ANALYSIS OF EXPERIMENTAL DESIGNS
WITH UNEQUAL GROUP VARIANCES

M. A. H. TALUKDER

Submitted in fulfilment of the requirements
for the degree of Ph.D. in Statistics

at the

University of Kent at Canterbury

1976

ABSTRACT

This thesis deals with weighted (generalised) least squares estimation and analysis for some common experimental designs with the error variance heteroscedastic with respect to the levels of one factor, namely, the treatments or (for split-plot designs) sub-plot treatments. The simple regression model with error variance heteroscedastic with respect to the values of the independent variable, is also considered briefly. The observations in any of the analyses considered are grouped in such a way that the error variance is constant within groups but varies from group to group.

On the assumption that the group variances are known, the weighted least squares estimators of the linear parameters and the corresponding analysis (Aitken, 1934-35; Plackett, 1960, pp. 47-49) are provided for each design or model. An expression for joint confidence intervals of parametric contrasts for the heteroscedastic models is also obtained. The estimators of the linear parameters and other statistics usually involve actual weights, the reciprocals of the group variances.

The actual weights are not usually known. The estimators of the group variances are therefore derived for each design or model. For some designs, the minimum norm quadratic unbiased estimators (Rao, 1970; 1973, pp. 303-305) of group variances are independently distributed as multiples of χ^2. For other designs, almost unbiased estimators (Horn et al., 1975) of group variances have negligible bias and are approximately independently distributed as multiples of χ^2. Reciprocals of

these estimators are used as the estimated weights.

The weighted least squares estimators of the linear parameters or variance components and other statistics including test-statistics using estimated weights, are generally biased. It is shown in the thesis how a major part of the bias can be removed; the procedure stems from a theorem due to Meier (1953). The estimators and other statistics using estimated weights are adjusted accordingly. A modified form of this theorem is also proved for correlated estimators of the group variances. A small Monte Carlo study conducted for completely randomised designs showed that the performances of the adjusted statistics are more or less satisfactory.

The designs and models covered in this thesis are: completely randomised designs, the general two-way model with proportional cell frequencies, general block designs, randomised complete block designs, latin square designs, split-plot designs with two treatment factors and the linear regression model. For the first three designs, both the fixed-effects models and random or mixed models are considered whereas only the fixed-effects models are dealt with for the remaining three designs.

CONTENTS

CHAPTER 1

INTRODUCTION AND PRELIMINARIES

1.1 Introduction

In the classical least squares theory, the error variances are assumed to be equal. For linear homoscedastic models, the least squares estimators of the parameters have some optimality properties as given in the Gauss-Markoff theorem (see John, 1971, p. 34). When the error variances differ and their values or relative values are known, the same properties are satisfied by the corresponding generalised (weighted) least squares estimators.

For variable unknown error variances, when the mean is functionally related to the error variance, variance-stabilizing transformations can be used to remove heteroscedasticity (see Bartlett, 1947, and others). Hoyle (1973) gave a detailed account, with bibliography, of different types of transformation and their uses. It has been observed from experience that such transformations often normalise the data so that the F-test remains valid.

However, the error variances may sometimes be different even if there is no reason to believe that the errors are non-normal. In animal-breeding experiments, the litters may originate from different ~~species~~ breeds and the error variance may vary from ~~species to species.~~ breed to breed. If several persons having different skills take measurements on the same objects, then it is not unreasonable to assume that the errors of measurement have ~~the same variance for each person.~~ different variances for different persons. Batches of chemicals used by an experimenter may have come from different sources and the error variance may differ from source to source.

Sometimes the treatments may not be reproduced exactly for repetition. There are then treatment errors which may have different variances for different treatments. In the data given by Fisher (1966, pp. 67-69) for a set of variety trials, Yates and Cochran (1938) found that one variety, Trebi, of Barley accounted for much of the variation due to varieties. Snedecor and Cochran (1967, p. 324) gave some examples of unequal variances due to treatment errors. Zyskind and Kempthorne (1960) considered treatment errors having unequal variances and found expectations of sums of squares over permutation distributions for some designs.

The concept of inequality of group error variances is thus quite old. In the late thirties, Bartlett (1937) proposed a method for testing the homogeneity of group variances for one-way models. Later on, Hartley (1950) gave a short-cut test. Han (1968) suggested a few methods for testing homogeneity of correlated variances. Russell and Bradley (1958), Johnson (1962), Han (1969), Maloney and Rastogi (1970) and Shukla (1972) dealt with the test of homogeneity of group variances in two-way models and Curnow (1957) with that in split-plot designs for only two sub-plot treatments.

Box (1954a and 1954b) derived some results on distributions of quadratic forms in normal variates and applied these to study the effect of inequality of group error variances on the F-test in one-way and two-way classifications. He found that moderate differences in error variances did not seriously affect the test for equal replications while much larger discrepancies were observed for unequal sample sizes.

Draper and Guttman (1966) utilized Box's results from a Bayesian point of view in one-way fixed effects models when only two different group variances are suspected. For heteroscedastic models, Box showed that the usual ratio of the error mean square to the treatment mean square was approximately distributed as a constant times an F-variate. Assuming some prior distributions of the means and variances of the populations, Draper and Guttman obtained estimators of the constants of such test-statistics. Applying standard analysis to some examples of unequal group variances, they concluded that "serious errors can result if the effects of unequal variances are ignored".

The problem of testing equality of two means when group variances are unknown and unequal, was first discussed by Behrens (1929) and Fisher (1935, 1939); the latter provided a method for such a test with the help of fiducial distributions of the parameters concerned. Welch (1938) suggested an approximate test based on the assumption that a linear function of two independent χ^2-variates is approximately distributed as a constant multiple of a χ^2-variate. Scheffé (1943) gave an exact solution to the Behrens-Fisher problem, in terms of interval estimation on the basis of a t-distribution. Welch (1947) suggested an asymptotic solution in which error of the first kind was held approximately constant.

Ghosh (1961) considered estimation of parametric functions in one-way models with unequal group variances and obtained a generalisation of Scheffé's (1943) result. Using Ghosh's result, Ghosh and Behari (1965) derived expressions for point estimators and confidence intervals for treatment contrasts in randomised block designs with groups of treatments having different variances.

Approximate test-criteria for testing equality of several means when group variances are unequal, were first given by James (1951) and Welch (1951). Using two successive Taylor's series expansions, James derived the following approximate expression for the $\alpha\%$ point:

$$\chi^2(\alpha)\left[1+\{3\chi^2(\alpha)+t+1\}\,\Sigma\{1/(r_i-1)\}(1-r_i\hat{w}_i/\Sigma r_i\hat{w}_i)^2/2(t^2-1)\right];$$

the weighted treatment sums of squares, using estimated weights, are to be compared with this quantity for testing equality of treatment means. In this expression $\chi^2(\alpha)$ is the value of χ^2 with (t-1) degrees of freedom (d.f.) at the $\alpha\%$ level of significance, t is the number of treatments, r_i is the number of replications for the ith treatment, and the estimated weight $\hat{w}_i$ is the reciprocal of the variance of the ith sample. Proceeding in the same way, James (1954) obtained approximate test criteria, again based on the χ^2 distribution, for tests of linear hypotheses for univariate and multivariate heteroscedastic models.

Welch (1951) provided another asymptotic solution, based on an F-test, to the above problem. He obtained the cumulant generating function of $F = \{\chi_1^2/(t-1)\}/(\chi_2^2/f)$, the ratio of two mean χ^2's, and took the expectation of F over χ^2_2. He then compared the cumulants, up to order $\{1/(r_i-1)\}$, of the terms of the resulting series with the corresponding terms of the cumulant generating function of the weighted treatment sum of squares; he suggested that the statistic

$$\{\sum_1^t r_i\hat{w}_i y_{i.}^2-(\Sigma r_i\hat{w}_i y_{i.})^2/\Sigma r_i\hat{w}_i\}/(t-1)\{1+2(t-2)\sum_1^t(1/(r_i-1)\times(1-r_i\hat{w}_i/\Sigma r_i\hat{w}_i)^2/(t^2-1)\}$$

with $y_{i.}$ as the mean of the ith sample, is approximately distributed as a central F under the null hypothesis with d.f. (t-1) and $f = \{3\sum_1^t\{1/(r_i-1)\}(1-r_i\hat{w}_i/\Sigma r_i\hat{w}_i)^2/(t^2-1)\}^{-1}$.

Brown and Forsythe (1974a) proposed an approximate

d.f. solution to the same problem. As both the numerator and the denominator of the statistic $\sum_1^t r_i(y_{i.}-y_{..})^2 / \{\sum_1^t (1-r_i/n)/\hat{w}_i\}$ with $y_{..} = \Sigma y_{i.}/t$ and $n = \Sigma r_i$, have the same expectation, they suggested, following Satterthwaite (1941), that this statistic is approximately distributed as an F with (t-1) and f_o d.f. under the null hypothesis where $f_o = 1/[\Sigma c_i^2/(r_i-1)]$ with $c_i = \{(1-r_i/n)/\hat{w}_i\}/\{\sum_i (1-r_i/n)/\hat{w}_i\}$.
From a Monte Carlo study, they found that the performances of their test-statistic and that of Welch (1951) were satisfactory for more than 10 observations per group and were not unreasonable for samples of sizes down to 5. They also offered some suggestions for evolving an improved test-statistic which would be useful in all situations including small samples. Brown and Forsythe (1974b) showed that their test-statistic mentioned above could be derived by combining orthogonal contrasts of treatments. The method was extended to two-way designs with unequal cell variances. They also proposed a method of obtaining a joint confidence interval for contrasts between treatment means.

Chakravarti (1965) showed that Hotelling's T^2 statistic could be used to test the hypotheses in respect of linear contrasts of the treatments in one-way heteroscedastic models. Such tests are valid when the number of treatments does not exceed the minimum number of replications.

For one-way models with unknown group variances, Spjøtvoll (1972) derived an approximate expression for the joint confidence interval of all contrasts of the treatment means. If ψ is any such contrast, then this joint confidence interval is

$$\hat{\psi} - A\,\hat{\sigma}_{\hat{\psi}} \leq \psi \leq \hat{\psi} + A\,\hat{\sigma}_{\hat{\psi}} ,$$

with $\sigma_{\hat{\psi}}$ as the estimated standard error of the estimator $\hat{\psi}$ of ψ and $A = \{aF_{\alpha}(t,b)\}^{\frac{1}{2}}$. The expressions for a and b in terms of individual d.f. were obtained by equating the first two cumulants of $\sum_i F(1, r_i - 1)$ to those of a $F(t, b)$.

For two-way heteroscedastic models, some methods of testing hypotheses were suggested by several authors besides Brown and Forsythe (1974b) mentioned above. Graybill (1954) considered randomised complete block designs assuming the errors to be heteroscedastic between treatment effects and correlated within each block. Subtracting the data for any one treatment from the corresponding data for each of the other treatments, he showed that Hotelling's T^2 statistic could be used for testing the treatment differences. The test is valid when there are more blocks than treatments.

Siotani (1957) dealt with replicated randomised complete block designs assuming the errors in any one experiment to be correlated and heteroscedastic but independent between the designs. Following Graybill (1954), he obtained tests of significance for main effects and interactions based again on a T^2 statistic.

Robinson and Balaam (1967) considered the same model as that of Graybill (1954) for each of a number of replicated complete block designs and gave a method of analysis, based on likelihood ratio tests, that uses the independent contrasts of observations under each treatment.

Schlesselman (1973) proposed a procedure for choosing a power transformation of observations of the replicated two-way designs when the usual assumptions of analysis of variance are not satisfied. To obtain such transformations, he suggested a weighted combination of Tukey's statistic for removable non-additivity and the t-statistic for testing the slope of log (sample cell variance) on log (sample cell mean). His method was then empirically compared with that of Box and Cox (1964). Point estimates for both procedures were emprically found to be the same on the average over many sets of data obtained through simulation.

Duby et al. (1975) gave a method for analysing the data of two-way designs when the cell variances are functions of the cell means. The method is based on Wald's (1943) large sample test criterion.

For general heteroscedastic linear models, Williams (1967) derived approximate variances of weighted least squares estimators using estimated weights based on equal replications. Bement and Williams (1968) extended these results to the case of unequal replications.

Williams (1959, pp. 67-70) and Draper and Smith (1966, pp. 77-81) discussed the weighted least squares method for estimating the linear parameters of heteroscedastic regression models. Jacquez et al. (1968), Rao and Subrahmaniam (1971) and Jacquez and Norusis (1973) undertook Monte Carlo studies on the efficiency of the weighted estimators of the parameters of linear regression models with unequal group variances.

For the experimental designs considered in this thesis, it is assumed that the error variance is heteroscedastic with respect to the levels of only one factor, namely the treatments or (for split-plot designs) sub-plot treatments. For the regression models, the error variance is assumed to be heteroscedastic with respect to the values of the independent variable. Thus the error variance is constant for the group of observations under each level of treatments or each value of the independent variable and varies from group to group. The methods are also applicable when the error variance is heteroscedastic with respect to the levels of any other main effect.

When the error variance is the same within a group of observations but varies from group to group under a linear model, some methods are available for estimating the error variances from a sample. The estimators of the error variances may then be used for obtaining the weighted least squares estimators of the linear parameters. Such weighted estimators will generally be biased. Similarly, use of estimated weights introduces unknown bias in other statistics including test-statistics for the analysis of data with heteroscedastic models. In this situation, one method is to remove much of the resulting bias of such weighted estimators and statistics for these to be of practical use.

In this thesis, the weighted least squares analysis (Aitken, 1934-35; Plackett, 1960, pp. 47-49) is given for each of several common designs, assuming the group variances to be known. The estimators (Rao, 1970, 1973, pp. 303-305; Horn et al., 1975) of the error variances are obtained. The weighted least squares estimators of the linear parameters and other statistics using estimated weights are adjusted for removing a major portion of the bias with the help of a theorem due to Meier (1953). A report on a small Monte Carlo study on the adequacy of the adjusted statistics for one-way heteroscedastic models is also given.

1.2 General principle of weighted (generalised) least squares analysis when the error variances are known

Let us consider the heteroscedastic linear model

$$\underset{\sim}{Y} = \underset{\sim}{X}'\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad \ldots\ldots\ldots\ldots \quad (1)$$

where $\underset{\sim}{Y}$ is the vector of observations, $\underset{\sim}{X}'$ the design matrix, $\underset{\sim}{\beta}$ the vector of linear parameters and $\underset{\sim}{\varepsilon}$ the vector of errors such that $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and var $(\underset{\sim}{\varepsilon}) = \text{diag}\,(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2) = \underset{\sim}{V}$, say, the error variances, σ_i^2, being the diagonal elements and n the number of observations. The error variances may not be all distinct. The matrix $\underset{\sim}{V}$ is non-singular.

If the error variances are known, then the weighted least squares estimator of the parameter vector $\underset{\sim}{\beta}$ is obtained by minimising the quadratic form $\underset{\sim}{\varepsilon}'\,\underset{\sim}{V}^{-1}\,\underset{\sim}{\varepsilon} = (\underset{\sim}{Y} - \underset{\sim}{X}'\underset{\sim}{\beta})'\underset{\sim}{V}^{-1}(\underset{\sim}{Y} - \underset{\sim}{X}'\underset{\sim}{\beta})$. Taking the derivative of the right hand side with respect to $\underset{\sim}{\beta}$ and setting it equal to zero, we get

$$\underset{\sim}{X}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{X}'\,\tilde{\underset{\sim}{\beta}} = \underset{\sim}{X}\,\underset{\sim}{V}^{-1}\underset{\sim}{Y} \quad \ldots\ldots\ldots \quad (2)$$

as the normal equations for finding the weighted estimator $\tilde{\underset{\sim}{\beta}}$ of $\underset{\sim}{\beta}$. Such an equation was first given by Aitken (1934-35) and then the principle was further developed by others, e.g. Goldman and Zellen (1964), to cover different cases. When $\underset{\sim}{V} = \sigma^2\,\underset{\sim}{I}$, this reduces to the normal equations of the simple least squares procedure.

Now define the weight $w_i = 1/\sigma_i^2$, $i = 1, 2, \ldots, n$, and $\underset{\sim}{V}^{-1} = \underset{\sim}{W}^{\delta}$, a diagonal matrix with w_i as the diagonal elements. Also let $\underset{\sim}{W}^{\delta/2}\underset{\sim}{X}' = \underset{\sim}{A}'$ and $\underset{\sim}{W}^{\delta/2}\underset{\sim}{Y} = \underset{\sim}{Z}$ where $\underset{\sim}{W}^{\delta/2}$ is the diagonal matrix with $w_i^{\frac{1}{2}}$ as the diagonal elements. Then $\text{var}(\underset{\sim}{Z}) = \underset{\sim}{I}$ and the normal equations (2) become

$$\underset{\sim}{A}\,\underset{\sim}{A}'\,\tilde{\underset{\sim}{\beta}} = \underset{\sim}{A}\,\underset{\sim}{Z}\ .$$

These are the normal equations of the simple least squares in transformed data so that the estimators possess optimality properties as mentioned at the beginning of section 1.1.

It also follows that the sum of squares (SS) due to the estimates, namely SS $(\beta) = \tilde{\beta}' A Z = \tilde{\beta}' X w^{\delta} Y$ and the SS due to error, namely SS $(E) = Z'Z - \tilde{\beta}' A Z = Y' w^{\delta} Y - \tilde{\beta}' X w^{\delta} Y = \tilde{\varepsilon}' V^{-1} \tilde{\varepsilon}$, with $\tilde{\varepsilon} = Y - X' \tilde{\beta}$, are independent. Moreover, since $E\{SS(\beta)\} = \beta' X V^{-1} X' \beta + \text{rank}(X')$ and $E\{SS(E)\} = n - \text{rank}(X')$, the SS due to estimates and the SS due to error are distributed as non-central and central χ^2 variables respectively with the corresponding degrees of freedom given by rank (X') and n-rank (X'). Thus the usual F-test can be used to test the hypothesis:

$$\beta = 0 .$$

(See Plackett, 1960, pp. 47-49)

This hypothesis can also be tested by a χ^2-test using $SS(\beta)$ only.

1.3 Methods of estimation of weights

As we are considering group variances, the variance model of the error term in equation (1), when the observations are arranged treatment by treatment, can be written as

$$\text{var}(\varepsilon) = V = V_1 \sigma_1^2 + \ldots + V_m \sigma_m^2 \qquad (3)$$

where the quantities σ_i^2 are the group error variances, and the matrices V_i are diagonal matrices having the form $V_i = \text{diag}(0, \ldots, 0, 1, \ldots 1, 0, \ldots, 0)$. The matrices V_i are idempotent and orthogonal, and sum to I.

Such a model was given by Nelder (1965, 1968) for variance components under orthogonal block structures. There the matrices V_i defined m strata of the analysis. Similar variance component models were considered by Hartley and Rao (1967) and Patterson and Thompson (1971, 1975).

The following are the methods of estimating σ_i^2.

(i) The MINQUE method of Rao (1970, 1973)

Rao defined the minimum norm quadratic unbiased estimator (MINQUE) of σ_i^2 by the quadratic form $\underset{\sim}{Y}' \underset{\sim}{A}_i \underset{\sim}{Y}$ where $\underset{\sim}{A}_i$ are matrices chosen in such a way that $tr(\underset{\sim}{A}_i \underset{\sim}{U})^2$ is minimised for all i. Here

$$\underset{\sim}{U} = \alpha_1^2 \underset{\sim}{V}_1 + \cdots + \alpha_m^2 \underset{\sim}{V}_m$$

and the minimisation is subject to the condition that $E(\underset{\sim}{Y}' \underset{\sim}{A}_i \underset{\sim}{Y}) = \sigma_i^2$. In general the estimates of σ_i^2 depend on the choice of α_i. Rao (1973) recommended that α_i^2 should be chosen approximately proportional to σ_i^2 wherever possible. In the absence of any prior information about σ_i^2, α_i may be taken to be unity.

As $tr(\underset{\sim}{A}_i \underset{\sim}{U})^2$ is the square Euclidian norm, the method is called 'minimum norm'.

Rao (1970) gave a computational method for obtaining such estimates. Let the projection matrix be $\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{X}' (\underset{\sim}{X}\underset{\sim}{X}')^- \underset{\sim}{X} = (s_{ij})$, A^- being any generalised inverse of A. Further let $\underset{\sim}{v}$ be the vector of squares of the residuals given by $\underset{\sim}{S}\underset{\sim}{Y}$, $\underset{\sim}{\delta}$ the vector of variances $\sigma_1^2, \ldots, \sigma_n^2$ and $\underset{\sim}{F} = \{s_{ij}^2\}$. Then the MINQUE's of σ_i^2 are obtained from the equation $\underset{\sim}{F}\,\underset{\sim}{\delta} = \underset{\sim}{v}$ when $\underset{\sim}{F}$ is non-singular. He also suggested that the group error variances can be estimated by solving the reduced equations obtained by adding the set of equations which correspond to the same variance.

Mallela (1972) derived necessary and sufficient conditions for $\underset{\sim}{F}$ to be non-singular. In this thesis, the coefficient matrix of the reduced equations for estimating the group variances will always be non-singular.

Horn et al.(1975) suggested almost unbiased estimators of variances and showed how these could be obtained from corresponding MINQUE's.

(ii) The method of maximum likelihood

Under the assumption of normality of errors, the likelihood function of the observations is given by

$$L = (2\pi)^{-n/2} |\underset{\sim}{V}|^{-1/2} \exp \{ -\tfrac{1}{2} (\underset{\sim}{Y} - \underset{\sim}{X}' \underset{\sim}{\beta})' \underset{\sim}{V}^{-1} (\underset{\sim}{Y} - \underset{\sim}{X}' \underset{\sim}{\beta}) \} .$$

The maximum likelihood method of estimating the linear parameter vector $\underset{\sim}{\beta}$ gives the same normal equations as the weighted least squares procedure. Following Hartley and Rao (1967), we find the equations for obtaining the maximum likelihood estimators of σ_i^2 as

$$\mathrm{tr}\, (\underset{\sim}{V}^{-1} \frac{\partial \underset{\sim}{V}}{\partial \sigma_i^2}) + (\underset{\sim}{Y} - \underset{\sim}{X}' \underset{\sim}{\tilde{\beta}})' \frac{\partial (\underset{\sim}{V}^{-1})}{\partial \sigma_i^2} (\underset{\sim}{Y} - \underset{\sim}{X}' \underset{\sim}{\tilde{\beta}}) = \underset{\sim}{0}; \quad i=1,2,..,m,$$

where $\underset{\sim}{\tilde{\beta}}$ is the weighted least squares estimator of $\underset{\sim}{\beta}$. The estimated variances are usually in terms of the estimators of the linear parameters and may be evaluated by an iterative method when the process converges.

(iii) The method of modified maximum likelihood

Patterson and Thompson (1971; 1975, pp. 197-207) proposed the method of modified maximum likelihood for estimating variance components $\sigma_1^2, \ldots, \sigma_m^2$, as in (3), but with $\underset{\sim}{V}$ singular in general. They suggested partitioning of the data into two parts - one represented by the transformed observations (residuals) $\underset{\sim}{S}\ \underset{\sim}{Y}$ and the other by $\underset{\sim}{Q}\ \underset{\sim}{Y}$ where $\underset{\sim}{Q}$ is such that $\mathrm{cov}\,(\underset{\sim}{S}\underset{\sim}{Y}, \underset{\sim}{Q}\underset{\sim}{Y}) = \underset{\sim}{0}$. The variance components were then estimated by maximising the likelihood of $\underset{\sim}{S}\underset{\sim}{Y}$ and $\underset{\sim}{\beta}$ by maximising that of $\underset{\sim}{Q}\underset{\sim}{Y}$. Patterson and Thompson (1975) suggested that the estimate of σ_i^2 should be obtained by equating $\underset{\sim}{Y}' (\underset{\sim}{S}\underset{\sim}{V}\underset{\sim}{S})^{+} \underset{\sim}{V}_i (\underset{\sim}{S}\underset{\sim}{V}\underset{\sim}{S})^{+} \underset{\sim}{Y}$ to its expectation, $i=1,2,\ldots,m$.

An iterative method was suggested for finding the actual estimates. Here $\underset{\sim}{A}^+$ denotes the unique Moore-Penrose (Moore, 1920, 1935; Penrose, 1955) generalised inverse of $\underset{\sim}{A}$.

(iv) The method of Nelder (1968)

As proposed by Nelder (1968) for the same model (3) in a different context, σ_i^2 can be estimated by equating the sums of squares $\underset{\sim}{Y}'\underset{\sim}{R}'\underset{\sim}{V}_i\underset{\sim}{R}\underset{\sim}{Y}$ to their expectations, $i=1,2,\ldots,m$, where $\underset{\sim}{R}= \underset{\sim}{I}-\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{V}^{-1}\underset{\sim}{X}')^+\underset{\sim}{X}\underset{\sim}{V}^{-1}$ (see Patterson and Thompson, 1975). Almost all the authors cited above suggested feedback of information for estimating the linear parameters.

It was shown by Patterson and Thompson (1975) that a single iteration in the solution for their estimate is equivalent to the MINQUE procedure and that their method gives the same results as those of Nelder's method.

In view of this fact and also because of the simpler algebraic procedure for obtaining MINQUE possessing some desirable properties, we have considered only the MINQUE method of estimation of the group variances in most of the cases studied in this thesis. The method of maximum likelihood estimation is also considered in some cases where simple expressions could be obtained for such estimators. Almost unbiased estimators (Horn et al.,1975) of error variances are also obtained from corresponding MINQUE's for two designs.

CHAPTER 2

COMPLETELY RANDOMISED DESIGNS

For fixed-effects one-way models with known unequal group variances, estimation and analysis are dealt with by the weighted least squares method. The estimators of the group variances are obtained and the test-statistics, using estimated weights, are adjusted for removing a major part of the bias of such statistics. A formula for a joint confidence interval of all contrasts of treatments and a report on a small Monte Carlo study are provided for such models. Finally, estimation and analysis for mixed and random models with unequal group error variances are discussed.

2.1 <u>One-way fixed-effects models</u>

2.1.1 <u>Weighted (generalised) least squares analysis when the group variances are known</u>

It is assumed that there are t treatments of which the ith treatment is applied to r_i plots in an experiment. Let the observations of such an experiment be expressed by the linear model*:

$$y_{ij}=\mu_i+\varepsilon_{ij} \ ; \ j=1,2,\ldots,r_i, r_i>1; \quad i=1,2,\ldots,t. \quad \ldots \ldots \quad (4)$$

where μ_i is the population mean for the ith treatment and ε_{ij} the error term having mean zero and variance σ_i^2 which in general differs from treatment to treatment. The errors are assumed to be independent of one another. For the ith treatment, there are r_i observations, which are different in general.

* Suggested by Dr D. A. Preece

Let $n = \sum_{1}^{t} r_i$.

If $\underset{\sim}{Y} = (y_{11},\ldots,y_{1r_1},\ldots,y_{t1},\ldots,y_{tr_t})'$ is the column vector of observations arranged treatment by treatment, then the above model can be written as

$$\underset{\sim}{Y} = \underset{\sim}{X}'\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{\beta}$ is the column vector of treatment means, $\underset{\sim}{X}'$ the design matrix and $\underset{\sim}{\varepsilon}$ the column vector of errors. The design matrix is of full rank = t and

$$\text{var}\,(\underset{\sim}{\varepsilon}) = \text{diag}\,(\sigma_1^2,\ldots,\sigma_1^2,\ldots,\sigma_t^2,\ldots,\sigma_t^2) = \underset{\sim}{V},$$

say. The variance model can be written as

$$\underset{\sim}{V} = \sigma_1^2\underset{\sim}{V}_1 + \ldots + \sigma_t^2\underset{\sim}{V}_t$$

where $\underset{\sim}{V}_i = \text{diag}(0,\ldots,0,1,\ldots 1,0,\ldots,0)$ with unity occurring r_i times after $\sum_{k=1}^{i-1} r_k$ places in the main diagonal. The matrices $\underset{\sim}{V}_i$ are symmetric, idempotent and independent, and sum to $\underset{\sim}{I}$.

By (2) of section 1.2, the normal equation for estimating μ_i by the weighted least squares method is given by

$$r_i w_i \hat{\mu}_i = w_i Y_{i.} \quad ; \quad i = 1,2,\ldots,t.$$

Hence, $\hat{\mu}_i = y_{i.}$; $i = 1,2,\ldots t$.

Here, we have used the convention that the dot suffix of a small letter denotes the mean and that of a capital letter the total over the corresponding variable suffix. This convention will be followed all through. The weight

$$w_i = 1/\sigma_i^2,\ i = 1,2,\ldots,t .$$

The estimators of the treatment means are thus independent of the weights and also of each other.

The sum of squares (SS) due to the estimates is given by

$$\text{SS(Est.)} = \sum_{1}^{t} w_i Y_{i.}^2/r_i$$

with t degrees of freedom (d.f.) and that due to error by

$$\text{SS(E)} = \sum_{ij}\sum w_i y_{ij}^2 - \sum_{i} w_i Y_{i.}^2/r_i$$

$$= \sum_{ij}\sum w_i(y_{ij} - y_{i\cdot})^2$$

with (n-t)d.f. Under the hypothesis of the equality of the treatment means i.e. $\mu_i = \mu$, the model at (4) reduces to $y_{ij} = \mu + \varepsilon_{ij}$. The weighted least squares estimator of the general mean is then given by

$$\hat{\mu} = \Sigma w_i Y_{i\cdot}/\Sigma r_i w_i$$

and the corresponding sum of squares by

$$\text{SS due to mean} = (\Sigma w_i Y_{i\cdot})^2/\Sigma r_i w_i$$

with 1 d.f. The sum of squares due to treatments corrected for the mean is thus obtained as

$$\begin{aligned} SS(\text{treat}) &= \Sigma w_i Y_{i\cdot}^2/r_i - (\Sigma w_i Y_{i\cdot})^2/\Sigma r_i w_i \\ &= \Sigma w_i r_i (y_{i\cdot} - \tilde{y}_{\cdot\cdot})^2 \end{aligned}$$

with (t-1)d.f., where $\tilde{y}_{\cdot\cdot} = \Sigma r_i w_i y_{i\cdot}/\Sigma r_i w_i$

Since $y_{i\cdot} = \mu_i + \varepsilon_{i\cdot}$ and $\tilde{y}_{\cdot\cdot} = \tilde{\mu} + \tilde{\varepsilon}_{\cdot\cdot}$ from the model at (4) with $\tilde{\mu} = \Sigma r_i w_i \mu_i/\Sigma r_i w_i$ and $\tilde{\varepsilon}_{\cdot\cdot} = \Sigma r_i w_i \varepsilon_{i\cdot}/\Sigma r_i w_i$,

we have, $$\begin{aligned} E\{SS(\text{treat})\} &= \Sigma r_i w_i (\mu_i - \tilde{\mu})^2 + \Sigma r_i w_i E(\varepsilon_{i\cdot} - \tilde{\varepsilon}_{\cdot\cdot})^2 \\ &= \Sigma r_i w_i (\mu_i - \tilde{\mu})^2 + (t-1) \end{aligned}$$

Moreover, $$\begin{aligned} E\{SS(E)\} &= E\{\sum_{ij}\sum w_i (y_{ij} - y_{i\cdot})^2\} \\ &= E\{\sum_{ij}\sum w_i (\varepsilon_{ij} - \varepsilon_{i\cdot})^2\} \\ &= n - t. \end{aligned}$$

Analysis of variance table

Source	d.f.	SS	E(MS)
Treat.	t-1	$\sum_i w_i r_i (y_{i\cdot} - \tilde{y}_{\cdot\cdot})^2$	$1 + \sum_1^t r_i w_i (\mu_i - \tilde{\mu})^2/(t-1)$
Error	n-t	$\sum_{ij}\sum w_i (y_{ij} - y_{i\cdot})^2$	1

Once an F-test or a χ^2-test has shown significant differences among the treatments, a normal test can be used to test the difference between the ith and jth treatment means using the fact that $z = (y_{i.}-y_{j.})/[1/r_i w_i + 1/r_j w_j]^{\frac{1}{2}}$ is a standardised normal variate under the null hypothesis. The ratio of this normal variate to the square root of the error mean square is the corresponding t-variate with n-t d.f.

2.1.2 <u>An exact test for equally replicated treatments when the group variances are not known.</u>

Let the tr observations be grouped arbitrarily into k replicates.

Let $\underset{\sim}{y}_k = (y_{1k}, \ldots, y_{tk})'$ be the vector of t observations at the kth replicate, $k = 1, 2, \ldots, r$. Then the vector $\underset{\sim}{y}_k$ is distributed as multivariate normal with mean vector $\underset{\sim}{\mu} = (\mu_1, \ldots, \mu_t)'$ and dispersion matrix $\underset{\sim}{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_t^2)$.

Let $\underset{\sim}{C}$ be any $(t-1) \times t$ matrix of rank $(t-1)$ such that $\underset{\sim}{C}\,\underset{\sim}{1} = \underset{\sim}{0}$ where $\underset{\sim}{1}$ is the vector with unity as its elements.

Let $\underset{\sim}{z}_k = \underset{\sim}{C}\,\underset{\sim}{y}_k$. Then $\underset{\sim}{z}_k$ is distributed as multivariate normal with mean vector, $\underset{\sim}{C\mu}$, and dispersion matrix $\underset{\sim}{C}\,\underset{\sim}{\Sigma}\,\underset{\sim}{C}'$ where $\underset{\sim}{C\Sigma C}'$ is non-diagonal. Hotelling's T^2-test is applicable here. The vector $\underset{\sim}{z}_k$ is the vector of $(t-1)$ independent contrasts of t observations of the vector $\underset{\sim}{y}_k$.

To test the hypothesis of equality of treatment means is the same as to test the hypothesis: $\underset{\sim}{C}\,\underset{\sim}{\mu} = 0$.

Thus,

$$T^2 = r\, \underset{\sim}{z}.'\, \underset{\sim}{S}^{-1}\, \underset{\sim}{z}.$$

is the Hotelling's generalised T^2-statistic with $(r-1)$d.f. for a $(t-1)$-dimensional distribution, where

$$\underset{\sim}{z}. = \sum_1^r \underset{\sim}{z}_k/r \text{ and } \underset{\sim}{S} = \sum_1^r (\underset{\sim}{z}_k - \underset{\sim}{z}.)(\underset{\sim}{z}_k - \underset{\sim}{z}.)'/(r-1)$$

Hence, $T^2(r-t+1)/(r-1)(t-1)$ is a central F-variate with $(t-1)$ and $(r-t+1)$d.f. under the null hypothesis.

The test-statistic is independent of the choice of the matrix $\underset{\sim}{C}$ (see Anderson, 1958, pp.110-111). The test is possible only when $r \geq t$.

The test was first given by Chakravarti (1965).

2.1.3 Estimation of error variances

For unknown error variances, the above test is not applicable when the replications r_i are not all equal and/or when $r < t$. In such situations, one may use estimators of the error variances in place of the actual ones. It is well-known that

$$s_i^2 = \sum_{j=1}^{r_i} (y_{ij}-y_{i.})^2/(r_i - 1)$$ is an unbiased estimator

of σ_i^2. It is shown below that s_i^2 is also the MINQUE.

(i) The maximum likelihood estimator

From section 1.3, we obtain the maximum likelihood estimator of σ_i^2 as

$$\hat{\sigma}_i^2 = \sum_{j=1}^{r_i}(y_{ij} - y_{i\cdot})^2/r_i \ , \quad i = 1,2,\ldots,t.$$

This is the familiar maximum likelihood estimator (MLE) of σ_i^2 for the ith population when considered singly. The estimators are independent of one another.

(ii) The MINQUE of error variances

Since $\underset{\sim}{X}'$ is of full rank=t, we have

$$\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X} = \text{diag}\ (\underset{\sim}{J}_{r_1}/r_1, \ldots, \underset{\sim}{J}_{r_t}/r_t)$$

where $\underset{\sim}{J}_{r_i}$ is the square matrix of order r_i with unity as its elements.

Hence, the resulting projection matrix $\underset{\sim}{S}$ is given by

$$\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{X}'\ (\underset{\sim}{X}\underset{\sim}{X}')^{-1}\ \underset{\sim}{X}$$

$$= \text{diag}\ (\underset{\sim}{I}_{r_1} - \underset{\sim}{J}_{r_1}/r_1, \ldots\ldots, \underset{\sim}{I}_{r_t} - \underset{\sim}{J}_{r_t}/r_t)$$

where $\underset{\sim}{I}_{r_i}$ is the identity matrix of order r_i. The elements of the vector $\underset{\sim}{S}\underset{\sim}{Y}$ are the observed residuals.

Then the normal equations for obtaining the MINQUE, $\tilde{\sigma}_i^2$ of σ_i^2 are given by $\underset{\sim}{F}\ \underset{\sim}{\delta} = \underset{\sim}{\nu}$ where $\underset{\sim}{F}$ is the matrix of the squares of the elements of the projection matrix, $\underset{\sim}{\delta}$ is the vector of variances with $\tilde{\sigma}_i^2$ repeated r_i times,

and $\underset{\sim}{\nu}$ is the vector of the squares of the residuals $(y_{ij} - y_{i\cdot})^2$. The ith set of equations involving $\tilde{\sigma}_i^2$ is given by

$$
\begin{array}{ccc}
(1 - {}^1/r_i)^2\, \tilde{\sigma}_i^2 & + (r_i - 1)\, \tilde{\sigma}_i^2/r_i^2 = & (y_{i1} - y_{i\cdot})^2 \\
\vdots & \vdots & \vdots \\
(1 - {}^1/r_i)^2\, \tilde{\sigma}_i^2 & + (r_i - 1)\tilde{\sigma}_i^2/r_i^2 = & (y_{ir_i} - y_{i\cdot})^2
\end{array}
$$

whence, on adding the equations,

$$
\begin{aligned}
\tilde{\sigma}_i^2 &= \sum_{j=1}^{r_i} (y_{ij} - y_{i\cdot})^2/(r_i - 1) \\
&= s_i^2 \quad ; \quad i = 1,2,\ldots,t.
\end{aligned}
$$

Thus the MINQUE of σ_i^2 is the familiar unbiased estimator, s_i^2, of σ_i^2 for the ith population when considered individually. We shall denote the MINQUE of σ_i^2 by s_i^2 . Like maximum likelihood estimators, the estimators s_i^2 are also independent of one another. As is well-known, the variate $(r_i-1)s_i^2/\sigma_i^2$ is distributed as χ^2 with (r_i-1)d.f.

Feedback of information is not necessary for the treatment estimators since these are independent of error variances. Bartlett's χ^2-test can be applied to test the homogeneity of error variances in any particular situation.

2.1.4 Adjustment of the test-statistics using estimated weights

The F-statistic in the analysis of variance of the weighted least squares and the normal test-statistic for testing the difference between any two treatment means, involve actual weights, the reciprocals of error variances. If the estimators of error variances are used in place of actual ones in these test statistics, then bias will be introduced. It is difficult to obtain the magnitudes of these biases analytically. But, since the estimators of error variances are independent, bias of order $\Sigma\{1/(r_i-1)\}$ can be eliminated by adjusting these statistics with the help of the following theorem due to Meier (1953).

Theorem 1. If $x_i, i = 1,2,\ldots t$, are independently distributed random variables with probability density functions

$$f_{n_i}(x_i) = \frac{(\frac{1}{2}n_i)^{\frac{n_i}{2}}}{\Gamma(\frac{n_i}{2})} \; x_i^{\left(\frac{n_i}{2} - 1\right)} \; e^{-\frac{1}{2} n_i x_i} \quad , \quad 0 \le x_i < \infty$$

and $R(x_1,\ldots,x_t)$ is a rational function with no singularities for $0 < x_1,\ldots,x_t < \infty$ then $E[R(x_1,\ldots,x_t)]$ can be expanded in an asymptotic series in the $1/n_i$. In particular

$$E\left[R(x_1,\ldots,x_t)\right] = R[1,\ldots,1] + \sum_{i=1}^{t} \frac{1}{n_i} \left[\frac{\partial^2 R}{\partial x_i^2}\right]_{\text{all } x_i=1} + O\left(\Sigma \frac{1}{n_i^2}\right) .$$

The result is based on Taylor's series expansion of the function $R(x_1,\ldots,x_t)$. This theorem implies that the adjusted statistic $R\left[x_1,\ldots,x_t\right] - \sum_{i=1}^{t} \left[\frac{\partial^2 R}{\partial x_i^2}\right]_{\text{all } x_i=1} \frac{1}{n_i}$, being free from terms of order $(\Sigma\frac{1}{n_i})$, approximates the actual value, $R\left[1,\ldots,1\right]$ of the function more closely than $R\left[x_1,\ldots,x_t\right]$ itself. In practice, actual weights are to be replaced by the corresponding estimated weights in the term $\sum_i \frac{1}{n_i}\left[\frac{\partial^2 R}{\partial x_i^2}\right]_{\text{all } x_i=1}$

In our case, $x_i = s_i^2/\sigma_i^2$ where s_i^2 is either the MINQUE or the MLE of σ_i^2, $i = 1,2,\ldots,t$. The estimated weights are: $\hat{w}_i = 1/s_i^2 = 1/x_i\sigma_i^2$, and $n_i = r_i - 1$.

(i) Adjusted F-statistic

The error SS using estimated weights based on the MINQUE's of error variances is $\sum_i \sum_j \hat{w}_i (y_{ij} - y_{i\cdot})^2$

$$= \sum_i \{\sum_j (y_{ij}-y_{i\cdot})^2 / \sum_j \frac{(y_{ij}-y_{i\cdot})^2}{r_i - 1}\} = n - t,$$

a constant. Similarly, the error SS using the estimated weights based on the MLE is also a constant. Thus, only the treatment SS is to be adjusted for adjusting the F-statistic.

The weighted treatment SS using estimated weights is

$$\Sigma r_i \hat{w}_i (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2$$

$$= \frac{r_i}{x_i \sigma_i^2} (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2 + \sum_{k \neq i} \frac{r_k}{x_k \sigma_k^2} (y_{k\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2$$

where $\hat{w}_i = 1/x_i\sigma_i^2$ and $\hat{\tilde{y}}_{\cdot\cdot} = \Sigma r_i \hat{w}_i y_{i\cdot} / \Sigma r_i \hat{w}_i$.

Now, we have $\dfrac{\partial (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2}{\partial x_i} = \dfrac{2r_i}{\sigma_i^2 x_i^2 \hat{w}_\cdot} (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2$

and $\dfrac{\partial (y_{k\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2}{\partial x_i} = \dfrac{2r_i}{\sigma_i^2 x_i^2 \hat{w}_\cdot} (y_{k\cdot} - \hat{\tilde{y}}_{\cdot\cdot})(y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})$

where $\hat{w}_\cdot = \sum_1^t r_i \hat{w}_i$. Thus

$$\frac{\partial (\text{Tr.SS})}{\partial x_i} = - \frac{r_i (y_i - \hat{\tilde{y}}_{\cdot\cdot})^2}{\sigma_i^2 x_i^2} + \frac{2 r_i^2 (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2}{\sigma_i^4 x_i^3 \hat{w}_\cdot} +$$

$$\frac{2 r_i (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})}{\sigma_i^2 x_i^2 \hat{w}_\cdot} \sum_{k \neq i} \frac{r_k (y_{k\cdot} - \hat{\tilde{y}}_{\cdot\cdot})}{x_k \sigma_k^2}$$

$$= - \frac{r_i (y_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})^2}{\sigma_i^2 x_i^2} .$$

Taking partial derivative of this again and putting $x_i = 1$ for all i and simplifying, we get,

$$\left[\frac{\partial^2(\text{Treat.SS})}{\partial x_i^{\,2}}\right]_{\text{all } x_i=1} = 2\, r_i w_i (y_{i.} - \tilde{y}_{..})^2 \left(1 - \frac{r_i w_i}{w.}\right)$$

where

$$w. = \Sigma\, r_i w_i \, .$$

Hence, by the above Theorem 1, treat. SS (adj)

$$= \sum_1^t r_i \hat{w}_i (y_{i}. - \hat{\tilde{y}}_{..})^2 - \sum_1^t \frac{2 r_i \hat{w}_i}{r_i - 1} (y_i. - \hat{\tilde{y}}_{..})^2 \left(1 - \frac{r_i \hat{w}_i}{\hat{w}.}\right)$$

$$= \sum_1^t r_i \hat{w}_i (y_i. - \hat{\tilde{y}}_{..})^2 \left\{ 1 - \frac{2}{r_i - 1}\left(1 - \frac{r_i \hat{w}_i}{\hat{w}.}\right) \right\} .$$

Thus, $\hat{F}$ (adj)

$$= \frac{(n-t)\left\{ \sum_i r_i \hat{w}_i (y_i. - \hat{\tilde{y}}_{..})^2 \right\} \left[1 - \frac{2}{r_i - 1}\left(1 - \frac{r_i \hat{w}_i}{\hat{w}.}\right)\right]}{(t-1)\left\{ \sum_i \sum_j \hat{w}_i (y_{ij} - y_i.)^2 \right\}}$$

with (t-1) and (n-t) d.f.

(ii) Adjusted approximate normal test-statistic

Let $\hat{z} = |y_{\ell .} - y_{k .}| \,/\, \{1/r_\ell \hat{w}_\ell + 1/r_k \hat{w}_k\}^{\frac{1}{2}}$

be the approximate normal test-statistic using estimated weights for testing the difference between the ℓth and kth treatment means.

Then the partial derivatives are given as

$$\frac{\partial \hat{z}}{\partial x_i} = |y_{\ell .} - y_{k .}| \left(-\tfrac{1}{2}\right)\left(\frac{x_\ell \sigma_\ell^{\,2}}{r_\ell} + \frac{x_k \sigma_k^{\,2}}{r_k} \right)^{-3/2} \frac{\sigma_i^{\,2}}{r_i}$$

and

$$\left[\frac{\partial^2 \hat{z}}{\partial x_i^{\,2}}\right]_{\text{all } x_i=1} = \left[|y_{\ell .} - y_{k .}| \left(-\tfrac{1}{2}\right)\left(-\tfrac{3}{2}\right)\left(\frac{x_\ell \sigma_\ell^{\,2}}{r_\ell} + \frac{x_k \sigma_k^{\,2}}{r_k} \right)^{-5/2} \frac{\sigma_i^{\,2}}{r_i} \frac{\sigma_i^{\,2}}{r_i} \right]_{\text{all} x_i=1}$$

$$= 3\, |y_{\ell .} - y_{k .}| \,(1/r_\ell w_\ell + 1/r_k w_k)^{-5/2} / 4 r_i^{\,2} w_i^{\,2}; \quad i = \ell \text{ or } k .$$

Hence, by the Theorem 1, we have, on simplification,

$$\hat{z}(\text{adj}) = \hat{z}\left[1 - 3\{ 1/r_\ell^2 (r_\ell - 1)\hat{w}_\ell^{\,2} + 1/r_k^{\,2}(r_k - 1)\hat{w}_k^{\,2} \} \,/\, 4(1/r_\ell \hat{w}_\ell + 1/r_k \hat{w}_k)^2 \right]$$

It is observed below from the Monte Carlo study that these adjusted test-statistics are more or less robust with respect to differences in error variances. The Ratio of the adjusted normal statistic to the square root of error mean squares is the corresponding adjusted t-variate with n-t d.f.

