

ESTIMATION OF VARIANCE COMPONENTS
IN A THREE-WAY CLASSIFICATION

by

HOU-PEN CHEN

B. A. National Taiwan University, 1954

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1965

Approved by:

A. M. Zeger
Major Professor

L13
2668
H4
1965
C518
C12

ii

TABLE OF CONTENTS

INTRODUCTION	1
BASIC THEORY	4
ESTIMATION OF VARIANCE COMPONENTS	8
Unadjusted sums of squares	8
"Correcting" the data for mixed model.	13
General least squares.	16
Method of fitting constants.	27
Weighted squares of means.	29
DISCUSSION	34
REFERENCES	35
ACKNOWLEDGEMENTS.	37

INTRODUCTION

The statistical technique known as "analysis of variance" is a commonly used statistical method by which estimates of a number of variances are made and by which the significance of the differences between these estimates is determined. This technique was introduced by R. A. Fisher in 1920 to partition the variance of the observations into components ascribable to different causes as well as to obtain the analysis and tests of significances of treatment effects in agricultural and biological research.

There are two main purposes of using analysis of variance: 1) to estimate a set of parameters in a model and to test the null hypothesis that a linear subset of those parameters is equal to zero; and 2) to be used as a basis for estimation of variance components.

Because of the different purposes of analysis of variance techniques, different fundamental assumptions are made. Based on these two kind of uses, Eisenhart (1947) denoted the two fundamentally different types of analysis of variance as fixed models and random models respectively, commonly known as Model I and Model II. With fixed effects the investigator is interested in the computation of means, regression coefficients, standard errors, tests of significance, orthogonal comparisons, interactions and mean separation procedures; whereas, with random effects one is interested in the estimation of variance components.

The elementary theory of variance components analysis has been discussed by Daniels (1939), Crump (1946) and Eisenhart (1947).

Cochran and Crump (1946) have considered briefly the variance components problem in a balanced one-way classification. In higher-way classifications, a number of different methods of analysis are in common use. Chief among these are the so called methods of fitting constants and weighted squares of means described by Yates (1934) for experiments based on Model I. Estimates of variance components may be obtained by the device of equating observed and expected mean squares.

Estimating variance components for a three-way classification with unequal subclass numbers has discussed by C. R. Henderson (1953). He assumed some of the first order interaction effects and second order interaction effects negligible and considered the mixed model case because a random model is not always appropriate. For example, data related to year effects should be regarded as fixed rather than random. Henderson discussed three different methods of calculating variance components. The method which is simplest in computation leads to biased estimates. One of the others yields unbiased estimates, but the computation is laborious.

More recently, Norman Bush and R. L. Anderson (1963) described and illustrated an analytic procedure for obtaining variances of estimates of variance components for the two-way classification and developed this procedure in a general multiway classification, when there are unequal numbers of observations in the subclasses. Three unbiased estimating procedures for a two-way classification are discussed and compared by them. Two of the procedures are based on Yates's methods of fitting constants and weighted squares of means.

The third procedure which uses unadjusted sums of squares was developed by C. R. Henderson (1953), and extended by LeRay and Gluckowski (1961) explicitly for experiments based on Model II.

The purpose of this report is to describe and illustrate, with the aid of a numerical example, some different unbiased methods for estimating variance components by using a complete three-way classification linear model with unequal subclass numbers.

$$y_{hijk} = \mu + a_h + b_i + c_j + (ab)_{hi} + (ac)_{hj} + (bc)_{ij} \\ + (abc)_{hij} + e_{hijk}$$

where μ : the population mean;

a, b, c : the population main effects;

$(ab), (ac), (bc), (abc)$: the population two factor and three factor interaction effects;

e : a random observation error assumed normally and independently distributed with a mean zero and σ_e^2 .

$h = 1, 2, \dots, p$;

$i = 1, 2, \dots, l$;

$j = 1, 2, \dots, m$;

$k = 1, 2, \dots, n_{hij}$

BASIC THEORY

Consider the general-linear-hypothesis model of full rank. It can be written in vector form:

$$Y = X B + e$$

where

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix},$$

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

In detail, each random variable y_i depends on p known quantities $x_{i1}, \dots, x_{ip}$ and on p unknown parameters $b_1, \dots, b_p$ where $x_{ij} = 1$ or 0 for all i, j .

The method of least squares may be used to estimate B and σ^2 , where $\sigma^2 = \text{var}(y_i)$ ($i = 1, 2, \dots, n$). Thus, the vector B is to be chosen so as to minimize

$$e_1 = e'e = (Y - XB)'(Y - XB),$$

which is accomplished by letting

$$\frac{\partial(e'e)}{\partial B} = 2X'Y - 2X'XB = 0.$$

Now $X'XB = X'Y$ is called the normal equation. The least-squares estimate of B is

$$\hat{B} = S^{-1} X' Y$$

where $S = X'X$. The unbiased estimate of σ^2 based on the least squares estimate of B is given by

$$\hat{\sigma}^2 = \frac{(Y - XB)'(Y - XB)}{n - p} = \frac{Y'(I - X S^{-1} X')Y}{n - p}$$

By the Gauss-Markoff Theorem, the best (minimum-variance) linear (linear functions of the Y_1) unbiased estimate of B is given by least squares; that is, $\hat{B} = S^{-1}X'Y$ is the best linear unbiased estimate of B .

Consider again the model $Y = XB + e$. It may be partitioned so that

$$Y = X_1 r_1 + X_2 r_2 + e$$

where r_1 is of dimension $r \times 1$, r_2 of dimension $(p - r) \times 1$.

In the analysis of variance, the total sum of squares is $Y'Y$ and $\hat{B}'X'Y$ is the reduction due to $\hat{B}$ (denoted as $R(B)$). Thus $R(B)$ consists of the product of the vector $\hat{B}'$, which is the solution of the normal equations, and the corresponding elements of the right-hand side; i. e., $X'Y$.

For the partitioned model, the normal equations are:

$$\begin{pmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} X_1' Y \\ X_2' Y \end{pmatrix}$$

If only the elements in r_2 are used in the model from which we obtain the normal equations, the reduced normal equations

$$X_2' X_2 \tilde{r}_2 = X_2' Y$$

comes out, from which we get $\tilde{r}_2' X_2' Y = R(r_2)$ (unadjusted). The quantity $\hat{B}'X'Y - \tilde{r}_2' X_2' Y$ will be called reduction due to r_1 adjusted for r_2 (denoted as $R(r_1 | r_2)$), such that

$$R(r_1 | r_2) = R(B) - R(r_2).$$

If $X_1' X_2 = 0$, the $X'X$ matrix is diagonal in blocks and r_1 will be said to be orthogonal to r_2 . Generally speaking, if $b_1, \dots, b_p$ are all orthogonal, then $X'X$ is a diagonal matrix, the reduction due to any set of the B adjusted for any other set is simply the reduction due to that particular set of the B ignoring the other set.

In the case where X is not of full rank, this is equivalent to saying $X'X$ is singular and has no inverse, and an examination of the system to see whether a solution exists is necessary. Since $X'X$ is of dimension $p \times p$ and rank $k < p$, there are an infinite number of different vectors B that satisfy $X'XB = X'Y$. It might be interesting to see whether there are any linear functions of Y that give rise to unbiased estimates of B . The linear combination $\lambda'B$ (λ is a vector of known constants) is estimable if and only if there exists a solution $Sr = \lambda$. Also it should be noted that there are exactly K linearly independent estimable functions, where K is the rank of X . Therefore, in the light of least squares normal equations, some restrictions have to be imposed in order to make $X'X$ of full rank. (Graybill (1961)).

