special case and is exactly the same as the one defined by Friedman (1937) (cf. (10.2) with n=k). In the notation of this chapter it becomes:

$$T = 12(k+1)^{-1} \sum_{h=1}^{k} (\bar{R}_h - \bar{r}_{..})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (11.3)$$

$$= 12[k^2(k+1)]^{-1} \sum_{h=1}^{k} R_h^2 - 3k(k+1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (11.4)$$

where all the symbols have the same meaning as in §11.2.1.

For k=3 and 4, the exact probabilities calculated for Friedman's (1937) (cf. Siegel, 1956, Table N) test statistic must be used. For $k > 4$ the X^2-approximation at (k-1) degrees of freedom may be used.

If ties are present among the observations within a block the CT defined in (10.8) must be calculated and T (equation (11.4)) must be divided by this value to yield T(corr).

The application of the analysis according to Method II is illustrated in Examples 11.2.

11.2.3 Method III - The latin square design replicated several times

Suppose we have a k x k latin square design replicated n times. Each latin square **is** now considered as a block. Ranks are now allocated to the observations within each latin square from 1 for the smallest observation to k^2 for the largest observation. If tied observations are present among the observations within a latin square design the tied observations within each tie are given the mean rank of the ranks they would have had, had they been different.

The sum of the ranks for the h-th treatment (h=1,...,k) is denoted by R_h.

The test statistic to test the null hypothesis that the effect of all treatments is the same is obtainable from (6.13) as a special case and is exactly the same as T_A defined in equation (5.66), or from equation (10.5) with $k=k^2$ and $p=k$. Rewritten in our notation it becomes;

$$T = 12[nk^3(k^2+1)]^{-1} \sum_{h=1}^{k} R_h^2 - 3n(k^2+1) \quad \text{.................(11.5)}$$

For k=2 (m in Table 10.1 and k^2 for k in Table 10.1) and n=2 exact probabilities are available in Table 10.1 of this thesis. For all other values of k and n the X^2-approximation at (k-1) degrees of freedom may be used.

If ties are present among the observations in a latin square the value CT defined in equation (10.8), with k substituted by k^2, must be calculated and T (equation

(10.5)) must be divided by this correction term to yield T(corr).

Attention is drawn to the following:

1. Method I can also be used to analyse a design as mentioned above. In this case ranks should be allocated from 1 for the smallest to nk^2 for the largest;

2. Method II can also be used to analyse the above design. Ranks should again be allocated within rows (or columns). This time however there will be nk such rows (or columns); and

3. which of the three methods is the most efficient for the above design will depend on the size of the squares, row and/or column effects.

An illustration of the application of Method III is given in Example 11.3

11.3 <u>Application</u>

<u>Example 11.1</u> - An example given in Cochran & Cox (1957) p. 121 is used to illustrate the application of the formulae of Method I.

This experiment was conducted to demonstrate the difficulty in selecting, by personal judgment, unbiased samples even from relatively small populations. In this experiment each population consisted of a small area of wheat containing about 80 tillers, the shoots being slightly over two feet in length. There were six samplers. Each sampler inspected each area and measured the lengths of eight tillers as a sample from that area. The quantity that will be analysed is the difference between the mean height of eight selected tillers and the true mean height in the corresponding area, i.e. the sampler's error.

The samplers represent experimental treatments, the six areas the columns and the order of sampling by the samplers the rows, of a 6 x 6 latin square design.

Table 11.1 gives the field plan (alphabetic letters); the height of the shoots from which a constant was subtracted ~~thus~~ hence the negative values (values without brackets); and the ranks corresponding to the heights (values between brackets).

Table 11.1 - Samplers error in tillers heights(cm) and ranks (r) obtained in a 6 x 6 latin square design

Rows	columns								
	I			II			III		
		cm	r		cm	r		cm	r
I	F	3.5	(8.5)	B	4.2	(18)	A	6.7	(28)
II	B	8.9	(33)	F	1.9	(4)	D	5.8	(23.5)
III	C	9.6	(34)	E	3.7	(11)	F	-2.7	(2)
IV	D	10.5	(36)	C	10.2	(35)	B	4.6	(20)
V	E	3.1	(7)	A	7.2	(30)	C	4.0	(15.5)
VI	A	5.9	(25)	D	7.6	(31)	E	-0.7	(3)

Rows	columns								
	IV			V			VI		
		cm	r		cm	r		cm	r
I	D	6.6	(27)	C	4.1	(17)	E	3.8	(17.5)
II	A	4.5	(19)	E	2.4	(5)	C	5.8	(23.5)
III	B	3.7	(11)	D	6.0	(26)	A	7.0	(29)
IV	E	3.7	(11)	A	5.1	(22)	F	3.8	(13.5)
V	F	-3.3	(1)	B	3.5	(8.5)	D	5.0	(21)
VI	C	3.0	(6)	F	4.0	(15.5)	B	8.6	(32)

Next the rank and yield totals for each treatment were obtained. The latter is not necessary but is obtained for interest sake:

	A	B	C	D	E	F
Rank(R_h)	153	122.5	131	164.5	50.5	44.5
Yield	364	335	367	415	160	72

From equation (11.2) we now calculate:

$$T = 12[N(N+1)\bar{k}]^{-1} \sum_{h=1}^{k} R_h^2 - 3(N+1)$$

$$= 12(36 \times 37 \times 6)^{-1}(153^2 + \ldots + 44.5^2) - 3 \times 37$$

$$= 19.88 \quad \text{with 5 d. f.}$$

From the tables for X^2 we obtain $X^2_{.05}(5) = 11.1$ and $X^2_{.01}(5) = 15.1$. Thus the value for treatments calculated above is highly significant.

Although there are a few tied observations we do not calculate CT since if we divide T by CT, T(corr) will be larger than T which is already highly significant.

The mean squares for the different components according to the parametric method of analysis are presented in Table 11.2.

Table 11.2 - Parametric analysis of variance

Source of variation	DF	MS
Rows	5	5.72
Columns	5	15.77**
Treatments	5	31.12**
Error	20	3.33

** significant at the 1% level

In the above example the conclusion drawn from both analyses is the same namely that the treatments differ significantly from each other. This will not always be the case.

Example 11.2 - To illustrate the analysis according to Method II the yields in Table 11.1 have been ranked within rows (by tossing a coin to decide whether it should be rows or columns) from the smallest to the largest. The corresponding ranks are set out in Table 11.3.