2.1.5 Multiple comparison

Scheffé (1959, pp. 68-70) developed a method of multiple comparison assuming the error variances to be constant. For the heteroscedastic models if we proceed in the same way, we find that the probability is $(1 - \alpha)$ that the values of all contrasts, ψ, of the population means, simultaneously satisfy the inequalities

$$(\hat{\psi} - S\, s\, \sigma_{\hat{\psi}}) \leq \psi \leq (\psi + S\, s \sigma_{\hat{\psi}})$$

where $S = \{(t-1)F_{\alpha}(t-1, n-t)\}^{\frac{1}{2}}$,

s is the square root of the error mean square of the weighted least squares analysis, $\hat{\psi} = \Sigma c_i y_{i.} (\Sigma c_i = 0)$ is an unbiased estimate of $\psi = \Sigma c_i \mu_i$ and $\sigma_{\hat{\psi}}$, is the standard error of $\hat{\psi}$.

This follows from the fact that if $\hat{\underset{\sim}{\psi}} = (\hat{\psi}_1, \ldots, \tilde{\psi}_q)'$ is an unbiased estimate of $\underset{\sim}{\psi} = (\psi_1, \ldots, \psi_q)'$, the vector of q independent contrasts of the population means, then the estimates, $\hat{\psi}_i$, are independent of s^2 and

$$(\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi})' \underset{\sim}{B}^{-1} (\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi})/qs^2 = F(q, n-t)$$

where $\underset{\sim}{B} = \text{var}(\hat{\underset{\sim}{\psi}})$.

From this it follows that the probability is $(1 - \alpha)$ that for all $\underset{\sim}{h}$

$$|\underset{\sim}{h}' \hat{\underset{\sim}{\psi}} - \underset{\sim}{h}' \underset{\sim}{\psi}| \leq \{qF_{\alpha}(q, n-t)\}^{\frac{1}{2}} s \, (\underset{\sim}{h}' \underset{\sim}{B}\underset{\sim}{h})^{\frac{1}{2}}$$

This can be written as $|\hat{\psi} - \psi| \leq \{qF_{\alpha}(q, n-t)\}^{\frac{1}{2}} s\, \sigma_{\hat{\psi}}$

where $\hat{\psi} = \underset{\sim}{h}' \underset{\sim}{\psi}$ so that $\sigma_{\hat{\psi}} = \{\text{var}(\tilde{\psi})\}^{\frac{1}{2}} = (\underset{\sim}{h}' \underset{\sim}{B}\underset{\sim}{h})^{\frac{1}{2}}$.

In actual practice, if $\hat{\psi} = \sum_1^t c_i y_{i\cdot}$ with $\sum_1^t c_i = 0$, is an estimator of the contrast $\psi = \Sigma c_i \mu_i$, then

$$\sigma_{\hat{\psi}} = \{\Sigma(c_i^2 \sigma_i^2 / r_i)\}^{\frac{1}{2}}.$$

If we replace σ_i^2 by an estimator, s_i^2 or $\hat{\sigma}_i^2$, then the resulting quantity, $\hat{\sigma}_{\hat{\psi}}$, will not be unbiased for $\sigma_{\hat{\psi}}$. Once again the bias of order $\{\sum_i 1/(r_i - 1)\}$ can be removed from $\hat{\sigma}_{\hat{\psi}}$ with the help of Meier's theorem.

Since

$$\left[\frac{\partial^2 \hat{\sigma}_{\hat{\psi}}}{\partial x_i^2}\right]_{\text{all } x_i = 1} = -\frac{1}{4} \frac{c_i^4 \sigma_i^4}{r_i^2} \left[\sum_1^t \frac{c_i^2 \sigma_i^2}{r_i}\right]^{-3/2},$$

$$\hat{\sigma}_{\hat{\psi}}(adj) = \{\Sigma(c_i^2 s_i^2 / r_i)\}^{\frac{1}{2}} \{1 + \frac{1}{4} \sum_1^t \frac{C_i^4 s_i^4}{r_i^2 (r_i - 1)} \left[\sum_i (c_i^2 s_i^2 / r_i)\right]^{-2}\}$$

using the MINQUE of σ_i^2 as the estimator. Since the mean square error s^2, computed from sample, is a constant, no adjustment is necessary for that. Thus, the expression for the estimator of the joint confidence interval of all contrasts ψ is given by

$$\hat{\psi} - Ss\, \hat{\sigma}_{\hat{\psi}}(adj) \leq \psi \leq \hat{\psi} + Ss\, \hat{\sigma}_{\hat{\psi}}(adj) \quad \cdots\cdots \quad (5)$$

For the example considered by Spjøtvoll (1972), the joint confidence interval at the 10% level of significance, for the contrast $\mu_1 - \mu_2$, is [19.3, 33.3] obtained by the above method using MINQUE of σ_i^2. The corresponding joint confidence intervals obtained by Spjøtvoll and by the method of Brown and Forsythe (1974b) are [17.5, 35.1] and [19.8, 32.8] respectively. The MLE of σ_i^2 produces a slightly larger confidence interval.

2.1.6 Summary dispersion measures of the estimators of the linear parameters

Dispersions of the individual treatment estimators are not comparable because of the differences in error variances. In order to have an idea about the overall dispersion of all the estimators, we consider summary measures of dispersion.

The weighted least squares (WLS) estimators of the treatment means are the same as those of the least squares (LS) method but their variances differ between the two procedures. The estimators are uncorrelated in both the methods so that the dispersion matrix of the estimated treatment means is a diagonal one in both the procedures.

Since the covariances are zero, three measures of location of the variances of the estimators may be taken as summary dispersion measures. These are the arithmetic mean (AM), geometric mean (GM) and harmonic mean (HM). All three measures take the variance of each estimator into account and represent dispersion per treatment. The AM is the (1/t)th part of the trace of the dispersion matrix of the estimators and GM the tth root of their generalised variance.

The measures and their estimators for the two methods are as follows:

(a) Weighted least squares estimation

Here, $\mathrm{var}\,(y_i) = \sigma_i^2/r_i \quad ; \quad i = 1,2,\ldots,t.$

Hence,

$$AM = \Sigma \frac{\sigma_i^2}{r_i}/t, \quad GM = (\pi \frac{\sigma_i^2}{r_i})^{1/t} \text{ and } HM = t/\Sigma r_i w_i$$

with $w_i = {}^1/\sigma_i^2$.

Since $AM \geq GM \geq HM$ on the assumption that each $\sigma_i^2 > o$, the last measure i.e., HM is the smallest of the three in the presence of differences in error variances.

All the measures have the same value when the replications r_i are proportional to the corresponding population variances σ_i^2 i.e., each σ_i^2/r_i is the same constant.

The estimated AM = $(1/t) \sum_1^t s_i^2/r_i$ is an unbiased estimator of the AM. The estimated GM = $\left(\pi \frac{s_i^2}{r_i} \right)^{1/t}$ is not unbiased for GM. Since $\left[\frac{\partial^2(\text{Est.GM})}{\partial x_i^2} \right]_{\text{all } x_i=1} = -\frac{t-1}{t^2} \left(\pi \frac{\sigma_i^2}{r_i} \right)^{1/t}$,

the estimated GM with the adjustment for bias is given by

$$\text{Est. GM (adj)} = \left(\pi \frac{s_i^2}{r_i}\right)^{1/t} \left(1 + \frac{t-1}{t^2} \sum \frac{1}{r_i - 1}\right).$$

Also since $\left[\frac{\partial^2(\text{Est. HM})}{\partial x_i^2} \right]_{\text{all } x_i=1} = -2tf_i(1-f_i)/ \sum_i r_i w_i$, the

estimated HM with adjustment for bias is

$$\text{Est. HM (adj)} = (t/\Sigma\, r_i\hat{w}_i) \{ 1 + 2 \Sigma\, \hat{f}_i\, (1 - \hat{f}_i)/(r_i - 1)\}$$

where $f_i = r_i w_i / \Sigma\, r_i w_i$ and $\hat{f}_i = r_i \hat{w}_i / \Sigma\, r_i \hat{w}_i$.

(b) Least squares estimation

Here var $(y_{i.}) = \sigma^2/r_i$, where σ^2 is assumed to be the constant variance of all the populations. Hence AM = $\frac{\sigma^2}{t} \Sigma \frac{1}{r_i}$, GM $= \sigma^2 \left(\pi \frac{1}{r_i}\right)^{1/t}$ and HM $= t\sigma^2/n$ where $n = \Sigma\, r_i$.

If MSE $= \Sigma\, (r_i - 1)\, s_i^2/(n-t)$ is the mean square error of the LS analysis then the estimated AM $= \frac{\text{MSE}}{t} \Sigma \frac{1}{r_i}$, estimated GM = MSE $\left(\pi \frac{1}{r_i}\right)^{\frac{1}{t}}$ and estimated HM = t(MSE)/n are the unbiased estimators of AM, GM and HM respectively.

When the treatments are equally replicated, the estimated AM of the WLS method equals that of the LS method and the leading terms of the estimated GM(adj) and the estimated HM (adj) of the WLS method do not exceed the esimated GM and the estimated HM respectively of the LS method.

2.1.7 The Monte Carlo study

In order to observe the adequacy of the theoretical results, a small Monte Carlo study was conducted. Combinations of some sets of values of replications and error variances were considered for each of 3, 5 and 8 treatments. The results on all possible combinations of the following 3 replication groups, 3 error variance groups and 3 treatment mean groups for 5 treatments are given below. The 3 replication groups, (6,6,$\cdots$6), (3,5,6,7,9) and (9,7,6,5,3), will be denoted by R(1), R(2) and R(3) respectively, the 3 treatment mean groups, (10,10,$\cdots$,10),(12,11,10,9,8) and (9,10,12,10,11) by T(1), T(2) and T(3) respectively and the 3 error variance groups, (1,1,$\cdots$,1), (3,2,1,$\frac{1}{2}$,$\frac{1}{3}$) and ($\frac{1}{2}$,1,4,1,$\frac{1}{2}$), by V(1), V(2) and V(3) respectively.

Only one table contains results on the probability of exceeding percentage points of the main tests for each of 3,5 and 8 treatments.

2.1.7.1 Sampling experiments

For the linear model (4) of section 2.1.1, the observation, y_{ij} was assumed to be normal with mean, μ_i, and variance, σ_i^2. For each set of values of r_i, μ_i and σ_i^2, 1000 distinct sample realisations were made at each run and the analysis was carried out for each sample in double precision on the University of London computer, CDC 7600, in FORTRAN. The normal samples were obtained with the help of the subroutines, G05AEF(A,B) and G05BBF, developed in package forms by the Numerical Algorithm Group (NAG).

2.1.7.2 Power of Bartlett's chisquared test on the homogeneity of error variances

Monte Carlo powers of this test were calculated over 1000 samples each and are given in Table 2 which shows that the powers are almost independent of the treatment differences as is expected. Data of the first row of the table show that the probabilities of exceeding the percentage points in the absence of differences in the error variances, are close to the nominal values. The power of the test is rather small even when the differences in error variances are quite large. The power appears to be larger in the equi-replicate case.

2.1.7.3 Confidence intervals of orthogonal contrasts

In order to investigate the behaviour of the joint confidence intervals of contrasts, two sets of four possible orthogonal contrasts stated in Table 1 below, were considered.

Table 1. Two sets of orthogonal contrasts

Set	Contrasts	
I	(i) $\mu_1 - \mu_2$	(ii) $\mu_1 + \mu_2 - 2\mu_3$
	(iii) $\mu_1+\mu_2+\mu_3 - 3\mu_4$	(iv) $\mu_1+\mu_2+\mu_3+\mu_4-4\mu_5$
II	(i) $\mu_1-\mu_5$	(ii) $\mu_2-\mu_4$
	(iii) $\mu_1+\mu_5-(\mu_2+\mu_4)$	(iv) $4\mu_3-(\mu_1+\mu_2+\mu_4+\mu_5)$

For computing confidence intervals of the treatment contrasts, the expression in (5) of section 2.1.5 was used for WLS method and that given by Scheffé (1959, p.69) used for LS method.

It has been observed from the sampling experiments that the width of the mean confidence interval is virtually independent of the treatment differences. Table 3 gives the average confidence intervals of the above contrasts over 1000 samples each when all the treatment means are the same for both the LS and WLS methods. The table shows that the mean intervals by LS procedure are more or less the same as those by WLS method using MINQUE for all contrasts in the absence of differences in error variances as is to be expected. For the WLS method, the MLE always produces somewhat larger mean confidence interval than the MINQUE. Mean confidence intervals involving fewer means are usually smaller than those involving larger numbers of means except that the last 3 contrasts of set I have approximately the same mean confidence interval by WLS method for most of the replication groups when group variances differ.

In presence of differences in error variances, the WLS method often produces smaller mean confidence intervals than the LS method especially when larger samples are associated with larger variances.

It is observed from the last 3 columns of the table that if the sample sizes are such that the ratios of the error variances and the corresponding replications are the same, then the mean WLS confidence intervals are almost always substantially smaller than those of the LS method. The effect of such proportional replications on the WLS method appears to be the virtual elimination of the inequality of the error variances and of the replications as is evident from comparison of the second and third columns with the last two columns.

2.1.7.4 Empirical size and power of some tests of significance

In order to observe the empirical size (Brown and Forsythe, 1974a) under the null hypothesis, and power under the alternative hypotheses, the following tests were considered:

(i) The usual LS F-test ignoring differences in error variances

(ii) The usual t-test for testing the difference between μ_1 and μ_2

(iii) The weighted least squares F-test (adjusted and unadjusted) using both MINQUE and MLE of group variances

(iv) The normal test (adjusted and unadjusted) using both MINQUE and MLE of group variances

Table 4 presents the results of these tests over 1000 samples at 5% and 1% nominal sizes; it gives the empirical sizes under the null hypothesis and the maximum and minimum powers under the alternative hypotheses. As is well-known the usual LS F-test shows marked discrepancies between the empirical and nominal sizes under the null hypothesis. The empirical size is much larger than the nominal one when smaller numbers of replications are associated with larger variances but the former is somewhat smaller than the latter in the opposite situation. The observed sizes of the WLS F-test (unadjusted) using either MINQUE or MLE of variances are always much larger than the corresponding nominal sizes. For equally replicated treatments, and for situations where larger samples are associated with larger variances, the differences are negligible when the test is adjusted by Meier's Theorem (Theorem 1). In other cases, there are slight variations especially for a nominal size of 1%. Both the methods of estimation of variances produce the same size in the equi-replicate case but the MLE produces slightly larger sizes than MINQUE when sample sizes are not the same.

Like the LS F-test, the usual t-test for testing the difference between μ_1 and μ_2 shows large discrepancies between the empirical and nominal sizes. For the normal test (unadjusted), the discrepancies are even larger. Adjustment of the normal test using the MLE of variances, does not improve the situation to a satisfactory level. The performances of the normal test (adj) using MINQUE of variances are much better although there are still some differences especially for a nominal size of 1%.

Under the alternative hypotheses, the maximum powers of all the F-tests are as large as possible at both levels of significance. Their minimum powers are also large except that the last treatment group coupled with the last error variance group produced moderate minimum power for the WLS F-test (adj) at the 1% level of significance. Maximum powers of t and normal tests are also large. The minimum powers of these latter tests are small because the minimum difference between μ_1 and μ_2 is small and one sample size is small. In general, powers of the WLS tests with adjustment were found to be quite large although these are somewhat less than the corresponding LS tests in some cases.

Table 5 gives the probabilities of exceeding the percentage points, of the main tests under the null hypothesis for each of 3, 5 and 8 treatments. The table shows that the WLS F-test (adj) using either MINQUE or MLE is more or less robust with respect to variations in error variances and sample sizes. The performance of the normal test (adj) using MINQUE is also not far from robustness if the sample sizes are not too small. The usual F-test and t-test show wide differences between the nominal and empirical sizes.

2.1.7.5 Concluding remarks

The WLS F-test (adj) using either MINQUE or MLE of the group variances is more or less robust with respect to differences in error variances. The normal test (adj) using MINQUE of variances for testing differences between two treatment means is also not far from robustness. Performances of these tests are sometimes better if larger samples are associated with larger variances. These tests are therefore recommended for testing appropriate hypotheses when Bartlett's χ^2-test reveals that the group variances differ.

The WLS formula appropriate for heteroscedastic models, using either MINQUE or MLE of group variances, often gives smaller mean joint confidence intervals of treatment contrasts than the usual LS method, especially when larger samples are associated with larger variances. The WLS method is therefore recommended for estimating joint confidence intervals of treatment contrasts when there are different error variances.

A minimum sample size of 4 can usually be expected to give more or less satisfactory results especially when larger samples are associated with larger variances.

2.2 One-way mixed models and random models with unequal group variances

Let the mixed model be

$$y_{ij} = \alpha + \tau_i + \varepsilon_{ij} \qquad j=1,2,\ldots\ldots,r_i; \quad i=1,2,\ldots\ldots,t,$$

where α is the general constant, τ_i the random effect of the ith treatment having mean zero and variance σ_τ^2 and ε_{ij} the error term having mean zero and variance σ_i^2. Treatment effects τ_i are assumed to be independent of one another and of the errors which are also assumed to be independent of one another. This means that the observations y_{ij} are correlated within a treatment and independent between treatments.
Let $n = \sum_1^t r_i$ as before.

2.2.1 Estimation of variance components and the analysis when error variances are known

From the above model, we have,

$$y_{i.} = \alpha + \tau_i + \varepsilon_{i.}$$

$$\tilde{y}_{..} = \alpha + \tilde{\tau} + \tilde{\varepsilon}_{..}$$

under the notation of section 2.1.1 with $\tilde{\tau} = \Sigma r_i w_i \tau_i / \Sigma r_i w_i$.

$$\text{Since } E\{\sum_j (y_{ij} - y_{i.})^2\}$$
$$= E\{\sum_j (\varepsilon_{ij} - \varepsilon_{i.})^2\}$$
$$= (r_i - 1)\sigma_i^2,$$

$s_i^2 = \sum_{j=1}^{r_i} (y_{ij} - y_{i.})^2/(r_i-1)$ is still unbiased for σ_i^2 for the mixed model stated above.

To obtain an estimate of σ_τ^2, let us consider the weighted treatment sum of squares, $\sum_i r_i w_i (y_{i.} - \tilde{y}_{..})^2$, which was obtained in section 2.1.1.

Since τ_i and ε_{ij} are independent, we have,

$$E\{ \Sigma\, r_i w_i (y_{i.} - \tilde{y}..)^2 \}$$

$$= E\,\{\Sigma\, r_i w_i (\tau_i - \tilde{\tau}) + (\varepsilon_{i.} - \tilde{\varepsilon}\,..)\}^2$$

$$= (t - 1) + \sigma_\tau^{\,2} \left(w. - \frac{\Sigma r_i^2 w_i^{\,2}}{w.} \right)$$

with $w. = \Sigma\, r_i w_{i.}$.

Hence, an unbiased estimator of $\sigma_\tau^{\,2}$ is given by

$$\tilde{\sigma}_\tau^2 = \{\Sigma\; r_i w_i (y_{i.} - \tilde{y}..)^2 - t + 1\} \,/\, \left(w. - \frac{\Sigma\, r_i^{\,2} w_i^{\,2}}{w.}\right)$$

when the actual weights $w_i = 1/\sigma_i^{\,2}$ are known.

Also, E (weighted within sum of squares)

$$= E\,\{\Sigma\Sigma\, w_i (y_{ij} - y_{i.})^2\}$$

$$= E\,\{\Sigma\Sigma\, w_i (\varepsilon_{ij} - \varepsilon_{i.})^2\}$$

$$= (n - t)$$

as before.

To show that the above two sums of squares are independent, we need only show that $(\varepsilon_{ij} - \varepsilon_{i.})$ and $(\varepsilon_{i.} - \tilde{\varepsilon}..)$ are independent.

Now $\operatorname{cov}(\varepsilon_{ij} - \varepsilon_{i.})(\varepsilon_{i.} - \tilde{\varepsilon}..) = E\,\{(\varepsilon_{ij} - \frac{\Sigma_j \varepsilon_{ij}}{r_i})$

$$\left(\frac{\Sigma_j \varepsilon_{ij}}{r_i} - \frac{\Sigma\, r_i w_i \varepsilon_{i.}}{\Sigma\, r_i w_i} \right)$$

$$= \frac{\sigma_i^{\,2}}{r_i} - \frac{r_i w_i\, \sigma_i^{\,2}}{w.\; r_i} - \frac{\sigma_i^{\,2}}{r_i} + \frac{r_i w_i \sigma_i^{\,2}}{w.\; r_i}$$

$$= o.$$

Hence, under the assumption of normality of errors, the above two quantities are independent.

It follows that $\Sigma\, r_i w_i (y_{i.} - \tilde{y}..)^2$ is distributed as a central χ^2 with (t-1) d.f. under the hypothesis that $\sigma_\tau = o$

and that $\Sigma\Sigma w_i(y_{ij} - y_{i.})^2$ is always distributed as a central χ^2-variate with $(n - t)$ d.f. under the assumption of normality of errors and that the two sums of squares are independent.

Hence,

$$F = \frac{\Sigma\, r_i w_i (y_{i.} - \tilde{y}_{..})^2/(t-1)}{\Sigma\Sigma\ w_i (y_{ij} - y_{i.})^2/(n-t)}$$

is a central F-variate under the null hypothesis: $\sigma_\tau = 0$, with $(t-1)$ and $(n-t)$ d.f.

2.2.2 Adjustment of the F-test statistic and the estimator of σ_τ^2, using estimated weights

Since the estimators s_i^2 of error variances are again independently distributed as gamma variates, Theorem 1 due to Meier may be applied for adjustment of bias.

The expression of the F-statistic is the same as that of section 2.1.1 so that the adjusted F-statistic using estimated weights is also the same, namely

$$\hat{F}(adj) = \frac{(n-t)\{\Sigma\ r_i \hat{w}_i\ (y_{i.} - \tilde{y}_{..})^2\}\{1 - \frac{2}{r_i - 1}(1 - \frac{r_i \hat{w}_i}{\hat{w}_.})\}}{(t-1) \sum\limits_{ij}\Sigma\ \hat{w}_i\ (y_{ij} - y_{i.})^2}$$

with $(t-1)$ and $(n-t)$ d.f.

Now the estimator of σ_τ^2 using the estimated weights is

$$\hat{\tilde{\sigma}}_\tau^{\,2} = \{\Sigma\, r_i \hat{w}_i (y_{i.} - \hat{\tilde{y}}_{..})^2 - t + 1\}/(\hat{w}_. - \frac{\Sigma\, r_i^2 \hat{w}_i^2}{\hat{w}_.}) = A/B$$

say, with $\hat{w}_. = \Sigma\, r_i \hat{w}_i$. The adjusted estimator is

$$\hat{\tilde{\sigma}}_\tau^2\ (adj) = \hat{\tilde{\sigma}}_\tau^{\,2} - \sum_1^t \frac{1}{r_i - 1} \left[\frac{\partial^2 \hat{\tilde{\sigma}}_\tau^{\,2}}{\partial x_i^2}\right]_{\text{all } x_i = 1} \quad \text{using estimated weights,}$$

where

$$\left[\frac{\partial^2 \hat{\tilde{\sigma}}_\tau^{\,2}}{\partial x_i^{\,2}}\right]_{\text{all } x_i=1} = \left[\frac{1}{B^3}\left\{B^2\frac{\partial^2 A}{\partial x_i^{\,2}} - BA\frac{\partial^2 B}{\partial x_i^{\,2}} - 2B\frac{\partial A \partial B}{\partial x_i \partial x_i} + 2A\left(\frac{\partial B}{\partial x_i}\right)^2\right\}\right]_{\text{all } x_i=1}$$

the individual derivatives being $\left[\frac{\partial A}{\partial x_i}\right]_{\text{all } x_i=1} = - r_i w_i (y_i. - \tilde{y}..)^2,$

$$\left[\frac{\partial^2 A}{\partial x_i^{\,2}}\right]_{\text{all } x_i=1} = 2\, r_i w_i (1-f_i)(y_i. - \tilde{y}..)^2, \quad \left[\frac{\partial B}{\partial x_i}\right]_{\text{all } x_i=1} = -r_i w_i + r_i w_i (2f_i - \Sigma r_i^2 w_i^{\,2}/w.^2)$$

and

$$\left[\frac{\partial^2 B}{\partial x_i^{\,2}}\right]_{\text{all } x_i=1} = 2r_i w_i - 2r_i w_i\ \{ f_i + (1-f_i)(2f_i - \Sigma r_i^2 w_i^{\,2}/w.^2)\}$$

with $f_i = r_i w_i / w..$

For the random model:

$$y_{ij} = \tau_i + \varepsilon_{ij}.$$

with τ_i as random variables having mean zero and variance, $\sigma_\tau^{\,2}$, if we proceed in the same way as above, we get the same estimator of $\sigma_\tau^{\,2}$ and the same F-test for testing the significance of σ_τ. But the above analysis is not valid if τ_i have non-zero mean because a separate estimator of $\sigma_\tau^{\,2}$ is not available in that case.

Table 2. Monte Carlo powers of Bartlett's chi-squared test on the homogeneity of error variances

Error variance group	Treatment and replication groups																	
	T(1)						T(2)						T(3)					
	R(1)		R(2)		R(3)		R(1)		R(2)		R(3)		R(1)		R(2)		R(3)	
	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%
V(1)	.059	.010	.063	.015	.049	.012	.041	.007	.045	.011	.053	.014	.054	.004	.047	.015	.049	.009
V(2)	.553	.302	.502	.273	.414	.161	.557	.288	.529	.259	.401	.177	.540	.284	.502	.275	.414	.180
V(3)	.546	.322	.532	.321	.535	.314	.550	.335	.504	.301	.499	.305	.559	.316	.532	.309	.546	.315

Table 3. Widths of Mean confidence intervals of two sets of orthogonal contrasts; letters LS denote least squares method and WLS weighted least squares procedure; numbers 1 and 2 after WLS stand for MINQUE and MLE respectively of group variances

Set	Contrast number	V(1) R(1) LS	V(1) R(1) WLS1	V(1) R(1) WLS2	V(1) R(2) LS	V(1) R(2) WLS1	V(1) R(2) WLS2	V(1) R(3) LS	V(1) R(3) WLS1	V(1) R(3) WLS2	V(2) R(1) LS	V(2) R(1) WLS1	V(2) R(1) WLS2	V(2) R(2) LS	V(2) R(2) WLS1	V(2) R(2) WLS2	V(2) R(3) LS	V(2) R(3) WLS1	V(2) R(3) WLS2
I	1	3.81	3.85	4.22	4.84	4.86	5.32	3.32	3.37	3.69	4.43	6.10	6.68	4.78	8.02	8.79	4.38	5.26	5.76
	2	6.60	6.68	7.32	7.26	7.42	8.13	6.32	6.35	6.96	7.67	8.22	9.00	7.17	9.77	10.70	8.34	7.59	8.32
	3	9.33	9.45	10.35	9.34	9.41	10.30	9.82	9.99	10.94	10.85	8.83	9.67	9.22	10.08	11.04	12.95	8.65	9.48
	4	12.05	12.17	13.33	10.73	10.81	11.84	16.08	16.16	17.71	14.00	9.32	10.21	10.60	10.08	11.04	21.21	10.95	11.99
II	1	3.81	3.91	4.28	4.42	4.39	4.81	4.39	4.45	4.87	4.43	4.96	5.44	4.36	6.88	7.54	5.79	4.46	4.89
	2	3.81	3.82	4.19	3.88	3.94	4.32	3.86	3.93	4.31	4.43	4.29	4.70	3.83	4.58	5.02	5.09	4.16	4.56
	3	5.39	5.46	5.98	5.88	5.91	6.48	5.85	5.95	6.52	6.26	6.58	7.21	5.81	8.32	9.11	7.71	6.10	6.68
	4	12.05	12.19	13.35	12.32	12.64	13.85	12.25	12.30	13.48	14.00	12.81	14.03	12.17	13.86	15.19	16.15	12.50	13.70

Set	Contrast number	V(3) R(1) LS	V(3) R(1) WLS1	V(3) R(1) WLS2	V(3) R(2) LS	V(3) R(2) WLS1	V(3) R(2) WLS2	V(3) R(3) LS	V(3) R(3) WLS1	V(3) R(3) WLS2	V(2) (18,12,6,3,2) LS	V(2) (18,12,6,3,2) WLS1	V(2) (18,12,6,3,2) WLS2
I	1	4.48	3.38	3.70	5.71	4.09	4.48	3.86	2.92	3.19	3.53	3.73	3.99
	2	7.76	11.42	12.51	8.56	11.76	12.88	7.36	11.15	12.22	8.51	6.44	6.87
	3	10.98	10.31	11.29	11.01	10.20	11.18	11.43	10.93	11.97	17.24	9.00	9.61
	4	14.17	10.45	11.45	12.65	9.70	10.63	18.71	12.72	13.94	27.88	11.90	12.69
II	1	4.48	2.76	3.02	5.21	3.17	3.47	5.11	3.11	3.41	7.07	3.77	4.02
	2	4.48	3.86	4.22	4.57	3.91	4.29	4.49	3.89	4.26	6.12	3.73	3.98
	3	6.34	4.74	5.19	6.93	5.05	5.53	6.80	4.99	5.47	9.35	5.30	5.65
	4	14.17	22.31	24.45	14.52	22.61	24.77	14.25	22.09	24.20	18.09	11.72	12.51

Table 4. Probabilities of exceeding percentage points under null hypothesis and maximum and minimum powers under alternative hypotheses, for 5% and 1% nominal sizes, of some tests of significance; letters LS-F stand for the usual LS F-test, WLS-F for weighted least squares F-test using estimated weights, t for usual t-test and Nor for normal test using estimated weights; numbers 1 and 2 denote estimated weights based on MINQUE and MLE respectively of error variances

Test	Probabilities of exceeding percentage points under null hypothesis																		Power under alternative hypotheses			
	Error variance and replication groups																					
	V(1)						V(2)						V(3)									
	R(1)		R(2)		R(3)		R(1)		R(2)		R(3)		R(1)		R(2)		R(3)		5%		1%	
	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	Max.	Min.	Max.	Min.
LS-F	.043	.007	.048	.008	.049	.010	.073	.020	.182	.077	.037	.004	.095	.033	.081	.038	.099	.031	1.00	1.00	1.00	1.00
WLS-F1(unadj)	.096	.029	.153	.065	.152	.069	.114	.032	.201	.094	.122	.044	.136	.045	.164	.072	.162	.070	1.00	.91	1.00	.77
WLS-F1(adj)	.040	.011	.050	.020	.058	.028	.039	.012	.056	.028	.041	.012	.052	.020	.065	.026	.061	.018	1.00	.77	1.00	.53
WLS-F2(unadj)	.096	.029	.166	.077	.170	.078	.114	.032	.216	.105	.129	.050	.136	.045	.175	.079	.182	.076	1.00	.91	1.00	.77
WLS-F2(adj)	.040	.011	.054	.022	.062	.028	.039	.013	.062	.032	.045	.013	.052	.020	.069	.030	.065	.020	1.00	.77	1.00	.53
t	.044	.008	.042	.007	.049	.008	.143	.052	.247	.124	.093	.036	.024	.005	.007	0.0	.012	.001	.97	.15	.88	.05
Nor 1(unadj)	.070	.031	.116	.052	.068	.021	.075	.029	.133	.058	.070	.031	.093	.039	.107	.061	.075	.021	.98	.21	.81	.11
Nor 1(adj)	.049	.021	.061	.025	.057	.018	.062	.018	.075	.040	.062	.021	.069	.024	.076	.026	.060	.014	.97	.13	.86	.06
Nor 2(unadj)	.100	.041	.159	.084	.087	.029	.100	.041	.181	.100	.083	.038	.116	.052	.139	.082	.091	.033	.98	.28	.74	.15
Nor 2(adj)	.073	.030	.106	.044	.070	.023	.076	.028	.121	.053	.068	.031	.097	.038	.103	.053	.072	.022	.98	.19	.81	.10

Table 5. Probabilities of exceeding percentage points under the null hypothesis, for 5% and 1% nominal sizes, of the usual F- and t-tests, the WLS F-test (adj) using MINQUE or MLE of error variances and the normal test (adj) using MINQUE, for 3, 5 and 8 treatments.

No. of treat-ments	Error variances	Replications	L S F-test		WLS F-test(adj) using MINQUE		WLS F-test(adj) using MLE		t-test for $\mu_1 = \mu_2$		Normal test (adj) using MINQUE for $\mu_1 = \mu_2$	
			5%	1%	5%	1%	5%	1%	5%	1%	5%	1%
3	$(2,1,\frac{1}{2})$	(4,4,4)	.059	.015	.038	.008	.038	.008	.076	.019	.066	.027
		(8,6,4)	.040	.004	.040	.011	.039	.012	.059	.012	.061	.015
		(4,6,8)	.098	.030	.045	.015	.047	.018	.114	.042	.058	.027
5	$(3,2,1,\frac{1}{2},\frac{1}{3})$	(6,6,6,6,6)	.073	.020	.039	.013	.039	.013	.143	.052	.062	.018
		(9,7,6,5,3)	.037	.004	.041	.012	.045	.013	.093	.036	.062	.021
		(3,5,6,7,9)	.182	.077	.056	.028	.062	.032	.247	.124	.075	.040
8	$(4,3,2,1,1,\frac{1}{2},\frac{1}{3},\frac{1}{4})$	(6,6,6,6,6,6,6,6)	.077	.025	.056	.024	.056	.024	.223	.099	.053	.013
		(16,14,12,10,10,8,6,4)	.031	.010	.054	.022	.059	.024	.130	.055	.057	.021
		(4,6,8,10,10,12,14,16)	.234	.110	.053	.022	.064	.025	.293	.167	.063	.021

CHAPTER 3

GENERAL TWO-WAY MODEL WITH PROPORTIONAL CELL FREQUENCIES

In this chapter, two-way models having proportional cell frequencies and unequal group variances are considered. On the assumption that the error variances are known, estimators of the linear parameters of the fixed-effects models are obtained and the analysis is given for two sets of constraints on the linear parameters. The MLE and MINQUE of group variances are derived. The estimators and test-statistics using estimated weights are adjusted for bias. Formulae for estimating joint confidence intervals are provided for contrasts of both main effects and interactions.

Two-way random models with unequal group variances are also considered for estimation of variance components; the corresponding analysis is given for both known and unknown weights. Finally, some simpler tests are discussed for two-way fixed-effects models with equally replicated treatments.

3.1 Two-way fixed-effects model

3.1.1 The model

In order to keep uniformity with the general terminology of the thesis, we shall refer to one of the two factors as treatments and the other as blocks. The model will cover experiments where block effects constitute a factor in which the experimenter is interested in addition to the treatments. For example, in an experiment where several persons work with the same set of machines, the experimenter may be interested in observing differences between machines as well as persons and both factors may be of equal interest, even though one is designated "blocks".

When the block effects are meant to eliminate from observations heterogeneity in any direction, they will not usually be of interest. In a variety trial in the field, varieties are of prime importance and blocks are introduced mainly to remove the heterogeneity.

We shall consider the non-additive model:

$$\left.\begin{array}{l} y_{ijk} = \beta_i + \tau_j + \delta_{ij} + \varepsilon_{ijk} \\ i=1,2,\ldots,b \;;\; j=1,2,\ldots,t \;;\; k=1,2,\ldots,n_{ij} \end{array}\right| \qquad \ldots\ldots (6)$$

where β_i is the effect of the ith block, τ_j the effect of the jth treatment, δ_{ij} the effect of the interaction between the ith block and jth treatment and ε_{ijk} the error term having mean zero and variance σ_j^2. The errors are assumed to be independent of one another. The variances of the errors under the same treatment are assumed to be the same but differ from treatment to treatment. The number $n_{ij}(\geq 1)$ of observations in the (i,j)th cell is assumed to be proportional to the marginal totals, that is,

$$n_{ij} = N_{i\cdot}\, N_{\cdot j} / N_{\cdot\cdot}$$

where $N_{i\cdot} = \sum_j n_{ij}$, $N_{\cdot j} = \sum_i n_{ij}$ and $N_{\cdot\cdot} = \sum_{ij}\sum n_{ij}$. This includes the case of equal number of observations per cell.

Let there be two types of constraint on the linear parameters of the model:

Constraints (I)

$$0 = \sum_i \delta_{ij} \quad \text{for all } j$$
$$= \sum_j n_{ij} w_j \delta_{ij} \quad \text{for all } i$$
$$= \sum_i \sum_j n_{ij} w_j \delta_{ij}$$

Constraints (II)

$$0 = \sum_i \delta_{ij} \quad \text{for all } j$$
$$= \sum_j \delta_{ij} \quad \text{for all } i$$
$$= \sum_i \sum_j \delta_{ij}$$

where the quantities $w_j = 1/\sigma_j^2$ are the weights. The constraints (I) which are more arbitrary than the usual constraints (II), facilitate the test for block effects as is shown below. There is no constraint on the block effects β_i which include the general parameter. Different sets of constraints imply different values of the parameters.

3.1.2 Estimation and analysis when the group variances are known

Let $\underset{\sim}{Y}$ be the vector of observations arranged treatment by treatment. Then the model (6) above can be written, in matrix notation, as

$$\underset{\sim}{Y} = \underset{\sim}{X}' \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{X}'$ is the design matrix, $\underset{\sim}{\beta}$ the vector of all linear parameters and $\underset{\sim}{\varepsilon}$ the vector of errors. Rank $\underset{\sim}{X}' = bt$. The vector $\underset{\sim}{Y}$ is given by

$$\underset{\sim}{Y} = (y_{111}, \ldots, y_{11n_{11}}, y_{211}, \ldots, y_{21n_{21}}, \ldots, y_{bt1}, \ldots, y_{btn_{bt}})'$$

so that

$$\text{var}(\underset{\sim}{Y}) = \text{diag}\,(\sigma_1^2, \ldots, \sigma_1^2, \ldots, \sigma_t^2, \ldots, \sigma_t^2) = \underset{\sim}{V},$$

say. Then $\underset{\sim}{V}^{-1} = \text{diag}\,(w_1, \ldots, w_1, \ldots, w_t, \ldots, w_t)$.

From equation (2) of section 1.2, we get the normal equations given at (7) for estimating the linear parameters.

$$
\begin{bmatrix}
N_{\cdot 1}w_1 & & & n_{11}w_1 \cdots n_{b1}w_1 & n_{11}w_1 \cdots n_{b1}w_1 & & \\
& N_{\cdot 2}w_2 & & n_{12}w_2 \cdots n_{b2}w_2 & \cdots\cdots & n_{12}w_2 \cdots n_{b2}w_2 & \\
& & \ddots & \vdots \qquad \vdots & & \vdots & \ddots \\
& & N_{\cdot t}w_t & n_{1t}w_t \cdots n_{bt}w_t & & & n_{1t}w_t \cdots n_{bt}w_t \\
n_{11}w_1 \cdots & \cdots & n_{1t}w_t & \Sigma n_{1j}w_j & n_{11}w_1 & n_{12}w_2 & \cdots\cdots n_{1t}w_t \\
\vdots & & \vdots & \ddots & \ddots & \ddots & \ddots \\
n_{b1}w_1 \cdots & \cdots & n_{bt}w_t & & \Sigma n_{bj}w_j \quad n_{b1}w_1 & n_{b2}w_2 & n_{bt}w_t \\
n_{11}w_1 & & & n_{11}w_1 & n_{11}w_1 & & \\
\vdots & & & \ddots & \ddots & & \\
n_{b1}w_1 & & & n_{b1}w_1 & n_{b1}w_1 & & \\
& \ddots & & n_{12}w_2 & & n_{12}w_2 & \\
& & \ddots & \ddots & & \ddots & \\
& & & n_{b2}w_2 & & n_{b2}w_2 & \\
& & & \vdots & & & \ddots \\
& & n_{1t}w_t & n_{1t}w_t & & & n_{1t}w_t \\
& & \vdots & \ddots & & & \ddots \\
& & n_{bt}w_t & n_{bt}w_t & & & n_{bt}w_t
\end{bmatrix}
\begin{bmatrix}
\hat{\tau}_1 \\ \vdots \\ \hat{\tau}_t \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_b \\ \tilde{\delta}'_{11} \\ \vdots \\ \tilde{\delta}'_{b1} \\ \tilde{\delta}'_{12} \\ \vdots \\ \tilde{\delta}'_{b2} \\ \vdots \\ \tilde{\delta}'_{1t} \\ \vdots \\ \tilde{\delta}'_{bt}
\end{bmatrix}
=
\begin{bmatrix}
w_1 Y_{\cdot 1 \cdot} \\ \vdots \\ w_t Y_{\cdot t \cdot} \\ \Sigma w_j Y_{1j\cdot} \\ \vdots \\ \Sigma w_j Y_{bj\cdot} \\ w_1 Y_{11\cdot} \\ \vdots \\ w_1 Y_{b1\cdot} \\ w_2 Y_{12\cdot} \\ \vdots \\ w_2 Y_{b2\cdot} \\ \vdots \\ w_t Y_{1t\cdot} \\ \vdots \\ w_t Y_{bt\cdot}
\end{bmatrix}
\quad \cdots\cdots\cdots\cdots \quad (7)
$$

From these the individual normal equations are obtained as

$$\tau_j:\quad N_{\cdot j}w_j\hat{\tau}_j + w_j\sum_i n_{ij}\tilde{\beta}_i + w_j\sum_i n_{ij}\tilde{\delta}'_{ij} = w_jY_{\cdot j\cdot}\,;\quad j = 1,2,\ldots,t$$

$$\beta_i:\quad \sum_j n_{ij}w_j\hat{\tau}_j + \sum_j n_{ij}w_j\tilde{\beta}_i + \sum_j n_{ij}w_j\tilde{\delta}'_{ij} = \sum_j w_jY_{ij\cdot}\,;\quad i = 1,2,\ldots,b$$

$$\delta_{ij}:\quad n_{ij}w_j\hat{\tau}_j + n_{ij}w_j\tilde{\beta}_i + n_{ij}w_j\tilde{\delta}'_{ij} = w_jY_{ij\cdot}\qquad \begin{array}{l} i = 1,2,\ldots,b \\ j = 1,2,\ldots,t \end{array}$$

Using the constraints given by

$$\sum_j N_{\cdot j}w_j\hat{\tau}_j = 0 = \sum_i N_{i\cdot}\tilde{\beta}_i = \sum_j n_{ij}w_j\tilde{\delta}'_{ij}\quad \text{for all } i$$

$$= \sum_i n_{ij}\tilde{\delta}'_{ij}\quad \text{for all } j$$

along with the proportionality conditions, we get the estimators as

$$\hat{\tau}_j = y_{\cdot j\cdot}\,;\quad j = 1,2,\ldots,t$$

$$\tilde{\beta}_i = \frac{\sum_j w_jn_{ij}y_{ij\cdot}}{\sum_j n_{ij}w_j} = \tilde{y}_{i\cdot\cdot},\quad i = 1,2,\ldots,b$$

say, and

$$\tilde{\delta}'_{ij} = (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot})\,.$$

Finally, $\tilde{\varepsilon}_{ijk} = (y_{ijk} - \tilde{\beta}_i - \hat{\tau}_j - \tilde{\delta}'_{ij}) = (y_{ijk} - y_{ij\cdot})$ from the last normal equation. Here, we have used the usual convention that the dot suffix of a small letter denotes the mean and that of a capital letter the total over the corresponding variable suffix. This convention will be followed in the sequel.

The corresponding sums of squares for the above three types of estimators are $\sum_j(w_jY^2_{\cdot j\cdot}/N_{\cdot j})$, $\sum_i\{(\sum_j w_jY_{ij\cdot})^2/\sum_j(n_{ij}w_j)\}$ and $\sum_{ij}\sum n_{ij}w_jy_{ij\cdot}(y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot})$ in that order.

To obtain the sums of squares corrected for the mean, let us assume that $\beta_i = \beta$ for all i, $\tau_j = 0$ for all j and

$\delta_{ij} = 0$ for all i and j. Then the model reduces to

$$y_{ijk} = \beta + \varepsilon_{ijk} .$$

The weighted least squares estimator of the general mean β is given by

$$\tilde{\beta} = \Sigma w_j Y_{\cdot j \cdot} / \Sigma N_{\cdot j} w_j = \tilde{y}_{\cdots} ,$$

say, and the corresponding sum of squares by

$$(\sum_j w_j Y_{\cdot j \cdot})^2 / \sum_j N_{\cdot j} w_j .$$

with 1 d.f. Then the above three sums of squares (SS) corrected for the mean are:

$$\text{SS (Treat.)} = \sum_j (w_j Y^2_{\cdot j \cdot}/N_{\cdot j}) - (\Sigma w_j Y_{\cdot j \cdot})^2 / \Sigma N_{\cdot j} w_j$$
$$= \sum_j N_{\cdot j} w_j (y_{\cdot j \cdot} - \tilde{y}_{\cdots})^2$$

with (t-1) d.f.

$$\text{SS (Block)} = \Sigma\{\sum_j w_j Y_{ij\cdot})^2 / \sum_j n_{ij} w_j \} - (\Sigma w_j Y_{\cdot j \cdot})^2 / \sum_j N_{\cdot j} w_j$$
$$= \sum_i \sum_j n_{ij} w_j (\tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdots})^2$$

with (b-1) d.f. and

$$\text{SS(Int.)} = \Sigma\Sigma\, n_{ij} w_j y_{ij\cdot} (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j \cdot}) + (\sum_j w_j Y_{\cdot j \cdot})^2 / \Sigma N_{\cdot j} w_j$$
$$= \Sigma\Sigma\, n_{ij} w_j (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j \cdot} + \tilde{y}_{\cdots})^2$$

with (b-1)(t-1) d.f.

To get the corrected SS due to the interactions, we are to add the SS due to the mean because the SS due to all linear parameters is a fixed quantity.

Finally, the sum of squares due to error is given by

$$\text{SS(E)} = \underset{\sim}{\varepsilon}' \underset{\sim}{V}^{-1} \underset{\sim}{\varepsilon} = \Sigma\Sigma\Sigma\, w_j \tilde{\varepsilon}^2_{ijk} = \Sigma\Sigma\Sigma\, w_j (y_{ijk} - y_{ij\cdot})^2$$

with ($N_{\cdot\cdot}$ - bt) d.f.

It follows that the estimators of the linear parameters are not unbiased under any of the two given sets of constraints. If we define $\tilde{\delta}_{ij} = (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot} + \tilde{y}_{\cdot\cdot\cdot})$, then $\tilde{\delta}_{ij}$ is unbiased for the interaction effect δ_{ij} under constraints (I). The estimated treatment contrasts are unbiased for the corresponding parametric contrasts under both sets of constraints. The estimated block effects contrasts are unbiased for the corresponding parametric contrasts under constraints (I) only.

The variances of the estimators are:

$$\text{Var}(\hat{\tau}_j) = \sigma_j^2/N_{\cdot j} \quad , \quad \text{Var}(\tilde{\beta}_i) = 1/\sum_j n_{ij}w_j$$

and $\text{Var}(\tilde{\delta}_{ij}) = \sigma_j^2(1/n_{ij} - 1/N_{\cdot j}) - (1/\sum_j n_{ij}w_j - 1/\sum_j N_{\cdot j}w_j)$.