If the one-way classification is taken for example, the model will be

$$y_{ij} = \mu + a_i + e_{ij}$$

a direct method of obtaining estimates of the a_i as deviations from μ is to impose the restriction on the least-squares equations that $\sum_{i=1}^p \hat{a}_i = 0$, the coefficients of one of the $\hat{a}_i$, say $\hat{a}_p$, must be subtracted from the coefficients of the other $\hat{a}_i$. One of the simplest

restrictions would be to simply delete the $\hat{a}_p$ equation and the column of coefficients for $\hat{a}_p$, and solve the remaining equations to obtain estimates of the unknowns. (Harvey (1960)).

The techniques developed for estimating variance components, which we shall illustrate, is based on these concepts.

ESTIMATION OF VARIANCE COMPONENTS

1). Unadjusted sums of squares:

Under Eisenhart's Model II, it is assumed that, except for μ , all elements of the model are uncorrelated variables with mean zero and variances $\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_{ab}^2, \sigma_{ac}^2, \sigma_{bc}^2, \sigma_{abc}^2$, or σ_e^2 . Sums of squares can be computed as in the standard analysis of variance of corresponding orthogonal data. In the method of unadjusted sums of squares suggested by Henderson (1953), one equates these computed sums of squares to their expectations and solves for the unknown variances.

Thus, the following quantities are computed:

$$T = \sum_h \sum_i \sum_j \sum_k y_{hijk}^2$$

$$A = \sum_h y_{h...}^2 / n_{h..}$$

$$B = \sum_i y_{.i..}^2 / n_{.i.}$$

$$C = \sum_j y_{...j}^2 / n_{...j}$$

$$AB = \sum_h \sum_i y_{hi..}^2 / n_{hi.}$$

$$AC = \sum_h \sum_j y_{h.j.}^2 / n_{h.j}$$

$$BC = \sum_i \sum_j y_{.ij.}^2 / n_{.ij}$$

$$ABC = \sum_h \sum_i \sum_j y_{hij.}^2 / n_{hij}$$

$$N = \sum_h \sum_i \sum_j n_{hij}$$

$$CF = y_{....}^2 / N$$

Next, the expectations of the above quantities are computed. Under the assumptions of Model II, the coefficient of μ^2 and the variances in these expectations are as shown in the following table:

	μ^2	σ_a^2	σ_b^2	σ_c^2	σ_{ab}^2	σ_{ac}^2	σ_{bc}^2	σ_{abc}^2	σ_e^2
T	N	N	N	N	N	N	N	N	N
A	N	N	K_1	K_2	K_1	K_2	K_3	K_3	P
B	N	K_4	N	K_5	K_4	K_6	K_5	K_6	L
C	N	K_7	K_8	N	K_9	K_7	K_8	K_9	M
AB	N	N	N	K_{10}	N	K_{10}	K_{10}	K_{10}	S_1
AC	N	N	K_{11}	N	K_{11}	N	K_{11}	K_{11}	S_2
BC	N	K_{12}	N	N	K_{12}	K_{12}	N	K_{12}	S_3
ABC	N	N	N	N	K_{13}	K_{13}	K_{13}	N	S_4
CF	N	K_{14}	K_{15}	K_{16}	K_{17}	K_{18}	K_{19}	K_{20}	1

where $N = \sum_h \sum_i \sum_j n_{hij}$

S_1 = number of filled subclasses.

The quantities $K_1, K_2, \dots, K_{20}$ are computed as follows:

$$K_1 = \sum_h \sum_i n_{hi.}^2 / n_{h..} \quad K_{10} = \sum_h \sum_i \sum_j n_{hij}^2 / n_{hi.}$$

$$K_2 = \sum_h \sum_j n_{h.j}^2 / n_{h..} \quad K_{11} = \sum_h \sum_i \sum_j n_{hij}^2 / n_{h.j}$$

$$K_3 = \sum_h \sum_i \sum_j n_{hij}^2 / n_{h..} \quad K_{12} = \sum_h \sum_i \sum_j n_{hij}^2 / n_{.ij}$$

$$K_4 = \sum_h \sum_i n_{hi.}^2 / n_{.i.} \quad K_{13} = \sum_h \sum_i \sum_j n_{hij}^2 / n_{hij}$$

$$K_5 = \sum_i \sum_j n_{.ij}^2 / n_{.i.} \quad K_{14} = \sum_h n_{h..}^2 / N$$

$$K_6 = \sum_h \sum_i \sum_j n_{hij}^2 / n_{.i.} \quad K_{15} = \sum_i n_{.i.}^2 / N$$

$$K_7 = \sum_h \sum_j n_{h.j}^2 / n_{..j} \quad K_{16} = \sum_j n_{..j}^2 / N$$

$$K_8 = \sum_i \sum_j n_{.ij}^2 / n_{..j} \quad K_{17} = \sum_h \sum_i n_{hi.}^2 / N$$

$$K_9 = \sum_h \sum_i \sum_j n_{hij}^2 / n_{..j} \quad K_{18} = \sum_h \sum_j n_{h.j}^2 / N$$

$$K_{19} = \frac{\sum_i \sum_j n_{ij}^2}{N}$$

$$K_{20} = \frac{\sum_h \sum_i \sum_j n_{hij}^2}{N}$$

If the data were orthogonal, the sums of squares in the analysis of variance would be:

$$\text{Among A} = A - CF$$

$$" \quad B = B - CF$$

$$" \quad C = C - CF$$

$$AB = AB - A - B + CF$$

$$AC = AC - A - C + CF$$

$$BC = BC - B - C + CF$$

$$ABC = ABC - AB - AC - BC + A + B + C - CF$$

$$\text{Error} = T - ABC$$

This method will now be applied to data taken from Henderson (1953):

(B) Herd	(C) Sire	(A) Year				Total
		1	2	3	4	
1	1	3 - 1414	2 - 981	2 - 862	5 - 1609	5 - 2395
	2		4 - 1766			6 - 2628
	3					5 - 1609
2	1	1 - 404	3 - 1270	5 - 2109 4 - 1563	2 - 740	4 - 1674
	2					5 - 2109
	3					6 - 2303
3	1		3 - 1705	2 - 1134		3 - 1705
	2		4 - 2310			6 - 3444
4	1	3 - 1113	5 - 1951	3 - 1291	6 - 2457	8 - 3064
	3					9 - 3748
T o t a l		7 - 2931	21 - 9983	16 - 6959	13 - 4806	57 - 24679

The computations of K's and the expectations of these quantities are presented in the following tables:

Table 1

	μ^2	σ_a^2	σ_b^2	σ_c^2	σ_{ab}^2	σ_{ac}^2	σ_{bc}^2	σ_{abc}^2	σ_e^2	RHM
T	57	57	57	57	57	57	57	57	57	11,124,007
A	57	57	19.51	39.22	19.51	39.22	15.10	15.10	4	10,776,451
B	57	21.49	57	24.04	21.49	15.16	24.04	15.16	4	10,893,666
C	57	30.33	18.51	57	11.62	30.33	18.51	11.62	3	10,776,273
AB	57	57	57	46.46	57	46.46	46.46	46.46	14	10,968,029
AC	57	57	22.57	57	22.57	57	22.57	22.57	6	10,146,183
BC	57	37.35	57	57	37.35	37.35	57	37.35	10	10,970,369
ABC	57	57	57	57	57	57	57	57	17	10,973,517
CF	57	16.05	14.93	19.11	5.28	10.19	6.19	3.88	1	10,685,141

Table 2

[illegible]

where RSM is the total sum of squares (uncorrected) of corresponding factors.

The solution to these eight equations presented in Table 2 is:

$$\sigma_a^2 = 41,942.3135$$

$$\sigma_b^2 = 10,469.0888$$

$$\sigma_c^2 = 26,020.9606$$

$$\sigma_{ab}^2 = -39,824.5215$$

$$\sigma_{ac}^2 = -78,729.3598$$

$$\sigma_{bc}^2 = -23,208.6645$$

$$\sigma_{abc}^2 = 75,351.1685$$

$$\sigma_e^2 = 3,762.2500$$

2). "Correcting" the data for mixed model:

The bias in estimating variance components due to the assumption that fixed elements of the model are random variables can be eliminated by "correcting" the data for a mixed model. It was used by Hazel and Terrill (1945) on data which were orthogonal except for the fixed effects. By using this method, Henderson (1953) obtained least squares estimates of fixed effects, used the corrected data in place of the original data and proceeded as in the previously discussed method.