Table 11.3 - Ranks allocated (within rows) to the observations given in Table 11.1

F	1	B	4	A	6	D	5	C	3	E	2
B	6	F	1	D	4.5	A	3	E	2	C	4.5
C	6	E	3	F	1	B	2	D	4	A	5
D	6	C	5	B	3	E	1	A	4	F	2
E	2	A	6	C	4	F	1	B	3	D	5
A	4	D	5	E	1	C	2	F	3	B	6

The rank totals for the different treatments are:

	A	B	C	D	E	F
R_h	28	24	24.5	29.5	11	9

From equation (11.4) we obtain:

$$T = 12[k^2(k+1)]^{-1}\sum_{h=1}^{k} R_h^2 - 3k(k+1)$$

$$= 12(6^2 \text{ x } 7)^{-1}(28^2 + \ldots + 9^2) - 3 \text{ x } 6 \text{ x } 7$$

$$= 18.4 \;, \quad \text{d.f.}=5$$

This value is also highly significant but a little smaller than the one calculated in Example 11.1.

If it was known prior to seeing the data that the effects of the different columns would be higher than the effects of the different rows, or if by chance it was decided to allocate ranks within columns, the value of T would have been 18.69, again smaller than the one calculated in Example 1.

Attention is drawn to the fact that from theoretical considerations it is to be expected that Method I will be a more efficient procedure compared with Method II when the row and column effects are small or zero.

<u>Example 11.3</u> - The example of Hodges & Lehmann (1962) is used to illustrate the third method of analysis.

Three treatments A, B and C are compared in two 3 x 3 squares yielding the following observations.

Square 1		
B 4.461 (6)	C 2.798 (3)	A 7.402 (9)
A 3.412 (4)	B 2.405 (2)	C 5.227 (7)
C 3.454 (5)	A 2.169 (1)	B 6.717 (8)
Square 2		
C 5.424 (2)	B 9.670 (9)	A 9.669 (8)
B 5.062 (1)	A 9.368 (7)	C 5.710 (3)
A 6.605 (4)	C 7.786 (6)	B 7.427 (5)

The values in brackets are the ranks obtained by ranking the observations from 1 to 9 within each latin square. The rank totals for the treatments are as follows:

	A	B	C
R_h	33	31	26

In this example k=3, n=2 and $\sum_h R_h^2 = 2726$.

Substituting these values in equation (11.5) yields:

$$T = 12[nk^3(k^2+1)]^{-1}\sum_h R_h^2 - 3n(k^2+1)$$

$$= 12[2 \times 3^3 \times (9+1)]^{-1} \times 2726 - 3 \times 2 \times 10$$

$$= 0.58$$

This value is not significant.

The parametric analysis yielded the following result:

Source of variation	DF	MS	F
Rows	4	1.4725	
Columns	4	10.1906	
Squares	1	45.6840	
Treatments	2	2.9035	5.973*
Error	6	0.48605	

* significant at the 5% level

The high mean square obtained for columns depresses the value of T in the nonparametric analysis and this is the reason for the small value obtained for T.

11.4 Summary of the procedures

Only the special case of ranks is considered.

11.4.1 Method I - These are the steps in the analysis of a $k \times k$ latin square design: ~~which is replicated only once~~:

1. All the $N=k^2$ observations are ranked in a single series assigning ranks from 1 to N, allocating mean ranks to tied observations;
2. the total of the ranks R_h in the h-th treatment $(h=1,\ldots,k)$ is determined;
3. the value of the test statistic T defined in (11.2) is calculated;
4. if a large proportion of observations is tied, calculate CT given in (8.9) and divide the test statistic calculated in 3. above by CT to obtain T(corr); and
5. the significance of T or T(corr) may be assessed as follows:
   a. If $k=3$ the tables for the Kruskal-Wallis (1952) test statistic must be used; and
   b. for $k>3$ the X^2-approximation at $(k-1)$ degrees of freedom must be used.

11.4.2 Method II

1. If no prior knowledge is available as to the size of the effects of rows and columns a coin is tossed to decide whether ranks will be allocated within rows or columns.
2. The observations are ranked within rows (or columns) from 1 for the smallest to k for the largest, allocating mean ranks to tied observations.
3. The total of the ranks R_h in the h-th treatment $(h=1,\ldots,k)$ is determined.
4. The value of the test statistic T defined in (11.4) is calculated.
5. If a large proportion of observations is tied

calculate CT given in (10.8) and divide the test statistic calculated in 4. above by CT to obtain T(corr).

6. The significance of T or T(corr) may be assessed as follows:

a. For k=3 and 4 use the exact probabilities calculated for Friedman's (1937) (cf. Siegel, 1956, Table N) test statistic; and

b. for $k > 4$ use the X^2-approximation at (k-1) degrees of freedom.

1.4.3 Method III -These are the steps in the analysis of a k x k latin square design which is replicated n times.

1. The observations within each latin square are ranked from 1 for the smallest to k^2 for the largest, allocating mean ranks to tied observations.

2. The total of the ranks R_h (h=1,...,k) for the h-th treatment is then determined.

3. The value of the test statistic T defined in (11.5) is calculated.

4. If a large proportion of the observations in a latin square **is** tied, CT given in (10.8) is calculated and the test statistic calculated in 3. above divided by this CT to obtain T(corr).

5. The significance of T or T(corr) may be assessed as follows:

a. For k=2 and n=2 use the exact probabilities tabulated in Table 10.1; and

b. $k > 2$ and/or $n > 2$ the X^2-approximation at (k-1) degrees of freedom must be used.

1.5 Evaluation of the new methods

The three distribution-free methods of analysis presented in this chapter are easy to apply and are not dependant on the distribution of the population from which the observations come.

The efficiency is, however, unknown but is not very high when compared with the F-test if the requirements of the F-test are met, especially when large row and/or column effects exist. This statement is effectively illustrated in Example 11.3.

Hodges & Lehmann (1962) presented a rank method of analysis for a latin square design, but the application of the method is very cumbersome.

Thus it still remains a problem to find an efficient distribution-free method of analysis for a latin square design.

CHAPTER XII

DISTRIBUTION-FREE MULTIPLE COMPARISON PROCEDURES

12.1 Introduction

If in an analysis of variance the null hypothesis of equal population means is rejected (i.e. the test statistic for treatments is significant) one would like to make decisions as to which means are responsible for the significant result.

