ON ESTIMATES OF VARIANCE COMPONENTS

By

ALBERT WILLIAM WORTHAM

Bachelor of Arts
East Texas State Teachers College
Commerce, Texas
1947

Master of Science
Oklahoma Agricultural and Mechanical College
Stillwater, Oklahoma
1949

Submitted to the faculty of the Graduate School of
The Oklahoma Agricultural and Mechanical College
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
May, 1954

ON ESTIMATES OF VARIANCE COMPONENTS

Thesis Approved:

Franklin Graybill
Thesis Adviser

Joe W. Whiteman

O. H. Hamilton

L. Wayne Johnson

E. F. Allen

Carl E. Marshall

Robert MacVicar
Dean of the Graduate School

PREFACE

The ideas for this thesis evolved while I was employed in the Statistical Laboratory of The Oklahoma Agricultural and Mechanical College, and while I was working directly with Dr. F. A. Graybill on numerous projects involving the study of variance components. Undoubtedly the fact that Dr. Graybill has a primary interest in the validity of variance component estimates played a major role in the evolution of these ideas. My interest was further stimulated by the interest of the research staff of the Agricultural Experiment Station and their inquiries concerning the validity of the estimates in the Balanced and Hierarchal models.

The two main questions which arise in the study of variance components are:

(a) How do we estimate the variance components?

(b) What are the characteristics of the estimator?

The first of these has been studied extensively while comparatively little attention has been given the second. This apparent neglect is probably due to the dependence of the characteristics on the particular problem or model and the complexity of the properties of the estimator.

The scope of this study will be, in the main, an investigation of the properties of the analysis of variance estimators of the variance components in the General Balanced Model (Definition 1, pg. 5).

I wish to express my appreciation to Dr. F. A. Graybill of the Statistical Laboratory and to Dr. O. H. Hamilton of the Mathematics Department for their suggestions and helpful criticisms which have undoubtedly improved the quality of this paper. I would further like to express my thanks to Mr. Carl E. Marshall of the Statistical Laboratory for numerous valuable suggestions in the preparation of the manuscript.

TABLE OF CONTENTS

SECTION		PAGE
I.	INTRODUCTION	1
II.	REVIEW OF THE LITERATURE	3
III.	NOTATION AND DEFINITIONS	5
IV.	FUNDAMENTAL LEMMAS	7
V.	QUADRATIC ESTIMATORS OF THE VARIANCE COMPONENTS IN THE BALANCED MODELS	34
VI.	BALANCED MODELS.	38
VII.	BALANCED MODELS WITH NORMALITY ASSUMPTIONS .	52
VIII.	SUMMARY.	57
BIBLIOGRAPHY.		59

I. INTRODUCTION

There are three basic models which are being used in present day statistical methodology. These models are all linear models which differ in their mathematical composition. That is, the properties of the terms vary from model to model.

We will refer to Model I as the linear model with only fixed effects. For example, if we have a set of data classified according to two classes of characteristics A and B and if Y_{ij} denotes an observation in the ith A class and jth B class, then $Y_{ij} = \mu + b_i + t_j + e_{ij}$ $(i=1,2,\ldots n;\ j=1,2,\ldots m)$ will be of the Model I type provided that μ is a constant common to all observations, b_i is a constant common to all observations with a first subscript i, t_j is a constant common to all observations with a second subscript j, and e_{ij} is a random variable with mean zero and variance σ^2. Similarly, Model III will be the class of linear models with only random effect. That is, in the above example, let b_i be a random variable from a distribution with mean zero and variance σ_b^2, let t_j be a random variable with mean zero and variance σ_t^2, and let μ and e_{ij} remain as in Model I. Model II will be the class of models which are combinations of Models I and III. That is, Model II will be the class of linear models with fixed and random effects. For example, let b_i be fixed as in Model I, let t_j be random as in Model III, and let μ and e_{ij} remain as in Model I.

In this paper we will investigate the analysis of variance estimators of the variance components in several models of the Model III class. For example, we will investigate the analysis of variance estimates of

σ_a^2 and σ_b^2 in the above example under this model. In general, we will show that in the balanced models (Definition 1, pg. 5), which includes Randomized Blocks, Latin Squares, Split Plots, Graeco-Latin Squares, General Factorial arrangements, and other common designs, the best (minimum variance) unbiased quadratic estimates of the variance components are given by the analysis of variance procedure. Further it is shown that in the balanced models, if the effects are normally distributed, the estimates given by the analysis of variance procedure are the best (minimum variance) unbiased estimates of the variance components. The analysis of variance procedure of estimating variance components is to equate the observed and expected mean squares and solve the resulting system of equations for the variance components.

II. REVIEW OF THE LITERATURE

The use of variance components as a method of investigating the sources of variation in a measurement was initiated by H. E. Daniels (4) in a paper read before the Industrial and Agricultural Research Section of the Royal Statistical Society, April 29, 1938. In this paper, Daniels was successful in the estimation of the variance components by solving the system of equations formed by equating the observed and expected mean squares. Even though the reasoning was apparently based on intuition rather than mathematical expectation, it is nevertheless correct. The important result of this paper was that the variances of the sources were segregated and subjected to comparative study.

H. E. Daniels' (3) second paper brings forth the basis for variance component estimation as we know it today. Indeed, in this paper Daniels introduces the ideas of random and fixed effects, discriminates between the two with respect to their variance components, and demonstrates the use of mathematical expectation in variance component studies.

Concurrent with these papers, P. L. Hsu (6) presented his paper in the Statistical Research Memoirs. His interests were devoted to the investigation of the validity of the estimates of the error term in a linear model. Hsu was successful in establishing the fact that the least squares estimate was also the best (minimum variance) quadratic estimate. The fact that this estimate was unbiased followed from the Markoff theorem.

From 1938 until S. Lee Crump's (2) paper in 1946, there were no significant developments in the study of variance components although there were numerous applications of the analysis of variance method of estimation. Crump's paper set forth the basic ideas of variance component estimation as a field of study, gave a complete exposition of the method, and cited the existing literature. The only shortcoming of the paper was the omission of a discussion of the validity of the estimates.

The next major contribution to this field of study is perhaps O. Kempthorne's (7) text. This contribution is primarily the defining of the problem and the emphasising of its existance as an unsolved problem.

The most recent contribution in the general problem of investigating the validity of the variance component estimates is contained in the Doctoral Thesis of F. A. Graybill (5). In this paper, the variance component estimates are shown to be the best (minimum variance) unbiased quadratic estimates for the General Nested Model and the General Balanced Cross Classification with normality assumptions. Significant contributions are also presented for other models.

III. NOTATION AND DEFINITIONS

DEFINITION 1. We will define the general balanced model of the Model III type as follows: Let the random variable $Y_{i_1 i_2 \ldots i_n}$ be given as

$$Y_{i_1 i_2 \ldots i_n} = \sum_{k=1}^{n} A_{ki_k} + e_{i_1 i_2 \ldots i_n} + \mu, \tag{3.1}$$

where $i_j = 1,2,\ldots n_j$; $j = 1,2,\ldots n$; μ is a constant; and A_{ki_k} and $e_{i_1 i_2 \ldots i_n}$ are independent random variables with the following properties:

(a) $E(A_{ki_k}) = 0$, where E denotes mathematical expectation, $(k = 1,2,\ldots n)$,

(b) Variance $(A_{ki_k}) = \sigma_k^2$ $(k = 1,2,\ldots n)$,

(c) $E(A_{ki_k}^4) = \mu_{k4} < \infty$, $(k = 1,2,\ldots n)$,

(d) $E(e_{i_1 i_2 \ldots i_n}) = 0$,

(e) Variance $(e_{i_1 i_2 \ldots i_n}) = \sigma^2$, (3.2)

(f) $E(e_{i_1 i_2 \ldots i_n}^4) = \mu_4 < \infty$,

(g) $E(A_{rn_r} A_{pm_p}) = S_{pm}^{rn}$,

(h) $N = \prod_{j=1}^{n} n_j$.

(This will also be termed the Y model or Y system)

DEFINITION 2. We will define an orthogonal transformation U from the Y system to a Z system (where one exists) as follows: The Z system is given by

$$Z_{ij} = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_n} U_{iji_1 i_2 \ldots i_n} Y_{i_1 i_2 \ldots i_n}, \tag{3.3}$$

where $$\sum_{i_1}\sum_{i_2}\cdots\sum_{i_n} U_{iji_1 i_2 \ldots i_n} = \sqrt{N}\, S^{ij}_{01}, \tag{3.4}$$

$$U_{01i_1 i_2 \ldots i_n} = 1/\sqrt{N}, \text{ and} \tag{3.5}$$

$$\sum_{i_1}\sum_{i_2}\cdots\sum_{i_n} U_{iji_1 i_2 \ldots i_n} U_{kbi_1 i_2 \ldots i_n} \begin{array}{ll} = 0 & \text{if } i \neq k \text{ or } j \neq b \\ = 1 & \text{if } i = k \text{ and } j = b. \end{array} \tag{3.6}$$

In order to condense the notation, write

$$e_{i_1 i_2 \ldots i_n} = e_i$$

$$U_{kji_1 \ldots i_n} = U_{kji}$$

$$Y_{i_1 i_2 \ldots i_n} = Y_i$$

DEFINITION 3. A best quadratic unbiased estimate of σ_k^2 is a quadratic form Q_k which satisfies the following:

(a) $E(Q_k) = \sigma_k^2$, i.e. Q_k is unbiased.

(b) Variance of Q_k is less than or equal to the variance of Q_k^*, where Q_k^* is any other quadratic form which satisfies (a).

DEFINITION 4. $\hat{\sigma}_k^2$ will be called the analysis of variance estimate for σ_k^2. $\hat{\sigma}_k^2$ is a quadratic function of the observations.

DEFINITION 5. Kronecker S^{ij}_{km}

$$S^{ij}_{km} \begin{array}{ll} = 0 & \text{if } i \neq k \text{ or } j \neq m \\ = 1 & \text{if } i = k \text{ and } j = m. \end{array}$$

DEFINITION 6. $i,j = p,q$ will mean $i=p$ and $j=q$. $i,j \neq p,q$ will mean the three cases $i \neq p$, $j=q$; $i=p$, $j \neq q$; and $i \neq p$, $j \neq q$.

DEFINITION 7. All ranges on summations will be over the complete range of the indicated subscript unless otherwise specified.

IV. FUNDAMENTAL LEMMAS

In this section we will establish six fundamental Lemmas which will be used in the proof of the main theorem of this thesis in Section V.

LEMMA I. If Z_{ij}, Z_{pq}, and Z_{kb} are elements of the Z system as given by Definition 2, and if they are selected so that $i,j = p,q \neq k,b$ and neither i,j nor k,b equal $0,1$, then

$$E(Z_{ij}^2\ Z_{pq}\ Z_{kb}) = \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4) \right] \sum_m U_{ijm}^3\ U_{kbm}, \quad (4.1)$$

where E denotes mathematical expectation.

PROOF: Consider $E(Z_{ij}^3\ Z_{kb})$.

Replacing Z_{ij} and Z_{kb} by the Y set gives

$$E(Z_{ij}^3\ Z_{kb}) = E\left[\sum_m U_{ijm}\ Y_m \right]^3 \left[\sum_n U_{kbn}\ Y_n \right]. \quad (4.2)$$

Replacing Y_m and Y_n by their values in terms of A_{ki_k}, e_i, and μ gives

$$E(Z_{ij}^3\ Z_{kb}) = E\left[\sum_m U_{ijm} \left(\sum_r A_{rm_r} + e_m + \mu \right) \right]^3 \times$$

$$\left[\sum_n U_{kbn} \left(\sum_p A_{pn_p} + e_n + \mu \right) \right]. \quad (4.3)$$

Expanding 4.3 and using $\sum_m U_{ijm} = \sqrt{N}\ S_{01}^{ij}$, we have

$$E(Z_{ij}^3\ Z_{kb}) = E\left[\sum_m \sum_r U_{ijm}\ A_{rm_r} + \sum_m U_{ijm}\ e_m \right]^3 \times$$

$$\left[\sum_n \sum_p U_{kbn} A_{pn_p} + \sum_n U_{kbn} e_n \right]. \tag{4.4}$$

Expanding 4.4,

$$E(Z_{ij}^3 Z_{kb}) = S_1 + S_2, \tag{4.5}$$

where

$$S_1 = E \left[\sum_m \sum_r U_{ijm} A_{rm_r} + \sum_m U_{ijm} e_m \right]^3 \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \tag{4.6}$$

and

$$S_2 = E \left[\sum_m \sum_r U_{ijm} A_{rm_r} + \sum_m U_{ijm} e_m \right]^3 \left[\sum_n U_{kbn} e_n \right]. \tag{4.7}$$

Expanding the cubic term of 4.6, we have

$$\begin{aligned} S_1 &= E \left[\sum_m \sum_r U_{ijm} A_{rm_r} \right]^3 \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \\ &+ 3E \left[\sum_m \sum_r U_{ijm} A_{rm_r} \right]^2 \left[\sum_s U_{ijs} e_s \right] \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \\ &+ 3E \left[\sum_m \sum_r U_{ijm} A_{rm_r} \right] \left[\sum_s U_{ijs} e_s \right]^2 \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \\ &+ E \left[\sum_m U_{ijm} e_m \right]^3 \left[\sum_n \sum_r U_{kbn} A_{rn_r} \right]. \end{aligned} \tag{4.8}$$