Consider, in the general case, the linear model

$$y_a = \sum_{i=1}^p b_i x_{ia} + e_a \quad a = 1, 2, \dots, N$$

where x 's are 1 or 0; the e 's are uncorrelated with mean zero and variance σ_e^2 .

If the b 's are all fixed, the least squares equations for estimating them are:

$$\sum_{i=1}^p C_{1i} b_i = Y_1$$

$$\sum_{i=1}^p C_{2i} b_i = Y_2$$

...

$$\sum_{i=1}^p C_{pi} b_i = Y_p$$

where $C_{ij} = \sum_{a=1}^N x_{ia} x_{ja}$ and $Y_i = \sum_{a=1}^N x_{ia} y_a$.

It is sometime necessary to impose one or more linear restriction on the estimates in order to obtain a solution to these equations.

Now suppose the $b_1, \dots, b_s$ are fixed and also that for all $i = 1, 2, \dots, s$ and the least square estimates have the property that

$E(\hat{b}_1 - b_1)^2 = K_1 e_e^2$, then the data can be corrected as follows:

$$z_a = y_a - \sum_{i=1}^s \hat{b}_1 x_{ia}$$

To illustrate this method with our data, Henderson would assume that the a 's were fixed. First the least squares estimates of the a 's were computed. This is done most simply by estimating them jointly with a 's and d 's to form the equations, where a 's refer to the fixed effect (year, in this case) while d 's refer to the interaction of the other two random effects.

Looking at the data, these equations are:

$$\begin{aligned} 7\hat{a}_1 + 3\hat{d}_{11} + \hat{d}_{21} + 3\hat{d}_{41} &= 2931 \\ 21\hat{a}_2 + 2\hat{d}_{11} + 4\hat{d}_{12} + 3\hat{d}_{21} + 3\hat{d}_{31} + 4\hat{d}_{32} + 5\hat{d}_{41} &= 9983 \\ 16\hat{a}_3 + 2\hat{d}_{12} + 5\hat{d}_{22} + 4\hat{d}_{23} + 2\hat{d}_{32} + 3\hat{d}_{43} &= 6959 \\ 13\hat{a}_4 + 5\hat{d}_{13} + 2\hat{d}_{23} + 6\hat{d}_{43} &= 4806 \\ 3\hat{a}_1 + 2\hat{a}_2 + 5\hat{d}_{11} &= 2395 \\ 4\hat{a}_2 + 2\hat{a}_3 + 6\hat{d}_{12} &= 2628 \\ &\dots\dots\dots \\ 3\hat{a}_3 + 6\hat{a}_4 + 9\hat{d}_{43} &= 3748 . \end{aligned}$$

These equations reduce to the ones shown in Table 3.

Table 3

$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	
3.825	-3.825	0	0	-73.500
-3.825	6.492	-2.667	0	101.500
0	-2.667	6.000	-3.333	41.333
0	0	-3.333	3.333	-69.333

One restriction must be imposed before a solution is obtainable.

Let $\hat{a}_4 = 0$. Then the solution is:

$$\hat{a}_1 = 12.08 \quad \hat{a}_2 = 31.30$$

$$\hat{a}_3 = 20.80 \quad \hat{a}_4 = 0$$

Using the $\hat{a}$'s to correct the data, the two-factor tables would be as shown in Table 4 - 6.

Table 4

Herd	Year				Total
	1	2	3	4	
1	1377.76	2559.20	820.40	1609	6366.36
2	391.92	1176.10	3484.80	740	5792.82
3	0	3795.90	1092.40	0	4888.30
4	1076.76	1794.50	1228.60	2457	6556.86
Total	2846.44	9325.70	6626.20	4806	23604.34

Table 5

Sire	Year				Total
	1	2	3	4	
1	2846.44	5500.10	0	0	8346.54
2	0	3825.60	3917.80	0	7743.40
3	0	0	2708.40	4806	7514.40
Total	2846.44	9325.70	6626.20	4806	23604.34

Table 6

Herd	Sire			Total
	1	2	3	
1	2296.16	2461.20	1609.00	6366.36
2	1568.02	2005.00	2219.80	5792.82
3	1611.10	3277.20	0	4888.30
4	2871.26	0	3685.60	6556.86
Total	8346.54	7743.40	7514.40	23604.34

The method of unadjusted sums of squares, which can be applied to this new corrected data, will lead to new estimates.

3). General least squares

The mathematical model of a three-way classification can be expressed in matrix notation as

$$Y = X B + e$$

and the normal equations are

$$X'XB = X'Y = \begin{bmatrix} G & T_a & T_b & T_c & T_{ab} & T_{ac} & T_{bc} & T_{abc} \end{bmatrix}'$$

where G is the grand total and T 's are the subtotal of corresponding effects indicating by subscripts. Under the assumptions of Model II, the A , B , C , AB , AC , BC and ABC and e are independently distributed as $N(0, \sigma_a^2 I)$, $N(0, \sigma_b^2 I)$, ..., $N(0, \sigma_{abc}^2 I)$ and $N(0, \sigma_e^2 I)$ respectively, where I is the identity matrix of suitable order.

Before applying this method to the example, it seems more convenient to take two classes each with three variables of classification to illustrate the techniques with the effects a_2 , b_2 , c_2 and all interactions associated with them deleted. (In general, any one class in each variable of classification can be deleted.)

Thus, the deleted $X'X$, deleted B' and deleted $X'Y$ can be expressed as:

$$W = X'X_{del.} = \begin{bmatrix} N & n_{1..} & n_{.1.} & n_{..1} & n_{11.} & n_{.11} & n_{.11} & n_{111} \\ & n_{1..} & n_{11.} & n_{.11} & n_{11.} & n_{.11} & n_{111} & n_{111} \\ & & n_{.1.} & n_{.11} & n_{11.} & n_{111} & n_{.11} & n_{111} \\ & & & n_{..1} & n_{111} & n_{.11} & n_{.11} & n_{111} \\ & & & & n_{11.} & n_{111} & n_{111} & n_{111} \\ & & & & & n_{.1.} & n_{111} & n_{111} \\ & & & & & & n_{.11} & n_{111} \\ & & & & & & & n_{111} \end{bmatrix}$$

where D_i ($i = 1, \dots, 7$) indicates the deleted column(s) of corresponding effects shown in parentheses. This can equivalently be rewritten as

$$\begin{bmatrix} W & D \end{bmatrix} \begin{bmatrix} B'_{del.} & a_2 & b_2 & c_2 & (ab)_{12} & (ab)_{21} & (ab)_{22} & (ac)_{12} \\ (ac)_{21} & (ac)_{22} & (bc)_{12} & (bc)_{21} & (bc)_{22} & (abc)_{121} \\ (abc)_{112} & (abc)_{122} & (abc)_{211} & (abc)_{212} & (abc)_{221} \\ (abc)_{222} \end{bmatrix}' = \begin{bmatrix} W & D \end{bmatrix} B^*$$

where B^* is the row matrix in the previous step.