In this chapter multiple comparison procedures are presented (§12.2) and illustrated with examples (§12.3) for a completely random design and a randomized block design.

The theory for the multiple comparison procedures **was** derived in Chapter VII for functions of ranks. Only the formulae for the special case of ranks itself are considered.

12.2 Formulae

12.2.1 Completely random design

The one-way classification of treatments is considered first. Suppose there are k treatments each replicated an equal number of times, say n. Rank the N=nk observations X_{jh} (j=1,...,k; h=1,...,n) jointly giving rank 1 to the smallest observation. If ties are present among the observations, the tied observations in each tie are given the mean rank of the ranks they would have had, had they been different. Denote the rank of X_{jh} by r_{jh} and let $R_j = \sum_{h=1}^{n} r_{jh}$ be the total of the ranks belonging to the j-th treatment.

The test statistic for treatments defined by Kruskal & Wallis (1952) (cf. equation (2.12)) is next calculated. If this value is significant any contrast amongst the treatments defined by:

$$\hat{\psi} = \sum_{j=1}^{k} c_j R_j \qquad \ldots\ldots(12.1)$$

where the c_j's are constants subject to $\sum_{j=1}^{k} c_j = 0$ and not all the c's zero, may be compared with

$$T_R = [(12)^{-1}nN(N+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} (\tfrac{1}{2} \sum_{j=1}^{k} |c_j|) \qquad \ldots\ldots(12.2)$$

where $q_{\alpha,k,\infty}$ is the upper α point of the studentized range for k treatments at ∞ degrees of freedom for error variance.

If now $|\hat{\psi}|$ exceeds T_R the contrast $\hat{\psi}$ is declared significant at the $100\alpha\%$ level.

The formula for comparing any two treatments (i.e. by means of the absolute value of the difference of the rank totals of the two treatments) is directly obtainable from equation (12.2) above by **substituting $c_j=1$ and $c_{j'}=-1$** and all other c's = 0, i.e.

$$LSD_R = [(12)^{-1}nN(N+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} \qquad \ldots\ldots(12.3)$$

If any of the differences $|R_j - R_{j'}|$ $(j \neq j';\ j,j'=1,\ldots,k)$ exceed LSD_R we say the j-th and j'-th treatments differ.

In the case where the k treatments are composed of two components A and B at m and p levels each, respectively (cf. Chapter VIII) and if the test statistic calculated for, say component A, is significant a multiple comparison method for the different levels of component A, can be obtained directly by observing that the m levels of component A are each replicated n' = np times instead of n times. Equations (12.1) and (12.2) can now respectively be written as:

$$\hat{\psi} = \sum_{i=1}^{m} c_{i.} R_{i.} \quad , \quad \sum_{i=1}^{m} c_{i.} = 0$$

where $R_{i.}$ is the total of the ranks in the i-th level of A; and

$$T_{R_A} = [(12)^{-1}n'N(N+1)]^{\frac{1}{2}}q_{\alpha,m,\infty} \quad (\tfrac{1}{2}\sum_{i=1}^{m}|c_j|) \quad \ldots\ldots\ldots\ldots(12.4)$$

where $n' = np$ in this case.

The test statistic for pairwise comparison of the different levels of component A is obtained from equation (12.3).

$$LSD_{R_A} = [(12)^{-1}n'N(N+1)]^{\frac{1}{2}}q_{\alpha,m,\infty} \quad \ldots\ldots\ldots\ldots\ldots\ldots(12.5)$$

where $n' = np$ in this case.

The above procedure seems to be applicable if $n \geqslant 4$ and the total number of observations $N \geqslant 20$.

It should be kept in mind that a multiple comparison technique will usually only be applied when an overall significant difference has been found by means of an analysis of variance technique.

Since the analysis of variance procedure presented in Chapter VIII and the multiple comparison techniques presented above are based on the same underlying theory, the application of the corresponding distribution-free multiple comparison techniques seems justified when the corresponding distribution-free analysis of variance procedure is considered applicable.

12.2.2 Randomized block design

The one-way classification of treatments is considered first.

Suppose there are k treatments and n replications. Rank the k observations within each block (or replication) giving rank 1 to the smallest observation. If tied observations are present among the observations within a block the tied observations in each tie are given the mean rank of the ranks they would have had, had they been different.

Denote the rank of X_{jh} $(j=1,\ldots,k;\ h=1,\ldots,n)$ by r_{jh}. The total of the ranks in the j-th treatment is given by:

$$R_j = \sum_{h=1}^{n} r_{jh}$$

The test statistic for treatments defined by Friedman (1937) (cf. §2.3.7) must now be calculated. If this test statistic **is** significant, any contrast amongst the treatments defined by

$$\hat{\psi} = \sum_{j=1}^{k} c_j R_j$$

where the c_j's are constants subject to $\sum_{j=1}^{k} c_j = 0$ and not all c's = 0, may be compared with

$$T_{Fr} = [(12)^{-1}nk(k+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} \quad (\tfrac{1}{2}\sum_{j=1}^{k}|c_j|) \quad \text{............}(12.7)$$

where $q_{\alpha,k,\infty}$ is the upper α point of the studentized range for k treatments at ∞ degrees of freedom for error variance.

If now $|\hat{\psi}|$ (equation (12.6)) exceeds T_{Fr} (equation (12.7)) the contrast is declared significant at the $100\alpha\%$ level.

The formulae for comparing any two treatments (i.e. by means of the absolute value of the difference of the rank totals of the two treatments) with each other is directly obtainable from equation (12.7) above by **substituting** $c_j = 1$ and $c_{j'} = -1$ and all other c's = 0, i.e.

$$LSD_{Fr} = [(12)^{-1}nk(k+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} \quad \text{.....................}(12.8)$$

If any of the differences $|R_j - R_{j'}|$ $(j \neq j';\ j,j'=1,\ldots,k)$ exceed LSD_{Fr} it is concluded that the j-th and j'-th treatments differ.

In the case where the k treatments are composed of two components A and B at m and p levels each respectively (cf. Chapter X) and if the test statistic calculated for, say component A, **is** significant a multiple comparison method for the different levels of component A, can directly be obtained by observing that the m levels of component A are each replicated n' = np times instead of n times. Equations (12.6) and (12.7) can thus be written as:

$$\hat{\psi} = \sum_{i=1}^{m} c_{i.} R_{i.}, \qquad \sum_{i=1}^{m} c_{i.} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12.9)$$

where $R_{i.}$ is the total of the ranks in the i-th level of A; and

$$LSD_{Fr_A} = [(12)^{-1}n'k(k+1)]^{\frac{1}{2}} q_{\alpha,m,\infty} (\tfrac{1}{2} \sum_{i=1}^{m} |c_{i.}|) \quad \ldots\ldots\ldots(12.10)$$

where n' = np in this case.