Consider the second term in the expansion 4.8. In view of the independence of A_{rn_r} and e_m and the fact that $E(e_m) = 0$, we have

$$\begin{aligned} &3E \left[\sum_m \sum_r U_{ijm} A_{rm_r} \right]^2 \left[\sum_s U_{ijs} e_s \right] \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \\ &= 3E \left[\sum_m \sum_r U_{ijm} A_{rm_r} \right]^2 \left[\sum_n \sum_p U_{kbn} A_{pn_p} \right] \sum_s U_{ijs} E(e_s) \\ &= 0. \end{aligned} \tag{4.9}$$

Consider the fourth term in the expansion 4.8. In view of the independence of A_{rn_r} and e_m and the fact that $E(A_{rn_r}) = 0$, we have

$$E\left[\sum_m U_{ijm}\, e_m\right]^3 \left[\sum_n \sum_r U_{kbn}\, A_{rn_r}\right]$$

$$= E\left[\sum_m U_{ijm}\, e_m\right]^3 \left[\sum_n \sum_r U_{kbn}\, E(A_{rn_r})\right] = 0. \qquad (4.10)$$

Further, using the independence of e_m and A_{rn_r}, the third term of the expansion 4.8 becomes

$$3E\left[\sum_m \sum_r U_{ijm}\, A_{rm_r}\right]\left[\sum_s U_{ijs}\, e_s\right]^2 \left[\sum_n \sum_p U_{kbn}\, A_{pn_p}\right]$$

$$= 3E\left[\sum_s U_{ijs}\, e_s\right]^2 \sum_m \sum_r U_{ijm}\, U_{kbm}\, E(A_{rm_r})^2$$

$$= 3E\left[\sum_s U_{ijs}\, e_s\right]^2 \sum_r \sigma_r^2 \sum_m U_{ijm}\, U_{kbm} = 0, \text{ since}$$

$$\sum_m U_{ijm}\, U_{kbm} = 0 \text{ for } i,j \neq k,b. \qquad (4.11)$$

Then we have

$$S_1 = E\left[\sum_m \sum_r U_{ijm}\, A_{rm_r}\right]^3 \left[\sum_n \sum_p U_{kbn}\, A_{pn_p}\right]. \qquad (4.12)$$

Expanding the cubic term of 4.12 over r and reducing as above gives

$$S_1 = E\sum_r \left[\sum_m U_{ijm}\, A_{rm_r}\right]^3 \left[\sum_n \sum_p U_{kbn}\, A_{pn_p}\right]. \qquad (4.13)$$

Expanding the cubic term of 4.13, we have

$$S_1 = E\left[\sum_r \sum_m U_{ijm}^3\, A_{rm_r}^3 + 3\sum_r \sum_{\substack{m \\ m\neq s}} \sum_s U_{ijm}^2\, U_{ijs}\, A_{rm_r}^2\, A_{rs_r}\right.$$

$$+ 6 \sum_r \sum_m \sum_s \sum_t U_{ijm}\, U_{ijs}\, U_{ijt}\, A_{rm_r}\, A_{rs_r}\, A_{rt_r} \Big] \times$$
$$m \neq s \neq t$$

$$\left[\sum_p \sum_n U_{kbn}\, A_{pn_p} \right]. \qquad (4.14)$$

Distributing the expected value in 4.14 and using the independence of A_{rn_r}, we have

$$S_1 = \sum_r \sum_m U^3_{ijm}\, U_{kbm}\, E(A^4_{rm_r}) + 3 \sum_r \sum_{\substack{m \\ m\neq s}} \sum_s U^2_{ijm}\, U_{ijs}\, U_{kbs} \times$$

$$E(A^2_{rm_r})\, E(A^2_{rs_r}). \qquad (4.15)$$

Using $E(A^4_{rm_r}) = \mu_{r4}$ and $E(A^2_{rm_r}) = \sigma^2_r$, we have

$$S_1 = \sum_r \sum_m U^3_{ijm}\, U_{kbm}\, \mu_{r4} + 3 \sum_r \sum_{\substack{m \\ m\neq s}} \sum_s U^2_{ijm}\, U_{ijs}\, U_{kbs}\, \sigma^4_r$$

or $$S_1 = \sum_r \mu_{r4} \sum_m U^3_{ijm}\, U_{kbm} + 3 \sum_r \sigma^4_r \sum_{\substack{m \\ m\neq s}} \sum_s U^2_{ijm}\, U_{ijs}\, U_{kbs}. \qquad (4.16)$$

Adding and subtracting

$$3 \sum_r \sigma^4_r \sum_m U^3_{ijm}\, U_{kbm},$$

we have

$$S_1 = \sum_r (\mu_{r4} - 3\sigma^4_r) \sum_m U^3_{ijm}\, U_{kbm} + 3 \sum_r \sigma^4_r \sum_m \sum_s U^2_{ijm}\, U_{ijs}\, U_{kbs}. \qquad (4.17)$$

Consider $J = \sum_m \sum_s U^2_{ijm}\, U_{ijs}\, U_{kbs}$. (4.18)

Then $J = 0$, (4.19)

since $\sum_{s} U_{ijs} U_{kbs} = 0.$

Hence

$$S_1 = \sum_{r} (\mu_{r4} - 3\sigma_r^4) \sum_{m} U_{ijm}^3 U_{kbm}. \tag{4.20}$$

Consider now 4.7. Expanding the cubic term, we have

$$\begin{aligned} S_2 = {} & E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^3 \left[\sum_{n} U_{kbn} e_n\right] \\ & + 3E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \left[\sum_{s} U_{ijs} e_s\right]\left[\sum_{n} U_{kbn} e_n\right] \\ & + 3E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]\left[\sum_{s} U_{ijs} e_s\right]^2 \left[\sum_{n} U_{kbn} e_n\right] \\ & + E\left[\sum_{m} U_{ijm} e_m\right]^3 \left[\sum_{n} U_{kbn} e_n\right]. \end{aligned} \tag{4.21}$$

Consider the first term of the expansion 4.21. Since A_{rm_r} and e_n are independent and since $E(e_n) = 0$, we have

$$\begin{aligned} & E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^3 \left[\sum_{n} U_{kbn} e_n\right] \\ & = E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^3 \sum_{n} U_{kbn} E(e_n) = 0. \end{aligned} \tag{4.22}$$

Consider the second term of the expansion 4.21. Since A_{rm_r} and e_r are independent, and since $E(e_m^2) = \sigma^2$ and $\sum_{m} U_{ijm} U_{kbn} = 0$, we have

$$\begin{aligned} & 3E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \left[\sum_{s} U_{ijs} e_s\right]\left[\sum_{n} U_{kbn} e_n\right] \\ & = 3E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \sum_{s} U_{ijs} U_{kbs} E(e_s^2) \end{aligned}$$

$$= 3E\left[\sum_m \sum_r U_{ijm} A_{rm_r}\right]^2 \sigma^2 \sum_s U_{ijs} U_{kbs}$$

$$= 0. \qquad (4.23)$$

Consider the third term of the expansion 4.21. Since A_{rm_r} and e_r are independent, and since $E(A_{rm_r}) = 0$, we have

$$3E\left[\sum_m \sum_r U_{ijm} A_{rm_r}\right]\left[\sum_s U_{ijs} e_s\right]^2\left[\sum_n U_{kbn} e_n\right]$$

$$= 3 \sum_m \sum_r U_{ijm} E(A_{rm_r}) E\left[\sum_s U_{ijs} e_s\right]^2\left[\sum_n U_{kbn} e_n\right]$$

$$= 0. \qquad (4.24)$$

Hence

$$S_2 = E\left[\sum_m U_{ijm} e_m\right]^3\left[\sum_n U_{kbn} e_n\right]. \qquad (4.25)$$

Expanding the cubic term of 4.25, we have

$$S_2 = E\left[\sum_m U_{ijm}^3 e_m^3 + 3 \sum_{\substack{m \\ m\neq s}} \sum_s U_{ijm}^2 U_{ijs} e_m^2 e_s\right.$$

$$\left. + 6 \sum_{\substack{m \\ m\neq s\neq t}} \sum_s \sum_t U_{ijm} U_{ijs} U_{ijt} e_m e_s e_t\right]\left[\sum_n U_{kbn} e_n\right]. \qquad (4.26)$$

Using the independence of e_m and $E(e_m) = 0$, we have

$$S_2 = \sum_m U_{ijm}^3 U_{kbm} E(e_m^4) + 3 \sum_{\substack{m \\ m\neq s}} \sum_s U_{ijm}^2 U_{ijs} U_{kbs} E(e_m^2) E(e_s^2). \qquad (4.27)$$

Taking the expected values,

$$S_2 = \mu_4 \sum_m U_{ijm}^3 U_{kbm} + 3\sigma^4 \sum_{\substack{m \\ m\neq s}} \sum_s U_{ijm}^2 U_{ijs} U_{kbs}. \qquad (4.28)$$

Adding and subtracting $3\sigma^4 \sum_m U_{ijm}^3 U_{kbm}$, we have

$$S_2 = (\mu_4 - 3\sigma^4) \sum_m U_{ijm}^3 U_{kbm} + 3\sigma^4 \sum_m \sum_s U_{ijm}^2 U_{ijs} U_{kbs}. \qquad (4.29)$$

But $\sum_s U_{ijs} U_{kbs} = 0$, since $i,j \neq k,b$.

Hence

$$S_2 = (\mu_4 - 3\sigma^4) \sum_m U_{ijm}^3 U_{kbm}. \qquad (4.30)$$

Therefore

$$E(Z_{ij}^3 Z_{kb}) = S_1 + S_2$$

$$= \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_m U_{ijm}^3 U_{kbm}. \qquad (4.31)$$

LEMMA II. If Z_{ij}, Z_{pq}, Z_{kb} are elements of the Z system, as given by Definition 2, and if they are selected so that $i,j \neq p,q \neq k,b$, and neither i,j; p,q; nor k,b equal $0,1$; then

$$E(Z_{ij}^2 Z_{pq} Z_{kb}) = \left[\sum_m (\mu_{m4} - 3\sigma_m^4) + (\mu_4 - 3\sigma^4)\right] \times$$

$$\left[\sum_m U_{ijm}^2 U_{pqm} U_{kbm}\right]. \qquad (4.32)$$

PROOF: Consider $E(Z_{ij}^2 Z_{pq} Z_{kb})$. Replacing Z_{ij}, Z_{pq}, and Z_{kb} by the Y set gives

$$E(Z_{ij}^2 Z_{pq} Z_{kb}) = E\left[\left(\sum_m U_{ijm} Y_m\right)^2 \left(\sum_n U_{pqn} Y_n\right) \left(\sum_s U_{kbs} Y_s\right)\right], \qquad (4.33)$$

and replacing Y_m, Y_n, and Y_s by their values in terms of A_{ki_k}, e_i, and μ, we have

$$E(Z_{ij}^2 Z_{pq} Z_{kb}) = E\left[\sum_m U_{ijm} \left(\sum_r A_{rm_r} + e_m + \mu\right)\right]^2 \times$$

$$\left[\sum_n U_{pqn}\left(\sum_t A_{tn_t} + e_n + \mu\right)\right]\left[\sum_s U_{kbs}\left(\sum_v A_{vs_v} + e_s + \mu\right)\right]. \tag{4.34}$$

Using $\sum_r U_{ijr} = \sqrt{N}\, S_{01}^{ij}$, we have

$$E(Z_{ij}^2\, Z_{pq}\, Z_{kb}) = E\left[\sum_m \sum_r U_{ijm} A_{rm_r} + \sum_m U_{ijm} e_m\right]^2 x$$

$$\left[\sum_n \sum_t U_{pqn} A_{tn_t} + \sum_n U_{pqn} e_n\right]\left[\sum_s \sum_v U_{kbs} A_{vs_v} + \sum_s U_{kbs} e_s\right]. \tag{4.35}$$

Squaring, expanding, and using the independence of A_{rs_r} and e_s in 4.35, we have

$$E(Z_{ij}^2\, Z_{pq}\, Z_{kb}) = L_1 + L_2 + 2L_3 + 2L_4 + L_5 + L_6, \tag{4.36}$$

where

$$L_1 = E\left[\sum_m \sum_r U_{ijm} A_{rm_r}\right]^2 \left[\sum_n \sum_t U_{pqn} A_{tn_t}\right] x$$

$$\left[\sum_v \sum_s U_{kbs} A_{vs_v}\right], \tag{4.37}$$

$$L_2 = E\left[\sum_m \sum_r U_{ijm} A_{rm_r}\right]^2 \left[\sum_n U_{pqn} e_n\right]\left[\sum_s U_{kbs} e_s\right], \tag{4.38}$$

$$L_3 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_t U_{ijt} e_t\right]\left[\sum_v \sum_n U_{pqn} A_{vn_v}\right] x$$

$$\left[\sum_s U_{kbs} e_s\right], \tag{4.39}$$

$$L_4 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_t U_{ijt} e_t\right]\left[\sum_v \sum_s U_{kbs} A_{vs_v}\right] x$$

$$\left[\sum_{n} U_{pqn}\, e_n\right], \tag{4.40}$$

$$L_5 = E\left[\sum_{m} U_{ijm}\, e_m\right]^2 \left[\sum_{n}\sum_{r} U_{pqn}\, A_{rn_r}\right]\left[\sum_{v}\sum_{s} U_{kbs}\, A_{vs_v}\right], \tag{4.41}$$

and

$$L_6 = E\left[\sum_{m} U_{ijm}\, e_m\right]^2 \left[\sum_{n} U_{pqn}\, e_n\right]\left[\sum_{s} U_{kbs}\, e_s\right]. \tag{4.42}$$