To get the coefficients of the variance components, Harvey (1960) suggested to form the (abc) interaction. Consider

$$SS_{ABC_{adj. \text{ all effects}}} = SS_{R_{all}} - SS_{R_{\mu, a, b, c, ab, ac, bc.}}$$

The first term of the right hand side is the uncorrected sum of squares of the subclasses.

$$SS_{R_{all}} = SS_{Sub_{uncorrected}} = \sum_h \sum_i \sum_j y_{hij}^2 / n_{hij}$$

$$E(SS_{R_{all}}) = N\mu^2 + N\sigma_a^2 + N\sigma_b^2 + N\sigma_c^2 + N\sigma_{ab}^2 + N\sigma_{ac}^2 + N\sigma_{bc}^2 + N\sigma_{abc}^2$$

For the second term of right hand side, it can be shown that

$$W_1 = X'X_{del. \text{ for } abc} = \begin{bmatrix} N & n_{1..} & n_{.1.} & n_{..1} & n_{11.} & n_{.11} & n_{.11} \\ & n_{1..} & n_{11.} & n_{.11} & n_{11.} & n_{.11} & n_{111} \\ & & n_{.1.} & n_{.11} & n_{11.} & n_{111} & n_{.11} \\ & & & n_{..1} & n_{111} & n_{.11} & n_{.11} \\ & & & & n_{11.} & n_{111} & n_{111} \\ & & & & & n_{1.1} & n_{111} \\ & & & & & & n_{.11} \end{bmatrix}$$

$$D_1 = \begin{matrix} & D_{11} & D_{12} & D_{13} & & D_{14} & & D_{15} & & D_{16} \\ \begin{matrix} n_{2..} & n_{.2.} & n_{..2} & n_{12.} & n_{21.} & n_{22.} & n_{1.2} & n_{2.1} & n_{2.2} & n_{.12} & n_{.21} & n_{.22} \\ 0 & n_{12.} & n_{1.2} & n_{12.} & 0 & 0 & n_{1.2} & 0 & 0 & n_{112} & n_{121} & n_{122} \\ n_{21.} & 0 & n_{.12} & 0 & n_{21.} & 0 & n_{112} & n_{211} & n_{212} & n_{.12} & 0 & 0 \\ n_{2.1} & n_{.21} & 0 & n_{121} & n_{211} & n_{221} & 0 & n_{2.1} & 0 & 0 & n_{.21} & 0 \\ 0 & 0 & n_{112} & 0 & 0 & 0 & n_{112} & 0 & 0 & n_{112} & 0 & 0 \\ 0 & n_{121} & 0 & n_{121} & 0 & 0 & 0 & 0 & 0 & 0 & n_{121} & 0 \\ n_{211} & 0 & 0 & 0 & n_{211} & 0 & 0 & n_{211} & 0 & 0 & 0 & 0 \end{matrix} \end{matrix}$$

$$D_{17} = \begin{matrix} n_{112} & n_{121} & n_{122} & n_{211} & n_{212} & n_{221} & n_{222} \\ n_{112} & n_{121} & n_{122} & 0 & 0 & 0 & 0 \\ n_{112} & 0 & 0 & n_{211} & n_{212} & 0 & 0 \\ 0 & n_{121} & 0 & n_{211} & 0 & n_{221} & 0 \\ n_{112} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & n_{121} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & n_{211} & 0 & 0 & 0 \end{matrix} \quad U_1 = \begin{bmatrix} n_{111} \\ n_{111} \\ n_{111} \\ n_{111} \\ n_{111} \\ n_{111} \\ n_{111} \end{bmatrix}.$$

It is obvious that W_1 , U_1 are obtained from W by deleting the column and row corresponding to (abc); D_1 is obtained from D by deleting the last row. (In the general case we should delete all the rows and columns corresponding to the deleted effects).

Then, the normal equations are

$$\begin{bmatrix} W_1 & \hat{B}_1 \end{bmatrix}' = \begin{bmatrix} X'Y \end{bmatrix}_1'$$

$$\text{where } \begin{bmatrix} X'Y \end{bmatrix}_1' = \begin{bmatrix} X'Y \end{bmatrix}_{\text{del. for (abc)}} = \begin{bmatrix} G & T_a & T_b & T_c & T_{ab} & T_{ac} & T_{bc} \end{bmatrix}$$

$$\text{then } SS R_{\mu, a, b, c, ab, ac, bc} = \begin{bmatrix} X'Y \end{bmatrix}_1' W_1^{-1} \begin{bmatrix} X'Y \end{bmatrix}_1$$

$$\text{and } E(SS R_{\mu, a, b, c, ab, ac, bc}) = E \left(\begin{bmatrix} X'Y \end{bmatrix}_1' W_1^{-1} \begin{bmatrix} X'Y \end{bmatrix}_1 \right).$$

For $X'Y$ it is equivalent to write $\begin{bmatrix} W_1 & U_1 & D_1 \end{bmatrix} B^*$.

Hence, for calculating the expectation of $SS R_{,a,b,c,ab,ac,bc}$ the only interest is in the diagonal elements of $W_1, U_1' W_1^{-1} U_1, D_{11}' W_1^{-1} D_{11}, D_{12}' W_1^{-1} D_{12}, \dots, D_{17}' W_1^{-1} D_{17}$, the final result of $E(SS ABC_{adj. all})$ is

$$\begin{aligned} E(SS ABC_{adj. all}) = & (N - W_{122} - \text{tr } D_{11}' W_1^{-1} D_{11}) \sigma_a^2 \\ & + (N - W_{133} - \text{tr } D_{12}' W_1^{-1} D_{12}) \sigma_b^2 \\ & + (N - W_{144} - \text{tr } D_{13}' W_1^{-1} D_{13}) \sigma_c^2 \\ & + (N - W_{155} - \text{tr } D_{14}' W_1^{-1} D_{14}) \sigma_{ab}^2 \\ & + (N - W_{166} - \text{tr } D_{15}' W_1^{-1} D_{15}) \sigma_{ac}^2 \\ & + (N - W_{177} - \text{tr } D_{16}' W_1^{-1} D_{16}) \sigma_{bc}^2 \\ & + (N - \text{tr } U_1' W_1^{-1} U_1 - \text{tr } D_{17}' W_1^{-1} D_{17}) \sigma_{abc}^2 \end{aligned}$$

$E(MS ABC_{adj. all})$ is obtained by dividing this by the degree of freedom for the (abc) interaction and adding σ_e^2 to it. Note that the coefficient of $\sigma_a^2, \sigma_b^2, \dots, \sigma_{bc}^2$ should vanish in this equation.

Next, work out the coefficients of variance components for the first order interactions by considering that:

$$SS AB_{adj. \mu, a, b, c, ac, bc} = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ac, bc};$$

$$SS AC_{adj. \mu, a, b, c, ab, bc} = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ab, bc};$$

$$SS BC_{adj. \mu, a, b, c, ab, ac} = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ab, ac}.$$

The first terms of the right hand sides are already at hand. The second terms of the right sides are needed. It is enough to work out one of them since the others follow the same procedure. In

$$SS R_{\mu, a, b, c, ab, ac},$$

the necessary matrices are as follows:

$$W_2 = \begin{bmatrix} N & n_{1..} & n_{.1.} & n_{..1} & n_{11.} & n_{.11} \\ & n_{1..} & n_{11.} & n_{.11} & n_{11.} & n_{.11} \\ & & n_{.1.} & n_{.11} & n_{11.} & n_{111} \\ & & & n_{..1} & n_{111} & n_{.11} \\ & & & & n_{11.} & n_{111} \\ & & & & & n_{.11} \end{bmatrix} \quad U_2 = \begin{bmatrix} n_{.11} & n_{111} \\ n_{111} & n_{111} \\ n_{.11} & n_{111} \\ n_{.11} & n_{111} \\ n_{111} & n_{111} \\ n_{111} & n_{111} \end{bmatrix}$$

$$\text{and } B_2' = \begin{bmatrix} \mu & a_1 & b_1 & c_1 & (ab)_{11} & (ac)_{11} \end{bmatrix}$$

which is obtained from W and B_{del}' by deleting the corresponding elements for (bc) and (abc) .