The test statistic for pairwise comparisons of the different levels of component A is obtained from (12.8).

$$LSD_{Fr_A} = [(12)^{-1}n'k(k+1)]^{\frac{1}{2}} q_{\alpha,m,\infty} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(12.11)$$

where n' = np in this case.

The following corrections to equations (12.7), (12.8), (12.10) and (12.11) are recommended in the case of small samples (cf. 7.3.3):

1. If $k \leq 4$ a $\frac{1}{2}$ should be added to the values calculated by means of the formulae mentioned; and
2. for $k > 4$ no correction seems to be necessary.

A remark similar to that given at the end of §12.2.2 is also applicable here.

The formulae, for a completely random design with unequal sample sizes, are not quoted here. They can, however, be obtained from Chapter VII.

2.3 Application

In all the examples below the requirements regarding sample sizes, for the theory to be applicable, are met.

Example 12.1 - The application of the formulae for a completely random design is illustrated by means of an example used by Steel (1961).

The Kruskal-Wallis (1952) test statistic calculated on the data in Table 12.1 is highly significant, and we thus continue to calculate the LSD_R with the aid of equation (12.3).

Table 12.1 - Final weights (gm) and ranks (r) of chickens at six weeks for various sources of protein supplement

H		L		S_b		S_f		M		C	
gm	r	gm	r	gm	r	gm	r	gm	r	gm	r
108	1	141	5	193	14	226	21	153	8	216	18
124	2	148	7	230	24	295	36	206	16	222	20
136	3	169	11	243	26	320	42	242	25	260	31.5
140	4	181	13	248	27	322	43	257	29.5	318	41
143	6	203	15	250	28	334	47	263	33	352	52
160	9	213	17	267	34	339	48	303	37	359	53
168	10	229	23	271	35	340	49	315	39	368	54
179	12	257	29.5	316	40	341	50	325	44	379	55
217	19	260	31.5	327	45	392	58	344	51	390	57
227	22	309	38	329	46	423	60	380	56	404	59
R_j	88		190		319		454		338.5		440.5

From the tables of the studentized range we obtain:

$q_{.05,6,\infty} = 4.03$

Thus

$$LSD_R = [(12)^{-1}nN(N+1)]^{\frac{1}{2}} q_{.05,6,\infty}$$
$$= [(12)^{-1} \times 10 \times 60 \times 61]^{\frac{1}{2}} \times 4.03$$
$$= 221$$

The value of $q_{.05,6,\infty} = 4.03$ was obtained from Federer (1955) Table II-1 on page 22-23.

The results obtained can be presented as follows:*

H	L	S_b	M	C	S_f
88	190	319	338.5	440.5	454

H is declared significantly different from S_b, M, C and S_f while L is significantly different from C and S_f. This result is exactly the same as the one obtained by Steel (1961).

* see footnote at bottom of the next page (p. 249)

For the Tukey parametric procedure, we find means of 160.2 (H), 211.0 (L), 267.4 (S_b), 278.8 (M), 323.2 (S_f) and 326.8 (C). Also $S_{\bar{x}} = 18.03$. For Tukey's test the least significant difference is 75.2.

Using this LSD to compare the means , leads in this example to the same conclusions as obtained by the non-parametric procedures.

Example 12.2 - Suppose one wants to decide which of the levels of the nitrogen fertilizer differed from each other in Example 8.1.

From the studentized range table the following value is obtained.

$q_{.05,3,\infty} = 3.32$

With the aid of formula (12.5) we calculate:

$$\begin{aligned} LSD_{R_N} &= [(12)^{-1} nmN(N+1)]^{\frac{1}{2}} q_{\alpha,p,\infty} \\ &= [(12)^{-1} \times 6 \times 2 \times 36 \times 37]^{\frac{1}{2}} \times 3.32 \\ &= 111.18 \end{aligned}$$

The results obtained can be summarized as follows:

N_1	N_2	N_3
146	185	335

The highest level of nitrogen fertilizer differs from the other two levels.

Example 12.3 - To illustrate the application of the formulae for the set of all contrasts take the values of Example 12.1.

Suppose we want to compare H and L with the rest of the treatments. From (12.1) we have:

* The method of underlining treatments which are not declared significant used in multiple comparison procedures may be adopted for application to rank sum procedures.

$$\hat{\psi} = \sum_{j=1}^{k} c_j R_j = 2R_1 + 2R_2 - R_3 - R_4 - R_5 - R_6$$

$$= 2(88) + 2(190) - 319 - 454 - 338.5 - 440.5$$

$$= -996$$

where

$$\sum_{j=1}^{k} c_j = 2 + 2 - 1 - 1 - 1 - 1 = 0$$

$$\tfrac{1}{2} \sum_{j=1}^{k} |c_j| = \tfrac{1}{2}(2+2+1+1+1+1) = 4$$

From (12.2) we have:

$$T_R = [(12)^{-1} nN(N+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} \left(\tfrac{1}{2} \sum_{j=1}^{k} |c_j|\right)$$

$$= [(12)^{-1} \times 10 \times 60 \times 61]^{\frac{1}{2}} \times 4.03 \times 4 = 884$$

Since $|\hat{\psi}| > 884$ we conclude that the contrast is significant.

Example 12.4 - To illustrate the use of the distribution-free multiple comparison procedures derived for a randomized block design an experiment was used which consisted of five cultivars of maize which was tested over two years and replicated five times. The design ~~were~~ was a randomized block design. The yields and ranks are given in Table 12.2.