Consider equation 4.37. Expanding, we have

$$L_1 = E\left[\sum_{r}\sum_{m} U^2_{ijm}\, A^2_{rm_r}\right]\left[\sum_{n}\sum_{v} U_{pqn}\, A_{vn_v}\right]\left[\sum_{t}\sum_{s} U_{kbs}\, A_{ts_t}\right]$$

$$+ E\left[\sum_{r}\sum_{m}\sum_{p}\sum_{t}\limits_{r,m\neq p,t} U_{ijm}\, U_{ijt}\, A_{rm_r}\, A_{pt_p}\right] x$$

$$\left[\sum_{n}\sum_{v} U_{pqn}\, A_{vn_v}\right]\left[\sum_{t}\sum_{s} U_{kbs}\, A_{ts_t}\right]. \tag{4.43}$$

Using the independence of A_{rm_r} and $E(A_{rm_r}) = 0$, we may write L_1 in the form

$$L_1 = E\left[\sum_{r}\sum_{m}\sum_{s}\sum_{t}\limits_{r,m\neq s,t} U^2_{ijm}\, U_{kbt}\, U_{pqt}\, A^2_{rm_r}\, A^2_{st_s}\right]$$

$$+ E\left[\sum_{m}\sum_{r} U^2_{ijm}\, U_{kbm}\, U_{pqm}\, A^4_{rm_r}\right]$$

$$+ E\left[\sum_{r}\sum_{m}\sum_{s}\sum_{t}\limits_{r,m\neq s,t} U_{ijm}\, U_{ijt}\, U_{kbt}\, U_{pqm}\, A^2_{rm_r}\, A^2_{st_s}\right]$$

$$+ E\left[\sum_{r}\sum_{m}\sum_{s}\sum_{t}\limits_{r,m\neq s,t} U_{ijm}\, U_{ijt}\, U_{kbm}\, U_{pqt}\, A^2_{rm_r}\, A^2_{st_s}\right]. \tag{4.44}$$

Forming the expected values, we have

$$L_1 = \sum_{\substack{r \\ r,m \neq s,t}} \sum_{m} \sum_{s} \sum_{t} U^2_{ijm} U_{kbt} U_{pqt} \sigma^2_r \sigma^2_s$$

$$+ \sum_{m} \sum_{r} U^2_{ijm} U_{kbm} U_{pqm} \mu_{r4}$$

$$+ \sum_{\substack{r \\ r,m \neq s,t}} \sum_{m} \sum_{s} \sum_{t} U_{ijm} U_{ijt} U_{kbt} U_{pqm} \sigma^2_r \sigma^2_s$$

$$+ \sum_{\substack{r \\ r,m \neq s,t}} \sum_{m} \sum_{s} \sum_{t} U_{ijm} U_{ijt} U_{kbm} U_{pqt} \sigma^2_r \sigma^2_s. \tag{4.45}$$

Adding and subtracting

$$3 \sum_{s} \sum_{r} U^2_{ijs} U_{kbs} U_{pqs} \sigma^4_r, \tag{4.46}$$

gives

$$L_1 = \sum_{r} (\mu_{r4} - 3\sigma^4_r) \sum_{s} U^2_{ijs} U_{kbs} U_{pqs}$$

$$+ \sum_{r} \sum_{m} \sum_{s} \sum_{t} U^2_{ijm} U_{kbt} U_{pqt} \sigma^2_r \sigma^2_s$$

$$+ \sum_{r} \sum_{m} \sum_{s} \sum_{t} U_{ijm} U_{ijt} U_{kbt} U_{pqm} \sigma^2_r \sigma^2_s$$

$$+ \sum_{r} \sum_{m} \sum_{s} \sum_{t} U_{ijm} U_{ijt} U_{kbm} U_{pqt} \sigma^2_r \sigma^2_s . \tag{4.47}$$

Now the last three terms of L_1 vanish if we sum on t since

$$\sum_{t} U_{ijt} U_{pqt} = \sum_{t} U_{kbt} U_{pqt} = \sum_{t} U_{ijt} U_{kbt} = 0$$

Hence

$$L_1 = \sum_r (\mu_{r4} - 3\sigma_r^4) \sum_s U_{ijs}^2 \, U_{kbs} \, U_{pqs} \,. \qquad (4.48)$$

Consider equation 4.38. In view of the independence of e_n and A_{rm_r}, we have

$$L_2 = E\left[\sum_m \sum_r U_{ijm} \, A_{rm_r}\right]^2 \; E\left[\sum_n U_{pqn} \, e_n \sum_s U_{kbs} \, e_s\right] . \qquad (4.49)$$

In view of the independence of e_n, we have on expanding the last two products of 4.49

$$L_2 = E\left[\sum_m \sum_r U_{ijm} \, A_{rm_r}\right]^2 \sum_n U_{pqn} \, U_{kbn} \, E(e_n)^2 . \qquad (4.50)$$

Taking expected values on the last sum of 4.50, we have

$$L_2 = E\left[\sum_m \sum_r U_{ijm} \, A_{rm_r}\right]^2 \; \sigma^2 \sum_n U_{pqn} \, U_{kbn} = 0, \qquad (4.51)$$

since $\sum_n U_{pqn} \, U_{kbn} = 0$.

Consider equation 4.39. In view of the independence of A_{rm_r} and e_m, we have

$$L_3 = E\left[\sum_r \sum_m U_{ijm} \, A_{rm_r}\right]\left[\sum_v \sum_n U_{pqn} \, A_{vn_v}\right] E\left[\sum_t U_{ijt} \, e_t\right] \times$$

$$\left[\sum_s U_{kbs} \, e_s\right] . \qquad (4.52)$$

Expanding the last two products in 4.52 and using the independence of e_m,

$$L_3 = E\left[\sum_r \sum_m U_{ijm} \, A_{rm_r}\right]\left[\sum_v \sum_n U_{pqn} \, A_{vn_v}\right] \sum_t U_{ijt} \, U_{kbt} \, E(e_t)^2 . \qquad (4.53)$$

Taking expected values on the last sum in 4.53, we have

$$L_3 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_v \sum_n U_{pqn} A_{vn_v}\right] \sigma^2 \sum_t U_{ijt} U_{kbt}$$

$$= 0, \qquad (4.54)$$

since $\sum_m U_{ijm} U_{kbm} = 0$.

Consider equation 4.40. In view of the independence of A_{rm_r} and e_m,

$$L_4 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_v \sum_s U_{bks} A_{vs_v}\right] x$$

$$E\left[\sum_t U_{ijt} e_t\right]\left[\sum_n U_{pqn} e_n\right]. \qquad (4.55)$$

Expanding the last two products of 4.55 and using the independence of e_m, we have

$$L_4 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_v \sum_s U_{bks} A_{vs_v}\right] \sum_t U_{ijt} U_{pqt} E(e_t)^2. \qquad (4.56)$$

Taking expected values on the last sum of 4.56, we have

$$L_4 = E\left[\sum_r \sum_m U_{ijm} A_{rm_r}\right]\left[\sum_v \sum_s U_{bks} A_{vs_v}\right] \sigma^2 \sum_t U_{ijt} U_{pqt}$$

$$= 0, \qquad (4.57)$$

since $\sum_m U_{ijm} U_{pqm} = 0$.

Consider equation 4.41. In view of the independence of e_m and A_{rm_r},

$$L_5 = E\left[\sum_m U_{ijm}\, e_m\right]^2 E\left[\sum_n \sum_r U_{pqn}\, A_{rn_r}\right]\left[\sum_v \sum_s U_{kbs}\, A_{vs_v}\right]. \qquad (4.58)$$

Expanding the last two products of 4.58 and using the independence of A_{rn_r}, we have

$$L_5 = E\left[\sum_m U_{ijm}\, e_m\right]^2 \sum_n \sum_r U_{pqn}\, U_{kbn}\, E(A_{rn_r})^2. \qquad (4.59)$$

Taking expected values on the last sum of 4.59, we have

$$L_5 = E\left[\sum_m U_{ijm}\, e_m\right]^2 \sum_r \sigma_r^2 \sum_n U_{pqm}\, U_{kbn} = 0, \qquad (4.60)$$

since $\sum_n U_{pqn}\, U_{kbn} = 0.$

Consider equation 4.42. Expanding the square term, we have

$$L_6 = E\left[\sum_m U_{ijm}^2\, e_m^2 + \sum_{\substack{m \\ m \neq t}} \sum_t U_{ijm}\, U_{ijt}\, e_m\, e_t\right] x$$

$$\left[\sum_n U_{pqn}\, e_n\right]\left[\sum_s U_{kbs}\, e_s\right]. \qquad (4.61)$$

Distributing the expected values in 4.61, we have

$$L_6 = E\left[\sum_m U_{ijm}^2\, e_m^2\right]\left[\sum_n U_{pqn}\, e_n\right]\left[\sum_s U_{kbs}\, e_s\right]$$

$$+ E\left[\sum_{\substack{m \\ m \neq t}} \sum_t U_{ijm}\, U_{ijt}\, e_m\, e_t\right]\left[\sum_n U_{pqn}\, e_n\right]\left[\sum_s U_{kbs}\, e_s\right]. \qquad (4.62)$$

Expanding the products in 4.62 and using the independence of e_m, we have

$$L_6 = E\left[\sum_m U_{ijm}^2 U_{pqm} U_{kbm} e_m^4\right]$$

$$+ E\left[\sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm}^2 U_{pqn} U_{kbn} e_m^2 e_n^2\right]$$

$$+ E\left[\sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm} U_{ijn} U_{pqm} U_{kbn} e_m^2 e_n^2\right]$$

$$+ E\left[\sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm} U_{ijn} U_{pqn} U_{kbm} e_m^2 e_n^2\right] . \qquad (4.63)$$

Taking expected values in 4.63, gives

$$L_6 = \mu_4 \sum_m U_{ijm}^2 U_{pqm} U_{kbm} + \sigma^4 \left[\sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm}^2 U_{pqn} U_{kbn}\right.$$

$$\left. + \sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm} U_{ijn} U_{pqm} U_{kbn} + \sum_{\substack{m \\ m\neq n}} \sum_n U_{ijm} U_{ijn} U_{pqn} U_{kbm}\right] . \qquad (4.64)$$

Adding and subtracting

$$3\sigma^4 \sum_m U_{ijm}^2 U_{pqm} U_{kbm}, \text{ we have}$$

$$L_6 = (\mu_4 - 3\sigma^4) \sum_m U_{ijm}^2 U_{pqm} U_{kbm} + \sigma^4 \left[\sum_m \sum_n U_{ijm}^2 U_{pqn} U_{kbn}\right.$$

$$\left. + \sum_m \sum_n U_{ijm} U_{ijn} U_{pqm} U_{kbn} + \sum_m \sum_n U_{ijm} U_{ijn} U_{pqn} U_{kbm}\right] . \qquad (4.65)$$

The terms in brackets vanish since

$$\sum_n U_{pqn} U_{kbn} = \sum_m U_{ijm} U_{pqm} = \sum_m U_{ijm} U_{kbm} = 0.$$

Hence

$$L_6 = (\mu_4 - 3\sigma^4) \sum_m U_{ijm}^2 \, U_{pqm} \, U_{kbm}.$$

Therefore

$$E\left[Z_{ij}^2 \, Z_{pq} \, Z_{kb}\right] = \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_m U_{ijm}^2 \, U_{pqm} \, U_{kbm}, \tag{4.66}$$

and the Lemma is proved.

LEMMA III. If the orthogonal transformation of Definition 2 is such that the $\sum_j Z_{ij}^2$ is the reduction sum of squares due to A_{in_i} in an analysis of variance table of a balanced model, then $\sum_j U_{ijk}^2 = C_i$, where C_i is a constant changing only with i.

PROOF: Consider the subset of the transformation set given by

$$\begin{bmatrix} Z_{i1} \\ Z_{i2} \\ \cdot \\ \cdot \\ \cdot \\ Z_{in_i} \end{bmatrix} = \begin{bmatrix} U_{i11} & U_{i12} & \cdot & \cdot & \cdot & U_{i1m} \\ U_{i21} & U_{i22} & \cdot & \cdot & \cdot & U_{i2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ U_{in_i1} & U_{in_i2} & \cdot & \cdot & \cdot & U_{in_im} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_m \end{bmatrix} \tag{4.67}$$

where

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_m \end{bmatrix} = \begin{bmatrix} Y_{111\ldots 1} \\ Y_{111\ldots 2} \\ \cdot \\ \cdot \\ \cdot \\ Y_{n_1 n_2 \ldots n_n} \end{bmatrix},$$

or more briefly, $Z = UY$. Then

$$Z' Z = (UY)' (UY)$$
$$= Y' U' U Y \tag{4.68}$$

or

$$\sum Z_{ij}^2 = \begin{bmatrix} Y_1 \ldots Y_m \end{bmatrix} \begin{bmatrix} \sum U_{ij1}^2 & \sum U_{ij1} U_{ij2} & \cdots & \sum U_{ij1} U_{ijm} \\ \sum U_{ij1} U_{ij2} & \sum U_{ij2}^2 & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \sum U_{ij1} U_{ijm} & . & \cdots & \sum U_{ijm}^2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_m \end{bmatrix}$$

(All summations being over $j = 1$ to $j = n_i$). (4.69)

In view of the symmetry of the analysis of variance sum of squares in the balanced models, the quadratic form on the right of 4.69 must be symmetric in the Y_i^2's. Hence

$$\sum_{j=1}^{n_i} U_{ijk}^2 = \sum_{j=1}^{n_i} U_{ijp}^2 = C_i, \tag{4.70}$$

thus proving the Lemma.