The treatment estimators are independent of one another and also the estimated block parameters are independent of one another under the usual assumption of normality of errors. The interaction estimators $\tilde{\delta}_{ij}$ are not independent since

$$\text{Cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{ik}) = -(1/\sum_j n_{ij}w_j - 1/\sum N_{\cdot j}w_j) \quad \text{for } j \neq k$$

$$\text{Cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{\ell j}) = -(1/N_{\cdot j}w_j - 1/\sum N_{\cdot j}w_j) \quad \text{for } i \neq \ell$$

and

$$\text{Cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{\ell k}) = 1/\sum_j N_{\cdot j}w_j \quad \text{for } i \neq \ell \text{ and } j \neq k \, .$$

Expectations of the sums of squares under the two sets of constraints are given below.

Under constraints (I), we have from model (6)

$$y_{ij\cdot} = \beta_i + \tau_j + \delta_{ij} + \varepsilon_{ij\cdot}\,; \quad y_{\cdot j\cdot} = \beta_{\cdot} + \tau_j + \varepsilon_{\cdot j\cdot}\,;$$

$$\tilde{y}_{i\cdot\cdot} = \beta_i + \tilde{\tau} + \tilde{\varepsilon}_{i\cdot\cdot} \quad \text{and} \quad \tilde{y}_{\cdot\cdot\cdot} = \beta_{\cdot} + \tilde{\tau} + \tilde{\varepsilon}_{\cdot\cdot\cdot}$$

where $\beta_{\cdot} = \sum\beta_i/b$, $\tilde{\varepsilon}_{i\cdot\cdot} = \sum_j n_{ij}w_j\,\varepsilon_{ij\cdot}/\sum_j n_{ij}w_j$, $\tilde{\varepsilon}_{\cdot\cdot\cdot} = \sum_j N_{\cdot j}w_j\varepsilon_{\cdot j\cdot}/\sum N_{\cdot j}w_j$ and $\tilde{\tau} = \sum_j N_{\cdot j}w_j\tau_j/\sum_j N_{\cdot j}w_j$. Thus we have

$$E\{SS\ (\text{Treatments})\} = \Sigma N_{.j}w_j(\tau_j - \tilde{\tau})^2 + E\{\Sigma N_{.j}w_j(\varepsilon_{.j.} - \tilde{\varepsilon}_{...})^2\}$$

$$= (t-1) + \Sigma N_{.j}w_j(\tau_j - \tilde{\tau})^2$$

$$E\{SS\ (\text{Blocks})\} = \Sigma\Sigma n_{ij}w_j(\beta_i - \beta_.)^2 + E\{\sum_{ij}\Sigma n_{ij}w_j(\tilde{\varepsilon}_{i..} - \tilde{\varepsilon}_{...})^2$$

$$= (b-1) + \Sigma\Sigma n_{ij}w_j(\beta_i - \beta_.)^2$$

and

$$E\{SS\ (\text{Interactions})\} = \Sigma\Sigma n_{ij}w_j\delta_{ij}^2 + E\{\Sigma\Sigma n_{ij}w_j(\varepsilon_{ij.} - \tilde{\varepsilon}_{i..} - \varepsilon_{.j.} + \tilde{\varepsilon}_{...})^2\}$$

$$= (b-1)(t-1) + \Sigma\Sigma n_{ij}w_j\delta_{ij}^2$$

since $E\{\Sigma\Sigma n_{ij}w_j(\varepsilon_{ij.} - \tilde{\varepsilon}_{i..})^2\} = b(t-1)$ and

$E\{\Sigma\Sigma n_{ij}w_j(\varepsilon_{ij.} - \tilde{\varepsilon}_{i..})(\varepsilon_{.j.} - \tilde{\varepsilon}_{...})\} = (t-1)$. Finally,

$$E\{SS\ (\text{Error})\} = \sum_{ij}\Sigma w_j E\{\sum_{k=1}^{n_{ij}}(\varepsilon_{ijk} - \varepsilon_{ij.})^2\} = (N_{..} - bt).$$

Under constraints (II), we have from model (6)

$y_{ij.} = \beta_i + \tau_j + \delta_{ij} + \varepsilon_{ij.}$; $y_{.j.} = \beta_. + \tau_j + \varepsilon_{.j.}$;

$\tilde{y}_{i..} = \beta_i + \tilde{\tau} + \tilde{\delta}_{i.} + \tilde{\varepsilon}_{i..}$ and $\tilde{y}_{...} = \beta_. + \tilde{\tau} + \tilde{\varepsilon}_{...}$

where $\tilde{\delta}_{i.} = \sum_j n_{ij}w_j\delta_{ij}/\sum_j n_{ij}w_j$. Thus we have

$$E\{SS\ (\text{Treatments})\} = (t-1) + \Sigma N_{.j}w_j(\tau_j - \tilde{\tau})^2 \text{ as above,}$$

$$E\{SS\ (\text{Blocks})\} = \Sigma\Sigma n_{ij}w_j(\beta_i - \beta_. + \tilde{\delta}_{i.})^2 +$$

$$E\{\Sigma\Sigma n_{ij}w_j(\tilde{\varepsilon}_{i..} - \tilde{\varepsilon}_{...})^2\}$$

$$= (b-1) + \sum_{ij}\Sigma n_{ij}w_j(\beta_i - \beta_. + \tilde{\delta}_{i.})^2,$$

$$E\{SS\ (\text{Interactions})\} = \Sigma\Sigma n_{ij}w_j(\delta_{ij} - \tilde{\delta}_{i.})^2 +$$

$$E\{\Sigma\Sigma n_{ij}w_j(\varepsilon_{ij.} - \tilde{\varepsilon}_{i..} - \varepsilon_{.j.} + \tilde{\varepsilon}_{...})^2\}$$

$$= (b-1)(t-1) + \Sigma\Sigma n_{ij}w_j(\delta_{ij} - \tilde{\delta}_{i.})^2$$

and

$$E\{SS\ (\text{Error})\} = N_{..} - bt \text{ as above.}$$

The analysis of variance table is given below.

Analysis of variance table

Source of variation	d.f.	SS	E(MS) under constraints (I)	E(MS) under constraints (II)
Blocks	b-1	$\Sigma\Sigma n_{ij}w_j(\tilde{y}_{i..}-\tilde{y}_{...})^2$	$1+\Sigma\Sigma n_{ij}w_j(\beta_i-\beta_{.})^2/(b-1)$	$1+\Sigma\Sigma n_{ij}w_j(\beta_i-\beta_{.}+\tilde{\delta}_{i.})^2/(b-1)$
Treatments	(t-1)	$\Sigma N_{.j}w_j(y_{.j.}-\tilde{y}_{...})^2$	$1+\Sigma N_{.j}w_j(\tau_j-\tilde{\tau})^2/(t-1)$	$1+\Sigma N_{.j}w_j(\tau_j-\tilde{\tau})^2/(t-1)$
Interactions	(b-1)(t-1)	$\Sigma\Sigma n_{ij}w_j(y_{ij.}-\tilde{y}_{i..}-y_{.j.}+\tilde{y}_{...})^2$	$1+\Sigma\Sigma n_{ij}w_j\delta_{ij}^2/(b-1)(t-1)$	$1+\Sigma\Sigma n_{ij}w_j(\delta_{ij}-\tilde{\delta}_{i.})^2/(b-1)(t-1)$
Error	N..- bt	$\Sigma\Sigma\Sigma w_j(y_{ijk}-y_{ij.})^2$	1	1
Total (corrected)	N.. - 1	$\Sigma\Sigma\Sigma w_j y_{ijk}^2-(\Sigma w_j Y_{.j.})^2/\Sigma N_{.j}w_j$		

It is evident from this table that the differences in block effects cannot be tested in the presence of interactions under constraints (II).

When the F-test or χ^2-test indicates significant differences among the treatments, block effects or interaction effects, the difference between any two of the treatments, block effects or interaction effects can be tested by the normal test. In fact the variates

$$Z_1 = (\hat{\tau}_j - \hat{\tau}_k)/(1/N_{.j}w_j + 1/N_{.k}w_k)^{\frac{1}{2}}$$

$$Z_2 = (\tilde{\beta}_i - \tilde{\beta}_\ell)/\{N_{..}(1/N_{i.} + 1/N_{\ell .})/\Sigma N_{.j}w_j\}^{\frac{1}{2}}$$

and

$$Z_3 = \begin{cases} (\tilde{\delta}_{ij} - \tilde{\delta}_{ik})/\{(N_{..}/N_{i.} - 1)(1/N_{.j}w_j + 1/N_{.k}w_k)\}^{\frac{1}{2}} & \text{for } j \neq k \\ (\tilde{\delta}_{ij} - \tilde{\delta}_{\ell j})/\{(N_{..}/N_{i.} + N_{..}/N_{\ell .})(1/N_{.j}w_j - 1/\Sigma N_{.j}w_j)\}^{\frac{1}{2}} & \text{for } i \neq \ell \\ (\tilde{\delta}_{ij} - \tilde{\delta}_{\ell k})/\{(N_{..}/N_{i.} - 1)/N_{.j}w_j + (N_{..}/N_{\ell .} - 1)/N_{.k}w_k - (N_{..}/N_{i.} + N_{..}/N_{\ell .})/\sum_j N_{.j}w_j\}^{\frac{1}{2}} & \text{for } i \neq \ell \text{ and } j \neq k \end{cases}$$

are all standardised normal under the null hypotheses.

Ratios of these normal variates to the square root of the error mean square are the corresponding t-variates with $N_{..} - bt$ d.f.

3.1.3 Estimation of weights

The estimators of the linear parameters and the test-statistics involve weights, the reciprocals of the error variances which are usually unknown. One procedure in such a situation is to use the estimated weights in place of actual weights and remove the major part of the resulting bias of the estimators and other statistics as done for one-way models.

(i) Maximum likelihood estimators of the error variances

The likelihood function of the model (6) is given by

$$L = (2\pi)^{-N_{..}/2} \prod_j (\sigma_j^2)^{-N_{.j}/2} \exp\{-\tfrac{1}{2}\sum_j (1/\sigma_j^2) \sum_{ik} (y_{ijk}-\beta_i-\tau_j-\delta_{ij})^2\}.$$

Taking partial derivative of $\log_e L$ with respect to the linear parameters, we get the same normal equations as those for the weighted least squares procedure and hence the same estimators.

Also, we have

$$\frac{\partial \log_e L}{\partial \sigma_j^2} = -\frac{N_{.j}}{2}\frac{1}{\sigma_j^2} - \frac{(-1)}{2\sigma_j^4}\sum_{ik}(y_{ijk}-\hat{\beta}_i-\tilde{\tau}_j-\tilde{\delta}'_{ij})^2 = 0$$

whence the maximum likelihood estimator (MLE) of σ_j^2 is

$$\hat{\sigma}_j^2 = \frac{1}{N_{.j}} \sum_{ik} (y_{ijk} - y_{ij.})^2 \quad ; \quad j = 1,2,\ldots t,$$

since $\hat{\beta}_i+\tilde{\tau}_j+\tilde{\delta}'_{ij} = y_{ij.}$ from the last normal equation of the weighted least squares in section 3.1.2. For $j \neq j'$, $\hat{\sigma}_j^2$ and $\hat{\sigma}_{j'}^2$ are independent.

(ii) The MINQUE of error variances

From the model of section 3.1.2, we have

$$
(\underset{\sim\sim}{XX'}) = \begin{bmatrix}
N_{\cdot 1} & & & n_{11}\cdots n_{b1} & n_{11}\cdots n_{b1} & & \\
& \ddots & & \vdots \qquad \vdots & & \ddots & \\
& & N_{\cdot t} & n_{1t}\cdots n_{bt} & & & n_{1t}\cdots n_{bt} \\
n_{11}\cdots n_{1t} & & & N_{1\cdot} & n_{11} & \cdots\cdots & n_{1t} \\
\vdots \qquad \vdots & & & \ddots & \ddots & & \ddots \\
n_{b1}\cdots n_{bt} & & & N_{b\cdot} & n_{b1} & & n_{bt} \\
n_{11} & & & n_{11} & n_{11} & & \\
\vdots & & & \ddots & \ddots & & \\
n_{b1} & & & n_{b1} & n_{b1} & & \\
& \ddots & & \vdots & & \ddots & \\
& & n_{1t} & n_{1t} & & & n_{1t} \\
& & \vdots & \ddots & & & \ddots \\
& & n_{bt} & n_{bt} & & & n_{bt}
\end{bmatrix}
$$

To find a generalised inverse of $(\underset{\sim\sim}{XX'})$, we consider the bt x bt matrix obtained by deleting the first (b+t) rows and (b+t) columns of $(\underset{\sim\sim}{XX'})$. Let it be denoted by $\underset{\sim}{D}$. Then $\underset{\sim}{D}$ is diagonal and has full rank. Its inverse is given by

$$\underset{\sim}{D}^{-1} = \text{diag}\,(1/n_{11},\ldots,1/n_{b1},\ldots,1/n_{1t},\ldots,1/n_{bt}).$$

Then according to Rao (1973, p.225), a generalised inverse of $(\underset{\sim\sim}{XX'})$ is obtained if we increase the order of $\underset{\sim}{D}^{-1}$ by inserting rows and columns of zero from where the dependent rows and columns were removed. Thus, a generalised inverse of $(\underset{\sim\sim}{XX'})$ is given by

$$(\underset{\sim\sim}{XX'})^{-} = \begin{bmatrix} \underset{\sim}{0} & \vdots & \underset{\sim}{0} \\ \cdots & & \cdots \\ \underset{\sim}{0} & \vdots & \underset{\sim}{D}^{-1} \end{bmatrix}$$

From this, we have

$$\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-}\underset{\sim}{X} = \text{diag}(\underset{\sim}{J}_{n_{11}}/n_{11},\ldots,\underset{\sim}{J}_{n_{b1}}/n_{b1},\ldots,\underset{\sim}{J}_{n_{1t}}/n_{1t},\ldots,\underset{\sim}{J}_{n_{bt}}/n_{bt})$$

where $\underset{\sim}{J}_m$ is the square matrix of order m with unity as its elements.

The projection matrix $\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-}\underset{\sim}{X}$ is thus given by

$$\underset{\sim}{S} = \text{diag}(\underset{\sim}{I}_{n_{11}} - \underset{\sim}{J}_{n_{11}}/n_{11}, \underset{\sim}{I}_{n_{21}} - \underset{\sim}{J}_{n_{21}}/n_{21},\ldots,\underset{\sim}{I}_{n_{bt}} - \underset{\sim}{J}_{n_{bt}}/n_{bt})$$

where $\underset{\sim}{I}_n$ is the identity matrix of order n. The product $\underset{\sim}{S}\underset{\sim}{Y}$ gives the observed residuals.

Let $\underset{\sim}{F}$ be the matrix whose elements are the squares of the elements of the projection matrix, $\underset{\sim}{v}$ the vector of squared residuals and $\underset{\sim}{\delta}$ the vector of the variances (σ_j^2 being repeated $N._j$ times). Then according to Rao (1970), the MINQUE of σ_j^2 are obtained from the equation $\underset{\sim}{F}\,\underset{\sim}{\delta} = \underset{\sim}{v}$.

Adding the equations involving σ_j^2 and simplifying, we get

$$\sum_{i=1}^{b} (n_{ij}-1)\, s_j^2 = \sum_i \sum_k (y_{ijk} - y_{ij.})^2$$

or

$$s_j^2 = \sum_i^{b} \sum_k^{n_{ij}} (y_{ijk} - y_{ij.})^2/(N._j - b); \quad j = 1,2,\ldots,t.$$

Unlike the MLE, s_j^2 is unbiased for σ_j^2. Here also, the estimators s_j^2 and $s_{j'}^2$ are independent when $j \neq j'$.

If the number of observations in any cell is unity, then the contributions from that cell to the degrees of freedom and to the SS for calculating either MLE or MINQUE of σ_j^2, will be zero. Thus, in order to get an estimate of σ_j^2, the inequality $n_{ij} > 1$ must be satisfied for at least one cell for the jth treatment.

As the estimators s_j^2 are independent, Bartlett's χ^2-test can be applied for testing the homogeneity of error variances in this case also.

It is obvious that the variate $\sum_{j=1}^{n_{ij}} (y_{ijk} - y_{ij.})^2$ is distributed as $\chi^2 \sigma_j^2$ with $(n_{ij} - 1)$ d.f. so that $(N_{.j} - b)\, s_j^2/\sigma_j^2$ is distributed as χ^2 with $(N_{.j} - b)$ d.f.

3.1.4 Adjustment of the estimators of the linear parameters

Since the estimators of the treatment parameters do not involve weights, no adjustment is necessary for these. Estimators of the block effects involve weights which also occur in the expressions of the estimators of interactions. To remove a major portion of the bias when estimated weights are used in the estimators of the linear parameters, the estimators have to be adjusted by Theorem 1 due to Meier.

Let $x_j = s_j^2/\sigma_j^2$. Then the estimated weight $\hat{w}_j = 1/s_j^2 = 1/x_j\sigma_j^2$. The MLE of σ_j^2 may also be used in defining $\hat{w}_j$. The estimators, using estimated weights, of block and interaction effects are

$$\hat{\tilde{\beta}}_i = \hat{\tilde{y}}_{i..} = \sum_j n_{ij}\hat{w}_j\, y_{ij.}/\sum_j n_{ij}\hat{w}_j$$

and

$$\hat{\tilde{\delta}}_{ij} = (y_{ij.} - \hat{\tilde{y}}_{i..} - y_{.j.} + \hat{\tilde{y}}_{...})$$

with $\hat{\tilde{y}}_{...} = \sum_j N_{.j}\hat{w}_j y_{.j.}/\sum_j N_{.j}\hat{w}_j$.

$$\text{Since } \left[\frac{\partial^2 \hat{\tilde{\beta}}_i}{\partial x_j^{\,2}}\right]_{\text{all } x_j=1} = 2\, f_j\, (1 - f_j)(y_{ij\cdot} - \tilde{y}_{i\cdot\cdot}),$$

the adjusted estimator of β_i is given by

$$\hat{\tilde{\beta}}_i(\text{adj}) = \hat{\tilde{y}}_{i\cdot\cdot} - 2 \sum_{j=1}^{t} \hat{f}_j(1-\hat{f}_j)(y_{ij\cdot} - \hat{\tilde{y}}_{i\cdot\cdot})/(N_{\cdot j} - b)$$

with $f_j = n_{ij}w_j / \sum_j n_{ij}w_j = N_{\cdot j}w_j / N_{\cdot j}w_j$ by the proportionality condition, $(N_{\cdot j}-b)$ as the d.f. for the estimator of $\sigma_j^{\,2}$ and $\hat{f}_j = N_{\cdot j}\hat{w}_j/\Sigma N_{\cdot j}\hat{w}_j$. Similarly,

$$\hat{\tilde{y}}_{\cdot\cdot\cdot}\ (\text{adj}) = \hat{\tilde{y}}_{\cdot\cdot\cdot} - 2 \sum_{1}^{t} \hat{f}_j(1-\hat{f}_j)(y_{\cdot j\cdot}-\hat{\tilde{y}}_{\cdot\cdot\cdot})/(N_{\cdot j}-b),$$

so that $\hat{\tilde{\delta}}_{ij}(\text{adj}) = y_{ij\cdot} - \hat{\tilde{y}}_{i\cdot\cdot}(\text{adj}) - y_{\cdot j\cdot} + \hat{\tilde{y}}_{\cdot\cdot\cdot}(\text{adj})$.

3.1.5 Adjustment of the test-statistics

(i) Adjustment of F-statistics

The error sum of squares (SS) using estimated weights based on the MINQUE of error variances is

$$\sum_j \{\sum_{ik}\sum (y_{ijk} - y_{ij\cdot})^2 / \left[\sum_{ik}\sum (y_{ijk} - y_{ij\cdot})^2/(N_{\cdot j} - b)\right]\} = N_{\cdot\cdot}-bt,$$

a constant. Similarly, the SS due to error, using the estimated weights based on MLE of error variances is also a constant. Hence, no adjustment of the error SS is necessary for removal of bias of the F-statistics.

The SS(treat.) using estimated weights is $\Sigma N_{\cdot j}w_j(y_{\cdot j\cdot} - \hat{\tilde{y}}_{\cdot\cdot\cdot})^2$. This is exactly in the same form as that of the SS(treat.) using estimated weights in the one-way model with unequal group variances (section 2.1.4). Hence, the adjusted SS(treat.) using estimated weights will be of the same form as that in the one-way model and it is given by

$$\text{Adjusted SS(treat)} = \Sigma N._j \hat{w}_j (y._j. - \hat{\tilde{y}}...)^2 \ \{1 - 2(1-\hat{f}_j)/(N._j - b)\}$$

Thus, the adjusted F-statistic for testing treatment differences is given by

$$\hat{F}_1(adj) = (N.. - bt) \ \sum_{j=1}^{t} N._j \hat{w}_j (y._j. - \hat{\tilde{y}}...)^2 \{ 1 - \frac{2(1-\hat{f}_j)}{N._j - b} \} /(t-1)$$

$$\{\Sigma\Sigma\Sigma \hat{w}_j (y_{ijk} - y_{ij}.)^2\} \qquad ....(8)$$

with (t-1) and (N.. - bt) d.f.

To find the adjustment for the other two sums of squares, we see that

$$\frac{\partial \hat{\tilde{y}}_i..}{\partial x_j} = \frac{n_{ij}}{\sigma_j^2} (\sum_j n_{ij} \hat{w}_j y_{ij}. - y_{ij}. \sum_j n_{ij} \hat{w}_j)/x_j^2 (\sum_j n_{ij} \hat{w}_j)^2$$

and

$$\left[\frac{\partial^2 \hat{\tilde{y}}_i..}{\partial x_j^2} \right]_{\text{all } x_j = 1} = - 2 f_j (1 - f_j)(\tilde{y}_i.. - y_{ij}.)$$

Similarly,

$$\frac{\partial \hat{\tilde{y}}...}{\partial x_j} = \frac{N._j}{\sigma_j^2} (\sum_j N._j w_j y._j. - y._j. \sum_j N._j w_j)/x_j^2 (\sum_j N._j w_j)^2$$

and

$$\left[\frac{\partial^2 \hat{\tilde{y}}...}{\partial x_j^2} \right]_{\text{all } x_j = 1} = - 2 f_j (1 - f_j)(\tilde{y}... - y._j.) \ .$$

The estimated block SS is given by

$$\text{Est. SS(block)} = \sum_{ij}\sum n_{ij}\hat{w}_j(\hat{\tilde{y}}_{i..} - \hat{\tilde{y}}_{...})^2$$

so that

$$\frac{\partial[\text{Est.SS(bl)}]}{\partial x_j} = -\frac{1}{x_j^2\sigma_j^2}\sum_i n_{ij}(\hat{\tilde{y}}_{i..}-\hat{\tilde{y}}_{...})^2 + 2\sum\sum n_{ij}\hat{w}_j$$

$$(\hat{\tilde{y}}_{i..} - \hat{\tilde{y}}_{...})\frac{\partial(\hat{\tilde{y}}_{i..}-\hat{\tilde{y}}_{...})}{\partial x_j}$$

and

$$\left[\frac{\partial^2\{\text{Est.SS(bl)}\}}{\partial x_j^2}\right]_{\text{all } x_j=1}$$

$$= 2\sum_i n_{ij}w_j(\tilde{y}_{i..}-\tilde{y}_{...})^2 + 4\sum_i n_{ij}w_jf_j(\tilde{y}_{i..}-\tilde{y}_{...})(y_{ij.}-y_{.j.}-\tilde{y}_{i..}+\tilde{y}_{...})$$

$$+ 2\sum_{ij}\sum n_{ij}w_jf_j^2(y_{ij.}-\tilde{y}_{i..}-y_{.j.}+\tilde{y}_{...})^2 + 4\sum_{ij}\sum n_{ij}w_jf_j(1-f_j)(\tilde{y}_{i..}-\tilde{y}_{...})$$

$$(y_{ij.}-\tilde{y}_{i..}-y_{.j.}+\tilde{y}_{...}),$$

Thus

$$\text{Est. SS(bl.)(adj)} = \sum_{ij}\sum n_{ij}\hat{w}_j(\hat{\tilde{y}}_{i..}-\hat{\tilde{y}}_{...})^2 - \sum_j\frac{1}{N_{.j}-b}$$

$$\left[\frac{\partial^2\text{Est.SS(bl)}}{\partial x_j^2}\right]_{\text{all } x_j=1} \quad \text{using estimated weights.}$$

$$= \sum_{ij}\sum n_{ij}\hat{w}_j(\hat{\tilde{y}}_{i..}-\hat{\tilde{y}}_{...})^2\left(1-\frac{2}{N_{.j}-b}\right) - 2\sum_{ij}\sum n_{ij}\hat{w}_j\hat{f}_j^2(y_{ij.}-\hat{\tilde{y}}_{i..}-$$

$$y_{.j.}+\hat{\tilde{y}}_{...})^2\left(\sum_j\frac{1}{N_{.j}-b}\right)$$

$$- 4\sum_{ij}\sum n_{ij}\hat{w}_j\hat{f}_j(\hat{\tilde{y}}_{i..}-\hat{\tilde{y}}_{...})(y_{ij.}-\hat{\tilde{y}}_{i..}-y_{.j.}+\hat{\tilde{y}}_{...})\left\{\frac{1}{N_{.j}-b}+(1-\hat{f}_j)\right.$$

$$\left.\left(\sum\frac{1}{N_{.j}-b}\right)\right\} \quad \ldots(9)$$

with (b-1) d.f.

The adjusted F-statistic for testing differences in block effects is thus given by

$$\hat{F}_2(\text{adj}) = \frac{(N_{..}-bt)\ \{\text{Est. SS(bl.)(adj)}\}}{(b-1)\{\ \Sigma\Sigma\Sigma\hat{w}_j(y_{ijk} - y_{ij\cdot})^2\ \}}$$

with (b-1) and (N..-bt) d.f.

The estimated SS due to interactions is given by

$$\text{Est. SS(Int)} = \sum_{ij}\sum n_{ij}\hat{w}_j y_{ij\cdot}(y_{ij\cdot}-\hat{\tilde{y}}_{i\cdot\cdot}-y_{\cdot j\cdot}+\hat{\tilde{y}}_{\cdots})$$

so that $$\frac{\partial\{\text{Est.SS(Int)}\}}{\partial x_j} = \sum_i n_{ij}\frac{(-1)}{x_j^2\sigma_j^2}\, y_{ij\cdot}(y_{ij\cdot}-\hat{\tilde{y}}_{i\cdot\cdot}-y_{\cdot j\cdot}+\hat{\tilde{y}}_{\cdots})$$

$$- \sum_{ij}\sum n_{ij}\hat{w}_j y_{ij\cdot}\ \frac{\partial(\hat{\tilde{y}}_{i\cdot\cdot}-\hat{\tilde{y}}_{\cdots})}{\partial x_j}$$

and $$\left[\frac{\partial^2\{\text{Est.SS(Int)}\}}{\partial x_j^2}\right]_{\text{all } x_j=1} = 2\sum_i n_{ij}w_j y_{ij\cdot}(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots})$$

$$+ 2\sum_i n_{ij}w_j y_{ij\cdot}\left[\frac{\partial\ (\hat{\tilde{y}}_{i\cdot\cdot}-\hat{\tilde{y}}_{\cdots})}{\partial x_j}\right]_{\text{all } x_j=1}$$

$$- \sum_{ij}\sum n_{ij}w_j y_{ij\cdot}\left[\frac{\partial^2(\hat{\tilde{y}}_{i\cdot\cdot}-\hat{\tilde{y}}_{\cdots})}{\partial x_j^2}\right]_{\text{all } x_j=1}$$

Hence, $$\text{Est. SS(Int)(adj)} = \sum_{ij}\sum n_{ij}\hat{w}_j(y_{ij\cdot}-\hat{\tilde{y}}_{i\cdot\cdot}-y_{\cdot j\cdot}+\hat{\tilde{y}}_{\cdots})^2$$

$$- \sum_j\frac{1}{N_{\cdot j}-b}\left[\frac{\partial^2\{\text{Est.SS(Int)}\}}{\partial x_j^2}\right]_{\text{all } x_j=1}$$

(using estimated weights)

$$= \Sigma\Sigma\, n_{ij}\hat{w}_j\; y_{ij.}(y_{ij.} - \hat{y}_{i..} - y_{.j.} + \hat{y}_{...})\left\{1 - 2(1-\hat{f}_j)\left(\frac{1}{N_{.j}-b} - \hat{f}_j \sum_j \frac{1}{N_{.j}-b}\right)\right\}, \quad \ldots(10)$$

on simplification, with $(b-1)(t-1)$ d.f.

The adjusted F-statistic for testing differences in interaction effects is thus given by

$$\hat{F}_3(\text{adj}) = \frac{(N_{..}-bt)\{\text{Est. SS (Int) (adj)}\}}{(b-1)(t-1)\{\Sigma\Sigma\Sigma\, \hat{w}_j\, (y_{ijk}-y_{ij.})^2\}}$$

with $(b-1)$ $(t-1)$ and $(N_{..}-bt)$ d.f.

(ii) Adjustment of the approximate normal test-statistics

The approximate normal test-statistic, using estimated weights, for testing the difference between the jth and kth treatment effects is

$$\hat{z}_1 = |\hat{\tau}_j - \hat{\tau}_k| \,/\, \{1/N_{.j}\hat{w}_j + 1/N_{.k}\hat{w}_k\}^{\frac{1}{2}}$$

This is in the same form as the corresponding normal test-statistic in the one-way model with unequal group variances. Hence, the adjusted normal test-statistic will be of the same form as that in one-way model (section 2.1.4) and it is given by

$$\hat{z}_1(\text{adj}) = \hat{z}_1\left\{1 - 3\left[1/N_{.j}^2(N_{.j}-b)\hat{w}_j^2 + 1/N_{.k}^2(N_{.k}-b)\hat{w}_k^2\right] / 4(1/N_{.j}\hat{w}_j + 1/N_{.k}\hat{w}_k)^2\right\}$$

The approximate normal test-statistic using estimated weights for testing the difference between the ith and ℓth block effects is

$$\hat{z}_2 = \left|\hat{\tilde{\beta}}_i(\text{adj}) - \hat{\tilde{\beta}}_\ell(\text{adj})\right| / (\Sigma N._j w_j)^{-\frac{1}{2}} \left(\frac{N..}{N_{i\cdot}} + \frac{N..}{N_{\ell\cdot}}\right)^{\frac{1}{2}}$$

$$= H\ (\Sigma N._j w_j)^{\frac{1}{2}}, \text{ say.}$$

Since
$$\left[\frac{\partial^2 \hat{z}_2}{\partial x_j^2}\right]_{\text{all } x_j = 1} = \frac{H\ N._j w_j}{(\Sigma N._j w_j)^{\frac{1}{2}}} \left(1 - \frac{N._j w_j}{4 \Sigma N_j w_j}\right),$$

it follows that

$$\hat{z}_2(\text{adj}) = \frac{|\hat{\tilde{\beta}}_i(\text{adj}) - \hat{\tilde{\beta}}_\ell(\text{adj})|}{\left\{\frac{N..}{\Sigma N_j \hat{w}_j}\left(\frac{1}{N_{i\cdot}} + \frac{1}{N_{\ell\cdot}}\right)\right\}^{\frac{1}{2}}} \left\{1 - \sum_{j=1}^{t} \frac{\hat{f}_j(1 - \hat{f}_j/4)}{N._j - b}\right\}.$$

Here, var $\{\hat{\tilde{\beta}}_i(\text{adj})\}$ has been approximated by var $(\tilde{\beta}_i) = 1/\sum_j n_{ij} w_j$ since the former is difficult to find analytically.

Similarly, if we approximate var $\{\hat{\delta}_{ij}(\text{adj})\}$ by var $(\tilde{\delta}_{ij})$, the approximate normal test-statistic using estimated weights for testing the difference between the (i,j)th and (ℓ,k)th interaction parameters is given by

$$\hat{z}_3 = \begin{cases} |\hat{\tilde{\delta}}_{ij}(\text{adj}) - \hat{\tilde{\delta}}_{ik}(\text{adj})| / \{(N../N_{i\cdot} - 1)(1/N._j\hat{w}_j + 1/N._k\hat{w}_k)\}^{\frac{1}{2}} & \text{for } j \neq k \\ |\hat{\tilde{\delta}}_{ij}(\text{adj}) - \hat{\tilde{\delta}}_{\ell j}(\text{adj})| / \{(N../N_{i\cdot} + N../N_{\ell\cdot})(1/N._j\hat{w}_j - 1/\Sigma N._j\hat{w}_j)\}^{\frac{1}{2}} & \text{for } i \neq \ell \\ |\hat{\tilde{\delta}}_{ij}(\text{adj}) - \hat{\tilde{\delta}}_{\ell k}(\text{adj})| / \{(N../N_{i\cdot} - 1)/N._j w_j + (N../N_{\ell\cdot} - 1) \\ \quad /N._k w_k - (N../N_{i\cdot} + N../N_{\ell\cdot})/\Sigma N._j w_j\}^{\frac{1}{2}} & \text{for } i \neq \ell \ \&\ j \neq k \end{cases}$$

The corresponding adjusted approximate normal test-statistic is

$$\hat{Z}_3(adj) = \begin{cases} \hat{Z}_3\left[1 - 3\sum_{u=j,k} 1/\{4N._u^2 w_u^2 (N._u - b)(1/N._j w_j + 1/N._k w_k)^2\}\right] & \text{for } j \neq k \\ \hat{Z}_3\left[1 + \{\hat{f}_j^2 - 3(1+\hat{f}_j^2)/4\}/(N._j - b) + \sum_{u \neq j}^{t} G_u\right] & \text{for } i \neq \ell \\ \hat{Z}_3\left[1 + \sum_{m=k,j} H_m + \sum_{u(\neq j \neq k)}^{t} (p_i + q_\ell + 2)\hat{f}_u \{1 - \hat{f}_u - 3\hat{f}_u (p_i + q_\ell + 2)/4P\}/P(N._u - b)\right] & \text{for } i \neq \ell \text{ and } j \neq k \end{cases}$$

where $G_u = \hat{f}_u \hat{f}_j \{1 - \hat{f}_u\left[1 + 3\hat{f}_j/4(1-\hat{f}_j)\right]\}/(1-\hat{f}_j)(N._u - b)$,

$$H_m = \left[\hat{f}_m (p_i + q_\ell + 2)(1 - \hat{f}_m) - 3\{p_i - \hat{f}_m^2 (p_i + q_\ell + 2)^2 / 4\hat{f}_m^2 P\}\right] / P(N._m - b),$$

$P = (p_i/\hat{f}_j + q_\ell/\hat{f}_k - p_i - q_\ell - 2)$, $p_i = (N../N_i. - 1)$

and

$q_\ell = (N../N._\ell - 1)$.

Ratios of these adjusted statistics to the square root of error mean square are the corresponding adjusted t-variates with $N..-bt$ d.f.

3.1.6 Multiple comparison

The inequality (5) of section 2.1.5, may be used to estimate the joint confidence intervals of contrasts of treatment, block or interaction effects.

(i) Treatment contrasts

Since the estimators of the treatment parameters and their variances are in the same forms as those for the one-way model, the formula with necessaty adjustment obtained in section 2.1.5 is also applicable here. If $\hat{\psi}_1 = \Sigma c_j \hat{\tau}_j$ is an estimate of the treatment contrast, $\psi_1 = \Sigma c_j \tau_j$, then the estimated joint confidence interval for all ψ_1 is given by

$$\hat{\psi}_1 - S_1 s\, \hat{\sigma}_{\hat{\psi}_1}(adj) \leq \psi_1 \leq \hat{\psi}_1 + S_1 s\, \hat{\sigma}_{\hat{\psi}_1}(adj)$$

where $S_1 = \{(t-1)F_\alpha(t-1, N..-bt)\}^{\frac{1}{2}}$; s = square root of mean square error and

$$\hat{\sigma}_{\psi_1}(adj) = \{ \sum_j (c_j^2 s_j^2/N._j) \}^{\frac{1}{2}} \{1 + \frac{1}{4} \sum_{j=1}^{t} \frac{c_j^4 s_j^4}{N._j^2(N._j - b)}$$

$$\left[\Sigma(c_j^2 s_j^2/N._j)\right]^{-2}\}$$ using the MINQUE of σ_j^2 as the estimator.

Approximating the variances of $\hat{\tilde{\beta}}_i(adj)$ and $\hat{\tilde{\delta}}_{ij}(adj)$ by those of $\tilde{\beta}_i$ and $\tilde{\delta}_{ij}$ respectively, we get the estimated joint confidence intervals of β-contrasts and interaction contrasts as follows.

(ii) β-contrasts

If $\psi_2 = \Sigma C_i \beta_i$, then var $(\tilde{\psi}_2)$ = var $(\Sigma C_i \tilde{\beta}_i)$ = $\sum_i (C_i^2/\sum_j n_{ij} w_j) = d/\sum_j N._j w_j$ where $d = \sum_i (C_i^2 N../N_{i.})$ by the proportionality condition. Thus $\hat{\sigma}_{\tilde{\psi}_2} = d^{\frac{1}{2}}(\sum_j N._j \hat{w}_j)^{-\frac{1}{2}}$.

The estimated joint confidence interval for all contrasts ψ_2 is then given by

$$\hat{\tilde{\psi}}_2 - S_2 s\, \hat{\sigma}_{\tilde{\psi}_2}(adj) \leq \psi_2 \leq \hat{\tilde{\psi}}_2 + S_2\, s\, \hat{\sigma}_{\tilde{\psi}_2}(adj)$$

where $S_2 = \{(b-1)F_\alpha(b-1, N..-bt)\}^{\frac{1}{2}}$, $\hat{\sigma}_{\tilde{\psi}_2}(adj)$ =

$$(d/\sum_j N._j \hat{w}_j)^{\frac{1}{2}}\{1 + \sum_{j=1}^{t} \frac{\hat{f}_j}{N._j - b}(1 - \tfrac{3}{4}\hat{f}_j)\}$$

and $\hat{\tilde{\beta}}_i(adj)$ are used in computing $\hat{\tilde{\psi}}_2$.

(iii) Interaction contrasts

If $\psi_3 = \sum_{ij}\sum c_{ij}\delta_{ij}$ is the interaction contrast, then

$$\text{var}(\hat{\psi}_3) = \text{var}(\sum_{ij}\sum c_{ij}\tilde{\delta}_{ij}) = \Sigma\Sigma c_{ij}^2 \text{var}(\tilde{\delta}_{ij}) + \sum_{ij}\sum_{\neq k}\sum c_{ij}c_{ik}$$

$$\text{cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{ik}) + \sum_{i \neq \ell}\sum_j\sum c_{ij}c_{\ell j}\,\text{cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{\ell j}) + \sum_{i \neq \ell}\sum\sum_{j \neq k}\sum c_{ij}c_{\ell k}$$

$\operatorname{cov}(\tilde{\delta}_{ij}, \tilde{\delta}_{\ell k}) = \Sigma Q_j (N._j w_j)^{-1} - Q(\Sigma N._j w_j)^{-1}$ where $Q_j = \sum\limits_i c_{ij}^2$ $(N../N_i.-1) - \sum\limits_{\ell \neq i} \Sigma c_{ij} c_{\ell j}$ and $Q = \sum\limits_i \sum\limits_j c_{ij}^2 (N../N_i.-1) + \sum\limits_i \sum\limits_{j \neq k} \Sigma$ $c_{ij} c_{ik} (N../N_i.-1) - \sum\limits_{i \neq \ell} \sum\limits_j \Sigma c_{ij} c_{\ell j} - \sum\limits_{i \neq \ell,} \sum\limits_{j \neq k} \Sigma \Sigma c_{ij} c_{\ell k}$. Thus the estimated standard error of $\hat{\psi}_3$ is given by

$$\hat{\sigma}_{\hat{\psi}_3} = \{\sum_j Q_j (N._j \hat{w}_j)^{-1} - Q(\Sigma N._j \hat{w}_j)^{-1}\}^{\frac{1}{2}} .$$

The joint confidence interval for all interaction contrasts ψ_3 is then estimated by

$$\hat{\psi}_3 - S_3 \, s \hat{\sigma}_{\hat{\psi}_3} (\text{adj}) \leqslant \psi_3 \leqslant \hat{\psi}_3 + S_3 \, s \hat{\sigma}_{\hat{\psi}_3} (\text{adj})$$

where $S_3 = \{ (b-1)(t-1) F_\alpha [(b-1)(t-1), N..-bt] \}^{\frac{1}{2}}$, $\hat{\tilde{\delta}}_{ij}(\text{adj})$ are used in computing $\hat{\psi}_3$ and

$$\hat{\sigma}_{\hat{\psi}_3} (\text{adj}) = h_1^{\frac{1}{2}} \Big[1 + \sum_{j \neq 1}^{t} N._j \hat{w}_j \{ h_{2j} (1 + N._j \hat{w}_j h_{2j}/4h_1) - N._j \hat{w}_j h_{3j} \} / h_1 (N._j - b) \Big]$$

with $h_1 = (\Sigma Q_j / N._j \hat{w}_j - Q / \Sigma N._j \hat{w}_j)$, $h_{2j} = Q_j / (N._j \hat{w}_j)^2 + Q/(\Sigma N._j \hat{w}_j)^2$

and

$$h_{3j} = Q_j / (N._j \hat{w}_j)^3 + Q/(\Sigma N._j \hat{w}_j)^3 .$$

3.2 Two-way random models.

Let the random model be

$$y_{ijk} = \beta_i + \tau_j + \delta_{ij} + \varepsilon_{ijk}$$

$$(i = 1,2,\ldots,b \; ; \; j = 1,2,\ldots,t \; ; \; k = 1,2,\ldots,n_{ij})$$

where β_i is the random effect of the ith block having mean μ and variance σ_β^2, τ_j the random effect of the jth treatment having mean zero and variance σ_τ^2, δ_{ij} the random effect of the interaction between the ith block effect and the jth treatment effect, having mean zero and variance σ_{δ}^{2}, and ε_{ijk} the error term having mean zero and variance σ_{j}^{2}. All the random effects and the errors are assumed to be independent of one another.

3.2.1 Estimation of the variance components and the analysis when error variances are known

From the above model we have,

$y_{ij\cdot} = \beta_i + \tau_j + \delta_{ij} + \varepsilon_{ij\cdot}$, $y_{\cdot j\cdot} = \beta_{\cdot} + \tau_j + \delta_{\cdot j} + \varepsilon_{\cdot j\cdot}$,

$\tilde{y}_{i\cdot\cdot} = \sum_j n_{ij} w_j y_{ij\cdot} / \sum_j n_{ij} w_j = \beta_i + \tilde{\tau} + \tilde{\delta}_{i\cdot} + \tilde{\varepsilon}_{i\cdot\cdot}$ and $\tilde{y}_{\cdots} = \beta_{\cdot} + \tilde{\tau} + \tilde{\delta}_{\cdot\cdot} + \tilde{\varepsilon}_{\cdots}$.

Since, $E\{\sum_{ik}\sum(y_{ijk}-y_{ij\cdot})^2\} = \sum_i E\{\sum_k(\varepsilon_{ijk}-\varepsilon_{ij\cdot})^2\} = \sum_i^b (n_{ij}-1)\sigma_j^2 = (N_{\cdot j}-b)\,\sigma_j^2$, the quantity $\sum_{ik}\sum(y_{ijk}-y_{ij\cdot})^2/(N_{\cdot j}-b)$ is still the unbiased estimator of σ_j^2 ; $j = 1,2,\ldots,t$.

Now E (Treatments SS) $= E\{\sum_{j=1}^{t} N_{\cdot j} w_j (y_{\cdot j\cdot} - \tilde{y}_{\cdots})^2\}$

$$= E\{\sum_j N_{\cdot j} w_j(\tau_j - \tilde{\tau})^2\} + E\{\sum_j N_{\cdot j} w_j(\delta_{\cdot j} - \tilde{\delta}_{\cdot\cdot})^2$$

$$+ E\{\sum N_{\cdot j} w_j(\varepsilon_{\cdot j\cdot} - \tilde{\varepsilon}_{\cdots})^2$$

$$= (t-1) + (w_{\cdot} - \sum N_{\cdot j}^2 w_j^2/w_{\cdot})\,(\sigma_\tau^2 + \sigma_\delta^2/b),$$

with $w_{\cdot} = \sum_j N_{\cdot j} w_j$,

$$E(\text{Blocks SS}) = E\{\sum_{ij}\sum n_{ij}\, w_j(\tilde{y}_{i}.. - \tilde{y}...)^2\}$$

$$= E\{\sum_{ij}\sum n_{ij}w_j(\beta_i - \beta.)^2\} + E\{\sum\sum n_{ij}w_j(\tilde{\delta}_i. - \tilde{\delta}..)^2\}$$

$$+ E\{\sum\sum n_{ij}w_j(\tilde{\varepsilon}_i.. - \tilde{\varepsilon}...)^2\}$$

$$= (b-1) + \sigma_\beta^2 \sum_j N._j w_j(1 - {}^1/_b) + \sigma_\delta^2(\sum_j N._j^2 w_j^2/w.)(1 - {}^1/_b)$$

since $E(\beta_i - \beta.)^2 = \sigma_\beta^2(1 - {}^1/_b)$ and $E(\tilde{\delta}_{i\cdot} - \tilde{\delta}..)^2$

$=\sigma_\delta^2(1 - {}^1/_b)(\sum_j N._j^2 w_j^2/w^2.)$

and

$$E(\text{Interactions SS}) = E\{\sum_{ij}\sum n_{ij}w_j(y_{ij}. - \tilde{y}_i.. - y._j. + \tilde{y}...)^2\}$$

$$= E\{\sum\sum n_{ij}w_j(\delta_{ij} - \tilde{\delta}_{i\cdot} - \delta._j + \tilde{\delta}..)^2\} + E\{\sum\sum n_{ij}w_j$$

$$(\varepsilon_{ij}. - \tilde{\varepsilon}_i.. - \varepsilon._j. + \tilde{\varepsilon}...)^2\}$$

$$= (b-1)(t-1) + \sigma_\delta^2(1 - {}^1/_b)(w. - \sum N._j^2 w_j^2/w.)$$

since $E(\delta_{ij} - \delta._j)^2 = \sigma_\delta^2(1 - {}^1/_b)$; $E(\tilde{\delta}_i. - \tilde{\delta}..)^2$

$=\sigma_\delta^2(1 - {}^1/_b)(\sum N._j^2 w_j^2/w^2.)$

and $E(\delta_{ij} - \delta._j)(\tilde{\delta}_i. - \tilde{\delta}..) = (N._j w_j/w.)\sigma_\delta^2(1 - {}^1/_b)$.