Table 12.2 - Yields (bpm) and ranks (r) of five cultivars of maize (C) tested over two years (Y)

Repl.	Treatment combinations									
	C_1Y_1		C_1Y_2		C_2Y_1		C_2Y_2		C_3Y_1	
	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r
1	41	2	38	1	52	9	50	7	44	4
2	44	3	35	1	61	10	51	5.5	54	7
3	45	2	41	1	58	8	48	4	51	5
4	45	2	39	1	66	10	64	7.5	52	4
5	44	1	45	2	48	5	63	8.5	60	6.5
Tot.	219	10	198	6	286	42	276	32.5	261	26.5

Repl.	treatment combinations									
	C_3Y_2		C_4Y_1		C_4Y_2		C_5Y_1		C_5Y_2	
	bpm	r	bpm	r	bpm	r	bpm	r	bpm	r
1	46	5	56	10	47	6	51	8	43	3
2	37	2	50	4	57	9	56	8	51	5.5
3	47	3	60	9	55	6	61	10	56	7
4	49	3	56	5.5	65	9	64	7.5	56	5.5
5	48	4	60	6.5	46	3	63	8.5	72	10
Tot.	227	17	282	35	270	33	295	42	278	31

The data in Table 12.2 were analysed according to the method described in Chapter X. The results obtained are given in Table 12.3

Table 12.3 - Distribution-free analysis of variance

Source of variation	DF	T
Cultivars	4	27.56**
Years	1	2.828
Interaction	4	0.681
Total	9	31.069

** significant at the 1% level

To decide which cultivars differ from the rest, calculate the least significant difference (LSD_{Fr}) with the aid of equation (12.11)

$$LSD_{Fr_C} = [(12)^{-1} n'k(k+1)]^{\frac{1}{2}} q_{\alpha,k,\infty} \quad \text{where } n'=5 \times 2 = 10$$

$$= [(12)^{-1} \times 10 \times 10 \times 11]^{\frac{1}{2}} \times 3.86$$

$$= 36.85$$

Applying this value to the totals of the ranks for the cultivars of maize we obtain:

C_1	C_3	C_4	C_5	C_2
16	43.5	68	73	74.5

For the parametric case $S_{\bar{x}} = 1.746$ and the means $C_1(41.7)$, $C_2(56.2)$, $C_3(48.8)$, $C_4(55.2)$ and $C_5(57.3)$. The LSD for the means calculated according to Tukey's w procedure is 7.05. The results obtained in the parametric case can be presented as follows:

C_1	C_3	C_4	C_2	C_5
41.7	48.8	55.2	56.2	57.3

Tukey's procedure in the parametric case showed that more varieties different than the distribution-free procedure.

12.4 Summary of the procedures

12.4.1 Completely random design

The steps in the application of the distribution-free multiple comparison procedures for the special case of ranks, are the following:

1. Rank all the observations for the k treatment combinations in a single series, assigning ranks from 1 to N, allocating mean ranks to tied observations;
2. the total of the ranks for every treatment (combination) is determined;
3. the Kruskal-Wallis (1952) test statistic or the appropriate test statistic (cf. Chapter VIII) is then calculated;
4. if the test statistic calculated in 3. above turns out to be significant proceed as follows:
    a. For any contrast among the treatments use equations (12.2) or (12.4) depending on whether the test statistic for the one- or more-way classification was calculated in 3. above;
    b. for pairwise comparisons use equations (12.3) or (12.5);

5. a contrast or the difference of any pair of treatments is declared significant if ~~their~~ its absolute value exceeds the value calculated in 4a. or 4b. respectively;

6. these methods of comparison are valid if $N \geqslant 20$ and the number of replications within a treatment $\geqslant 4$.

12.4.2 Randomized block design

The steps in the application of the distribution-free multiple comparison procedures, for the special case of ranks, are the following:

1. The observations within each replication are ranked from 1 for the smallest value to k for the largest value, allocating mean ranks to tied observations;

2. the total of the ranks for every treatment (combination) is determined;

3. the Friedman (1937) test statistic or the appropriate test statistic (cf. Chapter X) is calculated;

4. if the test statistic calculated above turns out to be significant proceed as follows:

   a. For any contrast among the treatments use equations (12.7) or (12.10) depending on whether the test statistic for the one- or more-way classifi-
c cation was calculated in 3. above;

   b. for pairwise comparisons use equations (12.8) or (12.11);

5. a contrast or a difference of any pair of treatments is declared significant if ~~their~~ its absolute value exceeds the value calculated in 4a. or 4b. respectively; and

6. for any contrast a $\frac{1}{2}$ should be added to the value calculated with equations (12.7), (12.8), (12.10) or (12.11) if $k \leq 4$. For $k > 4$ no correction is necessary.

12.5 Evaluation of the methods

The distribution-free multiple comparison procedures presented in this chapter are:

1. Easy to calculate;
2. applicable to the new methods of analysis presented in this thesis; and
3. independent of the distribution of the population from which the data come.

Their efficiency, however, is unknown but it can reasonably be expected that these methods will be conservative compared with the parametric methods if the requirements of the latter are met. It seems safe to use these methods in practice.

S U M M A R Y

The purpose of the work presented was firstly to develop distribution-free methods of analysis and multiple comparison procedures for basic experimental designs where the treatments form a multi-way classification, and secondly to present these methods in such a way that a research worker with an elementary knowledge of biometry would be able to apply these methods without much difficulty.

This thesis is divided into two parts. Part I deals with the mathematical derivation of test statistics and their exact and asymptotic distributions (under the null hypothesis) for different experimental designs, as well as the investigation of the applicability of the tests based on these test statistics under an alternative hypothesis by means of sampling studies; Part II deals with the application of the tests derived in Part I.

The new developments in Part I can be summarised as follows:

Lemmer (1964) derived test statistics based on a function of the ranks of the observations for the different components in a two-way classification of treatments when the design was a completely random design, as well as their asymptotic distributions (distribution-free) on the assumption that the null hypothesis was true. By means of sampling studies it was found that these tests for the different components, in the special case of ranks, tended to be conservative (i.e. did not reject the null hypothesis too soon) under different alternative hypotheses and due to this and certain theoretical considerations it was concluded that these tests were applicable under the null as well as under an alternative hypothesis. Distribution-free tests were also

developed for a factorial design, known as the single degree of freedom approach. The underlying theory and practical investigation which led to the above mentioned results, are given in Chapter III of this thesis.

The theoretical development of test statistics and their distributions under the null hypothesis for the different components of a multi-way classification of treatments when the design was a balanced incomplete block design is presented in Chapter IV. Since a balanced incomplete block design, where the treatments consisted of more than one component, is very seldom used, the formulae are of little use in practice. They are, however, of theoretical interest since the formulae for a randomized block design are obtainable from them as a special case.