LEMMA IV. If the Z_{ij} are the elements of the Z system as given by Definition 2, then

$$E(Z_{ij}^2 \, Z_{pq}^2) = \sum_m U_{ijm}^2 \, U_{pqm}^2 \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4) \right]$$

$$+ 2\mu \left[(\mu_3 + \sum_r \mu_{r3}) \sum_m (U_{ijm}^2 \, U_{pqm} \sqrt{N} \, S_{01}^{pq} + U_{ijm} \, U_{pqm}^2 \sqrt{N} \, S_{01}^{ij}) \right]$$

$$+ S_{pq}^{ij} \left[4\mu^2 \sigma^2 N \, S_{01}^{ij} \, S_{01}^{pq} + 4\sigma^2 \sum_r \sigma_r^2 + 4\mu^2 N \, S_{01}^{ij} \, S_{01}^{pq} \sum_r \sigma_r^2 \right.$$

$$+ 2 \left(\sum_{r} \sigma_r^2\right)^2 + 2\sigma^4 \Big] + \mu^2 N \left(S_{01}^{ij} + S_{01}^{pq}\right) \left(\sigma^2 + \sum_{t} \sigma_t^2\right)$$

$$+ \mu^4 N^2 S_{01}^{ij} S_{01}^{pq} + f,$$

where f does not depend on i, j, p, or q.

PROOF: Consider $E(Z_{ij}^2 Z_{pq}^2)$. Replacing Z_{ij} and Z_{pq} by the Y set gives

$$E(Z_{ij}^2 Z_{pq}^2) = E\left[\sum_{m} U_{ijm} Y_m\right]^2 \left[\sum_{n} U_{pqn} Y_n\right]^2 . \qquad (4.71)$$

Replacing Y_m and Y_n by their values in terms of A_{ki_k}, e_i, and μ, we have

$$E(Z_{ij}^2 Z_{pq}^2) = E\left[\sum_{m} U_{ijm} \left(\sum_{r} A_{rm_r} + e_m + \mu\right)\right]^2 x$$

$$\left[\sum_{n} U_{pqn} \left(\sum_{t} A_{tn_t} + e_n + \mu\right)\right]^2 . \qquad (4.72)$$

Expanding 4.72 and using the independence of A_{rm_r} and e_m, we have

$$E(Z_{ij}^2 Z_{pq}^2) = E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \left[\sum_{n}\sum_{t} U_{pqn} A_{tn_t}\right]^2$$

$$+ E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \left[\sum_{n} U_{pqn} e_n\right]^2$$

$$+ \mu^2 E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 N S_{01}^{pq}$$

$$+ 2\mu E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r}\right]^2 \left[\sum_{n}\sum_{t} U_{pqn} A_{tn_t}\right] \sqrt{N} S_{01}^{pq}$$

$$+ 4E\left[\sum_{m}\sum_{r} U_{ijm} A_{rm_r} \sum_{k} U_{ijk} e_k\right] x$$

$$\left[\sum_n \sum_t U_{pqn} A_{tn_t} \sum_s U_{pqs} e_s\right]$$

$$+ 4\mu^2 E\left[\sum_r \sum_m U_{ijm} A_{rm_r} \sum_n \sum_t U_{pqn} A_{tn_t}\right] N S_{01}^{ij} S_{01}^{pq}$$

$$+ 4\mu^2 E\left[\sum_m U_{ijm} e_m \sum_n U_{pqn} e_n\right] N S_{01}^{ij} S_{01}^{pq}$$

$$+ 2\mu E\left[\sum_m \sum_r U_{ijm} A_{rm_r}\right]\left[\sum_t \sum_n U_{pqn} A_{tn_t}\right]^2 \sqrt{N} S_{01}^{ij}$$

$$+ E\left[\sum_m U_{ijm} e_m\right]^2 \left[\sum_t \sum_n U_{pqn} A_{tn_t}\right]^2$$

$$+ E\left[\sum_m U_{ijm} e_m\right]^2 \left[\sum_n U_{pqn} e_n\right]^2 + \mu^2 E\left[\sum_m U_{ijm} e_m\right]^2 N S_{01}^{pq}$$

$$+ 2\mu E\left[\sum_m U_{ijm} e_m\right]^2 \left[\sum_n U_{pqn} e_n\right] \sqrt{N} S_{01}^{pq}$$

$$+ 2\mu E\left[\sum_m U_{ijm} e_m\right]\left[\sum_n U_{pqn} e_n\right]^2 \sqrt{N} S_{01}^{ij}$$

$$+ \mu^2 E\left[\sum_t \sum_n U_{pqn} A_{tn_t}\right]^2 N S_{01}^{ij}$$

$$+ \mu^2 E\left[\sum_n U_{pqn} e_n\right]^2 N S_{01}^{ij} + \mu^4 N^2 S_{01}^{ij} S_{01}^{pq}, \qquad (4.73)$$

or denoting the ith term of 4.73 by I_i

$$E(Z_{ij}^2 Z_{pq}^2) = I_1 + I_2 + \ldots + I_{16}. \qquad (4.74)$$

Consider I_1. Expanding and using the independence of A_{rm_r}, we have

$$I_1 = E\left[\sum_m \sum_r U^2_{ijm} A^2_{rm_r} + \sum_m \sum_r \sum_h \sum_k \underset{m,r\neq h,k}{} U_{ijm} A_{rm_r} U_{ijh} A_{kh_k}\right] \times$$

$$\left[\sum_n \sum_t U^2_{pqn} A^2_{tn_t} + \sum_n \sum_t \sum_s \sum_v \underset{n,t\neq s,v}{} U_{pqn} A_{tn_t} U_{pqs} A_{vs_v}\right]$$

$$= E\left[\sum_m \sum_r \sum_n \sum_t U^2_{ijm} A^2_{rm_r} U^2_{pqn} A^2_{tn_t}\right]$$

$$+ E\left[\sum_m \sum_r \sum_h \sum_k \underset{m,r\neq h,k}{} U_{ijm} A_{rm_r} U_{ijh} A_{kh_k}\right] \times$$

$$\left[\sum_n \sum_t \sum_s \sum_v \underset{n,t\neq s,v}{} U_{pqn} A_{tn_t} U_{pqs} A_{vs_v}\right].$$

Taking expected values, we have

$$I_1 = \sum_m \sum_r U^2_{ijm} U^2_{pqm}(\mu_{r4} - 3\sigma^4_r) + 2\left[\sum_r \sigma^2_r\right]^2 S^{ij}_{pq} + \left[\sum_r \sigma^2_r\right]^2 \tag{4.75}$$

Consider I_2. Expanding and using the independence of A_{rm_r} and e_s, we have

$$I_2 = \sum_m \sum_r U^2_{ijm} \sigma^2_r \sum_n U^2_{pqn} \sigma^2 = \sigma^2 \sum_r \sigma^2_r. \tag{4.76}$$

Consider I_3. Expanding and using the independence of A_{rm_r}, we have

$$I_3 = N\, S^{pq}_{01}\, \mu^2 \sum_m \sum_r U^2_{ijm} \sigma^2_r = N\, S^{pq}_{01}\, \mu^2 \sum_r \sigma^2_r. \tag{4.77}$$

Consider I_4. Expanding and using the independence of A_{rm_r}, we have

$$I_4 = 2\mu \sqrt{N}\, S_{01}^{pq} \sum_m \sum_r U_{ijm}^2\, U_{pqm}\, \mu_{r3}. \tag{4.78}$$

Consider I_5. Expanding and using the independence of A_{rm_r} and e_s, we have

$$I_5 = 4 \sum_r \sum_m U_{ijm}\, U_{pqm}\, \sigma_r^2 \sum_k U_{ijk}\, U_{pqk}\, \sigma^2. \tag{4.79}$$

Consider I_6. Expanding and using the independence of A_{rm_r}, we have

$$I_6 = 4\, \mu^2\, N\, S_{01}^{ij}\, S_{01}^{pq} \sum_r \sum_m U_{ijm}\, U_{pqm}\, \sigma_r^2. \tag{4.80}$$

Consider I_7. Expanding and using the independence of e_s, we have

$$I_7 = 4\, \mu^2\, N\, S_{01}^{ij}\, S_{01}^{pq} \sum_m U_{ijm}\, U_{pqm}\, \sigma^2. \tag{4.81}$$

Consider I_8. Expanding and using the independence of A_{rm_r}, we have

$$I_8 = 2\, \mu \sqrt{N}\, S_{01}^{ij} \sum_m \sum_r U_{ijm}\, U_{pqm}^2\, \mu_{r3}. \tag{4.82}$$

Consider I_9. Expanding and using the independence of e_m and A_{rm_r}, we have

$$I_9 = \sum_m U_{ijm}^2\, \sigma^2 \sum_t \sum_n U_{pqn}^2\, \sigma_t^2 = \sigma^2 \sum_t \sigma_t^2\ . \tag{4.83}$$

Consider I_{10}. Expanding and using the independence of e_s, we have

$$I_{10} = E\left[\sum_m U_{ijm}^2\, e_m^2 + \sum_{t \neq m} U_{ijm}\, U_{ijt} + e_m\, e_t\right] x$$

$$\left[\sum_{n} U_{pqn}^2 \, e_n^2 + \sum_{n \neq s} U_{pqn} \, U_{pqs} \, e_n \, e_s \right].$$

Taking expected values, we have

$$I_{10} = \sum_{m} U_{ijm}^2 \, U_{pqm}^2 \, (\mu_4 - 3\sigma^4) + 2S_{pq}^{ij} \, \sigma^4 + \sigma^4. \tag{4.84}$$

Consider I_{11}. Expanding and using the independence of e_s, we have

$$I_{11} = \mu^2 \, \sigma^2 \, N \, S_{01}^{pq} \sum_{m} U_{ijm}^2 = \mu^2 \, \sigma^2 \, N \, S_{01}^{pq}. \tag{4.85}$$

Consider I_{12}. Expanding and using the independence of e_s, we have

$$I_{12} = 2\mu \sum_{m} U_{ijm}^2 \, U_{pqm} \, \mu_3 \, \sqrt{N} \, S_{01}^{pq}. \tag{4.86}$$

Consider I_{13}. Expanding and using the independence of e_s, we have

$$I_{13} = 2\mu \sum_{m} U_{ijm} \, U_{pqm}^2 \, \mu_3 \, \sqrt{N} \, S_{01}^{ij}. \tag{4.87}$$

Consider I_{14}. Expanding and using the independence of A_{tn_t}, we have

$$I_{14} = \mu^2 \sum_{t} \sum_{n} U_{pqn}^2 \, \sigma_t^2 = \mu^2 \sum_{t} \sigma_t^2 \, N \, S_{01}^{ij}. \tag{4.88}$$

Consider I_{15}. Expanding and using the independence of e_n, we have

$$I_{15} = \mu^2 \, N \, S_{01}^{ij} \sum_{n} U_{pqn}^2 \, \sigma^2 = \mu^2 \, \sigma^2 \, N \, S_{01}^{ij}. \tag{4.89}$$

Combining these results, we have

$$E(Z_{ij}^2 \, Z_{pq}^2) = \sum_{m} \sum_{r} U_{ijm}^2 \, U_{pqm}^2 (\mu_{r4} - 3\sigma_r^4) + 2\left[\sum_{r} \sigma_r^2 \right]^2 S_{pq}^{ij} + \left[\sum_{r} \sigma_r^2 \right]^2$$

$$+ \sigma^2 \sum_r \sigma_r^2 + \mu^2 N S_{01}^{pq} \sum_r \sigma_r^2 + 2\mu \sqrt{N} S_{01}^{pq} \sum_r \mu_{r3} \sum_m U_{ijm}^2 U_{pqm}$$

$$+ 4\sigma^2 S_{pq}^{ij} \sum_r \sigma_r^2 + 4\mu^2 N S_{01}^{ij} S_{01}^{pq} S_{pq}^{ij} \sum_r \sigma_r^2 + 4\mu^2 S_{pq}^{ij} \sigma^2 N S_{01}^{ij} S_{01}^{pq}$$

$$+ 2\mu \sqrt{N} S_{01}^{ij} \sum_r \mu_{r3} \sum_m U_{ijm} U_{pqm}^2 + \sigma^2 \sum_t \sigma_t^2 + \sum_m U_{ijm}^2 U_{pqm}^2 (\mu_4 - 3\sigma^4)$$

$$+ 2\sigma^4 S_{pq}^{ij} + \sigma^4 + \mu^2 N \sigma^2 S_{01}^{ij} + 2\mu \mu_3 \sqrt{N} S_{01}^{pq} \sum_m U_{ijm}^2 U_{pqm}$$

$$+ 2\mu \mu_3 \sqrt{N} S_{01}^{ij} \sum_m U_{ijm} U_{pqm}^2 + \mu^2 N S_{01}^{ij} \sum_t \sigma_t^2 + \mu^2 \sigma^2 N S_{01}^{pq}$$

$$+ \mu^4 N^2 S_{01}^{ij} S_{01}^{pq}$$

$$= \sum_m U_{ijm}^2 U_{pqm}^2 \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4) \right]$$

$$+ 2\mu \left[(\mu_3 + \sum_r \mu_{r3}) \sum_m (U_{ijm}^2 U_{pqm} \sqrt{N} S_{01}^{pq} + U_{ijm} U_{pqm}^2 \sqrt{N} S_{01}^{ij}) \right]$$

$$+ S_{pq}^{ij} \Bigg[4\mu^2 \sigma^2 N S_{01}^{ij} S_{01}^{pq} + 4\sigma^2 \sum_r + \sigma_r^2 + 4\mu^2 N S_{01}^{ij} S_{01}^{pq} \sum_r \sigma_r^2$$

$$+ 2(\sum_r \sigma_r^2)^2 + 2\sigma^4 \Bigg] + \mu^2 N S_{01}^{pq} \sum_r \sigma_r^2 + \mu^2 N S_{01}^{ij} \sigma^2$$

$$+ N \mu^2 S_{01}^{ij} \sum_t + \sigma_t^2 + \mu^2 N S_{01}^{pq} + \mu^4 N^2 S_{01}^{ij} S_{01}^{pq} + f,$$

where f does not depend on i, j, p, or q. Thus proving the Lemma.