From these expectations, it follows that the unbiased estimators of the other three variance components are given by

$$\tilde{\sigma}_\delta^2 = b\{\text{Interactions SS} - (b-1)(t-1)\}/(w. - \sum_j N._j^2 w_j^2/w.)$$
$$\times(b-1),$$

$$\tilde{\sigma}_\tau^2 = \{\text{Treatments SS} - \text{Interaction SS}/(b-1)\}/(w. - \sum N._j^2 w_j^2/w.)$$

and

$$\tilde{\sigma}_\beta^2 = b(\text{Blocks SS} - b+1)/(b-1)w. - \{\text{Interaction SS} - (b-1)(t-1)\} / (w.^2/\Sigma N._j^2 w_j^2 - 1)$$

when the actual weights, $w_j = 1/\sigma_j^2$, are known.

Finally $E\{\text{Within (Error) SS}\} = E\{\Sigma\Sigma\Sigma\ w_j(y_{ijk}-y_{ij\cdot})^2\}$

$$= E\{\Sigma\Sigma\Sigma\ w_j(\varepsilon_{ijk} - \varepsilon_{ij\cdot})^2\}$$

$$= (N.. - bt)$$

as before.

It can easily be shown that $(\varepsilon_{ijk} - \varepsilon_{ij\cdot})$ is uncorrelated with $(\varepsilon_{\cdot j\cdot} - \tilde{\varepsilon}_{\cdots})$, $(\tilde{\varepsilon}_{i\cdot\cdot} - \tilde{\varepsilon}_{\cdots})$ and $(\varepsilon_{ij\cdot} - \tilde{\varepsilon}_{i\cdot\cdot} - \varepsilon_{\cdot j\cdot} + \tilde{\varepsilon}_{\cdots})$. Hence, by the assumption of normality of errors, the error SS is independent of the treatments SS, the blocks SS and the interaction SS. Similarly, the last three sums of squares are also mutually independent. Furthermore, each of these three sums of squares is distributed as non-central χ^2 times a constant while the error SS is always distributed as a central χ^2.

The hypotheses can thus be tested in the following way.

To test the hypothesis, H_1: $\sigma_\delta = 0$, we see that

$$F = \frac{\text{Interaction SS}/(b-1)(t-1)}{\text{Error SS}/(N..-bt)}$$

is a central F-variate under the hypothesis H_1 with $(b-1)(t-1)$ and $(N..-bt)$ d.f.

For tests of significance of σ_τ and σ_β, we are to consider two cases.

Case I: $\sigma_\delta = 0$.

In this case, it follows that

(i) $\{\text{Treatments SS}/(t-1)\}/\{\text{Error SS}/(N_{..}-bt)\}$

and (ii) $\{\text{Blocks SS}/(b-1)\}/\{\text{Error SS}/(N_{..}-bt)\}$

are central F with corresponding d.f. under the hypotheses, $H_2: \sigma_\tau = 0$ and $H_3: \sigma_\beta = 0$ respectively.

Case II: $\sigma_\delta \neq 0$.

In this case, σ_δ occurs in the expectations of both the treatments SS and the blocks SS.

To test the hypothesis $H_2: \sigma_\tau = 0$, we find that

$$\{\text{Treatments SS}/(t-1)\}/\{\text{Interaction SS}/(b-1)(t-1)\}$$

$$= F\{t-1,(b-1)(t-1)\}\left[1+\{\sigma_\tau^2(w_. - \Sigma N_{.j}^2 w_j^2/w_.)/(t-1)\} / \{1+\sigma_\delta^2(w_. - \Sigma N_{.j}^2 w_j^2/w_.)/b(t-1)\}\right]$$

$= F\left[t-1,(b-1)(t-1)\right]$ under the hypothesis $H_2: \sigma_\tau = 0$.

This test is valid also when $\sigma_\delta = 0$.

On the other hand, since

$$\{\text{Blocks SS}/(b-1)\}/\{\text{Interaction SS}/(b-1)(t-1)\}$$

$$= F\{b-1,(b-1)(t-1)\}\left[1+\{w_.\sigma_\beta^2/b + \sigma_\delta^2 A\}/\{1+\sigma_\delta^2(w_. - \Sigma N_{.j}^2 w_j^2/w_.)/b(t-1)\}\right]$$

where $A = \{\Sigma N_{.j}^2 w_j^2 - (w_. - N_{.j}^2 w_j^2)/(t-1)\}/bw_.$,

it follows that the hypothesis $H_3: \sigma_\beta = 0$ cannot be tested in this way when $\sigma_\delta \neq 0$.

3.2.2 Adjustment of the test-statistics and the estimators of variance components

Since the estimated error variances are independently distributed as multiples of χ^2-variates, the test-statistics and the estimated variance components using estimated weights can be adjusted as before to remove the bias of order

$$\left(\sum_j \frac{1}{N._j - b} \right) .$$

(i) Test-statistics

The F-statistic for testing H_1: $\sigma_\delta = 0$ is the same as that for testing the significance of interaction effects in section 3.1.2. Hence, the adjusted F-statistic using estimated weights will also be the same (section 3.1.5) i.e.,

$$\hat{F}_4(\text{adj}) = \frac{(N..-bt) \ \{\text{Interaction SS(using estimated weights\&adj.)}\}}{(b-1)(t-1)\{\sum_i \sum_j \sum_k \hat{w}_j (y_{ijk} - y_{ij}.)^2\}}$$

with (b-1)(t-1) and (N..-bt) d.f. where the interaction SS (using estimated weights & adj.) is given by equation (10) of section 3.1.5.

Similarly, for testing H_2: $\sigma_\tau = 0$ when $\sigma_\delta = 0$, the adjusted F-statistic using estimated weights is given, as in equation (8) of section 3.1.5, by

$$\hat{F}_5(\text{adj}) = \frac{(N..-bt) \ \sum N._j \hat{w}_j.(y._j.-\hat{\hat{y}}...)^2 \ \{1 - \frac{2\,(1-\hat{f}_j)}{N._j - b}\}}{(t-1)\{\sum\sum\sum \ \hat{w}_j (y_{ijk} - y_{ij}.)^2\}}$$

with (t-1) and (N..-bt) d.f., where $\hat{f}_j = N._j \hat{w}_j / \sum N._j \hat{w}_j$.

When $\sigma_\delta \neq 0$, the adjusted F-statistic for testing H_2: $\sigma_\tau = 0$ is more complicated. The F-statistic using estimated weights is

$$\hat{F}_6 = \frac{(b-1) \ \{\text{Treatments SS using estimated weights}\}}{\text{Interaction SS using estimated weights}}$$

with (t-1) and (b-1)(t-1) d.f.

Both the numerator and the denominator of $\hat{F}_6$ depend on estimated weights. Hence

$$\hat{F}_6 \text{ (adj)} = \hat{F}_6 - \sum_j \frac{1}{N_{\cdot j} - b} \left[\frac{\partial^2 (\hat{F}_6)}{\partial x_j^2} \right]_{\text{all } x_j = 1} \quad \text{using estimated weights.}$$

Denoting the treatments SS and the interaction SS, both using estimated weights, by TSS and ISS respectively, we have

$$\frac{\partial^2 (\hat{F}_6)}{\partial x_j^2} = \frac{b-1}{(ISS)^3} \left\{ (ISS)^2 \frac{\partial^2 TSS}{\partial x_j^2} - ISS.TSS \frac{\partial^2 ISS}{\partial x_j^2} - \right.$$

$$\left. 2\, ISS \frac{\partial\, ISS}{\partial x_j} \frac{\partial\, TSS}{\partial x_j} + 2\, TSS \left[\frac{\partial\, ISS}{\partial x_j} \right]^2 \right\}$$

where $\left[\dfrac{\partial (TSS)}{\partial x_j} \right]_{\text{all } x_j = 1} = - N_{\cdot j} w_j (y_{\cdot j \cdot} - \tilde{y}_{\cdots})^2,$

$$\left[\frac{\partial (ISS)}{\partial x_j} \right]_{\text{all } x_j = 1} = - \sum_i n_{ij} w_j y_{ij\cdot} (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot} + \tilde{y}_{\cdots}) +$$

$$\sum_i \sum_j n_{ij} w_j y_{ij\cdot} f_j (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot} + \tilde{y}_{\cdots}) \quad \ldots\ldots\ldots (11)$$

$$\left[\frac{\partial^2 (TSS)}{\partial x_j^2} \right]_{\text{all } x_j = 1} = 2\, N_{\cdot j} w_j (y_{\cdot j\cdot} - \tilde{y}_{\cdots})^2 (1 - f_j)$$

and $\left[\dfrac{\partial^2 (ISS)}{\partial x_j^2} \right]_{\text{all } x_j = 1} = 2(1 - f_j) \sum_i n_{ij} w_j y_{ij\cdot} (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot} + \tilde{y}_{\cdots})$

$$- 2 \sum_i \sum_j n_{ij} w_j y_{ij\cdot} f_j (y_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - y_{\cdot j\cdot} + \tilde{y}_{\cdots})(1 - f_j) \ldots (12)$$

When $\sigma_\delta = 0$, the adjusted F-statistic using estimated weights for testing H_3: $\sigma_\beta = 0$, is given, as in section 3.1.5, by

$$\hat{F}_7(adj) = \frac{(N..-bt)\,\{\text{Blocks SS(using estimated weights \& adj.)}\}}{(b-1)\,\{\Sigma\Sigma\Sigma\,\hat{w}_j(y_{ijk} - y_{ij.})^2\}}$$

with $(b-1)$ and $(N..-bt)$ d.f. and the blocks SS(using estimated weights & adj.) is given by the equation (9) of section 3.1.5.

As shown in the previous section, the hypothesis H_3: $\sigma_\beta = 0$ cannot be tested in the presence of interaction variance σ_δ^2.

(ii) Adjustment of the estimators of variance components

The estimator using estimated weights, of $\sigma_\delta^{\,2}$ is

$$\hat{\tilde{\sigma}}_\delta^{\,2} = \frac{b}{b-1} \cdot \frac{I\ SS - (b-1)(t-1)}{\hat{w}. - (\Sigma N_j^{\,2}\hat{w}_j^{\,2}/\hat{w}.)}$$

with $\hat{w}. = \Sigma N._j\hat{w}_j$ so that

$$\hat{\tilde{\sigma}}_\delta^{\,2}(adj) = \hat{\tilde{\sigma}}_\delta^2 - \sum_{j=1}^{t} \frac{1}{N._j - b} \left[\frac{\partial^2(\hat{\tilde{\sigma}}_\delta^{\,2})}{\partial x_j^{\,2}} \right]_{x_j=1} \text{ using estimated weights}$$

where

$$\frac{\partial^2(\hat{\tilde{\sigma}}_\delta^{\,2})}{\partial x_j^{\,2}} = \frac{b}{(b-1)A^3}\left[A^2\,\frac{\partial^2 ISS}{\partial x_j^{\,2}} - A\,\{ISS - (b-1)(t-1)\}\right.$$

$$\left.\frac{\partial^2 A}{\partial x_j^{\,2}} - 2A\,\frac{\partial A}{\partial x_j}\,\frac{\partial ISS}{\partial x_j} + 2\,\{ISS - (b-1)(t-1)\}\left(\frac{\partial A}{\partial x_j}\right)^2\right]$$

with $A = \hat{w}. - \Sigma N._j^{\,2}\hat{w}_j^{\,2}/\hat{w}.$,

$$\left[\frac{\partial A}{\partial x_j}\right]_{\text{all } x_j=1} = - N._j w_j \left(1 - \frac{2N._j w_j w. - \Sigma N._j{}^2 w_j{}^2}{w.^2}\right), \quad \ldots\ldots \quad (13)$$

$$\left[\frac{\partial^2 A}{\partial x_j{}^2}\right]_{\text{all } x_j=1} = \frac{f_j}{w.} \{ 2\, w.^2 - 6\, N._j w_j w. + 4\, N._j{}^2 w_j{}^2 + 2\, \Sigma N._j{}^2$$

$$w_j^2 (1-f_j)\} \quad \ldots\ldots\ldots\ldots\ldots \quad (14)$$

and $\left[\frac{\partial (ISS)}{\partial x_j}\right]_{\text{all } x_j=1}$ and $\left[\frac{\partial^2 ISS}{\partial x_j{}^2}\right]_{\text{all } x_j=1}$ are given by the equations (11) and (12) respectively.

In the same way, the adjusted estimator using estimated weights, of $\sigma_\tau^{\;2}$ is given by

$$\hat{\tilde{\sigma}}_\tau{}^2(adj) = \frac{TSS - ISS/(b-1)}{\hat{w}. - \Sigma N._j{}^2 \hat{w}_j{}^2/\hat{w}.} - \sum_1^t \frac{1}{N._j - b} \left[\frac{\partial^2 (\hat{\tilde{\sigma}}_\tau{}^2)}{\partial x_j{}^2}\right]_{\text{all } x_j=1}$$

using estimated weights

where

$$\frac{\partial^2 (\hat{\tilde{\sigma}}_\tau{}^2)}{\partial x_j{}^2} = \frac{1}{A^3} \left\{ A^2 \frac{\partial^2 B}{\partial x_j{}^2} - AB \frac{\partial^2 A}{\partial x_j{}^2} - 2A \frac{\partial A}{\partial x_j} \frac{\partial B}{\partial x_j} + 2B(\partial A/\partial x_j)^2 \right\},$$

$$B = TSS - ISS/(b-1),$$

$$\left[\frac{\partial B}{\partial x_j}\right]_{\text{all } x_j=1} = - N._j w_j (y._j. - \tilde{y}...)^2 + \frac{1}{b-1} \Bigg[\sum_i n_{ij} w_j y_{ij}.$$

$$(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots}) - \sum_i\sum_j n_{ij}w_jy_{ij\cdot}f_j(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots})\Bigg],$$

$$\left[\frac{\partial^2 B}{\partial x_j}\right]_{\text{all } x_j=1} = 2N_{\cdot j}w_j(y_{\cdot j\cdot}-\tilde{y}_{\cdots})^2(1-f_j) - \frac{2}{b-1}\Bigg[\sum_i n_{ij}w_jy_{ij\cdot}$$

$$(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots}) - 2\sum_i n_{ij}w_jy_{ij\cdot}f_j(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots})$$

$$- 2\sum_i\sum_j n_{ij}w_jy_{ij\cdot}f_j(1-f_j)(y_{ij\cdot}-\tilde{y}_{i\cdot\cdot}-y_{\cdot j\cdot}+\tilde{y}_{\cdots})\Bigg]$$

and $\left[\frac{\partial A}{\partial x_j}\right]_{\text{all } x_j=1}$ and $\left[\frac{\partial^2 A}{\partial x_j^2}\right]_{\text{all } x_j=1}$ are given by the equations (13) and (14) respectively.

Finally, the estimator using estimated weights, of σ_β^2, is

$$\hat{\tilde{\sigma}}_\beta^2 = b(BSS - b+1)/(b-1)\hat{w}. - \{ISS - (b-1)(t-1)\}/(\hat{w}.^2/\sum N_{\cdot j}^2\hat{w}_j^2 - 1)$$

so that

$$\hat{\tilde{\sigma}}_\beta^2(adj) = \hat{\tilde{\sigma}}_\beta^2 - \sum_j \frac{1}{N_{\cdot j}-b}\left[\frac{\partial^2\hat{\tilde{\sigma}}_\beta^2}{\partial x_j^2}\right]_{\text{all } x_j=1}$$ using estimated weights

where

$$\frac{\partial^2\hat{\tilde{\sigma}}_\beta^2}{\partial x_j^2} = \frac{b}{(b-1)\hat{w}.}\left\{\frac{\partial^2(BSS)}{\partial x_j^2} + 2\hat{f}_j\frac{\partial(BSS)}{\partial x_j} - 2\hat{f}_j(1-\hat{f}_j)(BSS-t+1)\right\}$$

$$-\left\{\frac{1}{C^3}\frac{\partial^2(ISS)}{\partial x_j^2} - \frac{1}{C^2}\frac{\partial^2 C}{\partial x_j^2}(ISS-\overline{b-1}\;\overline{t-1}) - \frac{2}{C^2}\right.$$

$$\left.\frac{\partial C}{\partial x_j}\frac{\partial ISS}{\partial x_j} + \frac{2}{C^3}\left(\frac{\partial C}{\partial x_j}\right)^2(ISS - \overline{b-1}\;\overline{t-1})\right\}$$

with $C = (\hat{w}.^2/\Sigma N._j^2\hat{w}_j^2 - 1)$, $\left[\frac{\partial C}{\partial x_j}\right]_{\text{all } x_j=1} = 2 N._j w_j w.\,(w.N._j w_j - \Sigma N._j^2 w_j^2)/(\Sigma N._j^2 w_j^2)^2$,

$$\left[\frac{\partial^2 C}{\partial x_j^2}\right]_{\text{all } x_j=1} = -2 N._j w_j \{N._j w_j C/\Sigma N._j^2 w_j^2 + 2w.\,(w.N._j w_j - \Sigma N._j^2 w_j^2)(1 - 2N._j^2 w_j^2/\Sigma N._j^2 w_j^2) / (\Sigma N._j^2 w_j^2)^2\} ,$$

$$\left[\frac{\partial (BSS)}{\partial x_j}\right]_{\text{all } x_j=1} = -\sum_i n_{ij} w_j (\tilde{y}_{i}.. - \tilde{y}...)^2 - 2\sum_i\sum_j n_{ij} w_j f_j (\tilde{y}_{i}.. - \tilde{y}...)(y_{ij}. - \tilde{y}_{i}.. - y._j. - \tilde{y}...)$$

$$\left[\frac{\partial^2 (BSS)}{\partial x_j^2}\right]_{\text{all } x_j=1} = 2\sum_i n_{ij} w_j (\tilde{y}_{i}.. - \tilde{y}...)^2 + 4\sum_i n_{ij} w_j f_j (\tilde{y}_{i}.. - \tilde{y}...)(y_{ij}. - \tilde{y}_{i}.. - y._j. + \tilde{y}...) + 2\sum_i\sum_j n_{ij} w_j f_j^2 (y_{ij}. - \tilde{y}_{i}.. - y._j. + \tilde{y}...)^2$$
$$+ 4\sum_i\sum_j n_{ij} w_j f_j (1-f_j)(\tilde{y}_{i}.. - \tilde{y}...)(y_{ij}. - \tilde{y}_{i}.. - y._j. + \tilde{y}...)$$

and $\left[\frac{\partial (ISS)}{\partial x_j}\right]_{\text{all } x_j=1}$ and $\left[\frac{\partial^2 (ISS)}{\partial x_j^2}\right]_{x_j=1}$ are given by (11) and (12) respectively.

3.3 Fixed effects models with equal replication

While the results of sections 3.1.2 to 3.1.6 are entirely applicable, some simpler tests are available in this case. These were first discussed by Robinson and Balaam (1967) for correlated and heteroscedastic errors.

The model is the same as that of (6) in section 3.1 under the usual constraints (II), with the exception that the quantities n_{ij} are now all equal to r. The proportionality condition is thus satisfied.

3.3.1 Test of significance of treatment effects

Taking the mean of the observations of the model with respect to the suffix i, we get under constraint (II),

$$y_{\cdot jk} = \beta_{\cdot} + \tau_j + \varepsilon_{\cdot jk} = \mu_j + \varepsilon_{\cdot jk}, \quad j=1,2,\ldots,t \;;$$

$k = 1,2,\ldots,r$; say, where $\text{var}\cdot(\varepsilon_{\cdot jk}) = \sigma_j^2/b$, which differs from treatment to treatment. Hence, this model is the same as that of the one-way model with unequal group variances. Thus the methods of estimation and analysis described in Chapter 2 may be used.

The methods are also applicable when the number of observations per cell is constant for each treatment but varies from treatment to treatment.

3.3.2 Test of significance of block effects

Taking the mean of the observations under the model at (6) with respect to the suffix j, we get,

$$y_{i.k} = \beta_i + \varepsilon_{i.k} \qquad i = i,2,\ldots,b \;;$$

$k = 1,2,\ldots,r$, where $\mathrm{var}\,(\varepsilon_{i.k}) = \sum_{1}^{t} \sigma_j^2/t^2$ which is a constant so that this model is a homoscedastic one-way one. The usual least squares analysis can be used for testing the significance of block effects. The procedure holds good even if the number of observations is constant within the cells of each block but varies from block to block.

3.3.3 Likelihood ratio tests for significance of interactions and treatment effects

Let $\underset{\sim}{Y}_{ik}$ be the column vector of observations at the kth realisation within the ith block, i.e., $\underset{\sim}{Y}_{ik} = (y_{i1k},\ldots, y_{itk})'$; $i = 1,2,\ldots,b$; $k = 1,2,\ldots,r$. Let $\underset{\sim}{L}$ be a $(t-1) \times t$ matrix such that

$$\underset{\sim}{L}\,\underset{\sim}{1} = \underset{\sim}{0} \text{ and } \underset{\sim}{L}\underset{\sim}{L}' = \underset{\sim}{I}_{t-1} .$$

Then the elements of the vector $\underset{\sim}{Z}_{ik} = \underset{\sim}{L}\,\underset{\sim}{Y}_{ik}$ are $(t-1)$ orthogonal contrasts amongst the kth set of observations within the ith block. The matrix $\underset{\sim}{L}$ will be called the matrix of orthogonal contrasts.

Then the model at (6) of section 3.1 can be written, in vector notation, as

$$\underset{\sim}{Z}_{ik} = \underset{\sim}{\tau} + \underset{\sim}{\delta}_i + \underset{\sim}{e}_{ik} \;;\quad i = 1,2,\ldots,b \;;\quad k = 1,2,\ldots,r;$$

where $\underset{\sim}{\tau} = \underset{\sim}{L}(\tau_1,\ldots,\tau_t)'$, $\underset{\sim}{\delta}_i = \underset{\sim}{L}(\delta_{i1},\ldots,\delta_{it})'$ and $\underset{\sim}{e}_{ik} = \underset{\sim}{L}\,(\varepsilon_{i1k},\ldots,\varepsilon_{itk})'$. It then follows that $\underset{\sim}{e}_{ik}$ is distributed as multivariate normal with mean vector $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{\Sigma}$ where

$$\underset{\sim}{\Sigma} = \underset{\sim}{L}\,\mathrm{diag}\,(\sigma_1^2,\ldots,\sigma_t^2)\,\underset{\sim}{L}', \text{ which is non-diagonal.}$$

We can now use the likelihood ratio (LR) tests of the multivariate analysis of variance for testing the hypotheses (i) $\underset{\sim}{\tau} = \underset{\sim}{0}$ and (ii) $\underset{\sim}{\delta}_i = \underset{\sim}{0}$ for all i.

Robinson and Balaam (1967) considered independent contrasts of treatment observations instead of orthogonal ones as used here. One advantage of using the orthogonal contrasts is that the LR test-statistics are invariant under such transformation of data.

The LR test-statistics given by them are as follows.

(i) H_T: $\underset{\sim}{\tau} = 0$ i.e., $\tau_i = 0$ for all i.

LR test criterion for testing this hypothesis is

$$\lambda^{2/br} = \frac{|\underset{\sim}{A}|}{|\underset{\sim}{A} + br\, \underset{\sim}{Z}_{..}\underset{\sim}{Z}'_{..}|} = (1 + \frac{r}{r-1}\, \underset{\sim}{Z}_{..}\, \underset{\sim}{S}^{-1}\underset{\sim}{Z}'_{..})^{-1}$$

where $\underset{\sim}{A} = \sum_{i=1}^{b} \sum_{k=1}^{r} (\underset{\sim}{Z}_{ik} - \underset{\sim}{Z}_{i\cdot})(\underset{\sim}{Z}_{ik} - \underset{\sim}{Z}_{i\cdot})' = b(r-1)\,\underset{\sim}{S}$ and

$$\underset{\sim}{Z}_{i\cdot} = \sum_{k=1}^{r} \underset{\sim}{Z}_{ik}/r\ .$$

Since $(br\, \underset{\sim}{Z}_{..}\, \underset{\sim}{S}^{-1}\, \underset{\sim}{Z}'_{..})$ is Hotelling's T^2, this is an exact test, i.e.,

$$\lambda^{2/br} = (1 + \frac{t-1}{br-b-t+2}\, F_{t-1,br-b-t+2})^{-1}$$

under the hypothesis H_T.

(ii) H_{BT}: $\underset{\sim}{\delta}_i = \underset{\sim}{0}$ for all i = 1,2,...b i.e., $\delta_{ij} = 0$ for all i and j. The LR criterion for this test is given in the notation of Anderson (1958 , p. 208), by

$$U_{t-1,b-1,b(r-1)} = \frac{|\underset{\sim}{A}|}{|\underset{\sim}{A} + \underset{\sim}{B}|}$$

where $\underset{\sim}{B} = r \sum_{i=1}^{b} (\underset{\sim}{Z}_{i\cdot} - \underset{\sim}{Z}_{..})(\underset{\sim}{Z}_{i\cdot} - \underset{\sim}{Z}_{..})'$.

Now $-\{ br - 1 - \frac{1}{2} (b + t - 1)\} \log_e U_{t-1,b-1,b(r-1)}$ is distributed asymptotically as χ^2 with $(b-1)(t-1)$ d.f. For small sample, further approximations may be used.

To show the invariance of the LR test-statistics, let $\underset{\sim}{M}$ be another matrix of orthogonal contrasts of treatment observations.

Then $\underset{\sim}{M}$ is given by

$$\underset{\sim}{M} = \underset{\sim}{C}\ \underset{\sim}{L}$$

where $\underset{\sim}{C}$ is an orthogonal matrix. This was stated by Shukla (1972) without proof which may be as follows.

Since $\underset{\sim}{M}'$ is a $t\ x(t-1)$ matrix of rank $(t-1)$, there exists a non-singular $t \times t$ matrix $\underset{\sim}{C}_o$, and an orthogonal $(t-1) \times (t-1)$ matrix $\underset{\sim}{R}$ such that

$$\underset{\sim}{M}' = \underset{\sim}{C}_o \begin{pmatrix} \underset{\sim}{I}_{t-1} \\ \cdots \\ \underset{\sim}{0} \end{pmatrix} \underset{\sim}{R} , \qquad \text{(see Rao, 1973, p.20).}$$

or,

$$\underset{\sim}{M} = \underset{\sim}{R}' (\underset{\sim}{I}_{t-1} \vdots \underset{\sim}{0})\ \underset{\sim}{C}_o' = \underset{\sim}{E}\ \underset{\sim}{C}_1$$

say, where $\underset{\sim}{E} = \underset{\sim}{R}'$ is orthogonal and $\underset{\sim}{C}_1 = (\underset{\sim}{I}_{t-1} \vdots \underset{\sim}{o})\ \underset{\sim}{C}_o'$ is a $(t-1) \times t$ matrix of rank $(t-1)$.

Now by definition, $\underset{\sim}{0} = \underset{\sim}{M}\ \underset{\sim}{1}' = \underset{\sim}{E}\ \underset{\sim}{C}_1 \underset{\sim}{1}$ which implies that $\underset{\sim}{C}_1 \underset{\sim}{1} = \underset{\sim}{0}$, and $\underset{\sim}{I} = \underset{\sim}{M}\ \underset{\sim}{M}' = \underset{\sim}{E}\ \underset{\sim}{C}_1 \underset{\sim}{C}_1'\ \underset{\sim}{E}'$ which implies that $\underset{\sim}{C}_1 \underset{\sim}{C}_1' = \underset{\sim}{I}$. Thus $\underset{\sim}{C}_1$ is again a matrix of orthogonal contrasts. Applying this result once again we find that

$$\underset{\sim}{C}_1 = \underset{\sim}{F}\ \underset{\sim}{T}$$

where $\underset{\sim}{F}$ is orthogonal and $\underset{\sim}{T}$ is another matrix of orthogonal contrasts.

It then follows that by a suitable choice of the orthogonal matrix $\underset{\sim}{C}$, the matrix $\underset{\sim}{M}$ can be written as

$$\underset{\sim}{M} = \underset{\sim}{C}\ \underset{\sim}{L}.$$

Now let λ_1 be the value of λ when $\underset{\sim}{M}$ is used in place of $\underset{\sim}{L}$. Then the vector of newly transformed observations is given by

$$\underset{\sim}{X}_{ik} = \underset{\sim}{M}\,\underset{\sim}{Y}_{ik} = \underset{\sim}{C}\,\underset{\sim}{L}\,\underset{\sim}{Y}_{ik} = \underset{\sim}{C}\,\underset{\sim}{Z}_{ik}$$

so that $$\lambda_1^{2/br} = \frac{|\underset{\sim}{C}\,\underset{\sim}{A}\,\underset{\sim}{C}'|}{|\underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{C}' + br\,\underset{\sim}{C}\,\underset{\sim}{Z}_{..}\underset{\sim}{Z}'_{..}\underset{\sim}{C}'|} = \frac{|\underset{\sim}{A}|}{|\underset{\sim}{A}+br\,\underset{\sim}{Z}_{..}\underset{\sim}{Z}'_{..}|} = \lambda^{2/br}.$$

Similarly, the expression of the other LR criterion, $U_{t-1,b-1,b(r-1)}$, also remains unchanged.

The above method is easily generalised to multi-way factorial designs with equal numbers of observations per cell and with unequal group variances.

CHAPTER 4

GENERAL BLOCK DESIGNS

An additive fixed-effects model with unequal group variances is considered here for general block designs including both extended and incomplete block designs. Estimators of the linear parameters are obtained on the assumption that the group variances are known, and the corresponding analysis is provided. Canonical forms of two sums of squares are derived. When the group variances are not known, adjustment of estimators and test-statistics using estimated weights is suggested for removing bias. Finally, recovery of inter-block information is discussed.

4.1 Estimation and intrablock analysis when group variances are known

Let the additive fixed-effects model be:

$$\left.\begin{array}{c} y_{ijk} = \beta_i + \tau_j + \varepsilon_{ijk} \\ i = 1,2,\ldots,b \quad ; \quad j = 1,2,\ldots,t \quad ; \quad k = 1,2,\ldots,n_{ji} \geq 0; \end{array}\right] \quad \ldots(15)$$

where β_i is the effect due to the ith block, τ_j the effect due to the jth treatment and ε_{ijk} the error term having mean zero and variance, σ_j^2. The errors are assumed to be independent of one another as before. Both incomplete and extended block designs are included in this model. Block sizes are unequal in general.

Let $\underset{\sim}{Y}$ be the vector of observations arranged treatment by treatment; then the model can be written as

$$\underset{\sim}{Y} = \underset{\sim}{\Delta}' \underset{\sim}{\tau} + \underset{\sim}{D}' \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{\tau}$ and $\underset{\sim}{\beta}$ are the vectors of treatment and block effects respectively with the corresponding design matrices, $\underset{\sim}{\Delta}'$ and $\underset{\sim}{D}'$, and $\underset{\sim}{\varepsilon}$ is the vector of errors. $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and $\mathrm{var}(\underset{\sim}{\varepsilon}) = \mathrm{diag}(\sigma_1^2, .., \sigma_1^2, ..., \sigma_t^2, ..., \sigma_t^2) = \underset{\sim}{V}$, say.. The rank of the overall design matrix is $(b + t - 1)$.

Further notations:

Let $\underset{\sim}{r} = (r_1, .., r_t)'$, the vector of replications of the treatments,

$\underset{\sim}{k} = (k_1, ..., k_b)'$, the vector of block sizes,

$\underset{\sim}{n} = \underset{\sim}{\Delta}\,\underset{\sim}{D}' = (n_{ji})$, the incidence matrix of treatments with the blocks,

$\underset{\sim}{w} = (w_1, ..., w_t)'$, the vector of weights with $w_j = (1/\sigma_j^2)$,

$\underset{\sim}{T} = \underset{\sim}{\Delta}\underset{\sim}{Y}$, the vector of treatment totals,

$\underset{\sim}{\tilde{B}} = \underset{\sim}{D}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{Y}$, the vector of weighted block totals with elements,
$$\tilde{B}_i = \sum_j n_{ji} w_j y_{ij.} \quad ,$$

$\tilde{G} = \underset{\sim}{w}'\,\underset{\sim}{T} = \underset{\sim}{1}'\underset{\sim}{\tilde{B}}$, the weighted total of all observations and $N = \sum_i \sum_j n_{ji}$.

Then, $\underset{\sim}{D}'\underset{\sim}{1} = \underset{\sim}{1} = \underset{\sim}{\Delta}'\underset{\sim}{1}$, $\underset{\sim}{D}\,\underset{\sim}{1} = \underset{\sim}{k} = \underset{\sim}{n}'\,\underset{\sim}{1}$, $\underset{\sim}{\Delta}\underset{\sim}{1} = \underset{\sim}{r} = \underset{\sim}{n}\,\underset{\sim}{1}$, $\underset{\sim}{k}'\,\underset{\sim}{1} = N = \underset{\sim}{r}'\,\underset{\sim}{1}$, $\underset{\sim}{D}\,\underset{\sim}{D}' = \underset{\sim}{k}^{\delta}$ and $\underset{\sim}{\Delta}\underset{\sim}{\Delta}' = \underset{\sim}{r}^{\delta}$ where the superscript δ denotes a diagonal matrix with elements of the vector as the diagonal elements. The superscript, $-\delta$, will denote the inverse of such a diagonal matrix. Also $\sum_i n_{ji} = r_j$ and $\sum_j n_{ji} = k_i$.

By (2) of section 1.2, the normal equations for finding the weighted least squares estimators of the linear parameters are given by

$$\left(\frac{\underset{\sim}{\Delta}}{\underset{\sim}{D}}\right) \underset{\sim}{V}^{-1} \left(\underset{\sim}{\Delta}' \,\vdots\, \underset{\sim}{D}'\right) \left(\frac{\tilde{\underset{\sim}{\tau}}}{\tilde{\underset{\sim}{\beta}}}\right) = \left(\frac{\underset{\sim}{\Delta}}{\underset{\sim}{D}}\right) \underset{\sim}{V}^{-1} \underset{\sim}{Y} \ .$$

or,

$$\begin{bmatrix} \underset{\sim}{\Delta} \underset{\sim}{V}^{-1} \underset{\sim}{\Delta}' & \vdots & \underset{\sim}{\Delta} \underset{\sim}{V}^{-1} \underset{\sim}{D}' \\ \cdots\cdots & \vdots & \cdots\cdots \\ \underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{\Delta}' & \vdots & \underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{D}' \end{bmatrix} \begin{bmatrix} \tilde{\underset{\sim}{\tau}} \\ \cdots \\ \tilde{\underset{\sim}{\beta}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{\Delta} \underset{\sim}{V}^{-1} \underset{\sim}{Y} \\ \cdots\cdots \\ \underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{Y} \end{bmatrix}.$$

Now, $\underset{\sim}{\Delta} \underset{\sim}{V}^{-1} \underset{\sim}{\Delta}' = \underset{\sim}{w}^{\delta} \, r^{\delta}$, $\underset{\sim}{\Delta} \underset{\sim}{V}^{-1} \underset{\sim}{D}' = \underset{\sim}{w}^{\delta} \underset{\sim}{n}$, $\underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{\Delta}' = \underset{\sim}{n}' \underset{\sim}{w}^{\delta}$
$\underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{D}' = (\underset{\sim}{n}'\underset{\sim}{w})^{\delta}$, $\underset{\sim}{\Delta} \underset{\sim}{V}^{-1}\underset{\sim}{Y} = \underset{\sim}{w}^{\delta} \underset{\sim}{T}$ and $\underset{\sim}{D} \underset{\sim}{V}^{-1} \underset{\sim}{Y} = \underset{\sim}{\tilde{B}}$.

The two sets of normal equations then become

$$\underset{\sim}{w}^{\delta} \underset{\sim}{r}^{\delta} \tilde{\underset{\sim}{\tau}} + \underset{\sim}{w}^{\delta} \underset{\sim}{n} \tilde{\underset{\sim}{\beta}} = \underset{\sim}{w}^{\delta} \underset{\sim}{T}$$

and

$$\underset{\sim}{n}' \underset{\sim}{w}^{\delta} \tilde{\underset{\sim}{\tau}} + (\underset{\sim}{n}'\underset{\sim}{w})^{\delta} \tilde{\underset{\sim}{\beta}} = \underset{\sim}{\tilde{B}}.$$

Eliminating $\tilde{\underset{\sim}{\beta}}$ from the first set of equations, we get the reduced normal equations for the treatments as

$$\{ \underset{\sim}{r}^{\delta} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta} (\underset{\sim}{n}'\underset{\sim}{w}^{\delta}) \} \tilde{\underset{\sim}{\tau}} = \tilde{\underset{\sim}{Q}},$$

with $\tilde{\underset{\sim}{Q}} = \underset{\sim}{T} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta} \tilde{\underset{\sim}{B}}$ as the vector of adjusted treatment totals.

Since $\{ \underset{\sim}{r}^{\delta} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}(\underset{\sim}{n}'\underset{\sim}{w}^{\delta}) \} \underset{\sim}{1} = \underset{\sim}{r}^{\delta}\underset{\sim}{1} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\underset{\sim}{n}'\underset{\sim}{w}=0$, a unique solution for the treatment estimates is not possible.

Following Tocher (1952), the singular coefficient matrix may be replaced by a non-singular one in the following way.

We have, $\tilde{G} = \underset{\sim}{w}'\underset{\sim}{T} = \underset{\sim}{w}'(\underset{\sim}{r}^{\delta}\tilde{\underset{\sim}{\tau}} + \underset{\sim}{n} \tilde{\underset{\sim}{\beta}}) = \underset{\sim}{w}' \underset{\sim}{r}^{\delta} \tilde{\underset{\sim}{\tau}}$ assuming the constraint $(\underset{\sim}{n}'\underset{\sim}{w})' \tilde{\underset{\sim}{\beta}} = 0$.

Then, $\tilde{\underset{\sim}{Q}} + \underset{\sim}{r}(\tilde{G}/\underset{\sim}{w}'\underset{\sim}{r}) = \{ \underset{\sim}{r}^{\delta} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta} (\underset{\sim}{n}'\underset{\sim}{w}^{\delta}) + \underset{\sim}{r}\underset{\sim}{w}'\underset{\sim}{r}^{\delta}$
$(1/\underset{\sim}{w}'\underset{\sim}{r})\} \tilde{\underset{\sim}{\tau}}$

$= \underset{\sim}{\Omega}^{-1} \tilde{\underset{\sim}{\tau}}$

say, with $\underset{\sim}{\Omega}^{-1} = \{ \underset{\sim}{r}^{\delta} - \underset{\sim}{n}(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\underset{\sim}{n}'\underset{\sim}{w}^{\delta} + \underset{\sim}{r}\underset{\sim}{w}'\underset{\sim}{r}^{\delta} \, (1/\underset{\sim}{w}'\underset{\sim}{r})\}$.
It follows that $\underset{\sim}{\Omega}^{-1}\underset{\sim}{1} = \underset{\sim}{r}$ so that $\underset{\sim}{\Omega}\, \underset{\sim}{r} = \underset{\sim}{1}$.

Thus the treatment estimators are obtained as

$$\tilde{\tau} = \underset{\sim}{\tilde{\Omega}}\{\tilde{Q} + \underset{\sim}{r}\,(\tilde{G}/\underset{\sim}{w}'\underset{\sim}{r})\} = \underset{\sim}{\tilde{\Omega}}\,\tilde{Q} + \underset{\sim}{1}\,(\tilde{G}/\underset{\sim}{w}'\underset{\sim}{r})\,.$$

It follows from the second set of normal equations that the sum of squares due to all estimates is

$$SS(Est.) = \tilde{\tau}\,'\,\underset{\sim}{w}^{\delta}\,\underset{\sim}{T} + \tilde{\beta}\,'\,\tilde{B} = \tilde{\tau}\,'\,\underset{\sim}{w}^{\delta}\,\tilde{Q} + \tilde{B}'(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\tilde{B}$$

with (b+t-1) d.f.

Ignoring the treatment effects, the model reduces to

$$\underset{\sim}{Y} = \underset{\sim}{D}'\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

The weighted least squares estimator of $\underset{\sim}{\beta}$ is now given by

$$\tilde{\tilde{\beta}} = (\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\tilde{B}$$

and the SS due to blocks (uncorrected) ignoring treatment effects by $\tilde{\tilde{\beta}}\,'\,\tilde{B} = \tilde{B}'(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\tilde{B}$ with b d.f.

Similarly, the SS due to treatment (uncorrected) ignoring block effects is given by $\underset{\sim}{T}'\,\underset{\sim}{r}^{-\delta}\,\underset{\sim}{w}^{\delta}\,\underset{\sim}{T}$ with t d.f.

As $\underset{\sim}{1}'\,\underset{\sim}{w}^{\delta}\tilde{Q} = \tilde{Q}'\underset{\sim}{w}^{\delta}\,\underset{\sim}{1} = 0$, the adjusted treatment sum of squares is

$$\text{Adjusted SS (treat.)} = \tilde{\tau}\,'\underset{\sim}{w}^{\delta}\tilde{Q} = \{\underset{\sim}{\tilde{\Omega}}\,\tilde{Q} + \underset{\sim}{1}(\tilde{G}/\underset{\sim}{w}'\underset{\sim}{r})\}'\underset{\sim}{w}^{\delta}\,\tilde{Q}$$

$$= \tilde{Q}'\,\tilde{\Omega}'\,\underset{\sim}{w}^{\delta}\,\tilde{Q}$$

with (t-1) d.f., and the SS due to error is

$$SS(E) = \underset{\sim}{Y}'\,\underset{\sim}{V}^{-1}\underset{\sim}{Y} - \tilde{B}'(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\,\tilde{B} - \tilde{Q}'\,\tilde{\Omega}'\,\underset{\sim}{w}^{\delta}\,\tilde{Q}$$

with (N-b-t+1) d.f. The above results reduce to those of Tocher (1952) when $\underset{\sim}{w} = \underset{\sim}{1}$ for homoscedastic models.

The analysis of variance table is given below.

Analysis of variance table

Source	d.f.	SS	SS	d.f.	Source
Block & general mean (unadj)	b	$S_1 = \underset{\sim}{\tilde{B}}'(\underset{\sim}{n}'\underset{\sim}{w})^{-\delta}\underset{\sim}{\tilde{B}}$	$S_4 = \underset{\sim}{T}'\underset{\sim}{r}^{-\delta}\underset{\sim}{w}^{\delta}\underset{\sim}{T}$	t	Treatments and gen.mean(unadj)
Treatments (adj)	t-1	$S_2 = \underset{\sim}{\tilde{Q}}'\underset{\sim}{\tilde{\Omega}}'\underset{\sim}{W}^{\delta}\underset{\sim}{\tilde{Q}}$	$S_5 = S_1 + S_2 - S_4$	b-1	Block (adj)
Error	(N-b-t+1)	$S_3 = \underset{\sim}{Y}'\underset{\sim}{V}^{-1}\underset{\sim}{Y} - S_1 - S_2$	S_3	N-b-t+1	Error

4.2 A special case

Let us consider an experiment where the number of blocks is equal to the number of treatments and where the ith block contains $r(>1)$ plots for the ith treatment and only one plot for each of the other $(t-1)$ treatments ; $i = 1,2,\ldots,t$. Then

$$\underset{\sim}{n} = \begin{bmatrix} r & 1 & . & . & 1 \\ 1 & r & . & . & 1 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & 1 & . & . & r \end{bmatrix} = \underset{\sim}{n}', \qquad \underset{\sim}{n}'\underset{\sim}{w}^{\delta} = \begin{bmatrix} r_1w_1 & w_2 & . & . & w_t \\ w_1 & r_2w_2 & . & . & w_t \\ . & . & . & . & . \\ . & . & . & . & . \\ w_1 & w_2 & . & . & r_tw_t \end{bmatrix}$$

$\underset{\sim}{r} = (r+t-1)\ \underset{\sim}{1} = \underset{\sim}{k}$ and $\underset{\sim}{n}'\underset{\sim}{w} = \{(r-1)w_1+w.,(r-1)w_2+w.,\ldots,(r-1)w_t+w.\}'$.

Consequently, if $\underset{\sim}{\Omega}^{-1} = (a_{ij})$, the elements a_{ij} are given by

$$a_{ii} = (r+t-1)\left(1+\frac{w_i}{w.}\right) - w_i\left[\frac{r^2-1}{(r-1)w_i+w.} + \sum_{j=1}^{t}\frac{1}{(r-1)w_j+w.}\right],$$

$$a_{ij} = w_j\left[\frac{r+t-1}{w.} + \frac{r-1}{(r-1)w_i+w.} + \frac{r-1}{(r-1)w_j+w.} + \sum_{k=1}^{t}\frac{1}{(r-1)w_k+w.}\right]$$

and

$$a_{ji} = w_i\ a_{ij}/w_j$$

$$(i \neq j = 1,2,\ldots,t)$$

Moreover,

$$T_j = ry_{jj.} + \sum_{i\neq j}^{t} y_{ij}$$

and

$$\tilde{B}_i = rw_iy_{ii.} + \sum_{j\neq i}^{t} w_j\ y_{ij}$$

so that the adjusted treatment total for the jth treatment is

$$\tilde{q}_j = T_j - (r-1)\,\tilde{B}_j/\{(r-1)w_j + w.\} - \sum_{k=1}^{t} \tilde{B}_k/\{(r-1)$$

$$w_k + w.\}\ , \quad j = 1,\ldots,t.$$

Here $w. = \Sigma\, w_j$.

4.3 Canonical forms of the sums of squares

The adjusted SS (treat.) and the SS due to error can be expressed in the same sorts of canonical form as used by Pearce and Jeffers (1971) for homoscedastic models.

We have, $\tilde{Q} = \Delta Y - \Delta D'(n'w)^{-\delta} D V^{-1} Y$

$$= \Delta (I - D'(n'w)^{-\delta} D V^{-1})\, Y$$

$$= \Delta\, \tilde{\phi}\, Y$$

say, with $\tilde{\phi} = I - D'(n'w)^{-\delta} D V^{-1}$.

Since $D V^{-1} 1 = n'w$, it follows that $\tilde{\phi}\ 1 = 0 = \tilde{\phi} D'$ and $\tilde{\phi}\tilde{\phi} = \tilde{\phi}$. Thus $\tilde{\phi}$ is idempotent but not symmetric.

Then, SS (treat.) adj. $= \tilde{Q}'\, \tilde{\Omega}' w^{\delta}\, \tilde{Q}$

$$= Y'\, \tilde{\phi}' \Delta' \tilde{\Omega}'\, w^{\delta} \Delta\, \tilde{\phi}\, Y\ ,$$

$$\tilde{\tau} = \{\tilde{\Omega}\, \Delta\, \tilde{\phi} + 1\ w'\ \Delta\ (1/w'r)\}\, Y$$

and var $(\tilde{\tau})$ $= \{\tilde{\Omega}\, \Delta\, \tilde{\phi} V + 1\ w'\ \Delta\ V\ (1/w'r)\}\{\tilde{\phi}'\Delta'\tilde{\Omega}'$

$$+ \Delta' w\ 1'(1/w'r)\}\ .$$

Now let

$$\tilde{\psi} = V^{-1}\tilde{\phi} - \tilde{\phi}'\ \Delta'\ \tilde{\Omega}'\ w^{\delta}\ \Delta\, \tilde{\phi}$$

It follows that $\tilde{\psi}\, \tilde{\phi} = \tilde{\psi}$ and $\tilde{\psi}\ 1 = 0 = \tilde{\psi}\, \tilde{D}'$

so that

$$Y'\, \tilde{\Psi}\, Y = Y'\, V^{-1} \tilde{\phi} Y - Y' \tilde{\phi}'\ \Delta'\ \tilde{\Omega}'\ w^{\delta} \Delta\, \tilde{\phi}\ Y$$

$$= Y'\ V^{-1}\ Y - \tilde{B}'(n'w)^{-\delta}\tilde{B} - \text{SS(treat.) adj}$$

$$= \text{SS (E)}$$

and $Y'(V^{-1}\tilde{\phi} - \tilde{\psi})\, Y = Y'\ \tilde{\phi}'\ \Delta'\ \tilde{\Omega}'\ w^{\delta} \Delta\, \tilde{\phi}\ Y$

$$= \text{Adjusted SS (treatments)}\ .$$

Thus the adjusted treatments SS and the SS for error can be expressed in terms of matrices, $\tilde{\phi}$ and $\tilde{\psi}$, which reduce respectively to the matrices ϕ and ψ defined by Pearce and Jeffers (1971) when $w = 1$ for homoscedastic models.