In the case of a randomized block design (Chapter V) the exact probability of obtaining a value as large or larger than the calculated value for the test statistic was calculated in the case of small samples for different values of the test statistic. Next the distribution of the test statistic in the case of small samples, was approximated by the β- and X^2-distributions. The exceeding probabilities for the given test statistic, obtained by using these two approximations, were then compared with the exact value of the exceeding probability of the given test statistic for the different values of the test statistic. The results obtained are given in Chapter V. It was found that the X^2-distribution could be used to calculate probabilities in all cases where the exact probabilities were not available. By means of sampling studies the applicability of the tests for the different components was investigated under an alternative hypothesis in the special case of ranks. Similar to the completely random design it was found that these tests were applicable under the null as well as under an alternative hypothesis. Under the alternative hypothesis they tended to be conservative (i.e. did not reject the null hypothesis too soon).

The possibility of correcting the observations in a latin square design, by sutracting the estimated row and column effects from the observations, and then applying a test statistic for a completely random design, was investigated. The discussion of the problem and results obtained in the sampling studies, are presented in Chapter VI. It was found that such a procedure generally tended to reject the null hypothesis too soon. Three methods were then presented for the analysis of a latin square design. Two were for the case where the latin square was replicated once and one for the case where it was replicated several times. All these methods were found to be unsatisfactory since they tended to be very conservative, especially when the row and/or column effects were substantial.

Asymptotic multiple comparison procedures were derived for the completely random design and for the randomized block design and are presented in Chapter VII. Estimations were given as to the sample sizes needed for these formulae to be applicable.

Part II deals with the application of the tests derived in Part I. At the beginning of each Chapter in Part II the formulae for purpose of calculating were given for the special case of ranks, and their application then illustrated with an example.

However, in Examples 8.2 and 10.2 the applications of the tests were illustrated for the special case where the ranks were transformed to "normal values", a transformation suggested by Van der Waerden (1957). In these two examples the formulae for calculating purposes were obtained from Part I.

Most of the examples for purpose of illustration were taken from the field of agronomy. The methods presented in this thesis, however, are applicable to any field of research as long as the requirements of the specific design used are met.

Part II was written in such a way that a person with an elementary knowledge of biometry should be able to understand it and apply the methods.

In conclusion it may be said that this thesis presents distribution-free analysis of variance methods for the completely random, the balanced incomplete block, the randomized block and the latin square designs. Multiple comparison procedures were given for the completely random and the randomized block designs. These distribution-free methods, however, tended to be conservative when they were used under an alternative hypothesis.

ANDERSON, R.L. & BANCROFT[illegible]. Statistical theory in research. 1st ed. [illegible]raw-Hill.

ANDREWS, F.C., 1954. Asymptotic behaviour of some rank tests for analysis of variance. Ann. Math. Stat. 25, 724-736.

BENARD, A. & VAN ELTEREN, P., 1953. A generalization of the method of m rankings. Indag. Math. 15, 358-369.

BHAPKAR, V.P., 1961. Some nonparametric median procedures. Ann. Math. Stat. 32, 846-863.

BRADLEY, R.A. & TERRY, M.E., 1952. The rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika 39, 324-345.

BRADLEY, R.A., 1954. The rank analysis of incomplete block designs. II. Additional tables for the method of paired comparisons. Biometrika 41, 502-537.

BRADLEY, R.A., 1955. The rank analysis of incomplete block designs. III. Some large-sample results on estimation and power for a method of paired comparisons. Biometrika 42, 450-470.

BROWN, G.W. & MOOD, A.M., 1951. On median tests for linear hypotheses. Proceedings of the second Berkeley symposium on mathematical statistics and probability. Berkeley: Univ. of Calif. Press, 159-166.

BROWNLEE, K.A., 1960. Statistical theory and methodology in science and engineering. 1st ed. New York: Wiley.

CHERNOFF, H. & SAVAGE, I.R., 1958. Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Stat. 29, 972-994.

CLARK, C.E. & HOLZ, B.W., 1960. Exponentially distributed random numbers. 1st ed. Baltimore: The John Hopkins Press.

COCHRAN, W.G. & COX, G., 1957. Experimental designs. 2nd ed. New York: John Wiley.

CRAMER, H., 1962. Random variables and probability distributions. 2nd ed. Cambridge University Press.

CROUSE, C.F., 1960. Combinatorial tests, for differences in location and dispersion, for the case of m samples. Thesis. University of London.

DAVID, H.A., 1963. The method of paired comparisons. Griffin's statistical monographs and courses, no. 12. London: Charles Griffin.

DUNCAN, D.B., 1947. Significance tests for differences between ranked variates drawn from normal populations. Thesis. Iowa State College.

DUNCAN, D.B., 1951. A significance test for differences between ranked treatments in an analysis of variance. The Virginia J. of Sc. Vol. 2, no. 3, 171-189.

DUNCAN, D.B., 1952. On the properties of the multiple comparisons test. The Virginia J. of Sci. Vol. 3, no. 1, 49-67.

DUNCAN, D.B., 1955. Multiple range and multiple F-tests. Biometrics 11, 1-42.

DUNNETT, C.W., 1955. A multiple comparison procedure for comparing several treatments with a control. J. Am. Stat. Ass. 50, 1096-1121.

DURBIN, J., 1951. Incomplete blocks in ranking experiments. Brit. J. of Psych. 4, 85-90.

FEDERER, W.T., 1955. Experimental design. 1st ed. London: The Macmillan Co.

FISHER, R.A., 1935. The design of experiments. 1st ed. London: Oliver & Boyd.

FISHER, R.A. & YATES, M.A., 1938. Statistical tables. 1st ed. London: Oliver & Boyd.

FRIEDMAN, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. Ass. 32, 675-701.

FRIEDMAN, M., 1940. A comparison of alternative tests of significance for the problem of m rankings. Ann. Math. Stat. 11, 86-92.

GOULDEN, C.H., 1952. Methods of statistical analysis. 2nd ed. London: John Wiley.

HARTER, H.L., 1960. Tables of the range and studentized range. J. Am. Stat. Ass. 55, 1122-1147.

HODGES, J.L. & LEHMANN, E.L., 1956. The efficiency of some nonparametric competitors of the t-test. Ann. Math. Stat. 27, 324.

HODGES, J.L. & LEHMANN, E.L., 1962. Rank methods for combination of independent experiments in the analysis of variance. Ann. Math. Stat. 33, 482-497.

HODGES, J.L. & LEHMANN, E.L., 1963. Estimates of location based on rank tests. Ann. Math. Stat. 34, 598-611.