LEMMA V. If Z_{ij}, Z_{pq}, Z_{kb} are elements of the Z system, as given by Definition 2, and if they are selected so that $i,j \neq k,b$; $p,q \neq k,b$; and $k,b = 0,1$; then

$$E\,(Z_{ij}^2\, Z_{pq}\, Z_{01}) = \left[\sum_r (\mu_{r3} + \mu_3)\right] \sum_m U_{ijm}^2\, U_{pqm}\, \mu\sqrt{N}$$

$$+ \left[\sum_r (\mu_{r4} - \sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_m U_{ijm}^2\, U_{pqm}\, U_{01m}.$$

PROOF: Consider $E(Z_{ij}^2\, Z_{pq}\, Z_{01})$. Replacing Z_{ij}, Z_{pq}, and Z_{01} by the Y set, gives

$$E(Z_{ij}^2\, Z_{pq}\, Z_{01}) = E\left[\sum_m U_{ijm}\, Y_m\right]^2 \left[\sum_n U_{pqn}\, Y_n\right]\left[\sum_s U_{01s}\, Y_s\right]. \tag{4.91}$$

Replacing Y_m, Y_n, and Y_s by their values in terms of A_{ki_k}, e_i, and μ, we have

$$E(Z_{ij}^2\, Z_{pq}\, Z_{01}) = E\left[\sum_m U_{ijm}\,(\sum_r A_{rm_r} + e_m + \mu)\right]^2 x$$

$$\left[\sum_n U_{pqn}\,(\sum_r A_{rn_r} + e_n + \mu)\right]\left[\sum_s U_{01s}\,(\sum_r A_{rs_r} + e_s + \mu)\right]. \tag{4.92}$$

Using $\sum_r U_{ijr} = \sqrt{N}\, S_{01}^{ij}$, we have

$$E(Z_{ij}^2\, Z_{pq}\, Z_{01}) = E\left[\sum_m \sum_r U_{ijm}\, A_{rm_r} + \sum_m U_{ijm}\, e_m\right]^2 x$$

$$\left[\sum_n \sum_t U_{pqn}\, A_{tn_t} + \sum_n U_{pqn}\, e_n\right]\left[\sum_s \sum_v U_{01s}\, A_{vs_v} + \sum_s U_{01s}\, e_s + \sqrt{N}\,\mu\right]. \tag{4.93}$$

Expanding 4.92, using the independence of A_{rs_r} and e_j, and 3.2 (a) and (d), we have

$$E(Z_{ij}^2 Z_{pq} Z_{01}) = \sqrt{N}\,\mu \left[\sum_m U_{ijm}^2 U_{pqm} E(e_m^3) + \sum_m \sum_r U_{ijm}^2 U_{pqm} E(A_{rm_r}^3)\right]$$

$$+ E\left[\sum_m \sum_r U_{ijm} A_{rm_r} + \sum_m U_{ijm} e_m\right]^2 \left[\sum_n \sum_r U_{pqn} A_{rn_r}\right.$$

$$\left. + \sum_n U_{pqn} e_n\right]\left[\sum_s \sum_r U_{01s} A_{rs_r} + \sum_s U_{01s} e_s\right]. \qquad (4.94)$$

Using the same arguments on the second term of 4.94 as was used on 4.4 and 4.35, we have

$$E(Z_{ij}^2 Z_{pq} Z_{01}) = \sqrt{N}\,\mu \left[\sum_m U_{ijm}^2 U_{pqm} \mu_3 + \sum_m \sum_r U_{ijm}^2 U_{pqm} \mu_{r3}\right]$$

$$+ \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_m U_{ijm}^2 U_{pqm} U_{01m}. \qquad (4.95)$$

Thus proving the Lemma.

LEMMA VI. If Z_{ij} are the elements of the Z system as given by Definition 2 and if g_p and h_{pqkb} are arbitrary constants having the restrictions that $\sum_q h_{pqpq} = 0$; $h_{0101} = 0$; $q = 1,2,\ldots n_p$; $p = 1,2,\ldots n$; $b = 1,2,\ldots n_k$; and $k = 1,2,\ldots n$; then

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \sum_{k=0}^{n} \sum_{b=1}^{n_k} \frac{g_i}{n_i} h_{pqkb} E(Z_{ij}^2 Z_{pq} Z_{kb}) = 0$$

PROOF: Consider

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \sum_{k=0}^{n} \sum_{b=1}^{n_k} \frac{g_i}{n_i} h_{pqkb} E(Z_{ij}^2 Z_{pq} Z_{kb}).$$

By Lemma I

$$E(Z_{ij}^2\, Z_{pq}\, Z_{kb}) = \left[\sum_r (\mu_{r4}-3\sigma_r^4) + (\mu_4-3\sigma^4)\right] \sum_m U_{ijm}^2\, U_{pqm}\, U_{kbm},$$

if i,j = p,q $\neq$ k,b.

By Lemma II

$$E(Z_{ij}^2\, Z_{pq}\, Z_{kb}) = \left[\sum_r (\mu_{r4}-3\sigma_r^4) + (\mu_4-3\sigma^4)\right] \sum_m U_{ijm}^2\, U_{pqm}\, U_{kbm},$$

if i,j $\neq$ p,q $\neq$ k,b.

By Lemma IV

$$E(Z_{ij}^2\, Z_{pq}^2) = \left[(\mu_4-3\sigma^4) + \sum_r (\mu_{r4}-3\sigma_r^4)\right] \sum_m U_{ijm}^2\, U_{pqm}^2$$

$$+ 2\mu\left[(\mu_3 + \sum_r \mu_{r3}) \sum_m (U_{ijm}^2\, U_{pqm} \sqrt{N}\, S_{01}^{pq} + U_{ijm}\, U_{pqm}^2 \sqrt{N}\, S_{01}^{ij})\right]$$

$$+ S_{pq}^{ij}\left[4\mu^2\, \sigma^2\, N\, S_{01}^{ij}\, S_{01}^{pq} + 4\sigma^2 \sum_r \sigma_r^2 + 4\mu^2\, N\, S_{01}^{ij}\, S_{01}^{pq} \sum_r \sigma_r^2\right.$$

$$\left. + 2(\sum_r \sigma_r^2)^2 + 2\sigma^4\right] + \mu^2\, N(S_{01}^{ij} + S_{01}^{pq})\, (\sigma^2 + \sum_t \sigma_t^2)$$

$$+ \mu^4\, N^2\, S_{01}^{ij}\, S_{01}^{pq} + f,$$ where f does not depend on i, j, p, or q.

By Lemma V

$$E(Z_{ij}^2\, Z_{pq}\, Z_{01}) = \left[\sum_r \mu_{r3} + \mu_3\right] \sum_m U_{ijm}^2\, U_{pqm}\, \mu\sqrt{N}$$

$$+ \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_m U_{ijm}^2\, U_{pqm}\, U_{01m},$$

if i,j $\neq$ 0,1 and p,q $\neq$ 0,1.

Hence we may write

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \sum_{k=0}^{n} \sum_{b=1}^{n_k} \frac{g_i}{n_i}\, h_{pqkb} \left[\sum_r (\mu_{r4}-3\sigma_r^4) + (\mu_4-3\sigma^4)\right] x$$

$$\left[\sum_m U_{ijm}^2 \, U_{pqm} \, U_{kbm}\right] + \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \frac{g_i}{n_i} h_{pqpq} \times$$

$$\left[2\mu \sqrt{N} \, (\mu_3 + \sum_r \mu_{r3}) \sum_m U_{ijm}^2 \, U_{pqm} \, S_{01}^{pq} + \mu^2 N \, (S_{01}^{pq}) \, (\sigma^2\right.$$

$$\left. + \sum_t \sigma_t^2) + f + S_{pq}^{ij} \, (4\sigma^2 \sum_r \sigma_r^2 + 2 \sum_t \sum_r \sigma_r^2 \, \sigma_t^2 + 2\sigma^4)\right]$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=1}^{n} \sum_{q=1}^{n_p} \frac{g_i}{n_i} h_{pq01} \left[\sum_r \mu_{r3} + \mu_3\right] \sum_m U_{ijm}^2 \, U_{pqm} \, \mu \sqrt{N} \, . \tag{4.96}$$

or $I = J_1 + J_2 + J_3$.

Consider J_1. Summing on j, using Lemma III, $\sum_m U_{pqm} \, U_{kbm} = S_{pq}^{kb}$, and the hypothesis, we have

$$J_1 = \sum_{i=1}^{n} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \sum_{k=0}^{n} \sum_{b=1}^{n_b} \frac{g_i}{n_i} h_{pqkb} \, C_i \sum_m U_{pqm} \, U_{kbm} \times$$

$$\left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right]$$

$$= \sum_{i=1}^{n} \sum_{p=0}^{n} \frac{g_i}{n_i} C_i \left[\sum_r (\mu_{r4} - 3\sigma_r^4) + (\mu_4 - 3\sigma^4)\right] \sum_{q=1}^{n_p} h_{pqpq}$$

$$= 0 \, . \tag{4.97}$$

Consider J_2. Summing on j, using Lemma III, $U_{01i} = 1/\sqrt{N}$, $\sum_m U_{pqm} = \sqrt{N} \, S_{01}^{pq}$, and the hypothesis, we have

$$J_2 = 2\mu \sqrt{N} \sum_{i=1}^{n} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \frac{g_i}{n_i} C_i \, h_{pqpq} \, (\mu_3 + \sum_r \mu_{r3}) \sum_m U_{pqm} \, S_{01}^{pq}$$

$$+ \mu^2 N \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{p=0}^{n} \sum_{q=1}^{n_p} \frac{g_i}{n_i} h_{pqpq} \left[(\sigma^2 + \sum_t \sigma_t^2) + f\right]$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n_i} \frac{g_i}{n_i} h_{ijij} \left[4\sigma^2 \sum_r \sigma_r^2 + 2 \sum_t \sum_r \sigma_r^2 \sigma_t^2 + 2\sigma^4 \right]$$

$$= 0. \tag{4.98}$$

Consider J_3. Summing on j, using Lemma III, and $\sum_m U_{pqm} = \sqrt{N}\, S_{01}^{pq}$, we have

$$J_3 = \sum_{i=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n_p} \frac{g_i}{n_i} C_i\, h_{pq01} \left[\sum_r \mu_{r3} + \mu_3 \right] \sum_m U_{pqm}\, \mu \sqrt{N}$$

$$= 0. \tag{4.99}$$

Hence I = 0, and the Lemma is proved.

V. QUADRATIC ESTIMATORS OF THE VARIANCE COMPONENTS

COMPONENTS IN THE BALANCED MODELS. This section is devoted to the proof of the main theorem on the quadratic estimators of variance components in the balanced models.

THEOREM I. Let Z_{01} be distributed as $f(Z_{01})$ with mean μ and variance σ_0^2, let Z_{ij} be distributed as $f(Z_{ij})$ with mean zero and variance σ_i^2 $(j = 1,2,\ldots n_i;\ i = 0,\ 1,\ 2,\ldots n)$ $(\sigma_0^2 = \sum_{i=1}^{n} d_i\ \sigma_i^2)$, and let Z_{ij} be the transformed orthogonal uncorrelated variates from a balanced model with finite fourth moments.