$$D_2 = \begin{bmatrix} D_{21} & D_{22} & D_{23} & & D_{24} & & D_{25} \\ n_{2..} & n_{.2.} & n_{..2} & n_{12.} & n_{21.} & n_{22.} & n_{1.2} & n_{2.1} & n_{2.2} \\ 0 & n_{12.} & n_{.12} & n_{12.} & 0 & 0 & n_{1.2} & 0 & 0 \\ n_{21.} & 0 & n_{.12} & 0 & n_{21.} & 0 & n_{112} & n_{211} & n_{212} \\ n_{2.1} & n_{.21} & 0 & n_{121} & n_{211} & n_{221} & 0 & n_{2.1} & 0 \\ 0 & 0 & n_{112} & 0 & 0 & 0 & n_{112} & 0 & 0 \\ 0 & n_{121} & 0 & n_{121} & 0 & 0 & 0 & 0 & 0 \\ \\ & D_{26} & & & D_{27} \\ n_{.12} & n_{.21} & n_{.22} & n_{112} & n_{121} & n_{122} & n_{211} & n_{212} & n_{221} & n_{222} \\ n_{112} & n_{121} & n_{122} & n_{112} & n_{121} & n_{122} & 0 & 0 & 0 & 0 \\ n_{.12} & 0 & 0 & n_{112} & 0 & 0 & n_{211} & n_{212} & 0 & 0 \\ 0 & n_{.21} & 0 & 0 & n_{121} & 0 & n_{211} & 0 & n_{221} & 0 \\ n_{112} & 0 & 0 & n_{112} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & n_{121} & 0 & 0 & n_{121} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which is formed by deleting (bc) , (abc) from D and $[X'Y]_2' = [W_2 \ U_2 \ D_2] B^*$

$$\begin{aligned}
 SS R_{\mu,a,b,c,ab,ac} &= [X'Y]_2' W_2^{-1} [X'Y]_2 \\
 &= B^* [W_2 \ U_2 \ D_2]' W_2^{-1} [W_2 \ U_2 \ D_2] B^*
 \end{aligned}$$

$E(SS B_{adj. \mu,a,b,c,ab,ac})$ is obtained by taking the diagonal elements of W_2 , $U_{21}' W_2^{-1} U_{21}$, $U_{22}' W_2^{-1} U_{22}$, $D_{21}' W_2^{-1} D_{21}$..., $D_{27}' W_2^{-1} D_{27}$ as the coefficient of variance components of corresponding effects in the expectation form and subtracting from $E(SS R_{\mu,a,b,c,ab,ac,bc})$ and dividing by the degrees of freedom of the (bc) interaction, and adding σ_e^2 to it. Note that the coefficients of μ^2 , σ_a^2 , σ_b^2 , σ_c^2 , σ_{ab}^2 , σ_{ac}^2 vanish.

Last, to find the coefficient of the variance components for the main effects, the following formulas are applied to work with the same technique.

$$SS A_{adj. \mu,b,c,bc} = SS R_{\mu,a,b,c,bc} - SS R_{\mu,b,c,bc}$$

$$SS B_{adj. \mu,a,c,ac} = SS R_{\mu,a,b,c,ac} - SS R_{\mu,a,c,ac}$$

$$SS C_{adj. \mu,a,b,ab} = SS R_{\mu,a,b,c,ab} - SS R_{\mu,a,b,ab}.$$

Let us now illustrate this method with the data used before. It involves four classes for A (year), four for B (herd) and three for C (sire). When the restriction was imposed, suppose, a_4 , b_4 and c_3 and all their combinations are deleted. The form of $W = X'X_{del.}$ and D represented with the empty cells omitted is as follows:

W =

57	7	21	16	16	15	9	20	17	3	1	6	3	7	2	9	2	7	13	8	9
7	7	0	0	3	1	0	7	0	3	1	0	0	0	0	0	0	3	0	0	0
21	0	21	0	6	3	7	13	8	0	0	6	3	7	0	0	0	0	13	8	0
16	0	0	16	2	9	2	0	9	0	0	0	0	0	2	9	2	0	0	0	0
16	3	6	2	16	0	0	5	6	3	0	6	0	0	2	0	0	3	2	4	2
15	1	3	9	0	15	0	4	5	0	1	0	3	0	0	9	0	1	3	0	5
9	0	7	2	0	0	9	3	6	0	0	0	0	7	0	0	2	0	3	4	2
20	7	13	0	5	4	3	20	0	3	1	2	3	3	0	0	0	7	13	0	0
17	0	8	9	6	5	6	0	17	0	0	4	0	4	2	5	2	0	0	8	9
3	3	0	0	3	0	0	3	0	3	0	0	0	0	0	0	0	3	0	0	0
1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0
6	0	6	0	6	0	0	2	4	0	0	6	0	0	0	0	0	0	2	4	0
3	0	3	0	0	3	0	3	0	0	0	0	3	0	0	0	0	0	3	0	0
7	0	7	0	0	0	7	3	4	0	0	0	0	7	0	0	0	0	3	4	0
2	0	0	2	2	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0	2
9	0	0	9	0	9	0	0	5	0	0	0	0	0	0	9	0	0	0	0	5
2	0	0	2	0	0	2	0	2	0	0	0	0	0	0	2	0	0	0	0	2
7	3	0	0	3	1	0	7	0	3	1	0	0	0	0	0	0	7	0	0	0
13	0	13	0	2	3	3	13	0	0	0	2	3	3	0	0	0	0	13	0	0
8	0	8	0	4	0	4	0	8	0	0	4	0	4	0	0	0	0	0	8	0
9	0	0	9	2	5	2	0	9	0	0	0	0	0	2	5	2	0	0	0	9
5	3	2	0	5	0	0	5	0	3	0	0	0	0	0	0	0	3	2	0	0
6	0	4	2	6	0	0	0	6	0	0	4	0	0	2	0	0	0	0	4	2
4	1	3	0	0	4	0	4	0	0	1	0	3	0	0	0	0	1	3	0	0
5	0	0	5	0	5	0	0	5	0	0	0	0	0	0	5	0	0	0	0	5
3	0	3	0	0	0	3	3	0	0	0	0	0	3	0	0	0	0	3	0	0
6	0	4	2	0	0	6	0	6	0	0	0	0	4	0	0	2	0	0	4	2
3	3	0	0	3	0	0	3	0	3	0	0	0	0	0	0	0	3	0	0	0
1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0
2	0	2	0	2	0	0	2	0	0	0	2	0	0	0	0	0	0	2	0	0
4	0	4	0	4	0	0	0	4	0	0	4	0	0	0	0	0	0	0	4	0
3	0	3	0	0	3	0	3	0	0	0	0	3	0	0	0	0	0	3	0	0
3	0	3	0	0	0	3	3	0	0	0	0	0	3	0	0	0	0	3	0	0
4	0	4	0	0	0	4	0	4	0	0	0	0	4	0	0	0	0	0	4	0
2	0	0	2	2	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0	2
5	0	0	5	0	5	0	0	5	0	0	0	0	0	0	5	0	0	0	0	5
2	0	0	2	0	0	2	0	2	0	0	0	0	0	0	2	0	0	0	0	2

```

5 6 4 5 3 6 3 1 2 4 3 3 4 2 5 2
3 0 1 0 0 0 3 1 0 0 0 0 0 0 0 0
2 4 3 0 3 4 0 0 2 4 3 3 4 0 0 0
0 2 0 5 0 2 0 0 0 0 0 0 0 2 5 2
5 6 0 0 0 0 3 0 2 4 0 0 0 2 0 0
0 0 4 5 0 0 0 1 0 0 3 0 0 0 5 0
0 0 0 0 3 6 0 0 0 0 0 3 4 0 0 2
5 0 4 0 3 0 3 1 2 0 3 3 0 0 0 0
0 6 0 5 0 6 0 0 0 4 0 0 4 2 5 2
3 0 0 0 0 0 3 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0
0 4 0 0 0 0 0 0 2 4 0 0 0 0 0 0
0 0 3 0 0 0 0 0 0 0 3 0 0 0 0 0
0 0 0 0 3 4 0 0 0 0 0 3 4 0 0 0
0 2 0 0 0 0 0 0 0 0 0 0 0 2 0 0
0 0 0 5 0 0 0 0 0 0 0 0 0 0 5 0
0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 2
3 0 1 0 0 0 3 1 0 0 0 0 0 0 0 0
2 0 3 0 3 0 0 0 2 0 3 3 0 0 0 0
0 4 0 0 0 4 0 0 0 4 0 0 4 0 0 0
0 2 0 5 0 2 0 0 0 0 0 0 0 2 5 2
5 0 0 0 0 0 3 0 2 0 0 0 0 0 0 0
0 6 0 0 0 0 0 0 0 4 0 0 0 2 0 0
0 0 4 0 0 0 0 1 0 0 3 0 0 0 0 0
0 0 0 5 0 0 0 0 0 0 0 0 0 0 5 0
0 0 0 0 3 0 0 0 0 0 0 3 0 0 0 0
0 0 0 0 0 6 0 0 0 0 0 4 0 0 0
3 0 0 0 0 3 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0
0 4 0 0 0 0 0 0 4 0 0 0 0 0 0 0
0 0 3 0 0 0 0 0 0 3 0 0 0 0 0 0
0 0 0 0 3 0 0 0 0 0 3 0 0 0 0 0
0 0 0 0 4 0 0 0 0 0 4 0 0 0 0
0 2 0 0 0 0 0 0 0 0 0 0 2 0 0
0 0 0 5 0 0 0 0 0 0 0 0 5 0
0 0 0 0 0 0 0 0 0 0 0 0 0 2