4.4 Estimation and analysis when group variances are unknown

The quantity n_{ji} denotes the number of times the jth treatment occurs in block i. For each j, we assume that there is at least one value of i for which $n_{ji} > 3$. If each treatment is replicated in exactly one block, and each block has only one treatment occurring more than once, then $b = t$. Otherwise, b may be greater or less than t. This includes extended block designs and also the designs where some or all blocks may not contain all the treatments. Block sizes are unequal in general as before.

$$\text{Since } E\{\sum_i \sum_k (y_{ijk} - y_{ij\cdot})^2\} = E\{\sum_i \sum_k (\varepsilon_{ijk} - \varepsilon_{ij\cdot})^2\}$$

$$= (r_j - b)\,\sigma_j^2,$$

$$s_j^2 = \sum_i \sum_k (y_{ijk} - y_{ij\cdot})^2/(r_j - b) \text{ is an unbiased}$$

estimator of σ_j^2, $j = 1,2,\ldots,t$. For $j \neq j'$, s_j^2 and $s_{j'}^2$ are independent. When $n_{ji} = 0$ or 1 in a cell, the contribution to the SS for s_j^2 and to its d.f. from this cell will be zero. Bartlett's χ^2-test can be used to test the homogeneity of error variances.

For any experiment under the model (15), the estimators of the linear parameters and other statistics may be calculated with the help of the formulae given above using estimated weights $\hat{w}_j = 1/s_j^2$ in place of the actual weights. Such estimators and other statistics including test-statistics using estimated weights can then be adjusted for bias by

Theorem 1 (section 2.1.4).

4.5 Recovery of inter-block information

Patterson and Thompson (1971) provided a method of modified maximum likelihood for recovery of inter-block information for incomplete block designs when block sizes are unequal. The same method may be used for mixed heteroscedastic models with random block effects as stated below.

The model is the same as that in (15) with the exception that the block effects, β_i, are now random variables with mean, β, and variances, $\sigma_\beta^{\ 2}$. Then the variance of the observation vector $\underset{\sim}{Y}$ is given by

$$\mathrm{var}(\underset{\sim}{Y}) = \mathrm{diag}\,(\sigma_1^{\ 2}+\sigma_\beta^{\ 2},\ldots,\sigma_1^{\ 2}+\sigma_\beta^{\ 2},\ldots,\sigma_t^{\ 2}+\sigma_\beta^{\ 2},\ldots,\sigma_t^{\ 2}+\sigma_\beta^{\ 2})$$

$= \underset{\sim}{H}$, say. As

$$E\left\{\sum_i\sum_k (y_{ijk} - y_{ij\cdot})^2\right\} = (r_j - b)\,\sigma_j^{\ 2} \quad ; \text{ it}$$

follows that
$$s_j^{\ 2} = \sum_i\sum_k (y_{ijk} - y_{ij\cdot})^2/(r_j-b)$$

remains an unbiased estimator of $\sigma_j^{\ 2}$, $j = 1,2,\ldots,t$; and $s_j^{\ 2}$ and $s_{j'}^{\ 2}$ are independent when $j \neq j'$.

The estimator of $\sigma_\beta^{\ 2}$ is obtained from the logarithm of the likelihood function of $\underset{\sim}{S}\,\underset{\sim}{Y}$ where $\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{\Delta}'(\underset{\sim}{\Delta}\,\underset{\sim}{\Delta}')^{-1}\underset{\sim}{\Delta}$, which is given by

$$L = \text{const.} - \tfrac{1}{2}\sum_s \log \xi_s - \tfrac{1}{2}\,\underset{\sim}{Y}'(\underset{\sim}{S}\underset{\sim}{H}\,\underset{\sim}{S})^{-}\underset{\sim}{Y}\ .$$

Here the quantities ξ_s are the non-zero latent roots of $\underset{\sim}{H}\,\underset{\sim}{S}$
and $\underset{\sim}{A}^{-g}$ denotes a generalised inverse of $\underset{\sim}{A}$ as defined by the authors.

The modified maximum likelihood estimator of $\sigma_\beta^{\ 2}$ is then obtained by solving the equation

$$\frac{\partial L}{\partial \sigma_\beta^2} = -\tfrac{1}{2}\,\underset{\sim}{E} + \tfrac{1}{2}\,\underset{\sim}{B} = 0$$

where $\underset{\sim}{B} = \underset{\sim}{Y}'\ (\underset{\sim}{SHS})^{-g}\ \ (\underset{\sim}{SHS})^{-g}\underset{\sim}{Y}$ and $\underset{\sim}{E} = \mathrm{tr}\{\ (\underset{\sim}{SHS})^{-g}\ \}$

The solution, $\tilde{\sigma}_\beta^{\,2}$, will be in terms of $\sigma_j^{\,2}$. The estimator, $\hat{\tilde{\sigma}}_\beta^{\,2}$, using the estimated weights can be adjusted by using Theorem 1.

Finally, the treatment estimators using the interblock information are obtained by solving the weighted least squares equations:

$$\underset{\sim}{\tilde{\tau}} = (\ \underset{\sim}{\Delta}\ \underset{\sim}{\hat{H}}^{-1}\ \underset{\sim}{\Delta}'\ )^{-1}\ \underset{\sim}{\Delta}\ \underset{\sim}{\hat{H}}^{-1}\ \underset{\sim}{Y}$$

where $\underset{\sim}{\hat{H}}$ is $\underset{\sim}{H}$ when $\sigma_j^{\,2}$ and $\sigma_\beta^{\,2}$ are replaced by their corresponding estimators.

CHAPTER 5

RANDOMISED COMPLETE BLOCK DESIGNS

For known group variances, the weighted least squares estimators of the linear parameters and the corresponding analysis are given. The MINQUE and almost unbiased estimators (AUE) of the error variances are derived. A theorem on the expectation of functions of correlated χ^2-variates is proved. The covariance between any two of the AUE's is found to be negligible. The test-statistics using estimated weights are adjusted for removing bias. Finally, expressions for joint confidence intervals of contrasts of both the treatment and block effects are provided.

5.1 Estimation and anlysis when the error variances are known

Let the linear model be

$$y_{ij} = \beta_i + \tau_j + \varepsilon_{ij}$$

$$(i = 1,2,\ldots,b \ ; \quad j = 1,2,\ldots,t)$$

where β_i is the effect due to the ith block, τ_j the effect due to the jth treatment and ε_{ij} the error term having mean zero and variance σ_j^2. The errors are assumed to be independent of one another. This is a special case of the model (6) in section 3.1.1 with the restrictions that $n_{ij} = 1$ for all i and j and that the interaction term is now the error term.

The weighted least squares estimators (WLS) of the linear parameters and the sums of squares can therefore be obtained from the corresponding expressions of section 3.1.2 and are given below.

The WLS estimators $\hat{\tau}_j = y_{\cdot j}$ and $\tilde{\beta}_i = \sum_j w_j y_{ij} / \sum_j w_j = \tilde{y}_{i\cdot}$ are unbiased for the parameters τ_j and $(\beta_i + \Sigma w_j \tau_j / w_j)$ respectively. Thus $\tilde{\beta}_i$ is biased for β_i unless $\Sigma w_j \tau_j = 0$ in the population although any contrast $\Sigma c_i \tilde{\beta}_i$ is unbiased for the corresponding parametric contrast $\Sigma c_i \beta_i$.

Furthermore, var $(\hat{\tau}_j) = \sigma_j^2/b = 1/bw_j$ and var $(\tilde{\beta}_i) = (1/\Sigma w_j)$ which is a constant.

The three (corrected) sums of squares (SS) for the analysis of variance are

$$\text{SS (treatments)} = b\,\Sigma\, w_j (y_{\cdot j} - \tilde{y}_{\cdot\cdot})^2$$

$$\text{SS (blocks)} = w_{\cdot}\, \Sigma\, (\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot})^2$$

and

$$\text{SS (error)} = \Sigma\Sigma\, w_j (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})^2$$

with d.f. (t-1),(b-1) and (b-1)(t-1) respectively, where $\tilde{y}_{\cdot\cdot} = \Sigma w_j y_{\cdot j} / w_{\cdot}$ and $w_{\cdot} = \Sigma w_j$.

Analysis of variance table

Source of variation	d.f.	S.S	E(MS)
Blocks	b-1	$w_{\cdot}\, \Sigma(\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot})^2$	$1 + w_{\cdot}\, \Sigma(\beta_i - \beta_{\cdot})^2/(b-1)$
Treatments	t-1	$b\Sigma w_j (y_{\cdot j} - \tilde{y}_{\cdot\cdot})^2$	$1 + b\Sigma w_j (\tau_j - \tilde{\tau})^2/(t-1)$
Error	(b-1)(t-1)	$\Sigma\Sigma\, w_j (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})^2$	1
Total(corrected)	(bt-1)	$\Sigma\Sigma\, w_j y_{ij}^2 - (\Sigma w_j Y_{\cdot j})^2/bw_{\cdot}$	

When the F-test or χ^2-test indicates significant differences among the treatment or block effects, the difference between any two of the treatment or block effects can be tested by the normal test because

$$z_1 = (\hat{\tau}_j - \hat{\tau}_k)/\{(1/bw_j) + (1/bw_k)\}^{\frac{1}{2}}$$

and $z_2 = (\tilde{\beta}_i - \tilde{\beta}_\ell)/\{2/w.\}^{\frac{1}{2}}$ are both standardised normal variates under the null hypotheses.

5.2 Estimation of weights

If the error variances are not known, these have to be estimated from the sample for use in computing the required statistics.

The maximum likelihood estimator of σ_j^2 is given by

$$\tilde{\sigma}_j^2 = \sum_{i=1}^{b} (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})^2/b \quad ,$$

which involves the error variances. Russell and Bradley (1958) showed that the iterative solution to this equation converges for all j. The limiting solution is zero for any one j = p, say. The other estimators are

$$\tilde{\sigma}_j^2 = \sum_i (y_{ij} - y_{\cdot j} - y_{ip} + y_{\cdot p})^2/b \quad , \qquad j \neq p,\ j=1,2,..,t.$$

The non-zero estimators are thus correlated and their distributional properties are difficult to obtain.

The minimum norm quadratic unbiased estimator (MINQUE) of σ_j^2 is obtained below.

Let $\underset{\sim}{Y}$ be the vector of observations arranged treatment by treatment; then the model can be written in the form

$$\underset{\sim}{Y} = \underset{\sim}{\Delta}'\underset{\sim}{\tau} + \underset{\sim}{D}'\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$
$$= (\Delta' \vdots D') \begin{pmatrix} \underset{\sim}{\tau} \\ \cdots \\ \underset{\sim}{\beta} \end{pmatrix} + \underset{\sim}{\varepsilon}$$

as in section 4.1. The over-all design matrix is singular and $\underset{\sim}{1}' \underset{\sim}{\tau} = 0$ by the constraint. To obtain the projection matrix we need a generalised inverse of the matrix

$\begin{pmatrix} \underset{\sim}{\Delta} \\ \cdots \\ \underset{\sim}{D} \end{pmatrix} (\underset{\sim}{\Delta}' \vdots \underset{\sim}{D}')$, which can be obtained by a method given by Rao (1973, p.225) as used in section 3.1.3. But a simpler method is to re-parameterize the treatments by an orthogonal transformation and thereby transform the design matrix into one of full rank*. For this let us consider Helmert's transformation of treatment parameters given by

$$\underset{\sim}{\tau}_1 = \underset{\sim}{c}\underset{\sim}{\tau}$$

where

$$\underset{\sim}{c} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \; \cdots\cdots & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \; \cdots\cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{1}{\sqrt{t(t-1)}} & \frac{1}{\sqrt{t(t-1)}} & \frac{1}{\sqrt{t(t-1)}} \; \cdots & \frac{1}{\sqrt{t(t-1)}} & -\frac{(t-1)}{\sqrt{t(t-1)}} \\ \frac{1}{\sqrt{t}} & \frac{1}{\sqrt{t}} & \frac{1}{\sqrt{t}} \; \cdots\cdots & \frac{1}{\sqrt{t}} & \frac{1}{\sqrt{t}} \end{bmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{c}_1 \\ \cdots \\ \underset{\sim}{c}_2 \end{pmatrix} ,$$

say, with $\underset{\sim}{c}_2$ as the last row of $\underset{\sim}{c}$. Since $\underset{\sim}{c}_2 \underset{\sim}{\tau} = 0$ by the constraint, the last element of $\underset{\sim}{\tau}_1$ is zero so that

$\underset{\sim}{\tau}_1 = \begin{pmatrix} \underset{\sim}{\tau}_o \\ \cdots \\ 0 \end{pmatrix}$, say. The matrix $\underset{\sim}{c}$ is orthogonal so that

$\underset{\sim}{c}'\underset{\sim}{c} = \underset{\sim}{I} = \underset{\sim}{c}\,\underset{\sim}{c}'$ and furthermore $\underset{\sim}{c}_1 \underset{\sim}{1} = 0$. Hence,

* This was suggested by Professor S. C. Pearce.

$\underset{\sim}{\tau} = (\underset{\sim}{c}_1' \vdots \underset{\sim}{c}_2') \begin{pmatrix} \underset{\sim}{\tau_0} \\ \cdots \\ 0 \end{pmatrix} = \underset{\sim}{c}_1' \underset{\sim}{\tau}_o$. Thus the model reduces to

$$\underset{\sim}{Y} = \underset{\sim}{\Delta}' \underset{\sim}{c}_1' \underset{\sim}{\tau}_o + \underset{\sim}{D}' \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$
$$= \underset{\sim}{X}' \underset{\sim}{\Theta} + \underset{\sim}{\varepsilon},$$

say, with the design matrix $\underset{\sim}{X}' = (\underset{\sim}{\Delta}' \underset{\sim}{c}_1' \vdots \underset{\sim}{D}')$, now a matrix of full rank.

As $\underset{\sim}{D}\,\underset{\sim}{\Delta}'\,\underset{\sim}{c}_1' = \underset{\sim}{J}\,\underset{\sim}{c}_1' = \underset{\sim}{0}$ and $\underset{\sim}{c}_1 \underset{\sim}{c}_1' = \underset{\sim}{I}_{t-1}$, we have

$$(\underset{\sim}{X}\underset{\sim}{X}') = \left[\begin{array}{c|c} \underset{\sim}{c}_1 b \underset{\sim}{I}_t \underset{\sim}{c}_1' & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & t\underset{\sim}{I}_b \end{array}\right] = \left[\begin{array}{c|c} b\underset{\sim}{I}_{t-1} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & t\underset{\sim}{I}_b \end{array}\right]$$

whence

$$(\underset{\sim}{X}\underset{\sim}{X}')^{-1} = \left[\begin{array}{c|c} \frac{1}{b} \underset{\sim}{I}_{t-1} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \frac{1}{t} \underset{\sim}{I}_b \end{array}\right]$$

where $\underset{\sim}{I}_b$ is the identity matrix of order b and $\underset{\sim}{J}$ is a matrix with all its elements equal to unity. Thus, we have,

$$\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X} = \begin{bmatrix} \frac{1}{t}\underset{\sim}{I}_b + \frac{t-1}{bt}\underset{\sim}{J}_b & \frac{1}{t}\underset{\sim}{I}_b - \frac{1}{bt}\underset{\sim}{J}_b & \cdots\cdots & \frac{1}{t}\underset{\sim}{I}_b - \frac{1}{bt}\underset{\sim}{J}_b \\ \vdots & \vdots & & \vdots \\ \frac{1}{t}\underset{\sim}{I}_b - \frac{1}{bt}\underset{\sim}{J}_b & \frac{1}{t}\underset{\sim}{I}_b - \frac{1}{bt}\underset{\sim}{J}_b & \cdots\cdots & \frac{1}{t}\underset{\sim}{I}_b + \frac{t-1}{bt}\underset{\sim}{J}_b \end{bmatrix}$$

and the projection matrix $\underset{\sim}{S}$ is given by

$$\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X} = \frac{1}{bt}\begin{bmatrix} b(t-1)\underset{\sim}{I}_b - (t-1)\underset{\sim}{J}_b & -b\underset{\sim}{I}_b + \underset{\sim}{J}_b \cdots\cdots b\underset{\sim}{I}_b + \underset{\sim}{J}_b \\ \vdots & \vdots \\ -b\,\underset{\sim}{I}_b + \underset{\sim}{J}_b & -b\underset{\sim}{I}_b + \underset{\sim}{J}_b \cdots b(t-1)\underset{\sim}{I}_b - (t-1)\underset{\sim}{J}_b \end{bmatrix}$$

where $\underset{\sim}{J}_b$ is the square matrix of order b with all its elements equal to 1. It is easily observed that $\underset{\sim}{S}\underset{\sim}{Y}$ is the vector of residuals.

Now let $\underset{\sim}{F} = (f_{ij})$ with f_{ij} as the square of the (i,j)th element of the projection matrix $\underset{\sim}{S}$, $\underset{\sim}{\delta} = (\sigma_1^2, \ldots, \sigma_1^2, \ldots, \sigma_t^2, \ldots, \sigma_t^2)'$, the vector of error variances, each σ_j^2 being repeated b times, and $\underset{\sim}{v} = \{(y_{11} - y_{1\cdot} - y_{\cdot 1} + y_{\cdot\cdot})^2, \ldots, (y_{bt} - y_{b\cdot} - y_{\cdot t} + y_{\cdot\cdot})^2\}'$, the vector of squares of residuals. Then the MINQUE's of σ_j^2 are obtained from

$$\underset{\sim}{F} \ \underset{\sim}{\delta} = \underset{\sim}{v} .$$

Adding the b equations for σ_j^2, we have,

$$\frac{1}{b^2 \ t^2} \{ b^2(b-1)\sigma_1^2 + b^2(b-1)\sigma_2^2 + \ldots + b^2(b-1)(t-1)^2\sigma_j^2 + \ldots + b^2(b-1)\sigma_t^2\}$$

$$= \sum_{i=1}^{b} (y_{ij} - y_{i.} - y_{\cdot j} + y_{\cdot\cdot})^2$$

or,

$$\frac{b-1}{t^2} \{ \sigma_i^2 + \ldots + (t-1)^2\sigma_j^2 + \ldots + \sigma_t^2 \} \qquad = S_j^2,$$

say, $j = 1,2,\ldots,t$. All the t equations can be written together as

$$\{(t^2 - 2t)\underset{\sim}{I}_t + \underset{\sim}{J}_t\} \quad (\hat{\sigma}_1^2, \ldots, \hat{\sigma}_t^2)' =$$

$$\frac{t^2}{b-1} (S_1^2, \ldots, S_t^2)' .$$

If we write the inverse of the coefficient matrix as $\alpha \underset{\sim}{I}_t + \beta \underset{\sim}{J}_t$, then α and β are given by

$$\alpha = 1/t(t-2) \quad \text{and} \quad \beta = -1/t^2(t-1)(t-2).$$

The MINQUE of σ_j^2 is then obtained as

$$\hat{\sigma}_j^2 = \{1/(b-1)(t-1)(t-2)\} \{(t^2-t) \sum_{i=1}^{b} (y_{ij}-y_{i\cdot}- y_{\cdot j}+y_{\cdot\cdot})^2 - \Sigma\Sigma (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2\}$$

Ehrenberg (1950) mentioned two unbiased estimators of σ_j^2 and this is one of them. This was also obtained by Russell and Bradley (1958) in a different way.

These estimators are obviously correlated and difficult to handle algebraically.

A simpler estimator called an almost unbiased estimator (AUE) was provided by Horn et al. (1975). They gave a method of obtaining an AUE from a MINQUE. Later on, Horn and Horn (1975) showed that the AUE possessed a smaller mean square error than the MINQUE in a wide range of situations.

In this case, the method of Horn et al. gives the AUE of σ_j^2 as

$$s_j^2 = (S_j^2/b)\,(1-k_{jj})^{-1}$$

$$= (S_j^2/b)\,\{1-(b+t-1)/bt\}^{-1}$$

where $k_{jj} = (b+t-1)/bt$ is the jth diagonal element of $\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X}$. Unlike MINQUE, AUE is always positive. The covariance between s_j^2 and $s_{j'}^2 (j \neq j')$ is negligible as is shown in section 5.4.

If we let $u_i = y_{ij} - y_{i.}$ so that $u_. = \sum_1^b u_i/b$, then the random variables u_i are independently and normally distributed on the assumption of normality of errors, and

$$\text{var}\,(u_i) = (1-2/t)\,\sigma_j^2 + \bar{\sigma}^2/t \text{ where } \bar{\sigma}^2 = \sum_1^t \sigma_j^2/t.$$

Replacing $\bar{\sigma}^2$ by σ_j^2 as an approximation, we have var $(u_i) = \sigma_j^2(1-1/t)$ so that the distribution of $S_j^2 = \sum_1^b (u_i-u_.)^2$ may be approximated by that of $\chi^2\sigma_j^2(1-1/t)$ with (b-1) d.f.

Johnson (1962) recommended that $F = S_j^2/S_{j'}^2$ $(j \neq j')$ might be regarded as an F-statistic with (b-1) and (b-1) d.f. for testing the hypothesis: $\sigma_j = \sigma_{j'}$, when $b > 5$.

As $S_j^2/\sigma_j^2(1-1/t) = b\, s_j^2\{1-(b+t-1)/bt\}/\sigma_j^2$

$(1 - 1/t) = (b - 1)\, s_j^2/\sigma_j^2$,

we may assume that $(b-1)\, s_j^2/\sigma_j^2$ is approximately a χ^2-variate with (b-1) d.f.

5.3 A Theorem on the expectation of functions of correlated χ^2-variates

When the estimators of the error variances are mutually correlated, the Theorem 1 (section 2.1.4) due to Meier needs to be generalised for use in the adjustment of statistics. The generalised form is given in

Theorem 2. Let $\nu_j x_j$ be χ^2-variates with ν_j d.f., $j = 1,2,\ldots,t$. Let these variates be mutually correlated and ν_j be large. Let $f(x_1,\ldots,x_t)$ be a rational function with no singularities in the range $0 < x_1,\ldots,x_t < \infty$. Then asymptotically in ν_j,

$$E\,\{f(x_1,\ldots,x_t)\} = f(1,\ldots,1) + \sum_{j=1}^{t} \frac{1}{\nu_j} \left[\frac{\partial^2 f(x_1,\ldots,x_t)}{\partial x_j^2} \right]_{\text{all } x_j=1}$$

$$+ \tfrac{1}{2} \sum_{j \neq k}\sum E(x_j-1)(x_k-1) \left[\frac{\partial^2 f(x_1,\ldots,x_t)}{\partial x_j\, \partial x_k} \right]_{\text{all } x_j=1}$$

$$+ \text{ terms of order lower than } O\,\{\sum_{jk}\sum (1/\nu_j \nu_k)^{\frac{1}{2}}\}.$$

Proof: As a rational function, f is the quotient of two polynomials and as such admits partial derivatives of all orders. By the non-singularity assumption, these derivatives are all finite within the range $(0, \infty)$. The Taylor's series expansion of f in x_j about its expected value 1 is thus given by

$$f(x_1,\ldots,x_t) = f(1,\ldots,1) + \sum_{r=1}^{n} \frac{1}{r!}\left[(x_1-1)\frac{\partial}{\partial x_1} + \ldots + (x_t - 1)\frac{\partial}{\partial x_t}\right]^r f(1,1,\ldots,1) + R_n \quad \ldots\ldots\ldots\ldots (16)$$

The term R_n is the remainder given by

$$R_n = \frac{1}{n!}\left[(x_1-1)\frac{\partial}{\partial x_1} + \ldots + (x_t-1)\frac{\partial}{\partial x_t}\right]^n \{f(\xi_1,\ldots,\xi_t) - f(1,\ldots,1)\}$$

where $|\xi_j-1| < |x_j-1|$ and the differentiation is done before the resulting expressions are evaluated at $x_j=1$ and $x_j = \xi_j$ for all j. Using the multinomial expansion, the remainder term can be written as

$$R_n = \frac{1}{n!}\sum (x_h-1)(x_i-1)\ldots(x_j-1)(x_k-1)\left[f^{(h,\ldots,k)}(\xi_1,\ldots,\xi_t) - f^{(h,\ldots,k)}(1,\ldots,1)\right]$$

where $f^{(h,\ldots,k)}$ denotes the nth order partial derivative of f with repsect to the variables in some order (including repetitions) and the sum includes all possible pure and mixed n-factor terms in the x_j's. It is shown below that $E(R_n) \longrightarrow 0$ as $n \longrightarrow \infty$.

By the generalised triangle inequality, we have

$$\left| E(R_n) \right| \leq (1/n!)\ \Sigma \left| E \left[(x_h-1) \ldots (x_k-1) \right.\right.$$

$$\left.\left. \{ f(\overset{(h,\ldots,k)}{\xi_1,\ldots,\xi_t}) - f(\overset{(h,\ldots,k)}{1,\ldots,1}) \} \right] \right|$$

$$\leq (1/n!)\Sigma \left| \left[E\ \{(x_h-1)^2 \ldots (x_k-1)^2\} \right.\right.$$

$$\left.\left. E\ \{ f(\overset{(h,\ldots,k)}{\xi_1,\ldots,\xi_t}) - f(\overset{(h,\ldots,k)}{1,\ldots,1}) \}^2 \right]^{\frac{1}{2}} \right|$$

by the Cauchy-Schwarz inequality. Moreover, since $(x_j-1)^2 \geq 0$, we get

$$E\ \{ (x_h-1)^2 \ldots (x_k-1)^2 \} \leq \{ E(x_h-1)^{2n} \ldots E(x_k-1)^{2n} \}^{1/n},$$

by the generalised Hölder's inequality (see Rao, 1973, p.55). Consequently,

$$|E(R_n)| \leq (1/n!)\ \Sigma\ | \{E(x_h-1)^{2n}\}^{1/2n} \ldots \{ E(x_k-1)^{2n} \}^{1/2n}$$

$$\left[E\ \{f(\overset{(h,\ldots,k)}{\xi_1,\ldots,\xi_t}) - f(\overset{(h,\ldots,k)}{1,\ldots,1})\}^2 \right]^{1/2} \Bigg|$$

For large ν_j, it follows that x_j is normally distributed with mean = 1 and variance = $2/\nu_j$. The joint distribution of the x_j's is thus asymptotically multi-variate normal having the form

$$k \exp\{ - (\underset{\sim}{X} - \underset{\sim}{1})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{X} - \underset{\sim}{1})/2 \}$$

where k is a constant; $\underset{\sim}{X}$ is the vector of the variates x_j, $\underset{\sim}{1}$ the vector of unity and $\underset{\sim}{\Sigma}$ the dispersion matrix of $\underset{\sim}{X}$.

The last expectation in the right hand side of $|E(R_n)|$ can therefore be written as

$$k \int \ldots \int \exp\{-(\underset{\sim}{X}-\underset{\sim}{1})' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{X}-\underset{\sim}{1})/2\}\{f^{(h\ldots k)}(\xi_1,\ldots,\xi_t)-f^{(h\ldots k)}(1,\ldots,1)\}^2 \pi dx_j .$$

As all the partial derivatives of f exist, $f^{(h \ldots k)}(\xi_1,\ldots,\xi_t)$ does not exceed a finite quantity M within the range of integration. Hence , this integral cannot exceed

$$k\{M-f^{(h\ldots k)}(1,\ldots,1)\}^2 \int \cdots \int \exp\{-(\underset{\sim}{X}-\underset{\sim}{1})' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{X}-\underset{\sim}{1})/2\}\pi dx_j$$

which is a constant. Thus this expectation is bounded.

Again, by the formula for central moments of the normal distribution, we have,

$$E\{(x_j-1)^{2n}\} = (2n)!/\nu_j^n n!$$

$$= c(1/\nu_j^n)(2n)^n,$$

for some constant c, on neglecting terms of order (1/n). Thus,

$$|E(R_n)| \leq (c_0/n!)\, \Sigma\, (1/\nu_h \nu_i \ldots \nu_j \nu_k)^{1/2} n^{n/2}$$

$$= \begin{cases} [c_1/(n/2)!]\, \Sigma\, (1/\nu_h \ldots \nu_k)^{1/2} & \text{if n is even} \\ [c_2/n^{1/2}\{(n-1)/2\}!]\, \Sigma\, (1/\nu_h \ldots \nu_k)^{1/2} & \text{if n is odd,} \end{cases}$$

up to the same order of approximation, where c_o, c_1 and c_2 are positive constants. Hence $|E(R_n)| \longrightarrow 0$ as $n \longrightarrow \infty$. This means that $E(R_n) \longrightarrow 0$ as $n \longrightarrow \infty$.

It follows from above that the expectation of a term in the multinominal expansion is of order $(1/\nu_h \ldots \nu_k)^{1/2}$.

Hence, the theorem follows if we take the expectation of (16) and keep terms up to $r = 2$.

A consequence of the theorem is that the adjusted statistic,

$$f(x_1,\ldots,x_t) - \sum_j^t \frac{1}{\nu_j}\left[\frac{\partial^2 f(x_1,\ldots,x_t)}{\partial x_j^2}\right]_{\text{all } x_j=1} - \frac{1}{2}\sum\sum_{j\neq k} E(x_j-1)(x_k-1)\left[\frac{\partial^2 f(x_1,\ldots,x_t)}{\partial x_j\,\partial x_k}\right]_{\text{all } x_j=1},$$

is free from terms of order $\{1/(\nu_j\nu_k)^{1/2}\}$ and thus approximates its theoretical value $f(1,\ldots,1)$, more closely than the statistic $f(x_1,\ldots,x_t)$ itself. When $E\{(x_j-1)(x_k-1)\}$ is negligible, the adjustment reduces to that obtained by Theorem 1 due to Meier (1953).

5.4 Covariance between s_j^2 and s_k^2 $(j\neq k)$

We have, $S_j^2 = \sum_{i=1}^{b}(y_{ij}-y_{i\cdot}-y_{\cdot j}+y_{\cdot\cdot})^2 =$

$$\sum_i(\varepsilon_{ij}-\varepsilon_{i\cdot}-\varepsilon_{\cdot j}+\varepsilon_{\cdot\cdot})^2 \quad \text{and}$$

$$S_k^2 = \sum_{i=1}^{b}(\varepsilon_{ik}-\varepsilon_{i\cdot}-\varepsilon_{\cdot k}+\varepsilon_{\cdot\cdot})^2 \quad \text{so that}$$

$$E(S_j^2S_k^2) = E\{\sum_i(\varepsilon_{ij}-\varepsilon_{i\cdot}-\varepsilon_{\cdot j}+\varepsilon_{\cdot\cdot})^2\sum_i(\varepsilon_{ik}-\varepsilon_{i\cdot}-\varepsilon_{\cdot k}+\varepsilon_{\cdot\cdot})^2\}$$

$$= E\Big[\sum_i(\varepsilon_{ij}-\varepsilon_{\cdot j})^2\sum_i(\varepsilon_{ik}-\varepsilon_{\cdot k})^2 + \sum_i(\varepsilon_{ij}-\varepsilon_{\cdot j})^2\sum_i(\varepsilon_{i\cdot}-\varepsilon_{\cdot\cdot})^2$$

$$-2\sum_i(\varepsilon_{ij}-\varepsilon_{\cdot j})^2\sum_i(\varepsilon_{ik}-\varepsilon_{\cdot k})(\varepsilon_{i\cdot}-\varepsilon_{\cdot\cdot})$$

$$+ \sum_i (\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})^2 \sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})^2 + \{ \sum_i (\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})^2 \}^2 - 2$$

$$\sum_i (\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})^2 \sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})$$

$$- 2 \sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot}) \sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})^2 - 2 \sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})$$

$$+ 4 \sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot}) \sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot}) \Big]$$

To find the expectations of the individual terms we observe that

(i) $\sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})^2 = (1 - \frac{1}{b}) \sum_i \varepsilon_{ij}^2 - \frac{1}{b} \sum_{i \neq \ell} \sum \varepsilon_{ij} \varepsilon_{\ell j}$,

(ii) $\sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})^2 = (1 - \frac{1}{b}) \sum_i \varepsilon_{ik}^2 - \frac{1}{b} \sum_{i \neq \ell} \sum \varepsilon_{ik} \varepsilon_{\ell k}$,

(iii) $\sum_i (\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot})^2 = (\frac{1}{t^2} - \frac{1}{bt^2}) \sum_{ij}\sum \varepsilon_{ij}^2 + \frac{1}{t^2} \sum_i \sum_{j \neq k} \sum \varepsilon_{ij} \varepsilon_{ik} - \frac{1}{bt^2}$

$$\sum_{(i,j) \neq (\ell,k)} \sum \varepsilon_{ij} \varepsilon_{\ell k} \, ,$$

(iv) $\sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot}) = (\frac{1}{t} - \frac{1}{bt}) \sum_i \varepsilon_{ij}^2 + \frac{1}{t} \sum_i \varepsilon_{ij} \sum_{k \neq j} \varepsilon_{ik} - \frac{1}{bt} \sum_{i \neq \ell} \sum$

$\varepsilon_{ij} \varepsilon_{\ell j} - \frac{1}{bt} \sum_i \varepsilon_{ij} \sum_i \sum_{k \neq j} \varepsilon_{ik}$,

and (v) $\sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})(\varepsilon_{i\cdot} - \varepsilon_{\cdot\cdot}) = (\frac{1}{t} - \frac{1}{bt}) \sum_i \varepsilon_{ik}^2 + \frac{1}{t} \sum_i \varepsilon_{ik}$

$$\sum_{j \neq k} \varepsilon_{ij} - \frac{1}{bt} \sum_{i \neq \ell} \sum \varepsilon_{ik} \varepsilon_{\ell k} - \frac{1}{bt} \sum_i \varepsilon_{ik} \sum_i \sum_{j \neq k} \varepsilon_{ij} \, .$$

Expectations of the nine individual terms are then given as

(a) $E \{ \sum_i (\varepsilon_{ij} - \varepsilon_{\cdot j})^2 \sum_i (\varepsilon_{ik} - \varepsilon_{\cdot k})^2 \} = (b-1)^2 \sigma_j^2 \sigma_k^2$

(b) $E\{\sum_i(\varepsilon_{ij}-\varepsilon_{\cdot j})^2 \sum_i(\varepsilon_{i.}-\varepsilon_{..})^2\} = (1-\frac{1}{b})(\frac{1}{t^2}-\frac{1}{bt^2})$

$E(\sum_i \varepsilon_{ij}^4 + \sum_{i\neq\ell}\sum \varepsilon_{ij}^2\varepsilon_{\ell j}^2 + \sum_i \varepsilon_{ij}^2 \sum_i \sum_{k\neq j} \varepsilon_{ik}^2) + \frac{1}{b^2t^2} E(\sum_{i\neq\ell}\sum \varepsilon_{ij}^2 \varepsilon_{\ell j}^2)$

$$= \frac{(b^3-2b+1)}{bt^2}\sigma_j^4 + \frac{(b-1)^2}{t^2}\sigma_j^2 \sum_{k\neq j}\sigma_k^2 .$$

(c) $E\{\sum_i(\varepsilon_{ij}-\varepsilon_{\cdot j})^2 \sum_i(\varepsilon_{ik}-\varepsilon_{\cdot k})(\varepsilon_{i\cdot}-\varepsilon_{..})\} = (1-\frac{1}{b})(\frac{1}{t}-\frac{1}{bt})$

$E(\sum_i \varepsilon_{ij}^2 \sum_i \varepsilon_{ik}^2) = \frac{(b-1)^2}{t}\sigma_j^2\sigma_k^2$

(d) $E\{\sum_i(\varepsilon_{i\cdot}-\varepsilon_{..})^2 \sum_i(\varepsilon_{ik}-\varepsilon_{\cdot k})^2\} = \frac{b^3-2b+1}{bt^2}\sigma_k^4$

$+ \frac{(b-1)^2}{t^2}\sigma_k^2 \sum_{j\neq k}\sigma_j^2$ from (b) by interchanging the roles of j and k.

(e) $E\{\sum_i(\varepsilon_{i\cdot}-\varepsilon_{..})^2\}^2 = \frac{(b-1)^2}{b^2} E(\sum_i \varepsilon_{i\cdot}^4) + \frac{b^2-2b+3}{b^2}$

$E(\sum_{i\neq\ell}\sum \varepsilon_{i.}^2 \varepsilon_{\ell\cdot}^2)$

$$= \frac{(b-1)^2}{b^2}\cdot\frac{3b}{t^4}(\sum_j \sigma_j^2)^2 + \frac{b^2-2b+3}{b^2}(\sum_i \sigma_j^2)^2 b(b-1)/t^4$$

$$= \frac{b^2-1}{t^4}(\sum_j \sigma_j^2)^2$$

(f) $E\{\sum_i(\varepsilon_{i\cdot}-\varepsilon_{..})^2 \sum_i(\varepsilon_{ik}-\varepsilon_{\cdot k})(\varepsilon_{i\cdot}-\varepsilon_{..})\} = (\frac{1}{t}-\frac{1}{bt})(\frac{1}{t^2}-\frac{1}{bt^2})$

$E(\sum_i \varepsilon_{ik}^2)(\sum_{ij}\sum \varepsilon_{ij}^2) + (\frac{1}{t^3}-\frac{1}{bt^3})E(\sum_i \varepsilon_{ik}^2 \sum_{j\neq k}\varepsilon_{ij}^2)$

$$+ \frac{1}{b^2t^3} E(\sum_{i\neq \ell} \sum \varepsilon_{ik}^2 \varepsilon_{\ell k}^2) - \frac{1}{bt^3} E(\sum_i \varepsilon_{ik}^2 \sum_{j\neq k} \varepsilon_{ij}^2) +$$

$$\frac{1}{b^2t^3} E(\sum_i \varepsilon_{ik}^2 \sum_i \sum_{j\neq k} \varepsilon_{ij}^2)$$

$$= \frac{b^3-2b+1}{bt^3} \sigma_k^4 + \frac{b(b-1)}{t^3} \sigma_k^2 \sum_{j\neq k} \sigma_j^2$$

(g) $E\{ \sum_i (\varepsilon_{ij} - \varepsilon_{.j})(\varepsilon_{i.} - \varepsilon_{..}) \sum_i (\varepsilon_{ik} - \varepsilon_{.k})^2 \} = (1 - \frac{1}{b})$

$$(\frac{1}{t} - \frac{1}{bt}) E(\sum_i \varepsilon_{ik}^2 \sum_i \varepsilon_{ij}^2) = \frac{(b-1)^2}{t} \sigma_j^2 \sigma_k^2$$

(h) $E\{ \sum_i (\varepsilon_{ij} - \varepsilon_{.j})(\varepsilon_{i.} - \varepsilon_{..}) \sum_i (\varepsilon_{i.} - \varepsilon_{..})^2 \} = \frac{b^3-2b+1}{bt^3} \sigma_j^4$

$+ \frac{b(b-1)}{t^3} \sigma_j^2 \sum_{k\neq j} \sigma_k^2$ from (f) by interchanging the roles of

j and k.

(i) $E\{ \sum_i (\varepsilon_{ij} - \varepsilon_{.j})(\varepsilon_{i.} - \varepsilon_{..}) \sum_i (\varepsilon_{ik} - \varepsilon_{.k})(\varepsilon_{i.} - \varepsilon_{..}) \} =$

$$(\frac{1}{t} - \frac{1}{bt})^2 E(\sum_i \varepsilon_{ij}^2 \sum_i \varepsilon_{ik}^2) + (\frac{1}{t^2} - \frac{1}{bt^2} - \frac{1}{bt^2} + \frac{1}{b^2t^2})$$

$$E(\sum_i \varepsilon_{ij}^2 \varepsilon_{ik}^2)$$

$$= \frac{\sigma_j^2 \sigma_k^2}{t^2} (b^2 - b - 1 + \frac{1}{b}).$$

Utilizing the above nine expectations and simplifying, we get

$$\frac{1}{b^2} E(S_j^2 S_k^2) = \frac{\sigma_j^2\sigma_k^2}{b^2}\{(b-1)^2 - \frac{4(b-1)^2}{t} + \frac{4}{t^2}(b^2-b-1+\frac{1}{b})\}$$

$$+ \frac{(b^3-2b+1)(t-2)}{b^3t^3}(\sigma_j^4 + \sigma_k^4)$$

$$+\left\{\frac{(b-1)^2}{b^2t^2} - \frac{2b(b-1)}{b^2t^3}\right\}(\sigma_j^2 \sum_{k\neq j}\sigma_k^2 + \sigma_k^2 \sum_{j\neq k}\sigma_j^2) + \frac{b^2-1}{b^2t^4}(\sum_i \sigma_j^2)^2$$

$$= \sigma_j^2\sigma_k^2(1-\frac{2}{b}-\frac{4}{t}) + \frac{\bar{\sigma}^2}{t}(\sigma_j^2+\sigma_k^2),$$

neglecting the terms of order $1/t^2$, $1/b^2$ or $1/bt$.

Also,

$$\frac{1}{b^2} E(S_j^2)E(S_k^2) = \frac{(b-1)^2}{b^2}\{\sigma_j^2(1-\frac{2}{t}) + \frac{\bar{\sigma}^2}{t}\}\{\sigma_k^2(1-\frac{2}{t}) + \bar{\sigma}^2/t\}$$

$$= \sigma_j^2\sigma_k^2(1-\frac{2}{b}-\frac{4}{t}) + \frac{\bar{\sigma}^2}{t}(\sigma_j^2+\sigma_k^2)$$

so that

$$\frac{1}{b^2}\text{cov}(S_j^2, S_k^2) = \{E(S_j^2S_k^2) - E(S_j^2)E(S_k^2)\}/b^2 = 0,$$

up to the same order of approximation. Thus, we have

$$\text{cov}(s_j^2, s_k^2) = 0$$

up to the order $1/t^2$, $1/b^2$ or $1/bt$.

It follows from above and section 5.2 that Bartlett's χ^2-test using s_j^2 may be used for testing equality of group variances. The likelihood ratio test and sphericity test (Shukla, 1972) may also be used.

5.5 Adjustment of the test-statistics

Let $x_j = s_j^2/\sigma_j^2$ where s_j^2 is the AUE of σ_j^2, $j = 1,2,\ldots,t$. Then the estimated weights are:

$\hat{w}_j = 1/s_j^{\,2} = 1/x_j\sigma_j^{\,2}$, the d.f. $\nu_j = (b-1)$. Let $\hat{w}. = \Sigma\hat{w}_j$.

It follows from the previous section that Cov $(x_j, x_k) = 0$ for $j \neq k$ up to the order $1/b^2$, $1/t^2$ or $1/bt$. Hence, the adjustments of the statistics using estimated weights to remove the bias of order $1/(b-1)$, by using Theorem 2 will be the same as that by Theorem 1 due to Meier (section 2.1.4). Such adjustment for one test-statistic $(\hat{F}_6)$ was given in section 3.2.2 for the more general case of the two-way classification with proportional cell frequencies. The adjusted test-statistics for the special case of randomised block designs are stated below using some of the expressions derived in that section.

(i) Adjusted F-statistics

(a) Significance of treatment effects

The F-statistic using estimated weights for testing the significance of treatment effects is given by

$$\hat{F}_1 = \frac{b \sum_j \hat{w}_j (y._j - \hat{\hat{y}}..)^2 / (t-1)}{\Sigma\Sigma\,\hat{w}_j (y_{ij} - \hat{\hat{y}}_i. - y._j + \hat{\hat{y}}..)^2/(b-1)(t-1)}$$

$$= (b-1)\ \text{TSS/E SS},$$

say, with $(t-1)$ and $(b-1)(t-1)$ d.f., where TSS and ESS denote, respectively the treatments SS and error SS using estimated weights. Then the adjusted F-statistic is

$$\hat{F}_1(\text{adj}) = \hat{F}_1 - \frac{1}{b-1} \sum_{j=1}^{t} \left[\frac{\partial^2 \hat{F}_1}{\partial x_j^{\,2}}\right]_{\text{all } x_j = 1} \quad \text{using estimated weights}$$

with

$$\frac{\partial^2 \hat{F}_1}{\partial x_j^2} = \frac{b-1}{(ESS)^3} \{ (ESS)^2 \frac{\partial^2 TSS}{\partial x_j^2} - (ESS)(TSS) \frac{\partial^2 ESS}{\partial x_j^2}$$

$$-2(ESS) \frac{\partial ESS}{\partial x_j} \frac{\partial TSS}{\partial x_j} + 2 (TSS) \left(\frac{\partial ESS}{\partial x_j}\right)^2 \} \quad \ldots\ldots\ldots (17)$$

where

$$\left[\frac{\partial (T\ SS)}{\partial x_j}\right]_{\text{all } x_j = 1} = - b\, w_j (y_{\cdot j} - y_{\cdot\cdot})^2,$$

$$\left[\frac{\partial (ESS)}{\partial x_j}\right]_{\text{all } x_j=1} = - \sum_i w_j y_{ij}(y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot}) + \sum_i \sum_j w_j f_j y_{ij} (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot}),$$

$$\left[\frac{\partial^2 (TSS)}{\partial x_j^2}\right]_{\text{all } x_j=1} = 2bw_j(1-f_j)(y_{\cdot j} - \tilde{y}_{\cdot\cdot})^2$$

and

$$\left[\frac{\partial^2 (ESS)}{\partial x_j^2}\right]_{\text{all } x_j=1} = 2 \sum_i w_j(1-f_j) y_{ij}(y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})$$

$$-2 \sum_i \sum_j w_j f_j (1-f_j) y_{ij}(y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})$$

with $f_j = w_j / w_{\cdot\cdot}$

(b) Equality of block effects

The F-statistic using estimated weights

for testing the equality of block effects is given by

$$\hat{F}_2 = \frac{\hat{w}_{\cdot} \sum (\hat{\tilde{y}}_{i\cdot} - \hat{\tilde{y}}_{\cdot\cdot})/(b-1)}{\sum_i \sum_j \hat{w}_j (y_{ij} - \hat{\tilde{y}}_{i\cdot} - y_{\cdot j} + \hat{\tilde{y}}_{\cdot\cdot})^2/(b-1)(t-1)}$$

$$= (t-1)\ \mathrm{BSS}/\mathrm{ESS},$$

say, with (b-1) and (b-1)(t-1) d.f., where BSS denotes the block SS using estimated weights. The adjusted F-statistic is

$$\hat{\hat{F}}_2 = \hat{F}_2 - \frac{1}{b-1} \sum_j \left[\frac{\partial^2 \hat{F}_2}{\partial x_j^2} \right]_{\text{all } x_j = 1} \quad \text{using estimated weights,}$$

where $\dfrac{\partial^2 \hat{F}_2}{\partial x_j^2}$ is given by (17) above with (b-1) and TSS replaced by (t-1) and BSS respectively. The two additional partial derivatives are

$$\left[\frac{\partial (\mathrm{BSS})}{\partial x_j} \right]_{\text{all } x_j=1} = - \sum_i w_j (\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot})^2 - 2 \sum_i \sum_j w_j f_j (\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot}) (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot})$$

and

$$\left[\frac{\partial^2 (\mathrm{BSS})}{\partial x_j^2} \right]_{\text{all } x_j=1} = 2 \sum_i w_j (\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot}) \{ (\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot}) + 2 f_j (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot}) \} + 2 \sum \sum w_j f_j (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot}) \{ f_j (y_{ij} - \tilde{y}_{i\cdot} - y_{\cdot j} + \tilde{y}_{\cdot\cdot}) + 2(1-f_j)(\tilde{y}_{i\cdot} - \tilde{y}_{\cdot\cdot}) \} .$$

(ii) Adjustment of the normal test-statistics

(a) A Treatment difference

The approximate normal test-statistic using estimated weights for testing the difference between the jth and kth treatments is

$$\hat{z}_1 = |y._j - y._k| \,/\left\{\frac{1}{b\hat{w}_j} + \frac{1}{b\hat{w}_k}\right\}^{\frac{1}{2}}.$$ This is in the same form as that for testing the difference between two treatments in the one-way model. Hence, the adjusted normal test-statistic is given, from section 2.1.4, by

$$\hat{z}_1(adj) = \{|\, y._j - y._k \,| \,/(1/b\hat{w}_j + 1/b\hat{w}_k)^{\frac{1}{2}}\} \times$$

$$\{1 - \tfrac{3}{4}\frac{1}{(1/\hat{w}_j + 1/\hat{w}_k)^2}\left[1/(b-1)\hat{w}_j^{\,2} + 1/(b-1)\hat{w}_k^{\,2}\right]\}.$$

(b) Difference between block effects

The approximate normal test-statistic using estimated weights for testing the difference between the hth and ith block effects is $\hat{z}_2 = |\,\hat{\tilde{y}}_{h.}(adj) - \hat{\tilde{y}}_{i}.(adj)\,| \,/\,(2/\hat{w}.)^{\frac{1}{2}}$ where $\hat{\tilde{y}}_{i.}(adj) = \sum_j \hat{f}_j y_{ij} - \frac{2}{b-1}\sum_j \hat{f}_j(1-\hat{f}_j)(y_{ij} - \hat{\tilde{y}}_{i.})$ from section 3.1.4 with $\hat{f}_j = \hat{w}_j/\hat{w}.$. This statistic is a special case of the corresponding test-statistic of section 3.1.5 and so the adjustment of $\hat{z}_2$ is obtained as

$$\hat{z}_2(adj) = \hat{z}_2\{\,1 - \Sigma\, \hat{f}_j(1-\hat{f}_j/4)/(b-1)\,\}\ .$$

5.6 Multiple comparison

For this design, the error sum of squares depends on weights. Thus, the square root s of the mean square error as well as $\hat{\sigma}_{\hat{\psi}}$ depends on the estimated weights.