The best (minimum variance) unbiased homogeneous quadratic estimator of $L = \sum_{i=1}^{n} g_i\ \sigma_i^2$, where the g_i are constants independent of the σ_i^2, μ, and Z_{ij}, is given by

$$M' = \sum_{i=1}^{n} g_i\ \hat{\sigma}_i^2,\ \text{where}\ \hat{\sigma}_i^2 = \sum_{j=1}^{n_i} \frac{Z_{ij}^2}{n_i}\ .$$

PROOF: The general homogeneous quadratic estimator of L has the form

$$M = \sum_{i=1}^{n} \sum_{j=1}^{n_i} g_i \frac{Z_{ij}^2}{n_i} + \sum_{p} \sum_{q} \sum_{k} \sum_{b} h_{pqkb}\ Z_{pq}\ Z_{kb}, \qquad (5.1)$$

where h_{pqkb} are arbitrary constants independent of μ and σ_i^2. Since M is unbiased, its mathematical expectation is L. That is, using the properties that

$$E(Z_{ij}\ Z_{kb}) = S_{kb}^{ij}\ \sigma_i^2\ \text{and}\ E(Z_{01}^2) = \mu^2 + \sigma_0^2, \qquad (5.2)$$

we have

$$E(M) = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n_i} g_i \frac{Z_{ij}^2}{n_i} + \sum_p \sum_q \sum_k \sum_b h_{pqkb} Z_{pq} Z_{kb}\right]$$

$$= \sum_{i=1}^{n} g_i \sigma_i^2 + \sum_p \sum_q h_{pqpq} \sigma_p^2 + h_{0101} (\mu^2 + \sigma_0^2), \qquad (5.3)$$

or

$$\sum_{i=1}^{n} g_i \sigma_i^2 + \sum_p \sum_q h_{pqpq} \sigma_p^2 + h_{0101} (\mu^2+\sigma_0^2) = \sum_{i=1}^{n} g_i \sigma_i^2. \qquad (5.4)$$

Hence, equating coefficients

$$\sum_p \sum_q h_{pqpq} \sigma_p^2 = 0 \text{ and } h_{0101} = 0, \qquad (5.5)$$

and since $\sigma_p^2 \neq 0$, it follows that

$$\sum_q h_{pqpq} = 0. \qquad (5.6)$$

Consider now the variance of M, denoted V(M),

$$V(M) = V\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i} + \sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right], \qquad (5.7)$$

where the summation index R_0 indicates $j = 1,2,\ldots n_i$; $i = 0,1,2,\ldots n$; $b = 1,2,\ldots n_k$; $k = 0,1,2,\ldots n$; $q = 1,2,\ldots n_p$; and $p = 0,1,2,\ldots n$; $n_0 = 1$. Expanding 5.7, we have

$$V(M) = V\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i}\right] + V\left[\sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right]$$

$$+ 2 \text{ Covariance}\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i}; \sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right]. \qquad (5.8)$$

Consider 1/2 the last term of 5.8, equal to I, say. Then

$$I = \text{Cov.}\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i} ; \sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right], \qquad (5.9)$$

or

$$I = E\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i} - \sum_i g_i \sigma_i^2\right]\left[\sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right]. \qquad (5.10)$$

Since the Z_{ij} are uncorrelated, we have

$$E(Z_{ij} Z_{kb}) = 0. \qquad (5.11)$$

Expanding 5.10 and using 5.11, we have

$$I = E\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i} \sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right], \qquad (5.12)$$

or distributing the expected value,

$$I = \sum_i \sum_j \sum_{R_0} \frac{g_i}{n_i} h_{pqkb} E(Z_{ij}^2 Z_{pq} Z_{kb}). \qquad (5.13)$$

Since the Z_{ij} are transformed orthogonal variates from a balanced model, Lemma VI applies to 5.13 giving

$$I = 0. \qquad (5.14)$$

Then we have

$$V(M) = V\left[\sum_i \sum_j g_i \frac{Z_{ij}^2}{n_i}\right] + V\left[\sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right]. \qquad (5.15)$$

Both terms on the right of 5.15 are postive, the first is independent of h_{pqkb}, and since

$$E\left[\sum_{R_0} h_{pqkb} Z_{pq} Z_{kb}\right] = 0,$$

V(M) will be a minimum when

$$\sum_{R_0} h_{pqkb} Z_{pq} Z_{kb} \equiv 0. \qquad (5.16)$$

That is, when $h_{pqkb} \equiv 0$ for all p, q, k, b.

Therefore the best unbiased quadratic estimate in the balanced model is

$$M = \sum_{i=1}^{n} g_i \, \hat{\sigma}_i^2. \tag{5.17}$$

VI. BALANCED MODELS

In this section we will discuss several models which are balanced and will prove that some of them satisfy the hypothesis of Theorem I of Section V. The method of proof is general and the extension to other cases, though algebraically tedious, is obvious.

RANDOMIZED BLOCK DESIGN. This model is usually given in the form $Y_{ij} = \mu + a_i + b_j + e_{ij}$

where $i = 1,2,\ldots n_1$; $j = 1,2,\ldots n_2$;

μ is a fixed constant;

a_i are independent random variables with mean zero and variance σ_a^2;

b_j are independent random variables with mean zero and variance σ_b^2;

e_{ij} are independent random variables with mean zero and variance σ^2.

THEOREM I. There exists an orthogonal transformation $\underline{Y} = AZ$ such that the Z system satisfies the hypothesis of Theorem I of Section V for the Randomized Block Design. That is, there exists an orthogonal transformation $Y = AZ$ such that the Z_{ij} have the following properties:

$$\begin{aligned} &(1)\quad E(Z_{ij}) = 0, && \text{if } i \neq 0 \text{ and } j \neq 1 \\ &(2)\quad E(Z_{01}) = \sqrt{n_1 n_2}\,\mu, && \\ &(3)\quad E(Z_{ij}\,Z_{mn}) = 0, && \text{if } m,n \neq i,j, \text{ and} \\ &(4)\quad E(Z_{ij}^2) = \sigma_i^2. && \end{aligned} \qquad (6.1)$$

PROOF: Consider

$$\sum_{i,j} Y_{ij}^2 = \sum_{i,j} \left[(Y_{ij}-Y_{i.}-Y_{.j}+Y_{..}) + (Y_{i.}-Y_{..}) + (Y_{.j}-Y_{..}) + Y_{..} \right]^2, \tag{6.2}$$

where the dots indicate averages over the indicated subscripts. Expanding 6.2, we have

$$\sum_{i,j} Y_{ij}^2 = \sum_{i,j} (Y_{ij} - Y_{i.} - Y_{.j} + Y_{..})^2 + \sum_i n_2 (Y_{i.} - Y_{..})^2$$

$$+ \sum_j n_1 (Y_{.j} - Y_{..})^2 + n_1 n_2 Y_{..}^2, \tag{6.3}$$

since the crossproducts sum to zero.

The quadratic form on the left has rank $n_1 n_2$ since it can be written as $Y' IY$, where Y is a column matrix with $n_1 n_2$ elements and I is a $n_1 n_2$ by $n_1 n_2$ identity matrix. The ranks of the quadratic forms on the right are less than or equal to $(n_1-1)(n_2-1)$, (n_1-1), (n_2-1), and 1 respectively since there are $n_1 + n_2 - 1$ linear restrictions on the first term, 1 on the second, 1 on the third, and none on the fourth (1). Further, since the rank of a sum is less than or equal to the sum of the ranks,

$$n_1 n_2 \leq (n_1-1)(n_2-1) + (n_1-1) + (n_2-1) + 1. \tag{6.4}$$

But, 6.4 is impossible except with the equality holding. Therefore, the ranks of the terms on the right are $(n_1-1)(n_2-1)$, (n_1-1), (n_2-1), and 1 respectively.

By Cochran's Theorem (1), we have the existence of an orthogonal transformation, say $Y = AZ$, which if applied to 6.3 gives

$$\sum_i \sum_j Y_{ij}^2 = \sum_{k=1}^{(n_1-1)(n_2-1)} Z_{3k}^2 + \sum_{k=1}^{n_1-1} Z_{1k}^2 + \sum_{k=1}^{n_2-1} Z_{2k}^2 + Z_{01}^2, \tag{6.5}$$

where

$$\sum_{k} Z_{3k}^2 = \sum_{i,j} (Y_{ij} - Y_{i.} - Y_{.j} + Y_{..})^2,$$

$$\sum_{k} Z_{2k}^2 = \sum_{i} n_2 (Y_{i.} - Y_{..})^2,$$

$$\sum_{k} Z_{1k}^2 = \sum_{j} n_1 (Y_{.j} - Y_{..})^2, \text{ and} \qquad (6.6)$$

$$Z_{01}^2 = n_1 n_2 Y_{..}^2 .$$

We will now establish the properties of this transformation. Since the transformation is orthogonal

$$A A' = I. \qquad (6.7)$$

Using the notation of Definition 2, we have

$$Y_{ij} = \sum_{p} \sum_{q} U_{pqij} Z_{pq} \qquad (6.8)$$

and

$$Z_{ij} = \sum_{p} \sum_{q} U_{ijpq} Y_{pq} \qquad (6.9)$$

in view of the orthogonality. Further

$$\sum_{q} \sum_{r} U_{ijqr} U_{knqr} = S_{kn}^{ij} \qquad (6.10)$$

and

$$\sum_{q} \sum_{r} U_{qrij} U_{qrkn} = S_{kn}^{ij}. \qquad (6.11)$$

Since

$$Z_{01}^2 = n_1 n_2 Y_{..}^2 = \left[\sum_{ij} Y_{ij}\right]^2 \Big/ n_1 n_2, \qquad (6.12)$$

we have

$$U_{01ij} = 1 / \sqrt{n_1 n_2} \ . \tag{6.13}$$

Substituting 6.13 in 6.10 gives

$$\sum_q \sum_r U_{ijqr} = S_{01}^{ij} \sqrt{n_1 n_2} \ . \tag{6.14}$$

Consider the third equation of 6.6

$$\sum_k Z_{1k}^2 = \sum_j n_1 \ (Y_{.j} - Y_{..})^2 = \sum_j n_1 \ Y_{ij}^2 - n_1 n_2 \ Y_{..}^2$$

$$= \sum_j \frac{\left[\sum_i Y_{ij}\right]^2}{n_1} - \frac{\left[\sum_i \sum_j Y_{ij}\right]^2}{n_1 n_2}$$

$$= \frac{\sum_i \sum_j Y_{ij}^2 + \sum_j \sum_i \sum_{m,\ i \neq m} Y_{ij} \ Y_{mj}}{n_1}$$

$$- \frac{\sum_{i,j} Y_{ij}^2 + \sum_i \sum_j \sum_m \sum_{n,\ i,j \neq m,n} Y_{ij} \ Y_{mn}}{n_1 n_2} \ . \tag{6.15}$$

By 6.9 we get

$$\sum_k Z_{1k}^2 = \sum_k \left[\sum_p \sum_q U_{1kpq} \ Y_{pq}\right]^2$$

$$= \sum_k \sum_q \sum_p U_{1kpq}^2 \ Y_{pq}^2 + \sum_k \sum_p \sum_q \sum_r \sum_{s,\ p,q \neq r,s} U_{1kpq} \ U_{1krs} \ Y_{pq} \ Y_{rs} \ . \tag{6.16}$$

Equating coefficients of similar terms in 6.15 and 6.16, we have

$$\sum_k U_{1kpq}^2 = \frac{n_2 - 1}{n_1 n_2} \qquad \text{for } Y_{ij} \ Y_{pq} \text{ if } i = p \text{ and } j = q, \tag{6.17}$$

$$\sum_k U_{1kpq}\, U_{1krq} = \frac{n_2-1}{n_1 n_2} \quad \text{for } Y_{pq}\, Y_{rj} \text{ if } p\neq r,\ j=q, \tag{6.18}$$

and

$$\sum_k U_{1kpq}\, U_{1krs} = \frac{-1}{n_1 n_2} \text{ for } Y_{pq}\, Y_{rs}, \text{ if either } p=r,\ q\neq s, \text{ or}$$

$p\neq r$, and $q\neq s$. (6.19)

Summing 6.19 over s and adding to 6.18 gives

$$\sum_k \sum_s U_{1kpq}\, U_{1krs} = 0. \tag{6.20}$$

Summing 6.20 over q and letting r=p (letting r=p is justified below since U_{1kpq} is proven to be equal to U_{1krq}), we have

$$\sum_k \sum_s \sum_q U_{1kpq}\, U_{1kps} = \sum_k \left[\sum_s U_{1kps}\right]^2 = 0 \tag{6.21}$$

and

$$\sum_s U_{1kps} = 0 \text{ for all p and k.} \tag{6.22}$$

Consider 6.17 and 6.18. We may write

$$\sqrt{\sum_k U^2_{1krq} \sum_k U^2_{1kpq}} = \sum_k U_{1kpq}\, U_{1krq} \,. \tag{6.23}$$

This is the equality case of the Cauchy Schwartz inequality. Hence

$$U_{1kpq} = CU_{1krq} \,. \tag{6.24}$$

Squaring and summing over k and using 6.18

$$\sum_k U^2_{1kpq} = C^2 \sum_k U^2_{ijrq} = C \sum_k U_{1kpq}\, U_{1krq} \tag{6.25}$$

or

$$\frac{n_2-1}{n_1 n_2} = C^2\, \frac{n_2-1}{n_1 n_2} = C\, \frac{n_2-1}{n_1 n_2} \,.$$

Therefore C = 1 and

$$U_{1kpq} = U_{1krq} \; . \tag{6.26}$$

By 6.10, we have

$$\sum_p \sum_q U_{1kpq} \, U_{1jpq} = S_k^j \; . \tag{6.27}$$

Similarly,

$$\sum_k Z_{2k}^2 = \sum_i n_2 \, (Y_{i.} - Y_{..})^2$$

yields

$$\sum_k U_{2kpq}^2 = \frac{n_1 - 1}{n_1 n_2}, \tag{6.28}$$

$$\sum_k U_{2kpq} \, U_{2kpr} = \frac{n_1 - 1}{n_1 n_2}, \tag{6.29}$$

$$\sum_k U_{2kpq} \, U_{2krs} = - \frac{1}{n_1 n_2} \qquad p \neq r, \; pq \neq rs, \tag{6.30}$$

$$\sum_p U_{2kps} = 0, \tag{6.31}$$

$$U_{2kpq} = U_{2kpr}, \text{ and} \tag{6.32}$$

$$\sum_p \sum_q U_{2kpq} \, U_{2jpq} = S_j^k \; . \tag{6.33}$$

Consider

$$\sum_k Z_{3k}^2 = \sum_i \sum_j (Y_{ij} - Y_{i.} - Y_{.j} + Y_{..})^2$$

$$= \left[\sum_i \sum_j Y_{ij}^2 - n_1 n_2 \, Y_{..}^2 \right] - \left[\sum_i n_2 \, Y_{i.}^2 - n_1 n_2 \, Y_{..}^2 \right]$$

$$- \left[\sum_j n_1 Y_{.j}^2 - n_1 n_2 Y_{..}^2 \right]$$

$$= \sum_i \sum_j Y_{ij}^2 - \frac{\sum_i \sum_j Y_{ij}^2 + \sum_i \sum_j \sum_n \sum_m Y_{ij} Y_{nm}}{n_1 n_2} \quad {\scriptstyle i,j \neq n,m}$$

$$- \left[\sum_i n_2 Y_{i.}^2 - n_1 n_2 Y_{..}^2 \right] - \left[\sum_j n_1 Y_{.j}^2 - n_1 n_2 Y_{..}^2 \right] . \qquad (6.34)$$