```

 $X'Y =$

```

24679
2931
9983
6959
6632
6086
5149
8838
8181
1414
404
2747
1270
4015
862
3672
1134
2931
5907
4076
4105
2395
2628
1674
2109
1705
3444
1414
404
981
1766
1270
1705
2310
862
2109
1134

```


The difficult result is obtained when this technique is applied to this particular data. Because of the unsuited degree of freedom under this full three-way classification model, which is used here as an example, it is impossible to get the sums of squares and mean squares for their interactions.

The unhappy result for three effects comes out as below:

$$\begin{aligned} SS A_{adj. \mu, b, c, bc} &= SS R_{\mu, a, b, c, bc} - SS R_{\mu, b, c, bc} \\ &= 21,980,989.84 - 15,466,266.60 \\ &= 6,514,723.24 \end{aligned}$$

$$\begin{aligned} E(SS A_{adj. \mu, b, c, bc}) &= E(SS R_{\mu, a, b, c, bc}) - E(SS R_{\mu, b, c, bc}) \\ &= 14.63 \sigma_a^2 + 34.02 \sigma_{ab}^2 + 37.65 \sigma_{ac}^2 + 5,000,009.46 \sigma_{abc}^2, \end{aligned}$$

$$\begin{aligned} SS B_{adj. \mu, a, c, ac} &= SS R_{\mu, a, b, c, ac} - SS R_{\mu, a, c, ac} \\ &= 567,056,630,100,000100 - 29,760,589.27 \\ &= 567,056,600,339,410.73, \end{aligned}$$

$$\begin{aligned} E(SS B_{adj. \mu, a, c, ac}) &= E(SS R_{\mu, a, b, c, ac}) - E(SS R_{\mu, a, c, ac}) \\ &= -490.21 \sigma_b^2 - 845.83 \sigma_{ab}^2 - 441.49 \sigma_{bc}^2 - 456.81 \sigma_{abc}^2, \end{aligned}$$

$$\begin{aligned} SS C_{adj. \mu, a, b, ab} &= SS R_{\mu, a, b, c, ab} - SS R_{\mu, a, b, ab} \\ SS R_{\mu, a, b, c, ab} &= 15,195,614.17, \end{aligned}$$

$$E(SS R_{\mu, a, b, c, ab}) = 56.17 \sigma_c^2 + 83.04 \sigma_{ac}^2 + 40.19 \sigma_{bc}^2 + 42.85 \sigma_{abc}^2,$$

The $SS R_{\mu, a, b, ab}$ cannot be computed for this example.

4). Method of fitting constant

This method involves computation of mean squares by a conventional least squares analysis of non-orthogonal data. One equates the mean squares to their expectations and solves for the unknown variances. N. Bush and R. L. Anderson (1965) used the technique for a three factor design and applied it to the problem of estimating variance components. The sums of squares are computed as follows:

$$SS A = SS(A | B, C) = SS R_{\mu, a, b, c} - SS R_{\mu, b, c}$$

$$SS B = SS(B | A, C) = SS R_{\mu, a, b, c} - SS R_{\mu, a, c}$$

$$SS C = SS(C | A, B) = SS R_{\mu, a, b, c} - SS R_{\mu, a, b}$$

$$SS AB = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ac, bc}$$

$$SS AC = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ab, bc}$$

$$SS BC = SS R_{\mu, a, b, c, ab, ac, bc} - SS R_{\mu, a, b, c, ab, ac}$$

$$SS ABC = SS R_{all} - SS R_{\mu, a, b, c, ab, ac, bc}$$

With this method the sums of squares for error and the interaction remain the same as with the complete least-squares analysis (general least square method). Also the coefficient for the interaction mean squares remain the same. The difference is that in this method the sums of squares for three main effects are the same as when the interactions were assumed to be non-existent. The coefficients for σ_a^2 , σ_b^2 , σ_c^2 in the expectation of the mean squares for A, B, C, respectively, also remain the same as when interaction was disregarded.

Apparently, with the knowledge of the technique developed in

the last section, the result for our data can be easily obtained.
This result which is presented below seems a little better than the
previous one:

$$\begin{aligned} SS A_{adj. \mu, b, c} &= SS R_{\mu, a, b, c} - SS R_{\mu, b, c} \\ &= 10,708,039.83 - 22,664,122.40 \\ &= -11,956,082.57 , \end{aligned}$$

$$\begin{aligned} E(SS A_{adj. \mu, b, c}) &= E(SS R_{\mu, a, b, c}) - E(SS R_{\mu, b, c}) \\ &= 4.51 \sigma_a^2 + 18.33 \sigma_{ab}^2 + 17.56 \sigma_{ac}^2 + 5.62 \sigma_{abc}^2 , \end{aligned}$$

$$\begin{aligned} SS B_{adj. \mu, a, c} &= SS R_{\mu, a, b, c} - SS R_{\mu, a, c} \\ &= 10,708,039.83 - 12,266,277.05 \\ &= -1,558,237.22 , \end{aligned}$$

$$\begin{aligned} E(SS B_{adj. \mu, a, c}) &= E(SS R_{\mu, a, b, c}) - E(SS R_{\mu, a, c}) \\ &= 31.56 \sigma_b^2 + 8.14 \sigma_{ab}^2 + 12.81 \sigma_{bc}^2 + 9.19 \sigma_{abc}^2 , \end{aligned}$$

$$\begin{aligned} SS C_{adj. \mu, a, b} &= SS R_{\mu, a, b, c} - SS R_{\mu, a, b} \\ &= 10,708,039.83 - 16,162,807.63 \\ &= -5,454,767.80 , \end{aligned}$$

$$\begin{aligned} E(SS C_{adj. \mu, a, b}) &= E(SS R_{\mu, a, b, c}) - E(SS R_{\mu, a, b}) \\ &= 5.05 \sigma_c^2 + 17.46 \sigma_{ac}^2 + 2.89 \sigma_{bc}^2 + 5.29 \sigma_{abc}^2 . \end{aligned}$$

5). Weighted squares of means:

When all degrees of freedom among a set of subclasses are partitioned into a set of orthogonal comparisons, as is done in factorial analysis, the estimates of all constants are determined from the subclass means. In this case weights are easily determined that are useful for computing sums of squares, variances of estimates of constant and coefficients of variance components.