So, the expression (5) of section 2.1.5 for estimating joint confidence intervals of parametric contrasts, needs to be modified. The modified form is

$$\hat{\psi} - D\,(adj) \leq \psi \leq \hat{\psi} + D\,(adj)$$

where $D = Ss\,\hat{\sigma}_{\hat{\psi}}$, $\dfrac{\partial^2 D}{\partial x_j^2} = S\left\{ s\,\dfrac{\partial^2 \hat{\sigma}_{\hat{\psi}}}{\partial x_j^2} + \hat{\sigma}_{\hat{\psi}}\,\dfrac{\partial^2 s}{\partial x_j^2} + \dfrac{2\,\partial s}{\partial x_j}\,\dfrac{\partial \hat{\sigma}_{\hat{\psi}}}{\partial x_j}\right\}$

and

$$D(adj) = D - \Sigma\,\frac{1}{r_j - 1}\left[\frac{\partial^2 D}{\partial x_j^2}\right]_{\text{all } x_j = 1} \quad \text{using estimated weights.} \qquad \cdot\cdot(18)$$

(i) Treatment contrasts

Let $\hat{\psi}_1 = \Sigma\, c_j y._j$ with $\Sigma\, c_j = 0$ be an estimate of the treatment contrast $\psi_1 = \Sigma\, c_j\,\tau_j$. Then the joint confidence interval of all contrasts ψ_1 is given by (18) with $\psi = \psi_1$, $S = \left[(t-1)F_{\alpha}\{(t-1),(b-1)(t-1)\}\right]^{\frac{1}{2}}$, $r_j - 1 = b-1$, $s = \{ESS/(b-1)(t-1)\}^{\frac{1}{2}}$ and $\hat{\sigma}_{\hat{\psi}_1} = (\Sigma\, c_j^2 s_j^2/b)^{\frac{1}{2}}$. The individual derivatives are

$$\left[\frac{\partial s}{\partial x_j}\right]_{\text{all } x_k=1} = \left[\frac{\partial ESS}{\partial x_j} \Big/ 2\,\{ESS(b-1)(t-1)\}^{\frac{1}{2}}\right]_{\text{all } x_k=1},$$

$$\left[\frac{\partial^2 s}{\partial x_j^2}\right]_{\text{all } x_k=1} = \left[\frac{\left\{\dfrac{\partial^2 ESS}{\partial x_j^2} - \left(\dfrac{\partial ESS}{\partial x_j}\right)^2 \Big/ 2\,s^2(b-1)(t-1)\right\} \Big/ 2s}{(b-1)(t-1)}\right]_{\text{all } x_k = 1},$$

$$\left[\frac{\partial \hat{\sigma}_{\hat{\psi}_1}}{\partial x_j}\right]_{\text{all } x_j=1} = c_j^{\,2}\sigma_j^{\,2} \Big/ 2b\,(\Sigma\, c_j^{\,2}\,\sigma_j^{\,2}/b)^{\frac{1}{2}}$$

and

$$\left[\frac{\partial^2 \hat{\sigma}_{\hat{\psi}_1}}{\partial x_j^2}\right]_{\text{all } x_j=1} = -c_j^{\,4}\sigma_j^{\,4}/4b^2(\Sigma\, c_j^{\,2}\sigma_j^{\,2}/b)^{3/2},$$

$\left[\dfrac{\partial ESS}{\partial x_j}\right]_{\text{all } x_j=1}$ and $\left[\dfrac{\partial^2 ESS}{\partial x_j^2}\right]_{\text{all } x_j=1}$ being given in the previous section.

(ii) Block contrasts

In the same way, the joint confidence interval of all contrasts $\psi_2 = \Sigma\, c_i\, \beta_i$, of block effects is given by (18) with $\psi = \psi_2$, $S = \left[(b-1)\, F_{\alpha}\{b-1, (b-1)(t-1)\}\right]^{\frac{1}{2}}$, $r_j - 1 = b-1$, $s = \{ESS/(b-1)(t-1)\}^{\frac{1}{2}}$ and $\hat{\sigma}_{\hat{\psi}_2} = (\Sigma\, c_i^{\,2}/\hat{w}.)^{\frac{1}{2}}$. The two derivatives $[\partial s/\partial x_j]_{\text{all } x_j=1}$ and $[\partial^2 s/\partial x_j^2]_{\text{all } x_j=1}$ are given above and the other two derivatives are

$$\left[\partial \hat{\sigma}_{\hat{\psi}_2} / \partial x_j\right]_{\text{all } x_j=1} = (\Sigma\, c_i^{\,2})^{\frac{1}{2}}\, f_j \,/\, 2\, w.^{\frac{1}{2}}$$

and

$$\left[\partial^2 \hat{\sigma}_{\hat{\psi}_2} / \partial x_j^2 \right]_{\text{all } x_j = 1} = (\Sigma c_i^2)^{\frac{1}{2}} f_j (3f_j / 4 - 1) / w.^{\frac{1}{2}} .$$

The quantities $\hat{\tilde{y}}_{i.}(adj)$ are used in computing $\hat{\psi}_2$.

5.7 Summary measures of dispersion

Since the variances of the treatment estimators are in the same forms as those in the one-way model, the estimated summary measures of dispersion of the estimated treatments are obtained from section 2.1.6 as

$$\text{Estimated A.M.} = \frac{1}{bt} \sum_1^t s_j^2$$

$$\text{Estimated G.M.(adj)} = \frac{1}{b} \left(\prod_1^t s_j^2 \right)^{1/t} \left\{ 1 + \frac{t-1}{t(b-1)} \right\}$$

and

$$\text{Estimated H.M.(adj)} = \frac{t}{b\hat{w}.} \left\{ 1 + 2 \sum_1^t \hat{f}_j (1 - \hat{f}_j)/(b-1) \right\}$$

The estimated block effects have constant variance and so no summary measure of dispersion is needed for them.

CHAPTER 6

LATIN SQUARE DESIGNS

A method for solving the normal equations to find the weighted least squares estimators of the linear parameters, is given along with a procedure for the corresponding analysis on the assumption that the group variances are known. The treatment estimators are found to be orthogonal to those of other linear parameters whereas the estimated row and column effects are not orthogonal to one another. The MINQUE and AUE of group variances are obtained. The AUE's are found to be approximately independent of one another. Adjustment of the test-statistics using estimated weights, for testing hypotheses about the treatments is provided for removing bias. Similarly, other test-statistics can be adjusted. Finally, expressions for joint confidence intervals of treatment contrasts are given.

6.1 Estimation and analysis when the error variances are known

Let the model for a t x t latin square design be

$$y_{ijk} = \beta_i + \gamma_j + \tau_k + \varepsilon_{ijk}$$

where β_i is the effect of the ith row, γ_j the effect of the jth column, τ_k the effect of the kth treatment and ε_{ijk} the error term having mean zero and variance σ_k^2. The errors are assumed to be independent of one another.

The suffices, i,j and k, individually assume values from 1 to t but collectively assume only t^2 sets (triples) of values depending on the design chosen.

Let $\underset{\sim}{Y}$ be the vector of observations arranged treatment by treatment, the observations within each treatment being arranged row by row. Consequently, the column effects are randomly distributed among the observations in $\underset{\sim}{Y}$.

Then the above model can be written as

$$\underset{\sim}{Y} = \underset{\sim}{\Delta}' \underset{\sim}{\tau} + \underset{\sim}{D}_1' \underset{\sim}{\beta} + \underset{\sim}{D}_2' \underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} \qquad \ldots\ldots (19)$$

where $\underset{\sim}{\Delta}'$, $\underset{\sim}{D}_1'$ and $\underset{\sim}{D}_2'$ are the design matrices for the treatment, row and column parameters respectively, $\underset{\sim}{\tau}$ is the vector of treatment effects, $\underset{\sim}{\beta}$ the vector of row effects, $\underset{\sim}{\gamma}$ the vector of column effects and $\underset{\sim}{\varepsilon}$ the vector of errors. Then var $(\underset{\sim}{\varepsilon})$ = diag $(\sigma_1^2, \ldots, \sigma_1^2, \ldots, \sigma_t^2, \ldots, \sigma_t^2)$ and $\underset{\sim}{1}' \underset{\sim}{\tau} = 0 = \underset{\sim}{1}' \underset{\sim}{\gamma}$, $\underset{\sim}{1}$ being the vector all elements of which are unity.

The weighted least squares normal equations for estimating the linear parameters are given by (20) where

$$
\begin{bmatrix}
tw_1 & & & w_1 & \cdots & w_1 & w_1 & \cdots & w_1 \\
 & \ddots & & \vdots & & \vdots & \vdots & & \vdots \\
 & & tw_t & w_t & \cdots & w_t & w_t & \cdots & w_t \\
w_1 & \cdots & w_t & w_{\cdot} & & & w_{i_1} & \cdots & w_{i_t} \\
\vdots & & \vdots & & \ddots & & \vdots & & \vdots \\
w_1 & \cdots & w_t & & & w_{\cdot} & w_{\ell_1} & \cdots & w_{\ell_t} \\
w_1 & \cdots & w_t & w_{i_1} & \cdots & w_{\ell_1} & w_{\cdot} & & \\
\vdots & & \vdots & \vdots & & \vdots & & \ddots & \\
w_1 & \cdots & w_t & w_{i_t} & \cdots & w_{\ell_t} & & & w_{\cdot}
\end{bmatrix}
\begin{bmatrix}
\hat{\tau}_1 \\ \vdots \\ \hat{\tau}_t \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_t \\ \tilde{\gamma}_1 \\ \vdots \\ \tilde{\gamma}_t
\end{bmatrix}
=
\begin{bmatrix}
w_1 \, Y_{\cdot\cdot 1} \\ \vdots \\ w_t \, Y_{\cdot\cdot t} \\ \sum_{jk} w_k \, y_{1jk} \\ \vdots \\ \sum_{jk} w_k \, y_{tjk} \\ \sum_{ik} w_k y_{i1k} \\ \vdots \\ \sum_{ik} w_k y_{itk}
\end{bmatrix}
\quad \cdots\cdots (20)
$$

$w_j = 1/\sigma_j^2$, $w. = \Sigma w_j$ as before and $(i_1,\ldots,i_t)$ and $(\ell_1,\ldots,\ell_t)$ are random permutations of the numbers, 1,2, ...t, based on the random distribution of the column effects as mentioned above.

From (20), the individual normal equations are

$$\tau_k:\quad tw_k\hat{\tau}_k + w_k\Sigma\tilde{\beta}_i + w_k\Sigma\tilde{\gamma}_j = w_k Y_{\cdot\cdot k}, \qquad k = 1,2,\ldots,t.$$

$$\beta_i:\quad \Sigma w_k\hat{\tau}_k + w.\tilde{\beta}_i + \Sigma w_{h_j}\tilde{\gamma}_j = \sum_{jk} w_k y_{ijk}, \qquad i=1,2,\ldots,t.$$

$$\gamma_j:\quad \Sigma w_k\hat{\tau}_k + \Sigma w_{k_i}\tilde{\beta}_i + w.\tilde{\gamma}_j = \sum_{ik} w_k y_{ijk}, \qquad j=1,2,\ldots,t.$$

Here also $(h_1,\ldots,h_t)$ and $(k_1,\ldots,k_t)$ are some random permutations of the numbers, $1,2,\ldots,t$, depending on the design matrix.

Using the constraints $\Sigma w_k\hat{\tau}_k = 0 = \Sigma\tilde{\beta}_i = \Sigma\tilde{\gamma}_j$, the three sets of equations reduce, respectively, to

$$\hat{\tau}_k = y_{\cdot\cdot k} \qquad k = 1,2,\ldots,t$$

$$\tilde{\beta}_i = \tilde{y}_{i\cdot\cdot} - \Sigma w_{h_j}\tilde{\gamma}_j/w. \qquad i = 1,2,\ldots,t$$

and $$\tilde{\gamma}_j = \tilde{y}_{\cdot j\cdot} - \Sigma w_{k_i}\tilde{\beta}_i/w. \qquad j = 1,2,\ldots,t.$$

where $\tilde{y}_{i\cdot\cdot} = \sum_{jk} w_k y_{ijk}/w.$ and $\tilde{y}_{\cdot j\cdot} = \sum_{ik} w_k y_{ijk}/w.$.

Thus the treatment estimators are the ordinary least squares estimators and are orthogonal to those of row and column effects. The last two estimators are non-orthogonal.

The reduced normal equations for the column effects are given by

$$\tilde{\gamma}_j - w_{k_1}(\Sigma w_{h_j}\tilde{\gamma}_j)/w^2. - \quad \ldots \quad - w_{k_t}(\Sigma w_{\ell_j}\tilde{\gamma}_j)/w^2.$$

$$= \tilde{y}_{\cdot j \cdot} - \Sigma w_{k_i}\tilde{y}_{i\cdot\cdot}/w.$$

$j = 1,2,\ldots,t$. The coefficient matrix is of full rank and the solution can be obtained by the method of pivotal condensation. Similarly, the reduced normal equations for $\tilde{\beta}_i$ are

$$\tilde{\beta}_i - \left(\frac{w_{h_1}}{w^2.}\right)(\Sigma w_{k_i}\tilde{\beta}_i) - \ldots - \left(\frac{w_{h_t}}{w^2.}\right)(\Sigma w_{v_i}\tilde{\beta}_i)$$

$$= \tilde{y}_{i\cdot\cdot} - \Sigma w_{h_j}\tilde{y}_{\cdot j\cdot}/w.$$

$i = 1,2,\ldots,t$, and the solution can be obtained in a similar way.

The sums of squares (SS) are

$$\text{SS (treatments) (uncorrected)} = \sum_k \hat{\tau}_k w_k Y_{\cdot\cdot k} = t\sum_k w_k y_{\cdot\cdot k}^2 \quad \text{with } t \text{ d.f.}$$

$$\text{SS (rows \& cols.)} = \sum_i \tilde{\beta}_i \sum_{jk} w_k y_{ijk} + \sum_j \tilde{\gamma}_j \sum_{ik} w_k y_{ijk} \quad \ldots \quad (21)$$

with $(2t-2)$ d.f.

and

$$\text{SS (Error)} = \sum_{ijk} w_k y_{ijk}^2 - \text{SS(treatments)} - \text{SS(Rows \&}$$

cols.) with $(t-1)(t-2)$ d.f.

Putting $\beta_i = \beta$ for all i and $\gamma_j = 0 = \tau_k$ for all j and k and proceeding in the same way as in section 3.1.1, we get the corrected SS (treatments) to be equal to

$t \sum_k w_k(y_{\cdot\cdot k} - \tilde{y}_{\cdot\cdot\cdot})^2$ with $(t-1)$ d.f. where $\tilde{y}_{\cdot\cdot\cdot} = (\sum_k y_{\cdot\cdot k}\, w_k/w.)$.

To obtain the SS(columns) adjusted for row effects, we put $\gamma_j = 0$ for all j. Then the model reduces to

$y_{ijk} = \beta_i + \tau_k + \varepsilon_{ijk}$ with the suffix j playing no role. This model is the same as that of randomised block designs with unequal group variances. Hence, from section 5.1, we have,

SS (treatments) ignoring $\gamma_j = t \Sigma w_k(y_{..k} - \tilde{y}_{...})^2$

with (t-1)d.f. and

SS (Rows) ignoring $\gamma_j = w. \sum_i (\tilde{y}_{i..} - \tilde{y}_{...})^2$

with (t-1) d.f.

It follows that

SS (Columns) adjusted for rows = (21) - $w. \sum_i (\tilde{y}_{i..} - \tilde{y}_{...})^2$.

with (t-1) d.f. Similarly,

SS (rows) adjusted for columns = (21) - $w. \Sigma (\tilde{y}_{.j.} - \tilde{y}_{...})^2$

with (t-1) d.f.

Analysis of variance table

Source	d.f.	SS	SS	d.f.	Source
Treatments	t-1	$S_1 = t\,\Sigma w_k(y_{\cdot\cdot k}-\tilde{y}_{\cdot\cdot\cdot})^2$	S_1	t-1	Treatments
Row(ignoring cols.)	t-1	$S_2=w.\,\sum\limits_i(\tilde{y}_{i\cdot\cdot}-\tilde{y}_{\cdot\cdot\cdot})^2$	$S_4=w.\,\sum\limits_j(\tilde{y}_{\cdot j\cdot}-\tilde{y}_{\cdot\cdot\cdot})^2$	t-1	Col.(ignoring rows)
Col.(adj. for rows)	t-1	(21) - S_2	(21) - S_4	t-1	Rows(adj. for cols.)
Error	(t-1)(t-2)	S_3 (By subtraction)	S_3	(t-1)(t-2)	Error
Total (corr.)	t^2-1	$\Sigma\, w_k y^2_{ijk} - (\Sigma\, w_k y_{\cdot\cdot k})^2/w.$			

If an F-test or a χ^2-test indicates significant treatment effects, difference between any two treatments can be tested by the normal test as $z = (\hat{\tau}_\ell - \hat{\tau}_k)/(1/tw_\ell + 1/tw_k)^{\frac{1}{2}}$ is a standardised normal variate under the null hypothesis.

6.2 Estimation of weights

Assuming normality of errors, the maximum likelihood estimators of the linear parameters are obtained from the same normal equations as for the weighted least squares (WLS). The estimator of σ_k^2 is then given by

$$\tilde{\sigma}_k^2 = \sum_{ij} (y_{ijk} - \hat{\tau}_k - \tilde{\beta}_i - \tilde{\gamma}_j)^2/t$$

involving the WLS estimators of the linear parameters. An iterative method may be used if convergent. But such estimators are not likely to be of any use for our purpose.

The MINQUE of σ_k^2 is obtained below.

Since the overall design matrix of the model at (19) is singular, let us re-parameterize the treatments and column effects by Helmert's transformation given in section 5.2. Thus

$$\underset{\sim}{\tau}_1 = \underset{\sim}{C}\,\underset{\sim}{\tau} \quad \text{and} \quad \underset{\sim}{\gamma}_1 = \underset{\sim}{C}\,\underset{\sim}{\gamma} ,$$

say, where $\underset{\sim}{C}$ is the matrix of transformation defined in section 5.2. Then the model (19) reduces to

$$E(\underset{\sim}{Y}) = \left(\underset{\sim}{\Delta}'\underset{\sim}{C}_1' \;\vdots\; \underset{\sim}{D}_1' \;\vdots\; \underset{\sim}{D}_2'\underset{\sim}{C}_1' \right) \begin{bmatrix} \underset{\sim}{\tau}_o \\ \cdots \\ \underset{\sim}{\beta} \\ \cdots \\ \underset{\sim}{\gamma}_o \end{bmatrix} = \underset{\sim}{X}'\underset{\sim}{\theta} ,$$

say, in the extended form of the notation of section 5.2. It follows from the same section that

$$XX' = \left[\begin{array}{c:c:c} tI_{t-1} & 0 & 0 \\ \hdashline 0 & tI_t & 0 \\ \hdashline 0 & 0 & tI_{t-1} \end{array}\right] \text{ so that } (XX')^{-1} = \left[\begin{array}{c:c:c} \frac{1}{t}I_{t-1} & 0 & 0 \\ \hdashline 0 & \frac{1}{t}I_t & 0 \\ \hdashline 0 & 0 & \frac{1}{t}I_{t-1} \end{array}\right]$$

with I_t as the identity matrix of order t. Thus, we have

$$X'(XX')^{-1}X = \frac{1}{t}\{\Delta'\Delta - \frac{1}{t}\Delta' J_t \Delta + D_1'D_1 - \frac{1}{t}D_2'J_t D_2 + D_2'D_2\}$$

$$= \frac{1}{t}\begin{bmatrix} J_t & & \\ & \ddots & \\ & & J_t \end{bmatrix} + \frac{1}{t}\begin{bmatrix} I_t & \cdots & I_t \\ \vdots & & \vdots \\ I_t & \cdots & I_t \end{bmatrix} - \frac{2}{t^2}J_{t^2} + \frac{1}{t}\begin{bmatrix} I_t & M_{12} & \cdots & M_{1t} \\ M_{12} & I_t & \cdots & M_{2t} \\ \vdots & \vdots & & \vdots \\ M_{1t} & M_{2t} & \cdots & I_t \end{bmatrix}$$

where the off-diagonal elements M_{ij} in the last term are square symmetric matrices of order t with unity occurring randomly only once in every row and in every column and zero elsewhere on the basis of the random distribution of the column effects as stated in section 6.1. Here J_t is the square matrix of order t with all its elements equal to unity.

It follows that the projection matrix is given by

$$I - X'(XX')^{-1}X = \frac{1}{t^2}\begin{bmatrix} E & G_{12} & \cdots & G_{1t} \\ G_{12} & E & \cdots & G_{2t} \\ \vdots & \vdots & & \vdots \\ G_{1t} & G_{2t} & \cdots & E \end{bmatrix}$$

where E is the square matrix of order t with (t-1)(t-2) as the diagonal elements and (2-t) as the off-diagonal elements, and G_{ij} are the square symmetric matrices with (2-t) as the diagonal elements and 2 as the off-diagonal elements except that (2-t) occurs in place of 2 randomly once in every row and in every column in the positions of unity of the corresponding $\underset{\sim}{M}_{ij}$ matrices. It is easily verified that $(\underset{\sim}{I} - \underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X})\,\underset{\sim}{Y}$ is the vector of residuals.

Now let the matrix $\underset{\sim}{F} = (f_{ij})$ with f_{ij} as the square of the (i,j)th elements of the projection matrix, $\underset{\sim}{\delta} = (\sigma_1^2, \ldots, \sigma_1^2, \ldots, \sigma_t^2, \ldots, \sigma_t^2)'$ be the vector of error variances, each σ_k^2 being repeated t times, and $\underset{\sim}{v}$ the vector of squares of residuals. Then the MINQUE of σ_k^2 is obtained by solving the equation $\underset{\sim}{F}\,\underset{\sim}{\delta} = \underset{\sim}{v}$.

Adding the t equations for σ_k^2 we get

$$\frac{t}{t^4}\Big[\{2(t-2)^2\sigma_1^2+4(t-2)\sigma_1^2\} + \ldots + \{(t-2)^2(t-1)^2\sigma_k^2 +(t-2)^2(t-1)\sigma_k^2\} + \ldots + \{2(t-2)^2\sigma_t^2+4(t-2)\sigma_t^2\}\Big]$$

$$= \sum_{ij}(y_{ijk} - y_{i..} - y_{.j.} - y_{..k} + 2y_{...})^2$$

or,

$$\sigma_1^2 + \sigma_2^2 + \ldots + \frac{(t-2)(t-1)}{2}\sigma_k^2 + \ldots + \sigma_t^2 = \frac{t^2}{2(t-2)}$$

$$\sum_{ij}(y_{ijk} - y_{i..} - y_{.j.} + y_{..k} + 2y_{...})^2 = \frac{t^2 S_k^2}{2(t-2)},$$

say, k = 1,2,...,t. All the t equations can be written together as

$$\left(\frac{t^2-3t}{2}\underset{\sim}{I}_t + \underset{\sim}{J}_t\right)(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_t^2)' = \frac{t^2}{2(t-2)}(S_1^2, S_2^2, \ldots, S_t^2)'.$$

If we write $(\alpha\underset{\sim}{I}_t + \beta\underset{\sim}{J}_t)$ as the inverse of the coefficient matrix, then α and β are given by

$$\alpha = 2/t(t-3) \text{ and } \beta = -4/t^2(t-1)(t-3).$$

The MINQUE of $\hat{\sigma}_k^2$ is then obtained as

$$\hat{\sigma}_k^2 = \{1/(t-1)(t-2)(t-3)\}\{(t^2-t)\sum_{ij}(y_{ijk}-y_{i\cdot\cdot}-y_{\cdot j\cdot}-y_{\cdot\cdot k}+2y_{\cdot\cdot\cdot})^2 - 2\sum_{ijk}(y_{ijk}-y_{i\cdot\cdot}-y_{\cdot j\cdot}-y_{\cdot\cdot k}+2y_{\cdot\cdot\cdot})^2\}$$

As $E(S_k^2) = (1/t)\sigma_k^2(t-2)(t-3) + 2\bar{\sigma}^2(t-2)/t$ so that $E(\sum_1^t S_k^2) = \bar{\sigma}^2(t-1)(t-2)$, it follows that $E(\hat{\sigma}_k^2) = \sigma_k^2$ as is expected. Here $\bar{\sigma}^2 = \Sigma\sigma_k^2/t$.

These estimators are correlated and not in a convenient form for algebraic treatment. We therefore consider the almost unbiased estimators (AUE) proposed by Horn et al. (1975). The AUE of σ_k^2 is given by

$$s_k^2 = (1-h_{kk})^{-1}S_k^2/t$$

$$= (S_k^2/t)\ \{1-(3t-2)/t^2\}^{-1}$$

where $h_{kk} = (3t-2)/t^2$ is the kth diagonal element of $\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X}$.

Now let $u_{ij} = y_{ijk} - y_{i\cdot\cdot} - y_{\cdot j\cdot}$ so that $u_{\cdot\cdot} = y_{\cdot\cdot k} - 2y_{\cdot\cdot\cdot}$. Then the random variables u_{ij} are normally distributed on the assumption of normality of errors. Since the covariance between any two such variates is of order $(1/t^2)$, these variables may be considered to be approximately independent of one another for large t. Again var $(u_{ij}) = \sigma_k^2(1-4/t+2/t^2)+2\bar{\sigma}^2/t$.

If we replace $\bar{\sigma}^2$ by $\sigma_k^{\ 2}$ as an approximation, then var $(u_{ij}) = \sigma_k^{\ 2}(1-2/t)$ on neglecting a term of order $(1/t^2)$. Consequently, the distribution of $S_k^{\ 2} = \sum\limits_{ij} (u_{ij} - u_{..})^2$ may be approximated by that of $\chi^2\sigma_k^{\ 2}(1-2/t)$ with $(t-1)$ d.f.

As $S_k^{\ 2}/\sigma_k^{\ 2}(1-2/t) = t\ s_k^2 \{1-(3t-2)/t^2\} / \sigma_k^{\ 2}(1-2/t)$ $= (t-1)\ s_k^{\ 2}/\sigma_k^{\ 2}$, we may assume that $(t-1)s_k^{\ 2}/\sigma_k^{\ 2}$ is approximately a χ^2-variate with $(t-1)$ d.f.

6.3 Covariance between $s_k^{\ 2}$ and $s_m^{\ 2}$ $(k \neq m)$

We have, $(S_k^{\ 2}) = \sum\limits_{ij} (y_{ijk}-y_{i..}-y_{.j.}-y_{..k}+2y_{...})^2 =$

$$\sum_{ij} \{(\varepsilon_{ijk} - \varepsilon_{..k})-(\varepsilon_{i..}-\varepsilon_{...})-(\varepsilon_{.j.}-\varepsilon_{...})\}^2 \text{ and}$$

$$S_m^2 = \sum_{ij} \{(\varepsilon_{ijm} - \varepsilon_{..m})-(\varepsilon_{i..}-\varepsilon_{...})-(\varepsilon_{.j.}-\varepsilon_{...})\}^2 \text{ so that}$$

$$\frac{1}{t^2} E(S_k^{\ 2}S_m^{\ 2}) = \frac{1}{t^2} E\Big[\Big\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2 + \sum_i(\varepsilon_{i..}-\varepsilon_{...})^2$$

$$+\sum_j(\varepsilon_{.j.}-\varepsilon_{...})^2 - 2\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})(\varepsilon_{i..}-\varepsilon_{...})$$

$$-2\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})(\varepsilon_{.j.}-\varepsilon_{...})$$

$$+2\sum_{ij}(\varepsilon_{i..}-\varepsilon_{...})(\varepsilon_{.j.}-\varepsilon_{...})\Big\}\ \Big\{\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2$$

$$+\sum_u(\varepsilon_{u..}-\varepsilon_{...})^2+\sum_v(\varepsilon_{.v.}-\varepsilon_{...})^2$$

$$-2\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})(\varepsilon_{u..}-\varepsilon_{...})-2\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})$$

$$(\varepsilon_{.v.}-\varepsilon_{...})+2\sum_{uv}(\varepsilon_{u..}-\varepsilon_{...})(\varepsilon_{.v.}-\varepsilon_{...})\}\Big].$$

For derivation of the expectation, a break-up of the individual terms will be useful. This is given below.

(a) $$\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2=\sum_{ij}\varepsilon^2_{ijk}\left(1-\frac{1}{t}\right)-\frac{1}{t}\sum_{ij\neq uv}\sum\varepsilon_{ijk}\,\varepsilon_{uvk}$$

(b) $$\sum_i(\varepsilon_{i..}-\varepsilon_{...})^2=\sum_{ijk}\varepsilon^2_{ijk}\left(\frac{1}{t^2}-\frac{1}{t^3}\right)+\frac{1}{t^2}\sum_i\left(\sum_{jk\neq \ell w}\sum\varepsilon_{ijk}\,\varepsilon_{i\ell w}\right)$$

$$-\frac{1}{t^3}\sum_{ijk\neq r\ell w}\sum\varepsilon_{ijk}\,\varepsilon_{r\ell w}$$

(c) $$\sum_v(\varepsilon_{.v.}-\varepsilon_{...})^2=\sum_{ijm}\varepsilon^2_{ijm}\left(\frac{1}{t^2}-\frac{1}{t^3}\right)+\frac{1}{t^2}\sum_v\left(\sum_{ij\neq ku}\sum\varepsilon_{ivj}\,\varepsilon_{kvu}\right)$$

$$-\frac{1}{t^3}\sum_{ijk\neq uvw}\sum\varepsilon_{ijk}\,\varepsilon_{uvw}$$

(d) $$\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})(\varepsilon_{i..}-\varepsilon_{...})$$

$$=\left(\frac{1}{t}-\frac{1}{t^2}\right)\sum_{ij}\varepsilon^2_{ijk}+\frac{1}{t}\sum_{ij}\varepsilon_{ijk}\sum_{uv\neq jk}\varepsilon_{iuv}-\frac{1}{t^2}\sum_{ij\neq uv}\sum\varepsilon_{ijk}$$

$$\varepsilon_{uvk}-\frac{1}{t^2}\sum_{ij}\varepsilon_{ijk}\sum_{\substack{iju\\ iju\neq ijk}}\varepsilon_{iju}$$

(e) $$\sum_{ij}(\varepsilon_{i..}-\varepsilon_{...})(\varepsilon_{.j.}-\varepsilon_{...})$$

$$=\frac{1}{t^2}\sum_{ij}\varepsilon^2_{ijk}-\frac{1}{t^3}\sum_{ijk}\varepsilon^2_{ijk}+\frac{1}{t^2}\sum_{ij}\left(\sum_{\substack{jk,uw\\ ijk\neq ujw}}\sum\varepsilon_{ijk}\,\varepsilon_{ujw}\right)$$

$$-\frac{1}{t^3}\sum_{ijk\neq uvw}\sum\varepsilon_{ijk}\,\varepsilon_{uvw}$$

The break-up of any other term is equivalent to one of the above.

In all, expectations of 36 terms need to be evaluated. But the expectations of all but the nine terms listed below are of order $(1/t^2)$ and can be neglected for our purpose. The expectations of the nine terms that matter are as follows:

(i) $$E\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2 \sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2\} = (t-1)^2\sigma_k^2\sigma_m^2$$

(ii) $$E\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2 \sum_{u}(\varepsilon_{u..}-\varepsilon_{...})^2\}$$

$$= (1-\tfrac{1}{t})(\tfrac{1}{t^2}-\tfrac{1}{t^3})\{\sum_{ij} E(\varepsilon^4_{ijk}) + E(\sum_{ij}\varepsilon^2_{ijk}\sum_{uv\neq ij}\varepsilon^2_{uvk}) +$$

$$+ E(\sum_{ij}\varepsilon^2_{ijk}\sum_{\substack{iju\\u\neq k}}\varepsilon^2_{iju})\} + \frac{1}{t^4}E(\sum_{ij\neq rs}\sum \varepsilon^2_{ijk}\varepsilon^2_{rsk})$$

$$= (t-1)(t^2+t-1)\sigma_k^4/t^3 + (t-1)^2\sigma_k^2\sum_{u\neq k}\sigma_u^2/t^2$$

(iii) $$E\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2\sum_{v}(\varepsilon_{.v.}-\varepsilon_{...})^2\} = (t-1)(t^2+t-1)$$

$$\sigma_k^2/t^3+(t-1)^2\sigma_k^2\sum_{u\neq k}\sigma_u^2/t^2 \qquad \text{from (ii)}$$

(iv) $$E\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})(\varepsilon_{u..}-\varepsilon_{...})\}$$

$$=(1-\tfrac{1}{t})(\tfrac{1}{t}-\tfrac{1}{t^2})E(\sum_{ij}\varepsilon^2_{ijk}\sum_{uv}\varepsilon^2_{uvm})$$

$$= (t-1)^2\sigma_k^2\sigma_m^2/t$$

(v) $$E\{\sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})^2\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})(\varepsilon_{.v.}-\varepsilon_{...})\}$$

$$=(t-1)^2\sigma_k^2\sigma_m^2/t \qquad \text{from (iv)}$$

(vi) $E\{\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2 \sum_i (\varepsilon_{i..}-\varepsilon_{...})^2\} = (t-1)(t^2+t-1)$

$$\sigma_m^2/t^3 + (t-1)^2\, \sigma_m^2 \sum_{k\neq m} \sigma_k^2/t^2 \quad \text{from (ii)}$$

(vii) $E\{\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2 \sum_j (\varepsilon_{.j.}-\varepsilon_{...})^2\} = (t-1)(t^2+t-1)$

$$\sigma_m^2/t^3 + (t-1)^2\, \sigma_m^2 \sum_{k\neq m} \sigma_k^2/t^2 \quad \text{from (ii)}$$

(viii) $E\{\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2 \sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})(\varepsilon_{i..}-\varepsilon_{...})\}$

$$=(t-1)^2\sigma_k^2\, \sigma_m^2/t \quad \text{from (iv)}$$

(ix) $E\{\sum_{uv}(\varepsilon_{uvm}-\varepsilon_{..m})^2 \sum_{ij}(\varepsilon_{ijk}-\varepsilon_{..k})(\varepsilon_{.j.}-\varepsilon_{...})$

$$=(t-1)^2\sigma_k^2\, \sigma_m^2/t \quad \text{from (iv)}.$$

Utilizing the above expectations and simplifying, we get,

$$\frac{1}{t^2} E(S_k^2 S_m^2) = \frac{1}{t^2}\left[\sigma_k^2\sigma_m^2(t-1)^2\left(1-\frac{8}{t}\right) + \left(\sigma_k^2 \sum_{m\neq k}\sigma_m^2 + \sigma_m^2\sum_{k\neq m}\sigma_k^2\right) 2(t-1)^2/t^2\right]$$

$$+ \text{ terms of order } (1/t^2)$$

$$= \sigma_k^2\sigma_m^2(1-10/t) + 2\bar{\sigma}^2(\sigma_k^2+\sigma_m^2)/t,$$

neglecting terms of order $(1/t^2)$. Again,

$$\frac{1}{t^2} E(S_k^2)\, E(S_m^2) = \{ \sigma_k^2(1-5/t+6/t^2)+(2\bar{\sigma}^2/t)(1-2/t)\} \{ \sigma_m^2 (1-5/t+6/t^2)+(2\bar{\sigma}^2/t)(1-2/t) \}$$

$$= \sigma_k^2\, \sigma_m^2(1-10/t) + 2\bar{\sigma}^2(\sigma_k^2+\sigma_m^2)/t$$

up to the same order of approximation. Hence,

$$\frac{1}{t^2} \operatorname{cov}(S_k^2, S_m^2) = \frac{1}{t^2} \{ E(S_k^2 S_m^2) - E(S_k^2)E(S_m^2) \} = 0$$

and consequently

$$\operatorname{cov}(s_k^2, s_m^2) = 0$$

up to the order $(1/t^2)$.

It follows from above and the previous section that Bartlett's χ^2-test using s_k^2 may be used as an approximate test for equality of group variances.

6.4 Adjustment of the test-statistics

Let $x_k = s_k^2/\sigma_k^2$ where s_k^2 is the AUE of σ_k^2, k= 1,2, ..,t. Then the estimated weights are: $\hat{w}_k = 1/s_k^2 = 1/x_k\,\sigma_k^2$, the number of d.f. is $\nu_k = (t-1)$ and $E(x_k) = 1$ approximately. It follows from the previous section that $\operatorname{cov}(x_k,x_m) = 0$ up to the order $(1/t^2)$ for $k \neq m$. Hence, the adjustment of the statistics can be made with the help of Theorem 1 of section 2.1.4.

(i) Adjustment of F-statistics

The F-statistic using estimated weights for testing equality of treatment effects is given by

$$\hat{F} = \frac{t \; \Sigma \; \hat{w}_k \; (y_{\cdot\cdot k} - \hat{\tilde{y}}_{\cdots})^2/(t-1)}{ESS \; / \; (t-1)(t-2)} = (t-2)TSS/ESS$$

say, with (t-1) and (t-1)(t-2) d.f., where TSS and ESS denote, respectively, the treatments SS and error SS using estimated weights. The treatments sum of squares is in the same form as that for randomised block designs. The adjusted F-statistic is

$$\hat{F}\,(adj) = \hat{F} - \frac{1}{(t-1)} \sum_{k=1}^{t} \left[\frac{\partial^2 F}{\partial x_k^2}\right]_{\text{all } x_k=1} \quad \text{using estimated weights}$$

where $\dfrac{\partial^2 \hat{F}}{\partial x_k^2}$ is given by (17) of section 5.5 with (b-1) replaced by (t-2).

Also from section 5.5, $\left[\dfrac{\partial\,(TSS)}{\partial \; x_k}\right]_{\text{all } x_k=1} = -\; t \; w_k(y_{\cdot\cdot k} - \tilde{y}_{\cdots})^2$ and

$\left[\dfrac{\partial^2 (TSS)}{\partial \; x_k^2}\right]_{\text{all } x_k=1} = 2t \; w_k \; (1 - w_k/w_{\cdot})(y_{\cdot\cdot k} - \tilde{y}_{\cdots})^2$. When the expression for ESS is obtained for any particular experiment, those for

$\left[\dfrac{\partial (ESS)}{\partial \; x_k}\right]_{\text{all } x_k=1}$ and $\left[\dfrac{\partial^2 (ESS)}{\partial \; x_k^2}\right]_{\text{all } x_k=1}$ can be similarly found.

Finally, once the adjusted Rows SS and Columns SS are obtained for an experiment, we can proceed in the same way as above for adjusting the F-statistics in order to test equality of row effects and that of column effects.

(ii) Adjustment of approximate normal test-statistic for testing treatment differences

The approximate normal test-statistic using estimated weights for testing the difference between kth and mth treatments is $\hat{z} = |\, y_{\cdot\cdot k} - y_{\cdot\cdot\cdot m} \,| \,/(1/t\hat{w}_k + 1/t\hat{w}_m)^{\frac{1}{2}}$. This is in the same form as that for the randomised block design. Hence, from section 5.5, we have

$$\hat{z}\,(\text{adj}) = \hat{z}\left[1 - \{\, 3/4(t-1)\,\}\,(1/\hat{w}_k^{\,2} + 1/\hat{w}_m^{\,2})(1/\hat{w}_k + 1/\hat{w}_m)^{-2}\right].$$

6.5 Multiple comparison of treatment parameters

As the error sum of squares depends on weights, the joint confidence interval of all treatment contrast $\psi = \Sigma\, c_k \tau_k$ $(\Sigma\, c_k = 0)$ is given by (18) of section 5.6 with $\hat{\psi} = \Sigma\, c_k\, y_{\cdot\cdot k}$, $S = \left[(t-1)\, F_{\alpha}\{t-1, (t-1)(t-2)\}\right]^{\frac{1}{2}}$ and $s = \left[ESS/(t-1)(t-2)\right]^{\frac{1}{2}}$. The partial derivatives are

$$\left[\frac{\partial s}{\partial x_k}\right]_{\text{all } x_k=1} = \left[\frac{\partial\, ESS}{\partial x_k} \Big/ 2\,\{\, ESS\,(t-1)(t-2)\,\}^{\frac{1}{2}}\right]_{\text{all } x_k=1},$$

$$\left[\frac{\partial^2 s}{\partial x_k^2}\right]_{\text{all } x_k=1} = \left[\{\frac{\partial^2 ESS}{\partial x_k^2} - \left(\frac{\partial\, ESS}{\partial x_k}\right)^2 /2\, s^2(t-1)(t-2)\,\}\,/2s(t-1)(t-2)\right]_{\text{all } x_k=1},$$

$$\left[\frac{\partial\ \hat{\sigma}_{\hat{\psi}}}{\partial\ x_k}\right]_{\text{all } x_k=1} = c_k^{\ 2}\ \sigma_k^{\ 2}/2t(\ \Sigma\, c_k^{\ 2}\ \sigma_k^{\ 2}/t)^{\frac{1}{2}}$$

and

$$\left[\frac{\partial^2\hat{\sigma}\ \hat{\psi}}{\partial\ x_k^{\ 2}}\right]_{\text{all } x_k=1} = -\ c_k^{\ 2}\ \sigma_k^{\ 2}/4b^2(\ \Sigma\, c_k^{\ 2}\ \sigma_k^{\ 2}/t)^{3/2}\ .$$

Here $\hat{\sigma}_{\hat{\psi}} = (\ \Sigma c_k^{\ 2} s_k^{\ 2}/t)^{\frac{1}{2}}$.

Also from section 5.7, the three summary measures of dispersion of the treatment estimators are

$$\text{Estimated AM} = \sum_1^t s_k^{\ 2}/t^2$$

$$\text{Estimated GM (adj)} = (1/t)(\ \pi s_k^{\ 2})^{1/t}\ (1+1/t)$$

and

$$\text{Estimated HM (adj)} = \{\ 1 + 2\ \sum_1^t \hat{f}_k\ (1-\hat{f}_k)/(t-1)\ \}\ /\hat{w}.$$

with $\hat{f}_k = \hat{w}_k/\hat{w}.$.

CHAPTER 7

SPLIT-PLOT DESIGNS

We consider here the usual split-plot designs with error variance heteroscedastic with respect to the levels of the sub-plot treatments. The weighted least squares estimators of the linear parameters are derived and the corresponding analysis is given on the assumption that the group variances are known. Estimators of the group variances having negligible bias, are obtained. The covariance between any two such estimators is found to be negligible. The estimators of the linear parameters and test-statistics using estimated weights, are adjusted for bias. Expressions for joint confidence intervals of contrasts of linear parameters are provided for each factor and interaction separately.

7.1 Estimation and analysis when the error variances are known

Let us consider the following model for split-plot experiments having blocks each of which comprises a replicate of the whole-plot treatments; and whole plots each of which comprises a replicate of the sub-plot treatments:

$$y_{ijk} = \beta_i + \gamma_j + \eta'_{ij} + \tau_k + \delta_{kj} + \varepsilon'_{ijk}$$

$$(i = 1,2,\ldots,b; \quad j = 1,2,\ldots,c; \quad k = 1,2,\ldots,t)$$

where β_i is the effect due to the ith block, γ_j the effect due to the jth whole-plot treatment, η'_{ij} the whole-plot error, τ_k the effect due to the kth sub-plot treatment, δ_{kj} the interaction effect between the jth whole-plot treatment and kth sub-plot treatment and ε'_{ijk} the sub-plot error. The errors are assumed to be all independent of one another. It is also assumed that $E(\eta'_{ij}) = 0 = E(\varepsilon'_{ijk})$, var $(\eta'_{ij}) = \sigma'^2$ and var $(\varepsilon'_{ijk}) = \sigma'^2_k$. Thus the heteroscedasticity of the error variance is assumed to be associated with the levels of the sub-plot treatments.