By 6.9 we get

$$\sum_n Z_{3k}^2 = \sum_k \left[\sum_p \sum_q U_{3kpq} Y_{pq} \right]^2$$

$$= \sum_k \sum_p \sum_q U_{3kpq}^2 Y_{pq}^2$$

$$+ \sum_k \sum_p \sum_q \sum_r \sum_s U_{3kpq} U_{3krs} Y_{pq} Y_{rs} \quad {\scriptstyle p,q \neq r,s} . \qquad (6.35)$$

Equating coefficients of 6.34 and 6.35, we have

$$\sum_k U_{3kpq}^2 = 1 - \frac{1}{n_1 n_2} - \frac{(n_1-1)}{n_1 n_2} - \frac{(n_2-1)}{n_1 n_2}$$

$$= \frac{n_1 n_2 - n_1 - n_2 + 1}{n_1 n_2} = \frac{(n_1-1)(n_2-1)}{n_1 n_2} , \qquad (6.36)$$

$$\sum_k U_{3kpq} U_{3kps} = - \frac{n_1-1}{n_1 n_2} \qquad \text{if } q \neq s, \qquad (6.37)$$

$$\sum_k U_{3kpq} U_{3krq} = - \frac{n_2-1}{n_1 n_2} \qquad \text{if } p \neq r, \text{ and} \qquad (6.38)$$

$$\sum_k U_{3kpq} U_{3krs} = \frac{1}{n_1 n_2} \qquad \text{if } p \neq r, \; q \neq s. \qquad (6.39)$$

Consider

$$I = \sum_k \left[\sum_s U_{3krs} \right]^2 = \sum_k \sum_s U^2_{3krs} + \sum_k \sum_{\substack{s \\ s \neq t}} \sum_t U_{3krs} U_{3krt} . \tag{6.40}$$

Using equations 6.36 and 6.37

$$I = \frac{(n_1-1)(n_2-1)}{n_1} - \frac{(n_1-1)(n_2-1)}{n_1} = 0.$$

Hence

$$\sum_s U_{3krs} = 0 \quad \text{for all } r \text{ and } k. \tag{6.41}$$

Consider

$$J = \sum_k \left[\sum_s U_{3ksr} \right]^2 = \sum_k \sum_s U^2_{3ksr} + \sum_k \sum_{\substack{s \\ s \neq t}} \sum_t U_{3ksr} U_{3ktr} . \tag{6.42}$$

Using equations 6.36 and 6.38

$$J = \frac{(n_1-1)(n_2-1)}{n_2} - \frac{(n_1-1)(n_2-1)}{n_2} = 0.$$

Hence

$$\sum_s U_{3ksr} = 0 \quad \text{for all } r \text{ and } k. \tag{6.43}$$

We will now prove the set of equations 6.1 hold.

Considering $E(Z_{ij})$ and using 6.9 and 6.14, we have

$$E(Z_{ij}) = E\left[\sum_p \sum_q U_{ijpq} Y_{pq} \right]$$

$$= \mu \sum_p \sum_q U_{ijpq}$$

$$= \mu\, S^{ij}_{01} \sqrt{n_1 n_2} . \tag{6.44}$$

Thus proving (1) and (2) of 6.1.

Consider

$$E(Z_{ij}\ Z_{mn}) = E\left[\sum_p \sum_q U_{ijpq}\ Y_{pq} \sum_r \sum_s U_{mnrs}\ Y_{rs}\right]$$

$$= \sum_p \sum_q \sum_r \sum_s U_{ijpq}\ U_{mnrs}\ E(Y_{pq}\ Y_{rs}). \qquad (6.45)$$

Since $Y_{ij} = \mu + a_i + b_j + e_{ij}$, we have

$$E(Y_{pq}\ Y_{rs}) = \mu^2 + \sigma_a^2\ S_r^p + \sigma_b^2\ S_s^q + \sigma^2\ S_{rs}^{pq}. \qquad (6.46)$$

Hence

$$E(Z_{ij}\ Z_{mn}) = \sum_p \sum_q \sum_r \sum_s U_{ijpq}\ U_{mnrs}(\mu^2+\sigma_a^2\ S_r^p+\sigma_b^2\ S_s^q+\sigma^2\ S_{rs}^{pq}), \qquad (6.47)$$

or

$$E(Z_{ij}\ Z_{mn}) = \mu^2 \sum_p \sum_q \sum_r \sum_s U_{ijpq}\ U_{mnrs}+\sigma_a^2 \sum_p \sum_q \sum_s U_{ijpq}\ U_{mnps}$$

$$+ \sigma_b^2 \sum_p \sum_r \sum_s U_{ijps}\ U_{mnrs}+ \sigma^2 \sum_p \sum_q U_{ijpq}\ U_{mnpq}. \qquad (6.48)$$

We will now consider the various cases

(1) If $i,j = 0,1$ and $m,n = 0,1$; we have by 6.13

$$E(Z_{01}^2) = n_1 n_2\ \mu^2 + n_2\ \sigma_a^2 + n_1\ \sigma_b^2 + \sigma^2. \qquad (6.49)$$

(2) If $i,j = 0,1$ and $m,n \neq 0,1$; we have by 6.13 and 6.14

$$E(Z_{01}\ Z_{mn}) = 0. \qquad (6.50)$$

(3) If $i,j = 1,k$ and $m,n = 1,k$; we have by 6.10, 6.14, and 6.22

$$E(Z_{1k}^2) = \sigma_b^2 \sum_p \sum_r \sum_s U_{1kps}\ U_{1krs} + \sigma^2\ . \qquad (6.51)$$

By 6.26 and 6.10

$$\sum_p \sum_s U_{1kps}^2 = n_1 \sum_s U_{1kps}^2$$

$$= 1,$$

and we have

$$\sum_p \sum_r \sum_s U_{1kps} U_{1krs} = \sum_s \left[\sum_p U_{1kps} \right]^2$$

$$= n_1^2 \sum_s U_{1kps}^2$$

$$= n_1. \qquad (6.52)$$

(4) If $i,j = 1,k$ and $m,n \neq 1,k$; we have by 6.14, 6.22, 6.31, and 6.10

$$E(Z_{1k} Z_{mn}) = 0. \qquad (6.53)$$

(5) If $i,j = 2,k$ and $m,n = 2,k$; we have by 6.10, 6.14, and 6.31

$$E(Z_{2k}^2) = \sigma_a^2 \sum_p \sum_q \sum_s U_{2kpq} U_{2kps} + \sigma^2 . \qquad (6.54)$$

By 6.10 and 6.32

$$\sum_p \sum_s U_{2kps}^2 = n_2 \sum_p U_{2kps}^2$$

$$= 1, \qquad (6.55)$$

and we have

$$\sum_p \sum_q \sum_s U_{2kpq} U_{2kps} = \sum_p \left[\sum_s U_{2kps} \right]^2$$

$$= n_2^2 \sum_p U_{2kps}^2$$

$$= n_2. \qquad (6.56)$$

(6) If $i,j = 2,k$ and $m,n \neq 2,k$; we have by 6.10, 6.14, 6.22, and 6.31

$$E(Z_{2k} Z_{mn}) = 0. \qquad (6.57)$$

(7) If $i,j = 3,k$ and $m,n = 3,k$; we have by 6.10, 6.14, 6.41 and 6.43

$$E(Z_{3k}^2) = \sigma^2. \tag{6.58}$$

(8) If $i,j = 3,k$ and $m,n = 3,j$; we have by 6.10, 6.14, 6.41, and 6.43

$$E(Z_{3j}\, Z_{3k}) = 0.$$

Therefore properties (1), (2), (3), and (4) of 6.1 are satisfied and the theorem is proved.

GENERAL CROSS CLASSIFICATION WITH NO INTERACTION. This model is usually given in the form

$$Y_{i_1 i_2 \ldots i_n} = \mu + A_{1i_1} + A_{2i_2} + \ldots + A_{ni_n} + e_{i_1 \ldots i_n}$$

where $i_1 = 1,2,\ldots n_1$;

$i_2 = 1,2,\ldots n_2$;

...

$i_n = 1,2,\ldots n_n$;

μ is a constant;

A_{ki_k} are independent random variables with means zero and variances σ_k^2 $(k = 1,2,\ldots n)$;

$e_{i_1 i_2 \ldots i_n}$ are independent random variables with mean zero and variance σ^2.

COROLLARY. There exist an orthogonal transformation $Y = AZ$ such that the Z system satisfies the hypothesis of Theorem I of Section V for the General Cross Classification With No Interaction.

PROOF: The proof will not be given since it is an obvious generalization of Theorem I of this section.

GENERAL CROSS CLASSIFICATION WITH INTERACTION. This model is usually given in the form

$$Y_{i_1 i_2 \ldots i_n} = \mu + A_{1i_1} + A_{2i_2} + (A_1A_2)_{i_1 i_2} + A_{3i_3} + (A_1A_3)_{i_1 i_3} + (A_2A_3)_{i_2 i_3} + \ldots + e_{i_1 i_2 \ldots i_n} \quad (i_j = 1,2,\ldots n_j;$$

$j = i, 2, \ldots n)$,

where μ is a fixed constant,

A_{ji_j} and all interactions are independent random variables,

$E(A_{ji_j}) = 0$,

$E(A^2_{ji_j}) = \sigma^2_j$,

$E(A^4_{ji_j}) = \mu_{j4} < \infty$,

$e_{i_1 i_2 \ldots i_n}$ are independent random variables,

$E(e_{i_1 i_2 \ldots i_n}) = 0$,

$E(e^2_{i_1 i_2 \ldots i_n}) = \sigma^2$,

$E(e^4_{i_1 i_2 \ldots i_n}) = \mu_4 < \infty$,

THEOREM II. There exists an orthogonal transformation $Y = AZ$ such that the Z system satisfies the hypothesis of Theorem I of Section V for the General Cross Classification Model with Interaction. That is, there exists an orthogonal transformation such that the Z_{ij} have the following properties:

(1) $E(Z_{ij}) = \sqrt{n_1 n_2 \ldots n_n}\, \mu\, S^{ij}_{01}$,

(2) $E(Z_{ij}\, Z_{mn}) = \sigma^2_i\, S^{ij}_{mn}$,

(3) $\sum_j Z^2_{ij}$ is a reduction sum of squares due to the interaction $(A_{j_1} A_{j_2} \ldots A_{j_p})$ in the analysis of variance table, and

(4) the sum of squares in the analysis of variance table are symmetric in the $Y^2_{i_1 i_2 \ldots i_n}$.

The existence of a transformation satisfying (1), (2), and (3) above is well known and an existence proof will not be given here. This proof can be established in the same manner as the proof of Theorem I of this section.

The property that must be demonstrated is the symmetry property. This symmetry follows immediately from the well known fact that any reduction sum of squares in the General Cross Classification can be expressed as a linear combination of terms each of which is symmetric in the $Y^2_{i_1 i_2 \ldots i_n}$ and the additive properties of symmetric quadratic forms. This symmetry property is demonstrated by noting that the reduction sum of squares due to the interaction of A_{j_1}, A_{j_2}, ..., and A_{j_p} is

$$R(A_{j_1} A_{j_2} \ldots A_{j_p}) = \frac{n_{j_1} n_{j_2} \ldots n_{j_p}}{n_1 n_2 \ldots n_n} \sum_{i_{j_1} \ldots i_{j_p}} Y^2_{\ldots i_{j_1} \ldots i_{j_2} \ldots i_{j_p} \ldots} - \frac{n_{j_2} \ldots n_{j_p}}{n_1 n_2 \ldots n_n} \sum_{i_{j_2} \ldots i_{j_p}} Y^2_{\ldots i_{j_2} \ldots i_{j_p} \ldots} - \ldots - \frac{n_{j_1} \ldots n_{j_{p-1}}}{n_1 n_2 \ldots n_n} \times$$

$$\sum_{i_{j_1} \ldots i_{j_{p-1}}} Y^2_{\ldots i_{j_1} \ldots i_{j_{p-1}} \ldots} + \frac{n_{j_3} \ldots n_{j_p}}{n_1 n_2 \ldots n_n} \sum_{i_{j_3} \ldots i_{j_p}} Y^2_{\ldots i_{j_3} \ldots i_{j_p} \ldots}$$

$$+ \ldots + \frac{(-1)^p}{n_1 n_2 \ldots n_n} Y^2_{\ldots}, \qquad (6.59)$$

where the dots in the subscript indicate totals over the indicated subscripts. From equation 6.59, we see that the coefficient of $Y^2_{i_1 i_2 \ldots i_n}$ is

$$\frac{1}{\prod_{i=1}^{n} n_i}\Big[n_{j_1}n_{j_2}\cdots n_{j_p} - n_{j_2}n_{j_3}\cdots n_{j_p} - \cdots - n_{j_1}n_{j_2}\cdots n_{j_{p-1}}$$

$$+ n_{j_3}n_{j_4}\cdots n_{j_p} + \cdots + n_{j_1}n_{j_2}\cdots n_{j_{p-2}} - \cdots + (-1)^p\Big]$$

$$= \frac{1}{\prod_{i=1}^{n} n_i}(n_{j_1}-1)(n_{j_2}-1)\cdots(n_{j_p}-1)$$

for all i_1, i_2, ..., and i_n. Thus the reduction is symmetric in the $Y^2_{i_1 i_2 \cdots i_n}$.