Bush and Anderson (1963) suggested for this procedure that sums of squares required are in the form of

$$SS_j = \underset{1 \times t}{\hat{r}'} \underset{s \times s}{C_j'} \left[\underset{s \times s}{C_j'} (A'A)^{-1} \underset{s \times t}{C_j} \right]^{-1} \underset{s \times t}{C_j'} \underset{t \times 1}{\hat{r}}$$

where $\hat{r}$ is simply a vector of the sample means for the non-empty cells of the design; $(A'A)^{-1}$ is a diagonal matrix in which the diagonal elements are the reciprocals of the cell frequencies; the C-matrix reflects a contrast between effects and is of rank s ; t is the number of occupied cells and $j = 1, 2, \dots$

$$\text{Let } Q_j = \underset{t \times t}{C_j'} \left[\underset{s \times s}{C_j'} (A'A)^{-1} \underset{s \times t}{C_j} \right]^{-1} \underset{s \times t}{C_j'} = \underset{s \times t}{C_j'} B \underset{s \times t}{C_j'}$$

$$t \times t$$

$$SS_j = \hat{r}' Q_j \hat{r},$$

as j varies, only the Q-matrix is altered. Regardless of the total number of observations in the experiment, the dimension of the Q-matrix depends only on the number of non-empty cells in the experiment. So far as a three-way classification is concerned, a numerous possible C'-matrix can be constructed, and usually it is convenient for us to choose that each row in C' refers to a contrast. Although

these are not orthogonal contrasts, orthogonal contrasts can be derived from them. If any cell is empty, the column associated with the empty cell is omitted in each C_j^i .

To apply this method for our numerical example, we have

$$(\hat{A}'\hat{A})^{-1} = \begin{bmatrix} 0.333 & & & & & & \\ & 1.000 & & & & & \\ & & 0.333 & & & & \\ & & & 0.500 & & & \\ & & & & 0.333 & & \\ & & & & & 0.200 & \\ & & & & & & 0.250 \\ & & & & & & & 0.250 \\ & & & & & & & & 0.500 \\ & & & & & & & & & 0.200 \\ & & & & & & & & & & 0.500 \\ & & & & & & & & & & & 0.250 \\ & & & & & & & & & & & & 0.333 \\ & & & & & & & & & & & & & 0.200 \\ & & & & & & & & & & & & & & 0.500 \\ & & & & & & & & & & & & & & & 0.167 \end{bmatrix}$$

$$\hat{r} = \begin{bmatrix} 471.333 & 404.000 & 371.000 & 490.500 & 423.333 & 568.333 & 390.200 \\ 441.500 & 577.500 & 431.000 & 421.800 & 567.000 & 390.750 & 430.333 \\ 321.800 & 370.000 & 409.500 \end{bmatrix}.$$

The C'-matrices may be chosen as:

$$A = \begin{pmatrix} 111 & 121 & 141 & 211 & 221 & 231 & 241 & 212 & 232 & 312 & 322 & 332 & 323 & 343 & 413 & 423 & 443 \\ 1 & 1 & 1 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & -1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$AB = \begin{pmatrix} 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix}$$

AC : fail

$$BC = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The expected values of the sums of squares are shown as follows:

$$E(A) = \text{tr}(V_a Q_a) \sigma_a^2 + \text{tr}(V_{ab} Q_a) \sigma_{ab}^2 + \text{tr}(V_{ac} Q_a) \sigma_{ac}^2 + \text{tr}(Q_a) \sigma_{abc}^2 \\ + (a-1) \sigma_e^2$$

$$E(B) = \text{tr}(V_b Q_b) \sigma_b^2 + \text{tr}(V_{ab} Q_b) \sigma_{ab}^2 + \text{tr}(V_{bc} Q_b) \sigma_{bc}^2 + \text{tr}(Q_b) \sigma_{abc}^2 \\ + (b-1) \sigma_e^2$$

$$E(C) = \text{tr}(V_c Q_c) \sigma_c^2 + \text{tr}(V_{ac} Q_c) \sigma_{ac}^2 + \text{tr}(V_{bc} Q_c) \sigma_{bc}^2 + \text{tr}(Q_c) \sigma_{abc}^2 \\ + (c-1) \sigma_e^2$$

$$E(AB) = \text{tr}(V_{ab} Q_{ab}) \sigma_{ab}^2 + \text{tr}(Q_{ab}) \sigma_{abc}^2 + (ab - 1) \sigma_e^2$$

$$E(AC) = \text{tr}(V_{ac} Q_{ac}) \sigma_{ac}^2 + \text{tr}(Q_{ac}) \sigma_{abc}^2 + (ac - 1) \sigma_e^2$$

$$E(BC) = \text{tr}(V_{bc} Q_{bc}) \sigma_{bc}^2 + \text{tr}(Q_{bc}) \sigma_{abc}^2 + (bc - 1) \sigma_e^2$$

$$E(ABC) = \text{tr}(Q_{abc}) \sigma_{abc}^2 ,$$

$$\text{where } V_a = \sigma_a^2 J_c \otimes J_b \otimes I_a$$

$$V_b = \sigma_b^2 J_c \otimes I_b \otimes J_a$$

$$V_c = \sigma_c^2 I_c \otimes J_b \otimes J_a$$

$$V_{ab} = \sigma_{ab}^2 J_c \otimes I_b \otimes I_a$$

$$V_{ac} = \sigma_{ac}^2 I_c \otimes J_b \otimes I_a$$

$$V_{bc} = \sigma_{bc}^2 I_c \otimes I_b \otimes J_a$$

$$V_{abc} = \sigma_{abc}^2 I_{abc} .$$

Note that, $\otimes$ indicates the direct product of two matrices and $I_p = I(P \times P)$; $J_p = J(P \times P)$ a matrix of all ones. These derivations can be simplified by noting that for a three-way classification,

$$V = \sigma_a^2 V_a + \sigma_b^2 V_b + \sigma_c^2 V_c + \sigma_{ab}^2 V_{ab} + \sigma_{ac}^2 V_{ac} + \sigma_{bc}^2 V_{bc} \\ t \times t \quad + \sigma_{abc}^2 V_{abc} + \sigma_e^2 V_e .$$

The dimensions of the matrices involved in these computations are dependent only on the number of occupied cells (t) and not the number of observations (n).

The result is presented as follows:

$$SS A = \hat{r}' Q_a \hat{r} = 2,906.22$$

$$SS B = \hat{r}' Q_b \hat{r} = 62,154.26$$

$$SS C = \hat{r}' Q_c \hat{r} = 3,330.31$$

$$SS AB = \hat{r}' Q_{ab} \hat{r} = 425,314.10$$

$$SS BC = \hat{r}' Q_{bc} \hat{r} = 36,992.00$$

$$SS ABC = \hat{r}' Q_{abc} \hat{r} = 36,992.00$$

The coefficients in the expected sums of squares are:

$$\text{tr}(V_a Q_a) = 18.34$$

$$\text{tr}(V_{ab} Q_a) = 0.0$$

$$\text{tr}(Q_a) = 8.04$$

$$\text{tr}(V_b Q_b) = 6.76$$

$$\text{tr}(V_{ab} Q_b) = 0.0$$

$$\text{tr}(V_{bc} Q_b) = 10.88$$

$$\text{tr}(Q_b) = 7.66$$

$$\text{tr}(V_c Q_c) = 5.94$$

$$\text{tr}(V_{bc} Q_c) = 7.44$$

$$\text{tr}(Q_c) = 7.44$$

$$\text{tr}(V_{ab} Q_{ab}) = 1.26$$

$$\text{tr}(Q_{ab}) = 7.76$$

$$\text{tr}(V_{bc} Q_{bc}) = 4.00$$

$$\text{tr}(Q_{bc}) = 4.00$$

$$\text{tr}(Q_{abc}) = 4.00$$

DISCUSSION

Some points are worthy of mention. By applying other methods it becomes clear why Henderson (1953) chose the model:

$$y_{hijk} = \mu + a_h + b_i + c_j + (bc)_{ij} + e_{hijk}$$

to apply to his numerical example. Application of a complete three-way classification model to this particular data is destined to failure. This is not surprising when one looks at the table for the AC interaction:

(C) Sire	(A) Year				Total
	1	2	3	4	
1	7	13	0	0	20
2	0	8	9	0	17
3	0	0	7	13	20
Total	7	21	16	13	57

The degree of freedom of this interaction is obviously zero and this prevents one from working with a complete model. Furthermore, this example has 31 empty cells out of total $4 \times 4 \times 3 = 48$.