The above model can also be written as

$$y_{ijk} = \beta_i + \gamma_j + \tau_k + \delta_{kj} + \varepsilon_{ijk} \quad \ldots \ldots (22)$$

where $\varepsilon_{ijk} = \eta'_{ij} + \varepsilon'_{ijk}$ so that var $(\varepsilon_{ijk}) = \sigma'^2 + \sigma'^2_k = \sigma_k^2$, say. Curnow (1957) considered this model with only two sub-plot treatments; he showed how to test for the equality of the two consequent group variances.

Let the constraints on the linear parameters be:

$\sum_k w_k \tau_k = 0 = \sum_j \gamma_j = \sum_j \delta_{kj} = \sum_k w_k \delta_{kj} = \sum\sum w_k \delta_{kj}$ where the weight $w_k = 1/\sigma_k^2$.

Let $\underset{\sim}{Y}$ be the vector of observations arranged systematically such that

$$\underset{\sim}{Y}' = (y_{111}, \ldots, y_{b11}, \ldots, y_{1c1}, \ldots, y_{bc1}, \ldots, y_{11t}, \ldots,$$
$$y_{b1t}, \ldots, y_{1ct}, \ldots, y_{bct}).$$

Then the model (22) can be written in matrix notation as

$$\underset{\sim}{Y} = \underset{\sim}{X}' \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{X}'$ is the overall design matrix, $\underset{\sim}{\beta}$ the corresponding vector of all linear parameters and $\underset{\sim}{\varepsilon}$ the vector of all errors. Thus we have,

$$\text{var}\ (\underset{\sim}{\varepsilon}) = \text{diag}\ (\sigma_1^2,\ldots,\sigma_1^2,\ldots,\sigma_t^2,\ldots,\sigma_t^2) = \underset{\sim}{V}\ ,$$

say, and

$$\underset{\sim}{V}^{-1} = \text{diag}\ (w_1,\ldots,w_1,\ldots,w_t,\ldots,w_t).$$

By (2) of section 1.2, the weighted least squares normal equations are given by (23).

$$
\begin{bmatrix}
bcw_1 & & & bw_1 \cdots bw_1 & bw_1 \cdots bw_1 & \cdots & & cw_1 \cdots cw_1 \\
& \ddots & & \vdots & \vdots & & \ddots & \vdots \\
& & bcw_t & bw_t \cdots bw_t & & & bw_t \cdots bw_t & cw_t \cdots cw_t \\
bw_1 \cdots & & bw_t & bw. & bw_1 & & bw_t & w. \cdots w. \\
\vdots & & \vdots & \ddots & \ddots & \cdots & \ddots & \vdots \\
bw_1 \cdots & & bw_t & bw. & bw_1 & & bw_t & w. \cdots w. \\
bw_1 & & bw_1 & bw_1 & & & & w_1 \cdots w_1 \\
\vdots & & \ddots & & \ddots & & & \vdots \\
bw_1 & & bw_1 & bw_1 & & & & w_1 \cdots w_1 \\
& \ddots & \vdots & & & \ddots & & \vdots \\
& & bw_t & bw_t & & & bw_t & w_t \cdots w_t \\
& & \vdots & \ddots & & & \ddots & \vdots \\
& & bw_t & bw_t & & & bw_t & w_t \cdots w_t \\
cw_1 \cdots cw_t & & & w. \cdots w. & w_1 \cdots w_1 & \cdots & w_t \cdots w_t & cw. \\
\vdots & & & \vdots & \vdots & & \vdots & \ddots \\
cw_1 \cdots cw_t & & & w. \cdots w. & w_1 \cdots w_1 & \cdots & w_t \cdots w_t & cw.
\end{bmatrix}
\begin{bmatrix}
\hat{\tau}_1 \\ \vdots \\ \hat{\tau}_t \\ \tilde{\gamma}_1 \\ \vdots \\ \tilde{\gamma}_c \\ \tilde{\delta}'_{11} \\ \vdots \\ \tilde{\delta}'_{1c} \\ \vdots \\ \tilde{\delta}'_{t1} \\ \vdots \\ \tilde{\delta}'_{tc} \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_b
\end{bmatrix}
=
\begin{bmatrix}
w_1 Y_{..1} \\ \vdots \\ w_t Y_{..t} \\ \Sigma w_k Y_{.1k} \\ \vdots \\ \Sigma\, w_k Y_{.ck} \\ w_1 Y_{.11} \\ \vdots \\ w_1 Y_{.c1} \\ \vdots \\ w_t Y_{.1t} \\ \vdots \\ w_t Y_{.ct} \\ \Sigma w_k Y_{1.k} \\ \vdots \\ \Sigma\, w_k Y_{b.k}
\end{bmatrix}
\quad \ldots\ldots(23)
$$

From this the individual normal equations are obtained as

$$\tau_k: \quad bcw_k\hat{\tau}_k + bw_k \Sigma \tilde{\gamma}_j + b\, w_k \sum_j \tilde{\delta}'_{kj} + cw_k \Sigma\tilde{\beta}_i = w_k y_{\cdot\cdot k};$$

$k = 1,2,\ldots,t$

$$\gamma_j: \quad b\, \Sigma\, w_k \hat{\tau}_k + bw.\, \tilde{\gamma}_j + b \sum_k w_k \tilde{\delta}'_{kj} + w.\Sigma\tilde{\beta}_i = \sum_k w_k Y_{\cdot jk};$$

$j = 1,\ldots,c.$

$$\delta'_{kj}: \quad bw_k\hat{\tau}_k + bw_k \tilde{\gamma}_j + bw_k \tilde{\delta}'_{kj} + w_k \Sigma\tilde{\beta}_i = w_k Y_{\cdot jk}$$

$j = 1,2,\ldots,c$

$k = 1,2,\ldots,t$

$$\beta_i: \quad c\, \Sigma\, w_k \hat{\tau}_k + w.\, \Sigma\tilde{\gamma}_j + \Sigma\Sigma\, w_k \tilde{\delta}'_{kj} + cw.\, \tilde{\beta}_i = \sum_k w_k Y_{i\cdot k};$$

$i = 1,2,\ldots,b$

Using the constraints, $\Sigma\, w_k \hat{\tau}_k = 0 = \Sigma\tilde{\beta}_i = \sum_j \tilde{\gamma}_j = \sum_j \tilde{\delta}'_{kj} = \sum_k \omega_k \tilde{\delta}'_{kj}$ $= \sum_{kj}\Sigma\, w_k \tilde{\delta}'_{kj}$, we get the estimators as

$$\hat{\tau}_k = y_{\cdot\cdot k}, \quad \tilde{\gamma}_j = \sum_k w_k y_{\cdot jk}/w. = \tilde{y}_{\cdot j\cdot}, \quad \tilde{\beta}_i = \sum_k w_k y_{i.k}/w. =$$

$\tilde{y}_{i\cdot\cdot}$ and $\tilde{\delta}'_{kj} = y_{\cdot jk} - y_{\cdot\cdot k} - \tilde{y}_{\cdot j.}$ where $w. = \sum_k w_k$. The corresponding sums of squares are $bc \sum_k y^2_{\cdot\cdot k}$, $bw. \sum_j \tilde{y}^2_{\cdot j\cdot}$, $cw. \sum_i \tilde{y}^2_{i\cdot\cdot}$ and $b \sum_j \sum_k w_k(y_{\cdot jk} - y_{\cdot\cdot k} - \tilde{y}_{\cdot j\cdot})^2$ in that order.

To obtain the corrected sums of squares, let $\beta_i = \beta$ for all i and let us ignore all other main effects and interactions. Then the model reduces to $y_{ijk} = \beta + \varepsilon_{ijk}$. From this, the weighted least squares estimator of β is $\tilde{\beta}$ $= \Sigma w_k y_{\cdot\cdot k}/w. = \tilde{y}_{\cdot\cdot\cdot}$ and the corresponding sum of squares is $bcw.\tilde{y}^2_{\cdot\cdot\cdot}$. Consequently, the corrected sums of squares

(SS) are given by

SS (sub-plot treatments) $= bc \sum_k w_k y^2_{\cdot\cdot k} - bcw.\tilde{y}^2_{\cdots} =$

$bc \sum_k w_k (y_{\cdot\cdot k} - \tilde{y}^2_{\cdots})$ with (t-1) d.f.

SS (whole-plot treatments) $= bw. \ (\tilde{y}_{\cdot j \cdot} - \tilde{y}_{\cdots})^2$

with (c-1) d.f.

SS (interactions) $= b \sum \sum w_k (y_{\cdot jk} - \tilde{y}_{\cdot j \cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots})^2$

with (c-1)(t-1) d.f.

SS (blocks) $= cw. \sum_i \tilde{y}^2_{i\cdot\cdot} - bcw.\tilde{y}^2_{\cdots} = cw. \sum_i (\tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdots})^2$

with (b-1) d.f.

To find the SS for whole-plot error we consider the whole-plot weighted totals $\tilde{Y}_{ij\cdot} = \sum_k w_k y_{ijk} = w.\ \tilde{y}_{ij\cdot}$ where $\tilde{y}_{ij\cdot}$ is the weighted mean for the (i,j)th whole plot, $i = 1,2,\ldots,b$; $j = 1,2,\ldots,c$. These totals have constant variance as shown below. The whole-plot totals may therefore be considered to be the data from a simple randomised block design and so the SS for whole-plot error may be written as $\sum_i \sum_j (\tilde{Y}_{ij\cdot} - \sum_i \tilde{Y}_{ij\cdot}/b - \sum_j \tilde{Y}_{ij\cdot}/c + \sum_i \sum_j \tilde{Y}_{ij\cdot}/bc)^2$

$= w.^2 \sum_i \sum_j (\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdot j \cdot} + \tilde{y}_{\cdots})^2$. However, the whole-plot analysis in the above procedure is in sub-plot units and the whole-plot totals are the weighted totals. Hence SS for whole-plot error is given by

$$SSE_1 = w.^2 \sum \sum (\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdot j\cdot} + \tilde{y}_{\cdots})^2 / \sum w_k$$

$$= w. \sum \sum (\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot} - \tilde{y}_{\cdot j\cdot} + \tilde{y}_{\cdots})^2$$

with $(b-1)(c-1)$ d.f. This is the blocks×whole-plot treatments interaction SS (corrected).

Finally, the sub-plot error SS is obtained as

$$SSE_2 = \underset{\sim}{Y}' V^{-1} \underset{\sim}{Y} - SSE_1 - \text{SS due to all the estimates}$$

$$= \sum_i \sum_j \sum_k w_k (y_{ijk} - \tilde{y}_{ij\cdot} - y_{\cdot jk} + \tilde{y}_{\cdot j\cdot})^2,$$

on simplification, with $c(b-1)(t-1)$ d.f.

It follows that $\tilde{y}_{i\cdot\cdot}$ is unbiased for β_i under the constraints. The estimators of the other main effects are not unbiased but their contrasts are unbiased for the corresponding parametric contrasts. If we define $\tilde{\delta}_{kj} = (y_{\cdot jk} - y_{\cdot\cdot k} - \tilde{y}_{\cdot j\cdot} + \tilde{y}_{\cdot\cdot\cdot})$, then $\tilde{\delta}_{kj}$ is an unbiased estimator of δ_{kj}. The variances of the estimators are:

$$\text{var}(\hat{\tau}_k) = \sigma_k^2/bc, \ \text{var}(\tilde{\beta}_i) = 1/cw_\cdot, \ \text{var}(\gamma_j)$$

$= 1/bw_\cdot$ and $\text{var}(\tilde{\delta}_{kj}) = (1/w_k - 1/w_\cdot)(c-1)/bc$.

The estimators of the levels of each of the three factors are independent of one another. But the interaction estimators $\tilde{\delta}_{kj}$ are mutually correlated.

Expectations of the sums of squares under the constraints are as follows:

(a) Whole-plot analysis

In view of the constraints, the model for the weighted totals of the whole-plots is given by

$$\tilde{Y}_{ij\cdot} = w_\cdot(\beta_i + \gamma_j + \eta'_{ij} + \tilde{\varepsilon}'_{ij\cdot})$$

where $\tilde{Y}_{ij\cdot} = \sum_k w_k y_{ijk}$ and $\tilde{\varepsilon}'_{ij\cdot} = \sum_k \varepsilon'_{ijk} w_k / w_\cdot$.

Dividing both sides by w., we have

$$y_{ij.} = \beta_i + \gamma_j + \eta'_{ij} + \tilde{\varepsilon}'_{ij.} = \beta_i + \gamma_j + \eta_{ij} ,$$

say, where $\tilde{y}_{ij.} = \tilde{Y}_{ij.}/w.$ and $\Sigma \gamma_j = 0$. This is the model of ordinary randomised complete block designs with $\text{var}(\eta_{ij}) = \sigma'^2 + \Sigma \sigma_k'^2 w_k^2/w.^2 = \sigma^2$, say, which is a constant.

It therefore follows that

$$E\ \{SS\ (\text{blocks})\} = w.\ E\{c \sum_i (\tilde{y}_{i..} - \tilde{y}_{...})^2\}$$

$$= w.c \sum_i (\beta_i - \beta.)^2 + (b-1)\sigma^2 w.$$

$$E\ \{SS\ (\text{whole-plot treatments})\} = w.\ E\{b \sum_j (\tilde{y}_{.j.} - \tilde{y}_{...})^2\}$$

$$= w.b \sum_j \gamma_j^2 + (c-1)\sigma^2 w.$$

and

$$E\ (SSE_1) = w.\ E\{\sum_i \sum_j (\tilde{y}_{ij.} - \tilde{y}_{i..} - \tilde{y}_{.j.} + \tilde{y}_{...})^2\}$$

$$= w.\ (b-1)(c-1)\sigma^2 .$$

(b) Sub-plot analysis

From the model (22) we have, under the constraints,

$y_{.jk} = \beta. + \gamma_j + \tau_k + \delta_{kj} + \varepsilon_{.jk}$, $y_{..k} = \beta. + \tau_k + \varepsilon_{..k}$,

$\tilde{y}_{ij.} = \beta_i + \gamma_j + \tilde{\varepsilon}_{ij.}$, $\tilde{y}_{i..} = \beta_i + \varepsilon_{i..}$, $\tilde{y}_{.j.} = \beta. +$

$\gamma_j + \tilde{\varepsilon}_{.j.}$ and

$\tilde{y}_{...} = \beta. + \tilde{\varepsilon}_{...}$ where $\tilde{\varepsilon}_{ij.} = \Sigma w_k \varepsilon_{ijk}/w.$, $\tilde{\varepsilon}_{i..} =$

$\Sigma w_k \varepsilon_{i.k}/w.$, $\tilde{\varepsilon}_{.j.} = \Sigma w_k \varepsilon_{.jk}/w.$ and $\tilde{\varepsilon}_{...} = \Sigma w_k \varepsilon_{..k}/w.$

It then follows that

$$E\text{ (sub-plot treatments SS)} = bc\, E\left\{ \Sigma w_k (\tau_k + \varepsilon_{\cdot\cdot k} - \tilde{\varepsilon}_{\cdots})^2 \right\}$$

$$= bc\, \Sigma\, w_k \tau_k^{\,2} + (t-1)\ ,$$

$$E\text{ (Interaction SS)} = b\, E\left\{ \Sigma\Sigma\, w_k (\delta_{kj} + \varepsilon_{\cdot jk} - \tilde{\varepsilon}_{\cdot j \cdot} - \varepsilon_{\cdot\cdot k} + \tilde{\varepsilon}_{\cdots})^2 \right\}$$

$$= b\, \Sigma\, \Sigma w_k\, \delta_{kj}^{\,2} + (c-1)(t-1)$$

and

$$E\,(SSE_2) = E\left\{ \Sigma\, \Sigma\, \Sigma w_k (\varepsilon_{ijk} - \tilde{\varepsilon}_{ij\cdot} - \varepsilon_{\cdot jk} + \tilde{\varepsilon}_{\cdot j \cdot})^2 \right\}$$

$$= c\,(t-1)(b-1).$$

Analysis of variance table

Source	d.f.	SS	E(MS)
Blocks	b-1	$cw.\sum_{i} (\tilde{y}_{i..} - \tilde{y}_{...})^2$	$w.\sigma^2 + cw.\sum(\beta_i - \beta.)^2/(b-1)$
Whole-plot treatments	c-1	$bw.\sum_{j} (\tilde{y}_{.j.} - \tilde{y}_{...})^2$	$w.\sigma^2 + bw.\sum\gamma_j^2/(c-1)$
Error_1	(b-1)(c-1)	$w.\sum_{i}\sum_{j}(\tilde{y}_{ij.} - \tilde{y}_{i..} - \tilde{y}_{.j.} + \tilde{y}_{...})^2$	$w.\sigma^2$
Sub-plot treatments	t-1	$bc\sum_{k} w_k(y_{..k} - \tilde{y}_{...})^2$	$1 + bc\sum w_k\tau_k^2/(t-1)$
Interaction	(c-1)(t-1)	$b\sum\sum w_k(y_{.jk} - \tilde{y}_{.j.} - y_{..k} + \tilde{y}_{...})^2$	$1 + b\sum\sum w_k\delta_{kj}^2/(c-1)(t-1)$
Error_2	c(b-1)(t-1)	$\sum\sum\sum w_k(y_{ijk} - \tilde{y}_{ij.} - y_{.jk} + \tilde{y}_{.j.})^2$	1
Total (corr.)	bct-1	$\sum\sum\sum w_k y_{ijk}^2 - bcw.\tilde{y}_{...}^2$	

If the F-test or χ^2-test indicates significant main effects and interactions, the difference between any two levels of any one of the factors or between any two interaction parameters can be tested by the normal test. Because, the variates

$$z_1 = (\hat{\tau}_k - \hat{\tau}_\ell)/(1/bcw_k + 1/bcw_\ell)^{\frac{1}{2}},$$

$$z_2 = (\tilde{\gamma}_j - \tilde{\gamma}_h)/(2/bw.)^{\frac{1}{2}}, \quad z_3 = (\tilde{\beta}_i - \tilde{\beta}_m)/(2/cw.)^{\frac{1}{2}}$$

and

$$z_4 = \begin{cases} (\tilde{\delta}_{kj} - \tilde{\delta}_{uj})/\{(c-1)(1/w_k+1/w_u)/bc\}^{\frac{1}{2}} & \text{for } k\neq u \\ (\tilde{\delta}_{kj} - \tilde{\delta}_{kv})/\{2(1/w_k+1/w.)/b\}^{\frac{1}{2}} & \text{for } j\neq v \\ (\tilde{\delta}_{kj} - \tilde{\delta}_{uv})/\{(c-1)(1/w_k+1/w_u)/bc + 2/bw.\}^{\frac{1}{2}} & \\ & \text{for } k\neq u \text{ and } j\neq v \end{cases}$$

are all standardised normal under the null hypotheses.

7.2 Estimation of weights

Since there are no replicated observations in the cells, independent and unbiased estimators of the error variances are not available for the design. But we can obtain approximately independent estimators having negligible bias as follows.

The method of simple least squares yields the estimated error of the usual model as

$$\hat{\varepsilon}_{ijk} = (y_{ijk} - y_{i..} - y_{.jk} + y_{...}).$$

Let $S_k^2 = \sum_i \sum_j (y_{ijk} - y_{i..} - y_{.jk} + y_{...})^2$. Then

$$E(S_k^2) = E\left[\sum_i \sum_j \{(\varepsilon_{ijk} - \varepsilon_{i..}) - (\varepsilon_{.jk} - \varepsilon_{...})\}^2\right]$$

$$= bc \left[\sigma_k^2 \left(1 - \frac{1}{b} - \frac{2}{ct} + \frac{2}{bct}\right) + \frac{\bar{\sigma}^2}{ct}\left(1 - \frac{1}{b}\right) \right]$$

or

$$E(S_k^2/bc) = \sigma_k^2(1 - 1/b - 1/ct + 1/bct),$$

on replacing $\bar{\sigma}^2 = \Sigma\sigma_k^2/t$ by σ_k^2 as an approximation.

Let us now define

$$s_k^2 = (S_k^2/bc)(1-1/b-1/ct+1/bct)^{-1}; \quad k = 1,2,\ldots,t.$$

Then s_k^2 has a negligible bias as an estimator of σ_k^2. The bias is of order $(1/ct - 1/b^2)$. It has been verified that s_k^2 is the almost unbiased estimator (AUE, Horn et al., 1975) of σ_k^2.

To find the approximate distribution of s_k^2, let $u_i = y_{ijk} - y_{i\cdot\cdot}$ so that $u_\cdot = y_{\cdot jk} - y_{\cdot\cdot\cdot}$. Then the random variables u_i, $i = 1,2,\ldots,b$, are independently and normally distributed under the assumption of normality of errors. Moreover,

$$\begin{aligned} \text{var}\,(u_i) &= E\,(\varepsilon_{ijk} - \varepsilon_{i\cdot\cdot})^2 \\ &= \sigma_k^2 + \bar{\sigma}^2/ct - 2\,\sigma_k^2/ct \\ &= \sigma_k^2(1-1/ct) \end{aligned}$$

on replacing $\bar{\sigma}^2$ by σ_k^2 as an approximation as before.

Thus $\sum_{i=1}^{b} (y_{ijk} - y_{i\cdot\cdot} - y_{\cdot jk} + y_{\cdot\cdot\cdot})^2 = \sum_{1}^{b} (u_i - u_\cdot)^2$ is approximately distributed as $\chi^2\,\sigma_k^2(1-1/ct)$ with $(b-1)$ d.f. so that $S_k^2/\sigma_k^2(1-1/ct)$ is approximately distributed as χ^2 with $c(b-1)$ d.f.

Since $S_k^2/\sigma_k^2(1-1/ct) = bc\, s_k^2(1-1/b-1/ct+1/bct)/\sigma_k^2(1-1/ct) = \{c(b-1)s_k^2/\sigma_k^2\}\,\{bc/(bc-c)\}\,(1-1/b-1/c+1/bct)/(1-1/ct) = c(b-1)s_k^2/\sigma_k^2$, we may assume that

$c(b-1)s_k^2/\sigma_k^2$ is approximately a χ^2-variate with $c(b-1)$ d.f.

It is shown in the next section that the covariance between the two estimators, s_k^2 and s_m^2 $(k \neq m)$, is negligible so that, by the normal approximation for large d.f., they are approximately independent.

7.3 Covariance between s_k^2 and s_m^2 $(k \neq m)$

We have, $s_k^2 = \sum_i \sum_j \{(\varepsilon_{ijk} - \varepsilon_{.jk}) - (\varepsilon_{i..} - \varepsilon_{...})\}^2 =$

$\sum_i \sum_j (\varepsilon_{ijk} - \varepsilon_{.jk})^2 + c\sum_i (\varepsilon_{i..} - \varepsilon_{...})^2 - 2c\sum_i (\varepsilon_{i.k} - \varepsilon_{..k})$

$(\varepsilon_{i..} - \varepsilon_{...})$ and $s_m^2 = \sum_u \sum_v (\varepsilon_{uvm} - \varepsilon_{.vm})^2 + c\sum_u (\varepsilon_{u..}$

$- \varepsilon_{...})^2 - 2c\sum_u (\varepsilon_{u.m} - \varepsilon_{..m})(\varepsilon_{u..} - \varepsilon_{...})$.

The individual terms may be partitioned as follows:

(a) $\sum_i \sum_j (\varepsilon_{ijk} - \varepsilon_{.jk})^2$

$$= \sum_i \sum_j \varepsilon_{ijk}^2 \left(1 - \frac{1}{b}\right) - \frac{1}{b} \sum_j \left(\sum_{i} \sum_{\neq \ell} \varepsilon_{ijk}\varepsilon_{\ell jk} \right)$$

(b) Similarly, $\sum_u \sum_v (\varepsilon_{uvm} - \varepsilon_{.vm})^2 = \sum_u \sum_v \varepsilon_{uvm}^2 \left(1 - \frac{1}{b}\right)$

$$- \frac{1}{b} \sum_v \left(\sum_{u} \sum_{\neq r} \varepsilon_{uvm}\varepsilon_{rvm} \right)$$

(c) $c\sum_i (\varepsilon_{i..} - \varepsilon_{...})^2$

$$= \sum_i \sum_j \sum_k \varepsilon_{ijk}^2 \left(\frac{1}{ct^2} - \frac{1}{bct^2}\right) + \frac{1}{ct^2} \sum_i \left(\sum\sum_{(jk) \neq (rs)} \sum\sum \right.$$

$$\left. \varepsilon_{ijk}\varepsilon_{irs} \right) - \frac{1}{bct^2} \sum\sum\sum_{(ijk)\neq(uvm)}\sum\sum\sum \varepsilon_{ijk}\varepsilon_{uvm}$$

(d) Similarly, $c \sum_{u} (\varepsilon_{u..} - \varepsilon_{...})^2$

$$= \sum_{u}\sum_{v}\sum_{m} \varepsilon_{uvm}^2 \left(\frac{1}{ct^2} - \frac{1}{bct^2}\right) + \frac{1}{ct^2} \sum_{u}\Big(\sum\sum_{(vm) \neq (rs)} \varepsilon_{uvm}\,\varepsilon_{urs}\Big) - \frac{1}{bct^2} \sum\sum\sum\sum\sum\sum_{(uvm)\neq(ijk)} \varepsilon_{uvm}\,\varepsilon_{ijk}$$

(e) $\sum_{i}\sum_{j} (\varepsilon_{i.k} - \varepsilon_{..k})(\varepsilon_{i..} - \varepsilon_{...})$

$$= \sum_{i}\sum_{j} \varepsilon_{ijk}^2\left(\frac{1}{ct} - \frac{1}{bct}\right) + \frac{1}{ct} \sum_{i} \Big(\sum_{j} \varepsilon_{ijk} \sum_{v\neq j} \varepsilon_{ivk}\Big)$$

$$- \frac{1}{bct} \sum\sum\sum\sum_{(ij)\neq(uv)} \varepsilon_{ijk}\,\varepsilon_{uvk} - \frac{1}{bct} \sum_{i}\sum_{j} \varepsilon_{ijk} \sum_{i}\sum_{j}\sum_{m\neq k} \varepsilon_{ijm}$$

(f) Similarly, $\sum_{v}\sum_{u} (\varepsilon_{u.m} - \varepsilon_{..m})(\varepsilon_{u..} - \varepsilon_{...})$

$$= \sum_{u}\sum_{v} \varepsilon_{uvm}^2 \left(\frac{1}{ct} - \frac{1}{bct}\right) + \frac{1}{ct}\sum_{u} \Big(\sum_{v} \varepsilon_{uvm} \sum_{j\neq v} \varepsilon_{ujm}\Big)$$

$$- \frac{1}{bct} \sum\sum\sum\sum_{(uv)\neq(ij)} \varepsilon_{uvm}\,\varepsilon_{ujm} - \frac{1}{bct} \sum_{u}\sum_{v} \varepsilon_{uvm}$$

$$\sum_{u}\sum_{v}\sum_{k\neq m} \varepsilon_{uvk} .$$

In all, the expectations of nine terms are to be evaluated. But four of the expectations are negligible up to the order of approximation given below. The other five expectations are as follows

(i) $E \{ \sum_{i}\sum_{j} (\varepsilon_{ijk} - \varepsilon_{.jk})^2 \sum_{u}\sum_{v}(\varepsilon_{uvm} - \varepsilon_{.vm})^2 \}$

$$= c^2(b-1)^2\, \sigma_k^2\, \sigma_m^2$$

(ii) $E \{ \sum_{i}\sum_{j}(\varepsilon_{ijk} - \varepsilon_{.jk})^2 c \sum_{u} (\varepsilon_{u..} - \varepsilon_{...})^2 \}$

$$= (b-1)(b^2c-bc+2b-1)\ \sigma_k^{\ 2}/bt^2+c(b-1)^2\sigma_k^{\ 2} \sum_{m\neq k} \sigma_m^{\ 2}/t^2.$$

(iii) $E\ \{\sum_i \sum_j (\varepsilon_{ijk} - \varepsilon_{\cdot jk})^2 \sum_u \sum_v (\varepsilon_{uvm} - \varepsilon_{\cdot vm})(\varepsilon_{u\cdot\cdot} - \varepsilon_{\cdot\cdot\cdot})\}$

$$= c(b-1)^2 \sigma_k^{\ 2} \sigma_m^{\ 2}/t$$

(iv) $E\{ \sum_u \sum_v (\varepsilon_{uvm} - \varepsilon_{\cdot vm})^2 c \sum_i (\varepsilon_{i\cdot\cdot} - \varepsilon_{\cdot\cdot\cdot})^2\}$

$$= (b-1)(b^2c-bc+2b-1)\ \sigma_m^{\ 2}/bt^2 + c(b-1)^2 \sigma_m^{\ 2} \sum_{k\neq m} \sigma_k^{\ 2}/t^2$$

from (ii)

(v) $E\ \{\sum_u \sum_v (\varepsilon_{uvm} - \varepsilon_{\cdot vm})^2 \sum_i \sum_j (\varepsilon_{ijk} - \varepsilon_{\cdot jk})(\varepsilon_{i\cdot\cdot} - \varepsilon_{\cdot\cdot\cdot})\}$

$$= c(b-1)^2\ \sigma_k^{\ 2}\ \sigma_m^{\ 2}/t \qquad \text{from (iii)}$$

Thus $\frac{1}{b^2c^2}\ E\,(S_k^{\ 2}, S_m^{\ 2}) = \frac{1}{b^2c^2}\ \{\ c^2(b-1)^2 \sigma_k^{\ 2} \sigma_m^{\ 2}$

$$+\ c(b-1)^2\sigma_k^2 \sum_{m\neq k} \sigma_m^2\ /t^2$$

$$-\ 2\ c(b-1)^2\sigma_k^{\ 2}\sigma_m^{\ 2}/t + c(b-1)^2\sigma_m^{\ 2} \sum_{k\neq m} \sigma_k^{\ 2}/t\ -\ 2c(b-1)^2$$

$$\sigma_k^2 \quad \sigma_m^2\ /t^2\ \}$$

+ terms involving reciprocals of cubic expressions in b, c and/or t

$$= \sigma_k^{\ 2}\ \sigma_m^{\ 2}\left(1 + \frac{1}{b^2} - \frac{2}{b} - \frac{4}{ct}\right) + \frac{\bar{\sigma}^2}{ct}\ (\sigma_k^{\ 2} + \sigma_m^{\ 2}),\ \text{neglecting}$$

terms involving reciprocals of cubic expressions in b, c and/ or t.

Also $E\left(\frac{s_k^2}{bc}, \frac{s_m^2}{bc}\right) = \{ \sigma_k^2 (1 - \frac{1}{b} - \frac{2}{ct} + \frac{2}{bct})$

$$+ \frac{\bar{\sigma}^2}{ct}(1 - \frac{1}{b}) \} \{ \sigma_m^2 (1 - \frac{1}{b} - \frac{2}{ct} + \frac{2}{bct}) + \frac{\bar{\sigma}^2}{ct}(1 - \frac{1}{b}) \}$$

$$= \sigma_k^2 \sigma_m^2 (1 + \frac{1}{b^2} - \frac{2}{b} - \frac{4}{ct}) + \frac{\bar{\sigma}^2}{ct} (\sigma_k^2 + \sigma_m^2)$$

up to the same order of approximation. Hence $\mathrm{cov}\left(\frac{s_k^2}{bc}, \frac{s_m^2}{bc}\right) = 0$ and, consequently, $\mathrm{cov}\,(s_k^2, s_m^2) = 0$ to the same order of approximation.

Now let $x_k = s_k^2/\sigma_k^2$, $k = 1,2,\ldots t$. Then the estimated weights are $\hat{w}_k = 1/s_k^2 = 1/x_k \sigma_k^2$, the number of d.f. is $\nu_j = c(b-1)$, and $E(x_k) = 1$ approximately. Let $\hat{w}. = \sum_1^t \hat{w}_k$.

It follows from the above that $\mathrm{cov}\,(x_k, x_m) = 0$ for $k \neq m$ up to the order of reciprocals of cubic expressions in b, c and/or t. Hence, the use of Theorem 2 (section 5.3) for adjustment of the statistics concerned will produce the same results as those by using Theorem 1 due to Meier (section 2.1.4).

7.4 Adjustment of the estimators

To obtain the adjustment of the statistics concerned, we need the following derivatives:

$$\frac{\partial \hat{\tilde{y}}_{ij\cdot}}{\partial x_k} = \frac{\partial(\Sigma\, \hat{w}_k y_{ijk}/\hat{w}.)}{\partial x_k} = -\frac{1}{\sigma_k^2} (\hat{w}.y_{ijk} - \sum_k \hat{w}_k y_{ijk}) / x_k^2 \hat{w}.^2$$

and $\dfrac{\partial^2 \hat{\tilde{y}}_{ij.}}{\partial x_k^2} = -\dfrac{2}{\sigma_k^2 x_k^4 w._4}(\hat{w}.y_{ijk} - \Sigma\,\hat{w}_k y_{ijk})(w_k\hat{w}. - x_k^2\hat{w}.^2)$

so that

$$\left[\frac{\partial\,\hat{\tilde{y}}_{ij.}}{\partial\,x_k}\right]_{\text{all } x_k=1} = -f_k(y_{ijk} - \tilde{y}_{ij}.) \text{ and } \left[\frac{\partial^2\hat{\tilde{y}}_{ij.}}{\partial\,x_k^2}\right]_{\text{all } x_k=1}$$

$$= 2\,f_k\,(1 - f_k)(y_{ijk} - \tilde{y}_{ij.})$$

where $f_k = w_k/w.$ Similarly, $\left[\dfrac{\partial\hat{\tilde{y}}_{i..}}{\partial\,x_k}\right]_{\text{all } x_k=1} = -f_k(y_{i.k}-\tilde{y}...)$,

$$\left[\frac{\partial^2\hat{\tilde{y}}_{i..}}{\partial x_k^2}\right]_{\text{all } x_k=1} = 2\,f_k(1-f_k)(y_{i\cdot k}-\tilde{y}_{i}\cdot\cdot); \left[\frac{\partial\,\hat{\tilde{y}}._{j.}}{\partial\,x_k}\right]_{\text{all } x_k=1}$$

$$= -f_k(y._{jk} - \tilde{y}._{j}.)\ ,$$

$$\left[\frac{\partial^2\hat{\tilde{y}}._{j.}}{\partial\,x_k^2}\right]_{\text{all } x_k=1} = 2\,f_k(1-f_k)(y._{jk}-\tilde{y}._{j.}); \left[\frac{\partial\,y...}{\partial\,x_k}\right]_{\text{all } x_k=1}$$

$$= -f_k(y.._k - \tilde{y}...)$$

and $\left[\dfrac{\partial^2 y...}{\partial x_k^2}\right]_{\text{all } x_k=1} = 2\,f_k(1-f_k)(y.._k-\tilde{y}...)$ where $\hat{\tilde{y}}_i.. = \Sigma\,\hat{w}_k$

$y_{i.k}/\hat{w}.$, $\hat{\tilde{y}}._{j.} = \Sigma\hat{w}_k y._{jk}/\hat{w}.$ and $\hat{\tilde{y}}... = \Sigma\,\hat{w}_k y.._k/\hat{w}.$.

As the estimators of the parameters for the sub-plot treatments do not involve weights, no adjustment is necessary for these. The adjusted forms of the other estimators using estimated weights are

$$\hat{\hat{\beta}}_i\,(adj) = \hat{\tilde{y}}_{i\cdot\cdot} - \frac{2}{c(b-1)} \sum_1^t \hat{f}_k\,(1-\hat{f}_k)(y_{i.k} - \hat{\tilde{y}}_{i\cdot\cdot})\ ,$$

$$\hat{\hat{\gamma}}_j\,(adj) = \hat{\hat{y}}_{\cdot j.} - \frac{2}{c(b-1)} \sum \hat{f}_k\,(1-\hat{f}_k)(y_{\cdot jk} - \hat{\hat{y}}_{\cdot j.})$$

and

$$\hat{\hat{\delta}}_{kj}(adj) = \hat{\hat{\delta}}_{kj} + \frac{2}{c(b-1)} \sum_k \hat{f}_k(1-\hat{f}_k)\hat{\hat{\delta}}_{kj}$$

where $\hat{f}_k = \hat{w}_k/\hat{w}.$ and $\hat{\hat{\delta}}_{kj} = y_{\cdot jk} - \hat{\hat{y}}_{\cdot j.} - y_{\cdot\cdot k} + \hat{\hat{y}}_{\cdot\cdot\cdot}$.

7.5 Adjustment of the test-statistics

(i) Adjustment of the F-statistics

(a) Whole-plot analysis

The F-statistic using estimated weights for testing the significance of whole-plot treatment effects is given by

$$\hat{F}_1 = b(b-1) \sum_j (\hat{\hat{y}}_{\cdot j.} - \hat{\hat{y}}_{\cdot\cdot\cdot})^2 / \sum_i \sum_j (\hat{\hat{y}}_{ij\cdot} - \hat{\hat{y}}_{i\cdot\cdot} - \hat{\hat{y}}_{\cdot j.} + \hat{\hat{y}}_{\cdot\cdot\cdot})^2$$

$= b(b-1)WTSS/WESS$, say. The adjusted F-statistic is given by

$$\hat{F}_1(adj) = \hat{F}_1 - \frac{1}{c(b-1)} \sum_{k=1}^{t} \left[\frac{\partial^2 \hat{F}_1}{\partial x_k^2}\right]_{\text{all } x_k=1} \text{ using estimated weights,}$$

where

$$\left[\frac{\partial^2 \hat{F}_1}{\partial x_k^2}\right]_{\text{all } x_k=1} = \left[\frac{b(b-1)}{(WESS)^3} \{(WESS)^2 \quad \frac{\partial^2(WTSS)}{\partial x_k^2}\right.$$

$$- (WESS)(WTSS) \frac{\partial^2 (WESS)}{\partial x_k^2} - 2 (WESS) \frac{\partial (WESS)}{\partial x_k} \frac{\partial (WTSS)}{\partial x_k}$$

$$\left. + 2(WTSS) \left(\frac{\partial (WESS)}{\partial x_k} \right)^2 \} \right]_{\text{all } x_k=1} \quad \ldots\ldots\ldots \quad (24)$$

the individual derivatives being

$$\left[\frac{\partial (WTSS)}{\partial x_k} \right]_{\text{all } x_k=1} = -2 f_k \sum_j (\tilde{y}_{\cdot j \cdot} - \tilde{y}_{\cdots})(y_{\cdot jk} - \tilde{y}_{\cdot j \cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots}),$$

$$\left[\frac{\partial^2 (WTSS)}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2 \sum_j f_k^2 (y_{\cdot jk} - \tilde{y}_{\cdot j \cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots})^2 + 4 f_k (1-f_k) \sum_j (\tilde{y}_{\cdot j \cdot} - \tilde{y}_{\cdots})(y_{\cdot jk} - \tilde{y}_{\cdot j \cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots}),$$

$$\left[\frac{\partial (WESS)}{\partial x_k} \right]_{\text{all } x_k=1} = 2 f_k \sum_i \sum_j (\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdot j \cdot} + \tilde{y}_{\cdots})^2 - 2 f_k \sum_i \sum_j (y_{ijk} - y_{i\cdot k} - y_{\cdot jk} + y_{\cdot\cdot k})(\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdot j \cdot} + \tilde{y}_{\cdots})$$

and

$$\left[\frac{\partial^2 (WESS)}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2 f_k^2 \sum_i \sum_j (y_{ijk} - y_{i.k} - y_{\cdot jk} + y_{\cdot\cdot k} - \tilde{y}_{ij\cdot} + \tilde{y}_{i\cdot\cdot} + \tilde{y}_{\cdot j \cdot}$$

$$- \tilde{y}_{\cdots})^2 + 4 f_k (1-f_k) \sum_i \sum_j (\tilde{y}_{ij\cdot} - \tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdot j \cdot} + \tilde{y}_{\cdots})(y_{ijk} - y_{i.k} - y_{\cdot jk}$$

$+ y_{\cdot\cdot k} - \tilde{y}_{ij\cdot} + \tilde{y}_{i\cdot\cdot} + \tilde{y}_{\cdot j\cdot} - \tilde{y}_{\cdots})\cdot$

For testing the equality of block effects, the F-statistic using estimated weights, is

$$\hat{F}_2 = c(c-1)\sum_i (\hat{\tilde{y}}_{i\cdot\cdot} - \hat{\tilde{y}}_{\cdots})^2 / \sum_i \sum_j (\hat{\tilde{y}}_{ij\cdot} - \hat{\tilde{y}}_{i\cdot\cdot} - \hat{\tilde{y}}_{\cdot j\cdot} + \hat{\tilde{y}}_{\cdots})^2$$

$$= c(c-1)\ \text{WBSS}/\ \text{WESS},$$

say. The adjusted F-statistic is then obtained as

$$\hat{F}_2\ (\text{adj}) = \hat{F}_2 - \frac{1}{c(b-1)} \sum_1^t \left[\frac{\partial^2 \hat{F}_2}{\partial x_k^2} \right]_{\text{all } x_k=1} \quad \text{using estimated weights,}$$

where $\left[\partial^2 \hat{F}_2 / \partial x_k^2 \right]_{\text{all } x_k=1}$ is given by the right hand side of

(24) above with b(b-1) and WTSS replaced by c(c-1) and WBSS respectively, and with

$$\left[\frac{\partial\ (\text{WBSS})}{\partial x_k} \right]_{\text{all } x_k=1} = -2f_k \sum_i (\tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdots})(y_{i\cdot k} - \tilde{y}_{i\cdot\cdot} + y_{\cdot\cdot k} + \tilde{y}_{\cdots})$$

and

$$\left[\frac{\partial^2 (\text{WBSS})}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2f_k^2 \sum_i (y_{i.k} - \tilde{y}_{i\cdot\cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots})^2 + 4f_k(1-f_k) \sum_i (\tilde{y}_{i\cdot\cdot} - \tilde{y}_{\cdots})(y_{i\cdot k} - \tilde{y}_{i\cdot\cdot} - y_{\cdot\cdot k} + \tilde{y}_{\cdots}),$$

$\left[\partial (\text{WESS}/ \partial x_k\right]_{\text{all } x_k=1}$ and $\left[\partial^2 (\text{WESS}/\partial x_k^2\right]_{\text{all } x_k=1}$ being given above.

(b) Sub-plot analysis

For testing the significance of sub-plot treatment effects, the F-statistic using estimated weights is

$$\hat{F}_3 = c(b-1)\ bc \sum_k \hat{w}_k (y_{\cdot\cdot k} - \hat{\tilde{y}}_{\cdot\cdot\cdot})^2 / \sum_i \sum_j \sum_k w_k (y_{ijk} - \hat{\tilde{y}}_{ij\cdot} - y_{\cdot jk} + \hat{\tilde{y}}_{\cdot j\cdot})^2$$

$$= bc^2(b-1)\ TSS/ESS,$$

say. The adjusted F-statistic is

$$\hat{F}_3(adj) = \hat{F}_3 - \frac{1}{c(b-1)} \sum_1^t \left[\partial^2 \hat{F}_3 / \partial x_k^2 \right]_{\text{all } x_k=1} \text{ using estimated weights,}$$

where $\left[\partial^2 \hat{F}_3 / \partial x_k^2 \right]_{\text{all } x_k=1}$ is given by the right hand side of

(24) with $b(b-1)$, WTSS and WESS replaced by $bc^2(b-1)$, TSS and ESS respectively. The individual derivatives concerned are:

$$\left[\frac{\partial (TSS)}{\partial x_k} \right]_{\text{all } x_k=1} = - w_k (y_{\cdot\cdot k} - \tilde{y}_{\cdot\cdot\cdot})^2, \quad \left[\frac{\partial^2 TSS}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2\ w_k (1-f_k)(y_{\cdot\cdot k} - \tilde{y}_{\cdot\cdot\cdot})^2$$

$$\left[\frac{\partial (ESS)}{\partial x_k} \right]_{\text{all } x_k=1} = -w_k \sum_i \sum_j (y_{ijk} - \tilde{y}_{ij.} - y_{\cdot jk} + \tilde{y}_{\cdot j.})^2$$

$$+ 2\sum\sum\sum w_k f_k (y_{ijk} - \tilde{y}_{ij\cdot} - y_{\cdot jk} + \tilde{y}_{\cdot j\cdot})^2$$

and

$$\left[\frac{\partial^2 ESS}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2w_k(1-2f_k) \sum_i \sum_j (y_{ijk} - \tilde{y}_{ij\cdot} - y_{\cdot jk} + \tilde{y}_{\cdot j\cdot})^2$$

$$- 4w_k f_k (1-3f_k/2) \sum_i \sum_j \sum_k (y_{ijk} - \tilde{y}_{ij.} - y_{\cdot jk} + \tilde{y}_{\cdot j\cdot})^2.$$

For testing the significance of the interaction effects, the F-statistic using estimated weights, is

$$\hat{F}_4 = \frac{bc(b-1)}{c-1} \frac{\sum_j \sum_k \hat{w}_k (y._{jk}-\hat{\tilde{y}}._j.-\hat{\tilde{y}}.._k+\hat{\tilde{y}}...)^2}{\sum\sum\sum \hat{w}_k (y_{ijk}-\hat{\tilde{y}}_{ij}.-y._{jk}+\hat{\tilde{y}}._j.)^2}$$

$$= \frac{bc(b-1)}{c-1} \cdot ISS/ESS,$$

say. The adjusted F-statistic is

$$\hat{F}_4\ (adj) = \hat{F}_4 - \frac{1}{c(b-1)} \sum_{k=1}^{t} \left[\partial^2 \hat{F}_4 / \partial x_k^2 \right]_{\text{all } x_k=1} \text{ using estimated weights,}$$

where $\left[\partial^2 \hat{F}_4 / \partial x_k^2 \right]_{\text{all } x_k=1}$ is given by the right hand side of (24) with b(b-1), WTSS and WESS replaced by bc(b-1)/(c-1), ISS and ESS respectively. The individual derivatives concerned are

$$\left[\frac{\partial ISS}{\partial x_k} \right]_{\text{all } x_k=1} = -w_k \sum_j (y._{jk}-\tilde{y}._j.-y.._k+\tilde{y}...)^2 + 2 \sum_j \sum_k w_k f_k (y._{jk}-\tilde{y}._j.-y.._k+\tilde{y}...)^2,$$

$$\left[\frac{\partial^2 ISS}{\partial x_k^2} \right]_{\text{all } x_k=1} = 2w_k(1-2f_k) \sum_j (y._{jk}-\tilde{y}._j.-y.._k+\tilde{y}...)^2$$

$$- 4 \sum_j \sum_k w_k f_k (1-3f_k/2)(y._{jk}-\tilde{y}._j.-y.._k+\tilde{y}...)^2,$$

and $\left[\partial ESS / \partial x_k \right]_{\text{all } x_k=1}$ and $\left[\partial^2 ESS / \partial x_k^2 \right]_{\text{all } x_k=1}$ are given above.