We have thus proved the important Corollary that:

CORLLARY. For the General Balanced Cross Classification, the best unbiased quadratic estimate of any linear combination of the variance components is the same linear combination of the analysis of variance estimates of the variance components.

OTHER DESIGNS. Theorems similar to I and II are also true for

(1) Latin Squares

(2) Graeco-Latin Squares

(3) Split Plot

(4) Split ... Split Plot

(5) Factorial Arrangements

All follow the same general argument as Theorems I and II and will not be given here.

VII. BALANCED MODELS WITH NORMALITY ASSUMPTIONS

In this section we will consider the cases of the balanced model where the A_{ki_k} and $e_{i_1 i_2 \ldots i_n}$ are distributed as follows:

(1) A_{ki_k} are normally and independently distributed with means zero and variances σ_k^2,

(2) $e_{i_1 i_2 \ldots i_n}$ are normally and independently distributed with mean zero and variance σ^2.

This being the case, the balanced designs will admit a Z system having the properties:

(1) Z_{01} is normally distributed with mean μ and variance σ^2.

(2) Z_{ij} are normally and independently distributed with means zero and variance σ_i^2.

THEOREM I. Let Z_{01} be distributed normally with mean μ and variance σ_0^2 and let Z_{ij} be distributed normally with mean zero and variance σ_i^2 where

$$\sigma_0^2 = \sum_{i=1}^{m} k_i \sigma_i^2 ;$$

$j = 1, 2, \ldots n_i$;

$i = 1, 2, \ldots m$; and

all Z_{ij} are independent.

The best (minimum variance) unbiased estimate of

$$L = \sum_{i=1}^{m} g_i \sigma_i^2 + g_0 \mu \text{ (g_i are known constants) is}$$

$$L' = \sum_{i=1}^{m} g_i \hat{\sigma}_i^2 + g_0 Z_{01}, \text{ where } \hat{\sigma}_i^2 = \sum_{j=1}^{n_i} Z_{ij}^2 \Big/ n_i .$$

PROOF: The joint density of the Z_{ij} is

$$f = f(Z_{01}, Z_{11}, \ldots Z_{mn_m}) = \left(\frac{1}{2\pi}\right)^{\frac{N+1}{2}} \frac{1}{\sigma_0} \prod_{i=1}^{n} \sigma^{-n_i} \quad h \tag{7.1}$$

where

$$h = \exp - \frac{1}{2}\left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{Z_{ij}^2}{\sigma_i^2} + \frac{(Z_{01}-\mu)^2}{\sigma_0^2}\right], \tag{7.2}$$

and $N = \prod_{i=1}^{m} n_i$.

From the functional form of 7.1, it is clear that $\hat{\sigma}_i^2$ and Z_{01} form a set of jointly sufficient statistics for the σ_i^2 and μ respectively.

C. R. Rao (8) has proved that if a sufficient set of statistics $T_1, \ldots, T_q$ exists for the parameters $Q_1, Q_2, \ldots Q_q$, then the minimum variance estimator of a function of the parameters is an explicit function of the sufficient set of statistics.

Therefore the minimum variance estimate of L can be written as

$$G = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i Z_{ij}^2}{n_i} + g_0 Z_{01} + p, \tag{7.3}$$

where p is an arbitrary function of $(Z_{01}, \ldots Z_{mn_m})$ and the g_i are constants independent of the Z_{ij}, σ_i^2, and μ. Since G is an unbiased estimate of L, we must have (where E denotes mathematical expectation)

$$E(G) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i \sigma_i^2}{n_i} + g_0 \mu. \tag{7.4}$$

Taking the expected value of 7.3, we have

$$E(G) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} g_i \frac{E(Z_{ij}^2)}{n_i} + g_0 E(Z_{01}) + E(p)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n_i} g_i \frac{\sigma_i^2}{n_i} + g_0 \mu + E(p). \tag{7.5}$$

Hence $E(p) = 0$. (7.6)

Expressing 7.6 in integral form, we have

$$\int \cdots \int_R p\ f\ dZ = 0, \tag{7.7}$$

where dZ denotes the differential of density and R is the region over which the Z_{ij} are defined. We can write 7.7 as

$$\int \cdots \int_R p\ h\ dZ = 0. \tag{7.8}$$

Differentiating equation 7.8 with respect to σ_t^2 gives

$$\int \cdots \int_R p\ h \left[\frac{\sum_{j=1}^{n_t} Z_{tj}^2}{2\sigma_t^4} + \frac{(Z_{01} - \mu)^2}{2\sigma_0^4} \right] dZ = 0. \tag{7.9}$$

Differentiating equation 7.8 with respect to μ gives

$$\int \cdots \int_R p\ h \frac{(Z_{01} - \mu)}{\sigma_0^2}\ dZ = 0 \tag{7.10}$$

or

$$\int \cdots \int_R p\ h\ Z_{01}\ dZ - \mu \int \cdots \int_R p\ h\ dZ = 0.$$

Now the second integral in 7.10 vanishes by using 7.8. Hence

$$\int \cdots \int_R p\ h\ Z_{01}\ dZ = 0. \tag{7.11}$$

Differentiating equation 7.11 with respect to μ and expanding, we have

$$\int \cdots \int_R p\ h\ Z_{01}^2\ dZ - \mu \int \cdots \int_R p\ h\ Z_{01}\ dZ = 0. \tag{7.12}$$

The second integral in 7.12 vanishes by 7.11. Hence

$$\int \cdots \int_R p\, h\, Z_{01}^2\, dZ = 0. \tag{7.13}$$

Writing equation 7.9 as

$$\int \cdots \int_R p\, h \frac{\sum_{j=1}^{n_t} Z_{tj}^2}{\sigma_t^4}\, dZ + \frac{kt}{\sigma_0^2}\left[\int \cdots \int_R Z_{01}^2\, p\, h\, dZ \right.$$

$$\left. - 2\mu \int \cdots \int_R Z_{01}\, p\, h\, dZ + \mu^2 \int \cdots \int_R p\, h\, dZ \right]. \tag{7.14}$$

We see that the bracketed terms vanish in view of 7.8, 7.11, and 7.13. Thus

$$\int \cdots \int_R p\, h \sum_{j=1}^{n_t} Z_{tj}^2\, dZ = 0. \tag{7.15}$$

Consider now the variance of G. We have

$$\mathrm{Var}\ (G) = \mathrm{Var}\left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} Z_{ij}^2 + g_0\, Z_{01} + p \right]$$

$$= \mathrm{Var}\left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} Z_{ij}^2 + g_0\, Z_{01} \right] + 2\ \mathrm{Cov}\left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} Z_{ij}^2 \right.$$

$$\left. + g_0\, Z_{01};\ p \right] + \mathrm{Var}\ (p). \tag{7.16}$$

The covariance term is given by

$$\mathrm{Cov}\left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} Z_{ij}^2 + g_0\, Z_{01};\ p \right]$$

$$= \int \cdots \int_R \left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} Z_{ij}^2 + g_0\, Z_{01} - \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{g_i}{n_i} \sigma_i^2 - g_0 \mu \right] pf\, dZ. \tag{7.17}$$

Expanding equation 7.17 and using equations 7.8, 7.13, and 7.15, we have

$$\mathrm{Cov}\left[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{g_i}{n_i}Z_{ij}^2 + g_0 Z_{01};\ p\right] = 0. \tag{7.18}$$

Thus

$$\mathrm{Var}\ (G) = \mathrm{Var}\left[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{g_i}{n_i}Z_{ij}^2 + g_0 Z_{01}\right] + \mathrm{Var}\ (p). \tag{7.19}$$

Since both terms on the right are positive, the variance of G will be a minimum when Var (p) = 0. Thus, since E(p) = 0, we have

$$p = 0. \tag{7.20}$$

∴ the best unbiased estimate of

$$L = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\frac{g_i}{n_i}\sigma_i^2 + g_0\,\mu \tag{7.21}$$

is

$$L = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\frac{g_i}{n_i}Z_{ij}^2 + g_0 Z_{01}\,. \tag{7.22}$$

VIII. SUMMARY

This thesis is concerned primarily with the investigation of the properties of the analysis of variance estimates of the variance components in balanced linear models with random effects.

The analysis of variance estimates are obtained by equating the observed and expected mean squares and solving the resulting system of equations for the variance components. The balanced linear Y model with random effects is defined as the special case of Model III (i.e. the model having all effects random) which admits a transformation to an orthogonal uncorrelated linear Z model.

It has been shown in this thesis that:

(1) For a balanced model, the best (minimum variance) unbiased quadratic estimate of any linear combination of the variance components is the same linear combination of the analysis of variance estimates of the variance components.

(2) For a balanced model with normally distributed effects, the best unbiased estimate of any linear combination of the variance components is the same linear combination of the analysis of variance estimates of the variance components.

(3) The following are balanced models:

(a) Completely Randomized

(b) Randomized Block

(c) Latin Square

(d) Graeco-Latin Square

(e) General Hierarchal

(f) Split Plot

(g) Split ... Split Plot

(h) General Cross Classification

BIBLIOGRAPHY

1. Cramér, Harald. Mathematical Methods of Statistics, pp. 109-118, Princeton University Press, Princeton, N. J., 1946.

2. Crump, S. Lee. "The estimation of variance components in analysis of variance." Biometrics Bulletin, 2:91-104, 1938.

3. Daniels, H. E. "The Estimation of Variance Components," Supplement to The Journal of The Royal Statistical Society, 6:186-197, 1939.

4. ----- "Some Problems of Statistical Interest in Wool Research," Supplement to The Journal of The Royal Statistical Society, 5: 89-128, 1938.

5. Graybill, Franklin A. "On Quadratic Estimates of Variance Components," Unpublished Ph. D. Thesis, Iowa State College Library, Ames, Iowa, 1952.

6. Hsu, P. L. "On the Best Unbiased Quadratic Estimate of the Variance," Statistical Research Memoirs, 2:91-104, 1938.

7. Kempthorne, O. The Design and Analysis of Experiments, pp. 104-113, John Wiley and Sons, Inc., New York, 1952.

8. Rao, C. R. "Information and Accuracy Attainable in the Estimation of Statistical Parameters," Bulletin Calcutta Mathematical Society, Vols. 36-38, 1944-1946, pp. 81-91.

VITA

Albert William Wortham
candidate for the degree of
Doctor of Philosophy

Thesis: ON ESTIMATES OF VARIANCE COMPONENTS

Major: Mathematics

Biographical: The writer was born in Athens, Texas, January 18, 1927, the son of William Albert and Maggie Wortham. He attended grade school in Malakoff, Texas, junior high and high school in Athens, Texas.

Upon graduating from high school in May of 1944, he enlisted in the U. S. Navy and was called to active duty in September of 1944. He was assigned to the hospital corps as a Seaman 1st Class and was sent to the Hospital Corps School in San Diego, California, for eight weeks of intensive training. By July of 1946, the writer had received two promotions and was discharged to inactive duty as a Pharmacist Mate, Second Class.

In the fall of 1946, he returned to East Texas State Teachers College at Commerce, Texas, where three years of undergraduate work had been completed before graduating from high school and enlistment in the Navy. In May of 1947, he completed the work for the Bachelor of Arts degree at East Texas State Teachers College. The major subject for this degree was mathematics. After graduation, he accepted a fellowship in the Mathematics Department of Oklahoma Agricultural and Mechanical College and began graduate study in mathematics and statistics. He completed the Master of Science degree in May, 1949, and remained in school until March of 1951, continuing graduate study towards the Doctor of Philosophy degree. During this four-year period the writer taught mathematics under a fellowship and as an instructor, worked as a Research Assistant in the Mathematics Department in connection with an Office of Naval Research Contract, and taught and assisted in consulting in the Statistical Laboratory.

In March of 1951, the writer accepted a position of Senior Project Analytical Engineer with Chance Vought Aircraft in Dallas, Texas. While employed by Chance Vought Aircraft, he was an active partner in a statistical consulting business in the Dallas-Fort Worth area. After two and one-half years employment by Chance Vought, he was granted a leave of absence in September, 1953, and returned to Oklahoma Agricultural and Mechanical College to continue work on the Doctor of Philosophy degree. From September, 1953, to May, 1954, he was employed by the Mathematics Department as a temporary instructor.

He is a member of the American Mathematical Society, Mathematical Association of America, Institute of Mathematical Statistics, American Statistical Association, and the American Society for Quality Control. He is also a member of the following honorary societies: Sigma Xi, Pi Mu Epsilon, and Alpha Sigma Eta.

THESIS TITLE: ON ESTIMATES OF VARIANCE COMPONENTS

AUTHOR: Albert William Wortham

THESIS ADVISERS: Dr. F. A. Graybill
Dr. O. H. Hamilton

The content and form have been checked and approved by the author and thesis advisers. The Graduate School Office assumes no responsibility for errors either in form or content. The copies are sent to the bindery just as they are approved by the author and faculty advisers.

TYPIST: Jo Ellen Gurule