JONKHEERE, A.R., 1954. A distribution-free k-sample test against ordered alternatives. Biometrika 41, 133-145.

KEMPTHORNE, O., 1952. The design and analysis of experiments. 1st ed. London: John Wiley.

KENDALL, M.G., 1955. Rank correlation methods. 2nd ed. London: Charles Griffin.

KENDALL, M.G. & STUART, A., 1961. The advanced theory of statistics, Vol I. 1st ed. London: Charles Griffin.

KEULS, M., 1952. The use of the studentized range in connection with an analysis of variance. Euphytica 1, 112-122.

KRAMER, C.Y., 1956. Extension of multiple range tests to group means with unequal numbers of replications. Biometrics 12, 307-310.

KRAMER, C.Y., 1957. Extension of multiple range tests to group correlated adjusted means. Biometrics 13, 13-18.

KRUSKAL, W.H., 1952. A nonparametric test for the several sample problem. Ann. Math. Stat. 23, 525-540.

KRUSKAL, W.H. & WALLIS, W.A., 1952. Use of ranks in one-criterium variance analysis. J. Am. Stat. Ass. 47, 583-621.

LEHMANN, E.L., 1963a. Robust estimation in analysis of variance. Ann. Math. Stat. 34, 957-966.

LEHMANN, E.L., 1963b. Asymptotically nonparametric inference: An alternative approach to linear models. Ann. Math. Stat. 34, 1494-1506.

LEHMANN, E.L., 1963c. Nonparametric confidence intervals for a shift parameter. Ann. Math. Stat. 34, 1507-1512.

LEHMANN, E.L., 1964. Asymptotic nonparametric inference in some linear models with one observation per cell. Ann. Math. Stat. 35, 726-734.

LEMMER, H.H., 1964. Verdelingsvrye toetsingsmetodes vir die probleem van twee of meer steekproewe. Thesis. University of Pretoria.

LEMMER, H.H., & STOKER, D.J., 1961. 'n Klas van verdelingsvrye toetsingsgroothede vir die probleem van k onafhanklike steekproewe. Tydskr. Natuurwet. 1, 231.

MANN, H.B. & WHITNEY, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 18, 50-60.

MEHRA, K.L., 1964. Rank tests for paired-comparisons experiments involving several treatments. Ann. Math. Stat. 35, 122-137.

MOOD, A.M., 1950. Introduction to the theory of statistics. 1st ed. London: McGraw-Hill.

NEMENYI, P., 1963. Distribution-free multiple comparisons. Thesis. State University of New York Downstate Medical centre.

PEARSON, K., 1934. Tables of the incomplete Beta-function. 1st ed. Cambridge: The University Press.

NEWMAN, D., 1939. The distribution of range in samples from a normal population expressed in terms of an independant estimate of standard deviation. Biometrika 31, 20-30.

PENDERGROSS, R.N. & BRADLEY, R.A., 1960. Ranking in triple comparisons. Contributions to probability and statistics (Essays in honor of H. Hotelling) ed. by I. Olkin and others. Stanford University Press.

PURI, M.L., 1964. Asymptotic efficiency of a class of c-sample tests. Ann. Math. Stat. 35, 102-121.

RHYNE, A.L., 1964. Some multiple comparison sign tests. Thesis. Consolidated University of North Carolina.

RYAN, T.A., 1960. Significance tests for multiple comparison of proportions, variances, and other statistics. Psychol. Bull. 57, 318-328.

SCHEFFÉ, H., 1953. A method for judging all contrasts in the analysis of variance. Biometrika 40, 87-104.

SCHEFFE, H., 1959. The analysis of variance. 1st ed. London: John Wiley.

SIEGEL, S., 1956. Nonparametric statistics for the behavioral sciences. 1st ed. London: McGraw-Hill.

STEEL, R.G.D., 1959. A multiple comparison rank sum test: Treatments vs. control. Biometrics 15, 560-572.

STEEL, R.G.D., 1960. A rank sum test for comparing all pairs of treatments. Technometrics 2, 197-207.

STEEL, R.G.D., 1961. Some rank sum multiple comparisons tests. Biometrics 17, 539-552.

STUDENT, 1927. Errors of routine analysis. Biometrika 19, 151-164.

STOKER, D.J., 1955. Oor 'n klas van toetsingsgroothede vir die probleem van twee steekproewe. Thesis. University of Amsterdam.

TATE, M.W. & CLELLAND, R.C., 1959. Nonparametric and short-cut statistics. Danville, Illinois: Interstate printers & publishers.

TUKEY, J.W., 1953. The problem of multiple comparisons. Ditto Princeton Univ.

TERPSTRA, T.J., 1954. A generalization of Kendall's rank correlation statistic, I. Indag. Math. 17, 690-696.

TERRY, M.E., 1952. Some rank order tests, which are most powerfull against specific parametric alternatives. Ann. Math. Stat. 23, 346-366.

VAN DER WAERDEN, B.L., 1953. Ein neuer test für das problem der zwei stichproben. Math. Ann. 126, 93-107.

VAN DER WAERDEN, B.L., 1957. Mathematische statistik. Springer-Verlag, Berlyn.

VAN ELTEREN, P. & NOETHER, G.E., 1959. The asymptotic efficiency of the X_r^2-test for a balanced incomplete block design. Biometrika 46, 475.

WHITNEY, D.R., 1951. A bivariate extension of the U-statistic. Ann. Math. Stat. 22, 274.

WILCOXON, F., 1945. Individual comparisons by ranking methods. Biometrics 1, 80-83.

WILKS, S.S., 1962. Mathematical statistics. 1st ed. New York: John Wiley.

WOLD, H., 1948. Random normal deviates. Cambridge Univ. Press.

APPENDIX

The consistency problem when interactions are present

Undesirable consistency properties exist in the tests for main effects if there is an interaction effect (not necessarily significant).

The two examples below, based on a two-way classification of treatments, where the design is a completely random design, will illustrate this statement.

In the examples below it is assumed that the observations X_{ijh} is rectangularly distributed on the interval $[c_{ij}-\frac{1}{2};\ c_{ij}+\frac{1}{2}]$ where any c_{ij} $(i=1,\ldots,m;\ j=1,\ldots,p)$ is the combined effect of the corresponding two main effects and the interaction effect in the (i,j)-th cell. It is further assumed that n observations are obtained in each cell. The special case of ranks will be considered.

Example 1 - Suppose we have two components A and B at four and two levels each respectively with the effects of the levels of the B component all zero.