(ii) Adjustment of the approximate normal test-statistics

For testing the difference between 2 sub-plot treatment effects, the approximate normal test-statistic using estimated weights is in the same form as that for testing the difference between two treatment effects in the one-way model. Hence, from section 2.1.4, the adjusted approximate normal test-statistic is

$$\hat{z}_1(\text{adj}) = \hat{z}_1 \{ 1 - 3(1/\hat{w}_k^2 + 1/\hat{w}_\ell^2)/4c(b-1)(1/\hat{w}_k + 1/\hat{w}_\ell)^2 \}$$

where

$$\hat{z}_1 = \left| \hat{\tau}_k - \hat{\tau}_\ell \right| / (1/bc\hat{w}_k + 1/bcw_\ell)^{\frac{1}{2}}.$$

Also for testing the difference between either two whole-plot treatment effects or two block effects, the approximate normal test-statistic using estimated weights is in the same form as that for testing the difference between two block effects in randomised block designs. Hence, from section 5.5, we have

$$\hat{z}_2(\text{adj}) = \{ \left| \hat{\hat{\gamma}}_j(\text{adj}) - \hat{\hat{\gamma}}_h(\text{adj}) \right| /(2/b\hat{w}.)^{\frac{1}{2}} \} \{ 1 - \Sigma \hat{f}_k (1-\hat{f}_k)/4c(b-1) \}$$

and

$$\hat{z}_3(\text{adj}) = \{ \left| \hat{\hat{\beta}}_i(\text{adj}) - \hat{\hat{\beta}}_m(\text{adj}) \right| /(2/c\hat{w}.)^{\frac{1}{2}} \} \{ 1 - \Sigma \hat{f}_k (1-\hat{f}_k)/4c(b-1) \}$$

where $\hat{\hat{\gamma}}_j(\text{adj})$ and $\hat{\hat{\beta}}_i(\text{adj})$ are as given in the previous section.

Finally for testing the difference between two interaction effects, the approximate normal test-statistic using estimated weights is given, from section 7.1, by

$$\hat{z}_4 = \begin{cases} |\hat{\hat{\delta}}_{kj}(adj) - \hat{\hat{\delta}}_{uj}(adj)| / \{(c-1)(1/\hat{w}_k + 1/\hat{w}_u)/bc\}^{\frac{1}{2}} & \text{for } k \neq u \\ |\hat{\hat{\delta}}_{kj}(adj) - \hat{\hat{\delta}}_{kv}(adj)| / \{2(1/\hat{w}_k + 1/\hat{w}.)/b\}^{\frac{1}{2}} & \text{for } j \neq v \\ |\hat{\hat{\delta}}_{kj}(adj) - \hat{\hat{\delta}}_{uv}(adj)| / \{(c-1)(1/\hat{w}_k + 1/\hat{w}_u)/bc + 2/b\hat{w}.\}^{\frac{1}{2}} & \text{for } k \neq u \text{ and } j \neq v \end{cases}$$

and the adjusted form of this test-statistic by

$$\hat{z}_4(adj) = \begin{cases} \hat{z}_4\{1 - 3/4c(b-1)(1/\hat{w}_k + 1/\hat{w}_u)\} & \text{for } k \neq u \\ \hat{z}_4\{1 - A_k/c(b-1) - \sum_{m \neq k}^{t} B_m/c(b-1)\} & \text{for } j \neq v \\ \hat{z}_4\{1 - \sum_{i \neq k,u} L_i/c(b-1) - \sum_{m \neq k,u}^{t} H_m/c(b-1)\} & \text{for } k \neq u \text{ and } j \neq v \end{cases}$$

where

$$A_k = 3(1+\hat{f}_k^{\,2})^2/4(1+\hat{f}_k)^2 + \hat{f}_k^{\,2}(1-\hat{f}_k)/(1+\hat{f}_k)$$

$$B_m = 3\hat{f}_m^{\,2}\hat{f}_k^{\,2}/4(1+\hat{f}_k)^2 + \hat{f}_m\hat{f}_k(1-\hat{f}_m)/(1+\hat{f}_k)$$

$$L_i = 3P_i/4\hat{w}_i^{\,2}G^2 + 2\hat{f}_i(1-\hat{f}_i)/b\hat{w}.G$$

$$H_m = 3\hat{f}_m^{\,2}/b^2G^2\hat{w}.^2 + 2\hat{f}_m(1-\hat{f}_m)/bG\hat{w}.$$

$$G = (c-1)(1/\hat{w}_k + 1/\hat{w}_u)/bc + 2/b\hat{w}.$$

and

$$P_i = \{(c-1)/bc + 2\hat{f}_i^{\,2}/b\}^2.$$

7.6 Multiple comparison

As the error mean squares for both the whole-plot analysis and sub-plot analysis depend on weights, the formula (18) of section 5.6 is appropriate for finding the joint confidence interval of contrasts of the linear parameters.

(i) Whole-plot treatment contrasts

The joint confidence interval of all contrasts $\psi_1 = \Sigma d_j \gamma_j$ with $\Sigma d_j = 0$ of the whole-plot treatment parameters is estimated by the formula (18) of section 5.6 with $\psi = \psi_1$, $S = [(c-1) F_\alpha\{c-1,(b-1)(c-1)\}]^{\frac{1}{2}}$, $s = \{\hat{w}.(WESS)/(b-1)(c-1)\}^{\frac{1}{2}}$, $\nu_j - 1 = c(b-1)$ and $\hat{\sigma}_{\hat{\psi}_1} = (\Sigma d_j^2/b\hat{w}.)^{\frac{1}{2}}$.

The quantities $\hat{\hat{\gamma}}_j(adj)$ are to be used in computing ψ_1.

The partial derivatives concerned are

$$\left[\frac{\partial \hat{\sigma}_{\hat{\psi}_1}}{\partial x_k}\right]_{\text{all } x_k=1} = \tfrac{1}{2} f_k (\Sigma d_j^2/bw.)^{\frac{1}{2}} ,$$

$$\left[\frac{\partial^2 \hat{\sigma}_{\hat{\psi}_1}}{\partial x_k^2}\right]_{\text{all } x_k=1} = (\Sigma d_j^2)^{\frac{1}{2}} f_k (3f_k/4 - 1)/(bw.)^{\frac{1}{2}} ,$$

$$\left[\frac{\partial s}{\partial x_k}\right]_{\text{all } x_k=1} = \left[s \;\{(\partial \, WESS/ \partial x_k)/ 2 \, WESS - \hat{f}_k/2 \}\right]_{\text{all } x_k=1}$$

and

$$\left[\frac{\partial^2 s}{\partial x_k^2}\right]_{\text{all } x_k=1} = \left[s \{ \hat{f}_k(1-\hat{f}_k)/4) + \frac{\partial^2 WESS}{\partial x_k^2} / 2WESS \right.$$

$$\left. - \left(\frac{\partial \, WESS}{\partial x_k}\right)^2 /4 \; (WESS)^2 \} \right]_{\text{all } x_k= 1} ,$$

$$\left[\frac{\partial \text{ WESS}}{\partial x_k}\right]_{\text{all } x_k=1} \quad \text{and} \quad \left[\frac{\partial^2 \text{WESS}}{\partial x_k^{\;2}}\right]_{\text{all } x_k=1}$$ being given in the

previous section. Here $f_k = w_k/w.$ and $\hat{f}_k = \hat{w}_k/\hat{w}.$.

(ii) <u>β-contrasts</u>

Similarly, the joint confidence interval of all β-contrasts $\psi_2 = \Sigma g_i \beta_i$ with $\Sigma g_i = 0$ is given by (18) of section 5.6 with $\psi = \psi_2$, $S = [(b-1) F_\alpha\{ b-1, (b-1)(c-1)\}]^{\frac{1}{2}}$, $s = \{\hat{w}. (\text{WESS})/(b-1)(c-1)\}^{\frac{1}{2}}$, $r_j - 1 = c(b-1)$ and $\hat{\sigma}_{\hat{\psi}_2} = (\Sigma g_i^{\,2}/c\hat{w}.)^{\frac{1}{2}}$. The quantities $\hat{\hat{\beta}}_i(\text{adj})$ are to be used in computing $\hat{\psi}_2$. The two partial derivatives, $\left[\frac{\partial s}{\partial x_k}\right]_{\text{all } x_k=1}$ and $\left[\frac{\partial^2 s}{\partial x_k^{\;2}}\right]_{\text{all } x_k=1}$

are given above in (i) and the other two derivatives are

$$\left[\frac{\partial \hat{\sigma}_{\hat{\psi}_2}}{\partial x_k}\right]_{\text{all } x_k=1} = (\Sigma\; g_i^{\,2})^{\frac{1}{2}}\, f_k/2(cw.)^{\frac{1}{2}}$$

and

$$\left[\frac{\partial^2 \hat{\sigma}_{\hat{\psi}_2}}{\partial x_k^{\;2}}\right]_{\text{all } x_k=1} = (\Sigma\; g_i^{\,2})^{\frac{1}{2}} f_k(3f_k/4-1)/(cw.)^{\frac{1}{2}} .$$

(iii) Sub-plot treatment contrasts

The joint confidence interval of all contrasts $\psi_3 = \Sigma c_k \tau_k$ of the sub-plot treatment parameters is also given by formula (18) of section 5.6 with $\psi = \psi_3$, $\hat{\psi}_3 = \Sigma c_k y_{\cdot\cdot k}$, $S = [(t-1) F_\alpha \{t-1, c(b-1)(t-1)\}]^{\frac{1}{2}}$, $s = \{ESS/c(b-1)(t-1)\}^{\frac{1}{2}}$, $r_j - 1 = c(b-1)$ and $\hat{\sigma}_{\hat{\psi}_3} = \{\Sigma c_k^2 s_k^2/bc\}^{\frac{1}{2}}$.

The partial derivatives concerned are

$$\left[\frac{\partial \hat{\sigma}_{\hat{\psi}_3}}{\partial x_k}\right]_{\text{all } x_k=1} = c_k^2 \sigma_k^2/2bc(\Sigma c_k^2 \sigma_k^2/bc)^{\frac{1}{2}},$$

$$\left[\frac{\partial^2 \hat{\sigma}_{\hat{\psi}_3}}{\partial x_k^2}\right]_{\text{all } x_k=1} = -c_k^4 \sigma_k^4/4b^2c^2(\Sigma c_k^2 \sigma_k^2/bc)^{3/2},$$

$$\left[\frac{\partial s}{\partial x_k}\right]_{\text{all } x_k=1} = \left[\frac{\partial ESS}{\partial x_k} / 2\{ESS(b-1)(c-1)c\}^{\frac{1}{2}}\right]_{\text{all } x_k=1}$$

and

$$\left[\frac{\partial^2 s}{\partial x_k^2}\right]_{\text{all } x_k=1} = \left[\left\{\frac{\partial^2 ESS}{\partial x_k^2} - \left(\frac{\partial ESS}{\partial x_k}\right)^2/2s^2c(b-1)(t-1)\right\}/2sc(b-1)(t-1)\right]_{\text{all } x_k=1},$$

$$\left[\frac{\partial \, ESS}{\partial x_k}\right]_{\text{all } x_k=1} \quad \text{and} \quad \left[\frac{\partial^2 ESS}{\partial x_k^2}\right]_{\text{all } x_k=1} \quad \text{being given in}$$

the previous section.

(iv) Interaction contrasts

If $\psi_4 = \Sigma\Sigma\, c_{kj}\, \delta_{kj}$ is an interaction contrast,

then $\text{var}(\hat{\psi}_4) = \text{var}\,(\Sigma\,\Sigma\, c_{kj}\, \tilde{\delta}_{kj}) = \Sigma\,\Sigma\, c_{kj}^2\, \text{var}(\tilde{\delta}_{kj})$

$+ \sum_j \sum_{k \neq u} \Sigma\, c_{kj}\, c_{uj}\, \text{cov}\,(\tilde{\delta}_{kj}, \tilde{\delta}_{uj}) + \sum_k \sum_{j \neq v} \Sigma\, c_{kj}\, c_{kv}\, \text{cov}(\tilde{\delta}_{kj}, \tilde{\delta}_{kv})$

$+ \Sigma\, \Sigma \sum_{k \neq u,\, j \neq v} \Sigma\, c_{kj}\, c_{uv}\, \text{cov}\,(\tilde{\delta}_{kj}, \tilde{\delta}_{uv}) = \sum_k G_k/w_k - G/w.,$

say, where $G_k = \sum_j c_{kj}^2\, (c-1)/bc - \sum_{j \neq v} \Sigma\, c_{kj} c_{kv}/bc$ and G

$= \Sigma\, \Sigma\, c_{kj}^2\, (c-1)/bc + \sum_j \sum_{k \neq u} \Sigma\, c_{kj}\, c_{uj}\, (c-1)/bc + (\sum_k \sum_{j \neq v} \Sigma\, c_{kj}\, c_{kv}$

$+ \Sigma\, \Sigma\, \sum_{k \neq u} \sum_{j \neq v} c_{kj}\, c_{uv})(2/b - 1/bc).$

Thus, using estimated weights, the standard error of $\hat{\psi}_4$ is

$$\hat{\sigma}_{\hat{\psi}_4} = (\, \Sigma\, G_k/\hat{w}_k - G/\hat{w}.)^{\frac{1}{2}}.$$

The joint confidence interval of all interaction contrasts ψ_4 is given by (18) of section 5.6 with

$\psi = \psi_4, \quad S = \left[\, (c-1)(t-1)\, F_\alpha\{\, (c-1)(t-1), c(b-1)(t-1)\,\}\,\right]^{\frac{1}{2}},$

$s = \{\, ESS/c(b-1)(t-1)\,\}^{\frac{1}{2}}$, and $r_j - 1 = c(b-1)$. The quantities $\hat{\tilde{\delta}}_{kj}$ (adj) are to be used in computing $\hat{\psi}_4$. The two partial derivatives

$$\left[\partial s/\partial x_k\right]_{\text{all } x_k=1} \quad \text{and} \quad \left[\partial^2 s/\partial x_k^2\right]_{\text{all } x_k=1}$$

are given above in (iii). The other two partial derivatives concerned are

$$\left[\frac{\partial\hat{\sigma}_{\hat{\psi}4}}{\partial x_k}\right]_{\text{all } x_k=1} = (G_k/w_k - Gf_k/w.)/2T$$

and

$$\left[\frac{\partial^2\hat{\sigma}_{\hat{\psi}}}{\partial x_k^2}\right]_{\text{all } x_k=1} = \{Gf_k(1-f_k)/w. - (G_k - Gf_k^2)^2/4Tw_k^2\}/T^{3/2}$$

with $f_k = w_k/w.$ and $T = \Sigma G_k/w_k - G/w.$.

Finally, the three summary measures of dispersion given at the end of the previous chapter can be used as those for the estimators of the sub-plot treatments.

CHAPTER 8

LINEAR REGRESSION WITH UNEQUAL GROUP VARIANCES

A linear regression model with error variance heteroscedastic with respect to the levels of the independent variable is considered here. On the assumption that the group variances are known, the expressions for the weighted least squares estimators of the linear parameters and the corresponding analysis are given. The usual variance of a group of observations is taken as the estimator of the corresponding group variance in the population. The estimators of the linear parameters and test-statistics are then adjusted for bias.

8.1 Estimation and analysis when the error variances are known

Let the simple linear regression model be

$$y_{ij} = \alpha + \beta x_i + \varepsilon_{ij}$$

$$(j = 1,2,\ldots, r_i, r_i > 1; \quad i = 1,2,\ldots,k)$$

where α is the intercept, β the regression coefficient, the values x_i are the fixed values of the independent variable x and ε_{ij} is the error term having mean zero and variance σ_i^2. The errors are assumed to be independent of one another. Let $n = \Sigma r_i$.

By minimising $\sum_i \sum_j (y_{ij} - \alpha - \beta x_i)^2/\sigma_i^2$, we get the weighted least squares (WLS) estimators of the linear parameters as

$$\tilde{\alpha} = (\Sigma w_i y_{i.} \Sigma w_i x_i^2 - \Sigma w_i x_i \Sigma w_i x_i y_{i.})/\{ w.\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2\}$$

and

$$\tilde{\beta} = (w.\ \Sigma w_i x_i y_{i\cdot} - \Sigma w_i x_i\ \Sigma w_i y_{i.})/\{\ w.\Sigma\ w_i x_i^{\ 2} - (\Sigma\ w_i x_i)^2\}$$

where the weight $w_i = r_i/\sigma_i^{\ 2}$, $i = 1,2,\ldots,k$ and $w. = \Sigma\ w_i$. These are also given by Jacquez et al. (1968) for estimated weights. They also empirically compared the efficiency of such estimators with those of ordinary least squares and maximum likelihood estimators. Jacquez and Norusis (1973) empirically compared a few summary dispersion measures of these estimators with those of the least squares estimators.

The sum of squares (SS) due to the estimates is

$$\text{SS (Est.)} = \tilde{\alpha}\ \Sigma\ w_i y_{i\cdot} + \tilde{\beta}\Sigma\ w_i x_i y_{i.}$$

$$= \frac{(\Sigma\ w_i y_{i\cdot})^2}{w.} + \frac{(\Sigma w_i x_i y_{i\cdot} - \Sigma\ w_i x_i\ \Sigma\ w_i y_{i\cdot}/w.)^2}{\Sigma\ w_i x_i^{\ 2} - (\Sigma\ w_i x_i)^2/w.}$$

with 2 d.f. Assuming $\beta = 0$, the model reduces to $y_{ij} = \alpha + \varepsilon_{ij}$. The WLS estimator of α is $\tilde{\tilde{\alpha}} = \Sigma\ w_i y_{i\cdot}/w.$ and the corresponding SS $= (\Sigma\ w_i y_{i\cdot})^2/w.$ with 1 d.f. Subtracting this from SS (Est.) we get the SS for the regression coefficient as

$$\text{SS}(\beta) = (\Sigma\ w_i x_i y_{i\cdot} - \Sigma\ w_i x_i\ \Sigma\ w_i y_{i\cdot}/w.)^2/\{\Sigma\ w_i x_i^{\ 2} - (\Sigma\ w_i x_i)^2/w.\ \}$$

with 1 d.f. The SS due to error is given by

$$\text{SS (E)} = \sum_i \sum_j\ w_i y_{ij}^2/r_i - \text{SS (Est.)}$$

$$= \{\Sigma\ \Sigma\ w_i y_{ij}^2/r_i - (\Sigma\ w_i y_{i\cdot})^2/w.\ \} - \{\ w.\Sigma\ w_i x_i y_{i\cdot} -$$

$(\Sigma w_i x_i)(\Sigma w_i y_{i\cdot})\}^2 / \{w_\cdot^2 \Sigma w_i x_i^2 - (\Sigma w_i x_i)^2 w_\cdot\}$

with (n-2) d.f. As $E\{SS(\beta)\} = \beta^2\{\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w_\cdot\}$, we can test the significance of the regression coefficient by an F-test, that is

$$F = SS(\beta)(n-2) / SS(E)$$

with 1 and n-2 d.f.

Since $E(\tilde{\beta}) = \beta$ and $\operatorname{var}(\tilde{\beta}) = 1/\{\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w_\cdot\}$, the corresponding t-statistic for testing the hypothesis: $\beta = \beta_o$ is given by

$$t = (\tilde{\beta} - \beta_o)\{\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w_\cdot\}^{\frac{1}{2}}(n-2)^{\frac{1}{2}} / \{SS(E)\}^{\frac{1}{2}}$$

with (n-2) d.f.

This latter hypothesis can also be tested with the help of normal test-statistic because the variate $u = (\tilde{\beta} - \beta_o)\{\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w_\cdot\}^{\frac{1}{2}}$ is standardised normal under the null hypothesis.

8.2 Estimators of weights

Rao (1970) gave a set of equations for obtaining the MINQUE of σ_i^2 for this model as an example. Since such estimates may sometimes be negative, Rao and Subrahmaniam (1971) proposed replacement of the MINQUE of σ_i^2 by the corresponding estimate $s_i^2 = \sum_j (y_{ij} - y_{i\cdot})^2/(r_i-1)$ based on the observations of the ith group whenever the MINQUE was less than a small positive quantity. From a Monte Carlo study, they found that for a few replications at many points, the WLS estimators of the linear parameters, using MINQUE (with the above modification), were substantially more efficient than those using s_i^2. However, the gains diminished when many replicates (>8) were taken especially at fewer points.

It follows from Rao and Subrahmaniam (1971) that the almost unbiased estimator (AUE) of σ_i^2 is $\sum_j (y_{ij} - \hat{\alpha} - \hat{\beta} x_i)^2/$ $r_i(1 - k_{ii})$ where $k_{ii} = 1/n + (x_i - x.)^2/\Sigma\, r_i(x_i - x.)^2$ is the ith diagonal element of $\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X}$ with $\underset{\sim}{X}'$ as the design matrix of the regression model and where $\hat{\alpha}$ and $\hat{\beta}$ are the usual least squares estimators of α and β respectively.

The MINQUE of σ_i^2 is too complicated. Even the AUE does not possess the distributional property needed for adjustment of the statistics concerned. We shall therefore use s_i^2 as the estimator of σ_i^2. Jacquez et al. (1968) used this estimator for obtaining the estimated weights.

As is well-known, $(r_i-1)s_i^2/\sigma_i^2$ is distributed as χ^2 with (r_i-1) d.f., and s_i^2 and s_j^2 are independent when $i \neq j$.

8.3 Adjustment of the estimators and test-statistics

Let $z_i = s_i^2/\sigma_i^2$ and the estimated weight $\hat{w}_i = r_i/s_i^2$, $i = 1,2,\ldots,k$. Let $\hat{w}. = \Sigma\, \hat{w}_i$. Since the estimators s_i^2 of the error variances are independent, the adjustment of the statistics concerned for removing the major part of the bias, can be made with the help of the Theorem 1 (section 2.1.4) due to Meier.

(i) Adjustment of the estimators of the linear parameters

The estimated regression coefficient using estimated weights is

$$\hat{\tilde{\beta}} = (\hat{w}_{.}\Sigma\ \hat{w}_i x_i y_{i.} - \Sigma\ \hat{w}_i x_i\ \Sigma\ \hat{w}_i y_{i.})/\ \{\hat{w}_{.}\Sigma\ \hat{w}_i x_i^2 - (\Sigma\ \hat{w}_i x_i)^2\}$$

$$= G/H,$$

say. The adjusted estimator is

$$\hat{\tilde{\beta}}\,(adj) = \hat{\tilde{\beta}} - \sum_{1}^{k} \frac{1}{r_i - 1} \left[\frac{\partial^2 \hat{\tilde{\beta}}}{\partial z_i^2} \right]_{\text{all } z_i = 1}$$ using estimated weights,

where

$$\frac{\partial^2 \hat{\tilde{\beta}}}{\partial z_i^2} = \frac{1}{H^3} \{ H^2 \frac{\partial^2 G}{\partial z_i^2} - HG \frac{\partial^2 H}{\partial z_i^2} - 2H \frac{\partial G}{\partial z_i} \frac{\partial H}{\partial z_i}$$

$$+ 2\,G \left[\frac{\partial H}{\partial z_i} \right]^2 \} . \qquad \ldots\ldots\ldots\ldots\ (25)$$

The individual derivatives are:

$$\left[\frac{\partial G}{\partial z_i} \right]_{\text{all } z_i=1} = -\, w_i(\sum_i w_i x_i y_{i.} + w_{.} x_i y_{i.} - x_i\ \Sigma\, w_i y_{i.} - y_{i.} \Sigma\ w_i x_i),$$

$$\left[\frac{\partial^2 G}{\partial z_i^2} \right]_{\text{all } z_i=1} = w_i\ \{ w_i (x_i - 1) y_{i.} + 2(\Sigma\, w_i x_i y_{i.} + w_{.} x_i y_{i.} - x_i \Sigma\, w_i y_{i.} - y_{i.} \Sigma\, w_i x_i\},$$

$$\left[\frac{\partial H}{\partial z_i} \right]_{\text{all } z_i=1} = -\, w_i (\Sigma\ w_i x_i^2 + w_{.} x_i^2 - 2\,\Sigma\ w_i x_i)$$

and $$\left[\frac{\partial^2 H}{\partial z_i^2} \right]_{\text{all } z_i = 1} = 2\ w_i\ \{ w_i x_i (x_i - 1) + \Sigma\, w_i x_i^2 + w_{.} x_i^2 - 2\,\Sigma\ w_i x_i\ \} .$$

The estimated intercept using estimated weights is

$$\hat{\hat{\alpha}} = (\Sigma \hat{w}_i y_i . \Sigma \hat{w}_i x_i^2 - \Sigma \hat{w}_i x_i \; \Sigma \hat{w}_i x_i y_i .)/\{ \hat{w}. \Sigma \hat{w}_i x_i^2 - (\Sigma \hat{w}_i x_i)^2\}$$

$$= L/H,$$

say. The adjusted estimator is

$$\hat{\hat{\alpha}}\,(adj) = \hat{\hat{\alpha}} - \sum_1^k \frac{1}{r_i - 1} \left[\frac{\partial^2 \hat{\hat{\alpha}}}{\partial z_i^2} \right]_{\text{all } z_i=1} \text{ using estimated weights,}$$

where $\left[\dfrac{\partial^2 \hat{\hat{\alpha}}}{\partial z_i^2} \right]$ is given by the right side of (25) with G

replaced by L. The individual derivatives are:

$$\left[\frac{\partial L}{\partial z_i} \right]_{\text{all } z_i=1} = - w_i(y_i . \Sigma w_i x_i^2 + x_i^2 \Sigma w_i y_i . - x_i \Sigma w_i x_i y_i . - x_i y_i . \Sigma w_i x_i),$$

$$\left[\frac{\partial^2 L}{\partial z_i^2} \right]_{\text{all } z_i=1} = 2 w_i(y_i . \Sigma w_i x_i^2 + x_i^2 \Sigma w_i y_i . - x_i \Sigma w_i x_i y_i . - x_i y_i . \Sigma w_i x_i),$$

$$\left[\frac{\partial H}{\partial z_i} \right]_{\text{all } z_i=1} \text{ and } \left[\frac{\partial^2 H}{\partial z_i^2} \right]_{\text{all } z_i=1} \text{ are given above.}$$

(ii) Adjustment of the F-statistic

For testing the significance of the regression coefficient, the F-statistic using estimated weights, is given by

$$\hat{F} = (n-2)\Big/\left[\{ \hat{w}.\Sigma\,\Sigma\,\frac{\hat{w}_i}{r_i}\,y_{ij}^2 - (\Sigma\,\hat{w}_i y_{i}.)^2\}\{\hat{w}.\,\Sigma\,\hat{w}_i x_i^2 - (\Sigma\hat{w}_i x_i)^2\} / \{\hat{w}.\,\Sigma\hat{w}_i x_i y_i. - \Sigma\,\hat{w}_i x_i\,\Sigma\,\hat{w}_i y_i.\}^2 - 1\right]$$

$$= (n-2)/(T/R - 1)\ ,$$

say. Then the adjusted F-statistic is

$$\hat{F}\,(\mathrm{adj}) = \hat{F} - \sum_1^k \frac{1}{r_i - 1}\left[\frac{\partial^2 \hat{F}}{\partial z_i^2}\right]_{\text{all } z_i=1} \qquad \text{using estimated weights,}$$

where
$$\frac{\partial^2 \hat{F}}{\partial z_i^2} = -\frac{n-2}{R^3(T/R-1)^2}\left[2\left(R\frac{\partial T}{\partial z_i} - T\frac{\partial R}{\partial z_i}\right)^2 \Big/ R(T/R-1) + R^2\frac{\partial^2 T}{\partial z_i^2} - TR\frac{\partial^2 R}{\partial z_i^2} - 2R\frac{\partial T}{\partial z_i}\frac{\partial R}{\partial z_i} + 2T\left(\frac{\partial R}{\partial z_i}\right)^2\right].$$

The individual derivatives concerned are:

$$\left[\frac{\partial T}{\partial z_i}\right]_{\text{all } z_i=1} = -\,w_i\,M, \qquad \left[\frac{\partial R}{\partial z_i}\right]_{\text{all } z_i=1} = -\,2\,w_i\,P,$$

$$\left[\frac{\partial^2 T}{\partial z_i^2}\right]_{\text{all } z_i=1} = w_i\left[2\,w_i\left\{\left(\Sigma\,\Sigma\,\frac{w_i}{r_i}\,y_{ij}^2 + w.\,\sum_j \frac{y_{ij}^2}{r_i} - 2\,y_i.\Sigma\,w_i y_i.\right)\right.\right.$$
$$(\Sigma\,w_i x_i^2 + w.\,x_i^2 - 2\,x_i\,\Sigma w_i x_i) + \{w.\,\Sigma\,w_i x_i^2 - (\Sigma\,w_i x_i)^2\}$$
$$\left.\left.(\sum_j y_{ij}^2/r_i - y_{i.}^2)\right\} + 2M\right]$$

and

$$\left[\frac{\partial^2 R}{\partial z_i^{\,2}}\right]_{\text{all } z_i=1} = 2\, w_i \;\{ w_i\; (\Sigma\, w_i x_i y_{i\cdot} + w_{\cdot} x_i y_{i\cdot} - x_i\, \Sigma w_i y_{i\cdot} -$$

$$y_{i\cdot}\, \Sigma\, w_i x_i)^2 + 2\, P\; \}$$

where $M = \{w_{\cdot} \sum_i \sum_j w_i y_{ij}^{\,2}/r_i - (\Sigma\, w_i y_{i\cdot})^2\}\;\{\Sigma\;\; w_i x_i^{\,2} + w_{\cdot} x_i^{\,2} -$

$$2\, x_i\, \Sigma\, w_i x_i\,\} \;+\; \{w_{\cdot}\, \Sigma\, w_i x_i^{\,2} - (\Sigma\, w_i x_i)^2\}\;\; (\Sigma\, \Sigma\;\; w_i y_{ij}^{\,2}/$$

$$r_i + w_{\cdot} \sum_j y_{ij}^{\,2}/r_i - 2 y_{i\cdot} \Sigma\;\; w_i y_{i\cdot})$$

and

$$P = (w_{\cdot}\, \Sigma\, w_i x_i y_{i\cdot} - \Sigma\, w_i x_i\, \Sigma\, w_i y_{i\cdot})(\Sigma\, w_i x_i y_{i\cdot} + w_{\cdot} x_i y_{i\cdot}$$

$$- x_i \Sigma w_i y_{i\cdot} - y_{i\cdot} \Sigma w_i x_i).$$

(iii) Adjustment of the t-statistic

For testing the hypothesis: $\beta = \beta_o$, the t-statistic using estimated weights is

$$\hat{t} = \frac{|\,\hat{\tilde{\beta}}\,(adj) - \beta_o\,|\; (n-2)^{\frac{1}{2}} \{\, \Sigma\, \hat{w}_i x_i^{\,2} - (\Sigma\, \hat{w}_i x_i)^2/\hat{w}_{\cdot}\,\}^{\frac{1}{2}}}{\left[\{\Sigma\Sigma\, \hat{w}_i y_{ij}^{\,2}/r_i - (\Sigma \hat{w}_i y_{i\cdot})^2/\hat{w}_{\cdot}\,\} - \{\hat{w}_{\cdot}\, \Sigma \hat{w}_i x_i y_{i\cdot} - \Sigma\, \hat{w}_i x_i \Sigma\, \hat{w}_i y_{i\cdot}\}^2/\{\hat{w}_{\cdot}^2\, \Sigma\, w_i x_i^{\,2} - (\Sigma\, \hat{w}_i x_i)^2 \hat{w}_{\cdot}\,\}\right]^{\frac{1}{2}}}$$

$$= (n-2)^{\frac{1}{2}}\, |\; \hat{\tilde{\beta}}\,(adj) - \beta_o |\;\{\; \hat{w}_{\cdot}\, \Sigma\;\hat{w}_i x_i^{\,2} - (\Sigma\;\hat{w}_i x_i)^2\,\}/(T-R)^{\frac{1}{2}}$$

$$= (n-2)^{\frac{1}{2}}\, |\hat{\tilde{\beta}}\;\; (adj) \;-\; \beta_o|\;\, s/(T-R)^{\frac{1}{2}}\;,$$

say. The underlying assumption is that var{ $\hat{\tilde{\beta}}$ (adj)} is approximately equal to var ($\tilde{\beta}'$). The adjusted t-statistic has the form

$$\hat{t}\ (\text{adj}) = \hat{t} - \sum_{1}^{k} \frac{1}{r_i - 1} \left[\frac{\partial^2 \hat{t}}{\partial z_i^2} \right]_{\text{all } z_i = 1} \quad \text{using estimated weights,}$$

where

$$\frac{\partial^2 \hat{t}}{\partial z_i^2} = \hat{t} \left[\frac{\partial^2 s}{\partial z_i^2} - \frac{\partial s}{\partial z_i} \left(\frac{\partial T}{\partial z_i} - \frac{\partial R}{\partial z_i} \right) / (T-R)^2 - \frac{s}{2} \left\{ \frac{\partial^2 T}{\partial z_i^2} \right. \right.$$

$$\left. \left. - \frac{\partial^2 R}{\partial z_i^2} - \frac{3}{2} \left(\frac{\partial T}{\partial z_i} - \frac{\partial R}{\partial z_i} \right)^2 \right\} / (T-R) \right],$$

with $\left[\partial s / \partial z_i \right]_{\text{all } z_i = 1} = - w_i (\Sigma w_i x_i^2 + w.x_i^2 - 2x_i \Sigma w_i x_i)$ and

$$\left[\partial^2 s / \partial z_i^2 \right]_{\text{all } z_i = 1}$$

$= 2 w_i (\Sigma w_i x_i^2 + w.\ x_i^2 - 2\ x_i\ \Sigma\ w_i x_i)$ and other partial derivatives being given in (ii) above.

(iv) Adjustment of the approximate normal test-statistic

The approximate normal test-statistic using estimated weights is

$$\hat{u} = |\hat{\tilde{\beta}}\ (\text{adj}) - \beta_o|\ \{ \Sigma\ \hat{w}_i x_i^2 - (\Sigma\ \hat{w}_i x_i)^2 / \hat{w}. \}^{\frac{1}{2}}$$

and its adjusted form is

$$\hat{u}(\text{adj}) = \hat{u} - \sum_{1}^{k} \frac{1}{r_i - 1} \left[\frac{\partial^2 \hat{u}}{\partial z_i^2} \right]_{\text{all } z_i = 1} \quad \text{using estimated weights,}$$

where $\dfrac{\partial^2 \hat{u}}{\partial z_i^2} = \hat{u} \left[(\hat{w}_i x_i^2 - \hat{f}_i B_i)^2 / 4A + \hat{w}_i \{ x_i^2 - \hat{f}_i x_i^2 - (1 - \hat{f}_i) B_i \} \right] / A$

with $A = \Sigma \hat{w}_i x_i^2 - (\Sigma \hat{w}_i x_i)^2/\hat{w}.$, $B_i = (2x_i - \Sigma \hat{f}_i x_i)(\Sigma \hat{f}_i x_i)$

and $\hat{f}_i = \hat{w}_i/\hat{w}.$.

CHAPTER 9

CONCLUSIONS

In this chapter the main results of the thesis are summarised and areas for further work indicated.

9.1 Summary of the results

The error variance has been assumed to be heteroscedastic with respect to the levels of sub-plot treatments in split-plot designs and the treatments in all other designs. As a result, the treatment estimators as well as the corresponding sum of squares obtained by the weighted least squares method, have the same form for all designs excepting the non-orthogonal general block designs. Orthogonality of different kinds of estimators of the linear parameters is maintained for all designs except general block designs and latin square designs where the estimated row and column effects are not orthogonal to one another. Three summary dispersion measures are suggested for the treatment estimators.

The expression for computing joint confidence intervals of parametric contrasts depends on both weights and error mean squares of the weighted least squares analysis. The adjusted form of this expression for the first three designs is different from that for the remaining three because the error mean squares are independent of weights for the former designs but depends on them for the latter designs.

As the replicated observations are available for at least one cell under each treatment, the MINQUE of group
three
variances for the first two designs and their unbiased

estimators for the third design, are independently distributed as multiples of χ^2. This facilitates adjustment of the estimators of the linear and other parameters and other statistics using estimated weights, for removal of bias. For the other three designs, the AUE's of group variances have negligible bias and are approximately independently distributed as multiples of χ^2 and necessary adjustment of the statistics concerned has therefore been made.

For random models of the first two designs, the test of significance of a variance component is found to be the same as that of significance or equality of the corresponding fixed effects.

For split-plot designs if the weights are large, then the error mean square of the whole plot analysis is expected to be much larger than that of the sub-plot analysis.

The weighted constraints on some linear parameters facilitate certain tests especially for models with an interaction term.

9.2 Discussion and further work

Adjustment of the statistics using estimated weights based on replications is expected to yield better results than that of statistics using other types of weights. It is thus desirable that replicated observations should be taken wherever possible for at least one cell for each group.

The adjustment of most of the statistics using estimated weights has given rise to complicated expressions having limited practical application. Empirical work may reveal that some of the terms of such expressions are negligible in comparison with other terms, and this may lead to simpler expressions.

A Monte Carlo study for one-way heteroscedastic models showed that performances of the adjusted test-statistics are more or less satisfactory. Such study may be undertaken to observe the adequacy of the adjusted statistics of other designs.

Random or mixed models for the first three designs were considered in this thesis. Other types of mixed or random models may be investigated for these and other designs with unequal group variances. Similarly, multiple regression models with unequal group variances may be considered.

Missing-value techniques and covariance analysis have not been discussed in this thesis. These are other topics for which further work could be undertaken.

The problem of finding the optimum number of replications as a balance between cost and adequacy of the adjusted statistics may be investigated for some designs.

Finally, only a special kind of heteroscedasticity of linear models has been dealt with in this thesis for some common designs. Heteroscedasticity in general is yet to be explored.

ACKNOWLEDGEMENTS

I express my thanks and indebtedness to Professor S. C. Pearce and Dr D. A. Preece for advice and encouragement. I am grateful to the Ford Foundation for financial support. My thanks are also due to my wife who patiently reared a large family despite poor health.

REFERENCES

AITKEN, A.C. (1934-35). On least squares and linear combination of observations. Proc. Roy. Soc. Edin. A, 55, 42-48.

ANDERSON, T.W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley and Sons, N.Y.

BARTLETT, M.S. (1937). Properties of sufficiency and statistical tests. Proc. Roy. Soc. A, 160, 268-282.

BARTLETT, M.S. (1947). The use of transformations. Biometrics, 3, 39-52.

BEHRENS, W.U. (1929). Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen. Landw. Jb, 68, 807-837.

BEMENT, T.R. and WILLIAMS, J.S.(1969). Variance of weighted regression estimators when sampling errors are independent and heteroscedastic. J. Am. Stat. Assc., 64, 1369-1382.

BOX, G.E.P. (1954a). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. Ann. Math. Stat., 25, 290-302.

BOX, G.E.P. (1954b). Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two-way classification. Ann. Math. Stat., 25, 485-498.

BOX, G.E.P. and COX, D.R. (1964). An analysis of transformations. J. Roy. Stat. Soc. B, 26, 211-252.

BROWN, M.B. and FORSYTHE, A.B. (1974a). The small sample behaviour of some statistics which test the equality of means. Technometrics, 16, 129-132.

BROWN, M.B. and FORSYTHE, A.B. (1974b). The ANOVA and multiple comparisons for data with heterogeneous variances. Biometrics, 30, 719-724.

CHAKRAVARTI, S. (1965). The problem of testing the hypothesis of equality of means of several univariate normal populations when variances are different. Cal. Stat. Ass. Bull., 14, 127-150.

CURNOW, R.N. (1957). Heterogeneous error variances in split-plot experiments. Biometrika, 44, 378-383.

DRAPER, N.R. and GUTTMAN, I. (1966). Unequal group variances in the fixed-effects one-way analysis of variance - a Bayesian sidelight. Biometrika, 53, 27-35.

DRAPER, N.R. and SMITH, H. (1966). Applied Regression Analysis. Wiley and Sons, N.Y.

DUBY, C., MOUGEY, Y. and ULMO, J. (1975). Analyse de plans D'expériences a deux facteurs controlés quand l'écart type est une fonction connue de la moyenne. Revue de Statistique Appliquée, 23, 5-33.

EHRENBERG, A.S.C. (1950). The unbiased estimation of heterogeneous error variances. Biometrika, 37, 347-357.

FISHER, R.A. (1935). The fiducial argument in statistical inference. Ann. Eugen., 6, 391-398.

FISHER, R.A. (1939). The comparison of samples with possibly unequal variances. Ann. Eugen., 9, 174-180.

FISHER, R.A. (1966). The Design of Experiments. (8th Ed.) Oliver and Boyd, Edinburgh.

GHOSH, M.N. (1961). Linear estimation of parameters in the case of unequal variances. Cal. Stat. Ass. Bulletin, 10, 117-130.

GHOSH, M.N. and BEHARI, V. (1965). Analysis of randomised block experiments when treatments have different errors. Cal. Stat. Ass. Bull., 14, 93-105.

GOLDMAN, A.J. and ZELEN, M. (1964). Weak generalised inverses and minimum variance linear unbiased estimation. J. Res. Nat. Bur. Stand., 68B, 151-172.

GRAYBILL, F.A. (1954). Variance heterogen-eity in a randomised block design. Biometrics, 10, 516-520.

HAN, CHEIEN-PAI (1968). Testing homogen-eity of a set of correlated variances. Biometrika, 55, 317-328.

HAN, CHEIEN - PAI(1969). Testing homogen-eity of variances in a two-way classification. Biometrics, 25, 153-158.

HARTLEY, H.O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. Biometrika, 37, 308-312.

HARTLEY, H.O. and RAO, J.N.K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. Biometrika, 54, 93-108.

HORN, S.D. and HORN, R.A. (1975). Comparison of estimators of heteroscedastic variances in linear models. J. Am. Stat. Assoc., 70, 872-879.

HORN, S.D., HORN, R.A. and DUNCAN, D.B. (1975). Estimating heteroscedastic variances in linear models. J. Am. Stat. Assoc., 70, 380-385.

HOYLE, M.H. (1973). Transformations - An introduction and a bibliography. Int. Stat. Rev., 41, 203-223.

JACQUEZ, J.A., MATHER, F.J. and CRAWFORD, C.R. (1968). Linear regression with non-constant, unknown error variances: Sampling experiments with least squares, weighted least squares and maximum likelihood estimators. Biometrics, 24, 607-626.

JACQUEZ, J.A. and NORUSIS, M. (1973). Sampling experiments on the estimation of parameters in heteroscedastic linear regression. Biometrics, 29, 771-779.

JAMES, G.S. (1951). The comparison of several groups of observations when the ratios of the population variances are unknown. Biometrika, 38, 324-329.

JAMES, G.S. (1954). Tests of linear hypothesis in univariate and multivariate analysis when the ratios of the population variances are unknown. Biometrika, 41, 19-43.

JOHN, P.W.M. (1971). Statistical Design and Analysis of Experiments. Macmillan, N.Y.

JOHNSON, N.L. (1962). Some notes on the investigation of heterogeneity in interactions. Trabajos de Estadistica, 13, 183-199.

MALLELA, P. (1972). Necessary and sufficient conditions for MINQU - estimation of heteroscedastic variances in linear models. J. Am. Stat. Assoc., 67, 486-487.

MALONEY, C.J. and RASTOGI, S.C. (1970). Significance test for Grubb's estimators. Biometrics, 26, 671-676.

MATHEWS, J. and WALKER, R.L. (1965). Mathematical Methods of Physics. Benjamin Inc., N.Y.

MEIER, P. (1953). Variance of a weighted mean. Biometrics, 9, 59-73.

MOORE, E.H. (1920). On the reciprocal of the general algebraic matrix (abstract). Bull. Amer. Math. Soc., 26, 394-395.

MOORE, E.H. (1935). General Analysis. American Philosophical Society, Philadelphia.

NELDER, J.A. (1965). The analysis of randomised experiments with orthogonal block structure: Block structure and the null analysis of variance. Proc. Roy. Soc. A, 283, 147-162.

NELDER, J.A. (1968). The combination of information in generally balanced designs. J. Roy. Stat. Soc. B, 30, 303-311.

PATTERSON, H.D. and THOMPSON, R. (1971). Recovery of inter-block information when block sizes are unequal. Biometrika, 58, 545-554.

PATTERSON, H.D. and THOMPSON, R. (1975). Maximum likelihood estimation of components of variance. Proceedings of the 8th International Biometric Conference, Constanta, Romania. Editura Academiei Republicii Socialiste Romania, Bucharest.

PEARCE, S.C. and JEFFERS, J.N.R. (1971). Block designs and missing data. J. Roy. Stat. Soc. B, 33, 131-136.

PENROSE, R. (1955). A generalised inverse for matrices. Proc. Camb. Phil. Soc., 51, 406-413.

PLACKETT, R.L. (1960). Principles of Regression Analysis. Oxford University Press, London.

RAO, C.R. (1970). Estimation of heteroscedastic variances in linear models. J. Am. Stat. Assoc., 65, 161-172.

RAO, C.R. (1973). Linear Statistical Inference and its Applications (2nd Ed.). John Wiley and Sons, N.Y.

RAO, J.N.K. and SUBRAHMANIAM, K. (1971). Combining independent estimators and estimation in linear regression with unequal variances. Biometrics, 27, 971-990.

ROBINSON, J. and BALAAM, L.N. (1967). Variance heterogeneity and error correlation in factorial experiments. Aust. J. Stat., 9, 126-130.

RUSSELL, T.S. and BRADLEY, R.A. (1958). One-way variances in a two-way classification. Biometrika, 45, 111-129.

SATTERTHWAITE, F.E. (1941). Synthesis of variance. Psychometrika, 6, 309-316.

SCHEFFÉ, H. (1943). On solutions of the Behrens-Fisher problem, based on the t-distribution. Ann. Math. Stat., 14, 35-44.

SCHEFFÉ, H. (1959). The Analysis of Variance. Wiley and Sons, N.Y.

SCHLESSELMAN, J.J. (1973). Data transformation in two-way analysis of variance. J. Am. Stat. Assoc.,68, 369-378.

SHUKLA, G.K. (1972). An invariant test for the homogeneity of variances in a two-way classification. Biometrics, 28, 1063-1072.

SIOTANI, M. (1957). Note on the utilisation of the generalised student ratio in the analysis of variance or dispersion. Ann. Inst. Stat. Math., 9, 157-171.

SNEDECOR, G.W. and COCHRAN, W.G. (1967). Statistical Methods (6th Ed.). Iowa University Press, Iowa.

SPJØTVOLL, E. (1972). Joint confidence intervals for all linear functions of means in the one-way layout with unknown group variances. Biometrika, 59, 683-685.

TOCHER, K.D. (1952). The design and analysis of block experiments. J. Roy. Stat. Soc. B, 14, 45-91.

WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Am. Math. Soc., 54, 426-482.

WELCH, B.L. (1938). The significance of the difference between two means when population variances are unequal. Biometrika, 29, 350-362.

WELCH, B.L. (1947). The generalisation of 'Student's' problem when several different population variances are involved. Biometrika, 34, 28-35.

WELCH, B.L. (1951). On the comparison of several mean values: an alternative approach. Biometrika, 38, 330-336.

WILLIAMS, E.J. (1959). Regression Analysis. Wiley and Sons, N.Y.

WILLIAMS, J.S. (1967). The variance of weighted regression estimators. J. Am. Stat. Assoc., 62, 1290-1301.

YATES, F. and COCHRAN, W.G. (1938). The analysis of groups of experiments. J. Agri. Sc., 28, 556-580.

ZYSKIND, G. and KEMPTHORNE, O. (1960). Treatment errors in comparative experiments. Wright Air Development Tech. Note, 59, 19-45.