The methods presented herein have their own peculiar characteristics. The first technique depends upon unadjusted sums of squares. An increasing σ_c^2 , say, causes an increasing $\text{var}(\sigma_b^2)$ and $\text{var}(\sigma_a^2)$, if σ_b^2 , σ_a^2 , and their interactions as well as σ_e^2 are held fixed. In the last two methods, the $\text{var}(\sigma_c^2)$ remains invariant over any change in σ_a^2 and σ_b^2 , and vice versa.

Where possible one should use the complete model since Harvey (1960) indicated that estimates of variance components obtained from least-squares analysis, presumably have smaller sampling errors than others.

REFERENCES

- Anderson, R. L. & Bush, N. : A Comparison of Three Different Procedures for Estimating Variance Components, *Technometrics*, vol. 5, 1963. p.p. 421-440.
- Anderson, R. L. & Bush, N. : Further Contributions to the Problem of Estimating Variance Components in the General Multiway Classification. Unpublished paper. p.p. 1-8.
- Crump, S. L. : The Estimation of Variance Components in Analysis of Variance, *Biom.*, vol. 2, 1946. p.p. 7-11.
- Crump, S. L. : The Present Status of Variance Components Analysis, *Biom.*, vol. 7, 1951. p.p. 1-16.
- Daniels, H. E. : The Estimation of Components of Variance, *J. Roy. Stat. Soc.*, suppl., vol. 6, 1939. p.p. 186-197.
- Eisenhart, C. : The Assumptions Underlying Analysis of Variance, *Biom.*, vol. 3, 1947. p.p. 1-21.
- Graybill, F. A. : An Introduction to Linear Statistical Models, McGraw-Hill, 1961.
- Hartley, H. O. : Programming Analysis of Variance for General Purpose Computers, *Biom.*, vol. 12, 1956. p.p. 110-122.
- ✓ Harvey, W. R. : Least Square Analysis of Data with Unequal Subclass Numbers. Published by U. S. Depart. of Agriculture Biom. Services. 1960.
- Hazel, L. N. : The Covariance Analysis of Multiple Classification Tables with Unequal Subclass Numbers, *Biom.*, vol. 6, 1946. p.p. 21-25.
- Henderson, C. R. : Estimating of Variances and Covariances Components, *Biom.*, vol. 9, 1953. p.p. 226-252.

Kramer, C. Y. : On the Analysis of Variance of a Two-way Classification with Unequal Subclass Numbers, *Biom.*, vol. 11, 1955. p.p. 441-452.

Le Roy, H. L. & Gluckowski, W. : Die Bestimmung der Varianz Komponenten in a b c-faktoren Versuch Mit Ungleichen Klassenfrequenzen. *Biometrische Zeitschrift*. 3, 1961. p.p. 73-91.

Schultz, E. F. Jr. : Rules of Thumb for Determining Expectations of Mean Squares in AOV. *Biom.*, vol. 11, 1955. p.p. 123-135.

Scheffe, H. : Alternative Models for Analysis of Variance, *Ann. Math. Stat.*, vol. 27, 1956. p.p. 251-271.

Scheffe, H. : A Mixed Model for the Analysis of Variance, *Ann. Math. Stat.*, vol. 27, 1956. p.p. 23-36.

Searle, S. R. : Matrix Methods in Components of Variance & Covariance Analysis, *Ann. Math. Stat.*, vol. 27, 1956. p.p. 737-748.

Stevens, W. L. : Statistical Analysis of a Non-orthogonal Tri-factorial Experiment, *Biometrika*, 1948. p.p. 346-367.

Yates, F. : The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes, *J. Am. Stat. Ass.* 29, 1934. p.p. 51-66.

ACKNOWLEDGEMENTS

I am deeply indebted to Prof. A. M. Feyerherm for serving my advisor for five semesters and for guidance and encouragement.

Appreciation is extended to Dr. H. C. Fryer, Head of Department of Statistics, for providing financial assistance through a research assistantship.

My sincere thanks are offered to Dr. Gray F. Krause for his great many helps. Thanks also go to Dr. Stanley Wearden for valuable suggestions.

ESTIMATION OF VARIANCE COMPONENTS
IN A THREE-WAY CLASSIFICATION

by

HOU-PEN CHEN

B. A. National Taiwan University, 1954

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1965

Approved by:

C. M. Zeyerhous
Major Professor

AN ABSTRACT OF
ESTIMATION OF VARIANCE COMPONENTS
IN A THREE-WAY CLASSIFICATION

The purpose of this report is to describe some different unbiased methods for estimating variance components by using a complete three-way classification linear model with unequal subclass numbers. The methods are discussed by C. R. Henderson, W. R. Harvey and R. L. Anderson.

Under Eisenhart's Model II, it is assumed that, except for μ , all elements of the model are uncorrelated variables with mean zero and variances σ_a^2 , σ_b^2 , ..., or σ_e^2 . In the method of unadjusted sums of squares the sums of squares are computed as in the standard analysis of variance of corresponding orthogonal data. This method suggested by Henderson equates computed sums of squares to their expectations and solves for the unknown variances.

The bias in estimating variance components due to the assumption that fixed elements of the model are random variables can be eliminated by "correcting" the data for a mixed model. Henderson obtained least squares estimates of fixed effects, and used the corrected data in place of the original data for computation.

Harvey discussed the general least squares method for a two-way classification. The mathematical model can be expressed by matrix notation.

$$Y = X B + e .$$

For a three-way classification, the normal equations are

$$X'XB = X'Y = \begin{bmatrix} G & T_a & T_b & T_c & T_{ab} & T_{ac} & T_{bc} & T_{abc} \end{bmatrix}'$$

where G is the grand total and T 's are subtotals of corresponding effects indicated by subscripts. Under the assumptions of Model II, the A , B , C , AB , AC , BC , ABC and e are independently distributed as $N(0, \sigma_a^2 I)$, $N(0, \sigma_b^2 I)$, ..., $N(0, \sigma_{abc}^2 I)$ and $N(0, \sigma_e^2 I)$ respectively, where I is identity matrix of suitable order.

The method of fitting constants involves computation of mean squares by a conventional least squares analysis of non-orthogonal data. The mean squares are equated to their expectations and the resulting system solved for the unknown variances. N. Bush and R. L. Anderson developed the technique for a three factor design and applied it to the problem of estimating variance components. The difference between this and the general least square method is that in this method the sums of squares for the three main effects are computed are the same as when the interactions were assumed to be non-existent.

Bush and Anderson using weighted squares of means require that sums of squares be in the form

$$SS_j = \hat{r}' \underset{1 \times t \text{ txs}}{C_j} \left[\underset{s \times s}{C_j' (A' A)^{-1} C_j} \right]^{-1} \underset{s \times t \text{ txl}}{C_j'} \hat{r}$$

where $\hat{r}$ is simply a vector of the sample means for the non-empty cells of the design; $(A' A)^{-1}$ is a diagonal matrix in which the diagonal elements are the reciprocals of the cell frequencies; the C' -matrix reflects a contrast between effects and is of rank s ; t is the number of occupied cells and $j = 1, 2, \dots$