Table A - The effects and the total of the ranks in the cells of a 4 x 2 experiment

	B_0	B_1	
A_0	$c_{11}=-15$ $R_{11}=\frac{1}{2}n(n+1)$	$c_{12}=-3$ $R_{12}=\frac{1}{2}n(3n+1)$	$c_{1.}=-18$ $R_{1.}=\frac{1}{2}n(4n+2)$
A_1	$c_{21}=4$ $R_{21}=\frac{1}{2}n(11n+1)$	$c_{22}=-2$ $R_{22}=\frac{1}{2}n(5n+1)$	$c_{2.}=2$ $R_{2.}=\frac{1}{2}n(16n+2)$
A_2	$c_{31}=5$ $R_{31}=\frac{1}{2}n(13n+1)$	$c_{32}=2$ $R_{32}=\frac{1}{2}n(7n+1)$	$c_{3.}=7$ $R_{3.}=\frac{1}{2}n(20n+2)$
A_3	$c_{41}=6$ $R_{41}=\frac{1}{2}n(15n+1)$	$c_{42}=3$ $R_{42}=\frac{1}{2}n(9n+1)$	$c_{4.}=9$ $R_{4.}=\frac{1}{2}n(24n+2)$
	$c_{.1}=0$ $R_{.1}=\frac{1}{2}n(40n+4)$	$c_{.2}=0$ $R_{.2}=\frac{1}{2}n(24n+4)$	$R_{..}=\frac{1}{2}8n(8n+1)$

The test statistics for the different components are directly obtainable from equation (3.18) and are given by:

$$\left.\begin{aligned} T_A &= 12[npN(N+1)]^{-1}\sum_{i=1}^{m}[R_{i.} - \tfrac{1}{2}np(N+1)]^2 \\ T_B &= 12[nmN(N+1)]^{-1}\sum_{j=1}^{p}[R_{.j} - \tfrac{1}{2}nm(N+1)]^2 \\ &\text{and} \\ T_{AB} &= 12[nN(N+1)]^{-1}\sum_{i=1}^{m}\sum_{j=1}^{p}[R_{ij} - \tfrac{1}{2}n(N+1)]^2 - T_A - T_B \end{aligned}\right)\ldots(A)$$

where

n = number of observations per cell;

p = number of levels of component B;

m = number of levels of component A; and

$N = nmp$.

Substitution of the values in Table A in the formulae of equation (A) yields:

$T_A = 42n^2/(8n+1)$; $T_B = 12n^2/(8n+1)$; $T_{AB} = 9n^2/(8n+1)$.

Although the effects of the levels of component B are zero, the value of the test statistic T_B is >0 and tends to infinity as n tends to infinity and thus will become significant for n sufficiently large.

Example 2 - In this example both components A and B are at three levels with the effects of the levels of component B all zero.

Table B - The effects and totals of ranks in the cells of a 3 x 3 experiment

	A_o	A_1	A_2	
B_o	$c_{11}=-15$ $R_{11}=\frac{1}{2}n(n+1)$	$c_{12}=7$ $R_{12}=\frac{1}{2}n(15n+1)$	$c_{13}=8$ $R_{13}=\frac{1}{2}n(17n+1)$	$c_{1.}=0$ $R_{1.}=\frac{1}{2}n(33n+3)$
B_1	$c_{21}=6$ $R_{21}=\frac{1}{2}n(13n+1)$	$c_{22}=5$ $R_{22}=\frac{1}{2}n(11n+1)$	$c_{23}=-11$ $R_{23}=\frac{1}{2}n(3n+1)$	$c_{2.}=0$ $R_{2.}=\frac{1}{2}n(27n+3)$
B_2	$c_{31}=4$ $R_{31}=\frac{1}{2}n(9n+1)$	$c_{32}=-7$ $R_{32}=\frac{1}{2}n(5n+1)$	$c_{33}=3$ $R_{33}=\frac{1}{2}n(7n+1)$	$c_{3.}=0$ $R_{3.}=\frac{1}{2}n(21n+3)$
	$c_{.1}=-5$ $R_{.1}=\frac{1}{2}n(23n+3)$	$c_{.2}=5$ $R_{.2}=\frac{1}{2}n(31n+3)$	$c_{.3}=0$ $R_{.3}=\frac{1}{2}n(27n+3)$	$R_{..}=\frac{1}{2}9n(9n+1)$

With the aid of the formulae in equation (A) we obtain:

$T_A = 32n^2/9(9n+1)$; $T_B = 8n^2/(9n+1)$;

$T_{AB} = 616n^2/9(9n+1)$.

Once again the effects of the levels of component B are zero while the value of the test statistic for component B tends to infinity as $n \rightarrow$

The possibility that an interaction effect can cause the test statistic for a main effect to reach a significant value, although no main effect exists, is illustrated by the above two examples.

The most extreme case abserved thus far is Example 1 above. In this example the test statistic for T_B has a significant value if n=2. The ratio $T_{AB}/T_B = 3/4$ independent of n. In all other examples where an incorrect decision was obtained the ratio T_{AB}/T_B was $> 3/4$.

If no interaction effect is present an increase in the effects of the levels of one main component will generally decrease the value of the test statistic for the remaining main component. This was shown in the practical investigation which is reported in §3.7.

It seems thus advisable not to attach any value to the outcome of the test for a main component (B) if the ratio $T_{AB}/T_B > \frac{1}{2}$ and to be careful if it is in the vicinity of a $\frac{1}{2}$ provided the value of T_{AB} is not significant. If the value of T_{AB} is significant or if the ratio T_{AB}/T_B is $> \frac{1}{2}$ and the main effect is significant the treatment combinations are to be considered as a one-way classification and can be compared with each other by means of multiple comparison procedures (cf. Chapter XII).

Remark: A similar result holds for the randomized block design and it seems advisable* not to attach any value to the outcome of the test statistic for a main component (B) if the ratio $T_{AB}/T_B > 1$ in this case.

In the parametric case Finney D.J. (1948) (Main effects and interactions. Jnl. Am. Statist. Ass. Vol. 43, 566-571) and Elston R.C. & Bush N. (1964) (The hypothesis that can be tested when there are interactions in an analysis of variance model. Biometrics, Vol. 20, 681-698) discussed the problem of estimating main effects when interactions are present.

* This is to be considered as preliminary advice until further investigation of this phenomenon has been undertaken e.g. by means of sampling studies.