RELATIONSHIPS BETWEEN A THREE-WAY CLASSIFICATION
DISPROPORTIONATE NUMBERS ANALYSIS OF VARIANCE AND
SEVERAL TWO-WAY CLASSIFICATION AND NESTED ANALYSES*

BU-152-M

W. T. Federer

August 1962

Abstract

A study was made to determine what information on the sources of variation for a three-way classification analysis of variance was extractable from nested and two-way classification analyses of variance. The case of unequal numbers of observations in the subclasses was considered. Although various analyses of variance produced various and different sums of squares, it was not possible to extract all of the information via the analyses considered in this investigation. Thus, if all the sums of squares for a higher-way classification analysis of variance is desired, it will be necessary to perform the complete analyses; if only sums of various sums of squares are suitable then these may be obtained as described herein.

* Biometrics Unit, mimeograph series, Plant Breeding Department, Cornell University.

RELATIONSHIPS BETWEEN A THREE-WAY CLASSIFICATION
DISPROPORTIONATE NUMBERS ANALYSIS OF VARIANCE AND
SEVERAL TWO-WAY CLASSIFICATION AND NESTED ANALYSES*

BU-152-M

W. T. Federer

August 1962

It is a simple matter to set forth the sums of squares in the analysis of variance for an unequal numbers three-way (or higher-way) classification in symbolic form using standard regression notation. The form of the notation is undisturbed by the distribution of numbers of observations over the subclasses. The computational form for the analysis of variance is, however, quite different for equal and proportional numbers in the subclasses as contrasted to unequal numbers in the subclasses. The latter situation may involve tedious computations, especially when some of the interactions are present.

To illustrate, suppose that the symbolic procedure described by Federer in volume 13, pages 333-362, 1957, of Biometrics is followed and suppose that the restrictions on the solutions to the normal equations are added via Lagrangian multipliers. Now, as the number of levels of a classification (factor or category) and as the number of classifications increases, it becomes obvious that the matrix to be inverted to obtain solutions rapidly gets beyond the capacity of a high speed computer, regardless of its size. Because of this and because most high speed computing establishments have programs for nested and two-way analyses of variance with unequal numbers in the subclasses, it is instructive to study what quantities are obtainable from two-way analyses and nested classifications.

The purpose of this paper will be to set forth the various quantities obtainable from these analyses. The full analysis for the sums of squares described in Table 5 of the paper cited above is not obtained in this paper. However, much of the information for this table is obtainable through other analyses. The notation and formulae used in the present paper is that of the above reference. This was done in order to shorten the present discussion. Although the results are confined to a three-way

* Biometrics Unit, mimeograph series, Plant Breeding Department,
Cornell University.

classification, the extension to four-way and higher-way classifications are possible. Also, even though attention herein is confined to y^2 products and for each X set equal to zero, a straightforward extension yields results for the xy and x^2 products in the above reference.

Nested Classification

The simplest procedure to use for unequal numbers analysis is a nested classification analysis of variance and estimation procedures. However, if classification (or factors) are of equal importance and if interactions are present, not all of the information is extracted from the three nested classification analyses of variance in Table 1. From these results it is seen that the sum of squares for any main effect and its associated interactions after eliminating the other effects is available using the three analyses in Table 1. The extension to more than three classifications is straightforward.

Collapsing Two Classifications into One

If any two of the three classifications are collapsed into a single classification, then this becomes equivalent to a two-way classification for which analyses are straightforward and may fall into a Case I, Case II, or Case III analysis as cited in the above reference. If all desired information can be extracted from such an analysis then everything is satisfactory. However, if more information on individual effects is desired then some of the analyses in the following sections may be more appropriate. In Table II the A and C classifications are treated as a single classification. Likewise, A and D in Table III and C and D in Table III are treated as single classifications. This procedure may be used for any number of classifications.

Table I. Analyses of variance for a nested classification

Source of variation	df	Sum of Squares
Correction for mean	1	$Y^2.../n...$
A (elim. mean; ign. all else)	a-1	$\sum_{i=1}^a Y_i^2.../n_i... - Y^2.../n...$
C:A = C within A (ign. D, AxD, CxD, and AxCxD)	a(c-1)	$\sum_{i=1}^a \left\{ \sum_{j=1}^c Y_{ij}^2.../n_{ij}... - Y_i^2.../n_i... \right\}$
D:C:A = D within C:A = D+AxD+CxD+AxCxD (elim. all else)	ac(d-1)	$\sum_{i=1}^a \sum_{j=1}^c \left\{ \sum_{h=1}^d Y_{ijh}^2.../n_{ijh}... - Y_{ij}^2.../n_{ij}... \right\}$
Within subclasses	n....-acd	$\sum_i \sum_j \sum_h \left\{ \sum_f Y_{ijhf}^2 - Y_{ijh}^2.../n_{ijh} \right\}$
Correction for mean	1	$Y....^2/n...$
C (elim. mean; ign. all else)	c-1	$\sum_{j=1}^c Y_{.j}^2.../n_{.j}... - Y^2.../n...$
D:C = D+CxD (elim. C and mean but ign. all others)	c(d-1)	$\sum_{j=1}^c \left\{ \sum_{h=1}^d Y_{.jh}^2.../n_{.jh}... - Y_{.j}^2.../n_{.j}... \right\}$
A:C:D = A+AxC +AxD+AxCxD (elim. all others)	cd(a-1)	$\sum_{j=1}^c \sum_{h=1}^d \left\{ \sum_{i=1}^a Y_{ijh}^2.../n_{ijh}... - Y_{.jh}^2.../n_{.jh}... \right\}$
Within Subclasses	n....-acd	$\sum_i \sum_j \sum_h \left\{ \sum_f Y_{ijhf}^2 - Y_{ijh}^2.../n_{ijh} \right\}$

Table I. Continued

Source of variation	df	Sum of Squares
Correction for mean	1	$Y_{...}^2/n_{...}$
D (elim. mean; ign. all else)	$d-1$	$\sum_{h=1}^d Y_{..h}^2/n_{..h} - Y_{...}^2/n_{...}$
A:D = A+AxD (elim. D and mean; ign. all else)	$d(a-1)$	$\sum_{h=1}^d \left\{ \sum_{i=1}^a Y_{ih}^2/n_{i..} - Y_{..h}^2/n_{..h} \right\}$
C:A:D = C+AxC+CxD +AxCxD (elim. all others)	$ad(c-1)$	$\sum_{h=1}^d \sum_{i=1}^a \left\{ \sum_{j=1}^c Y_{ijh}^2/n_{ijh} - Y_{i..}^2/n_{i..} \right\}$
Within subclasses	$n_{...}-acd$	$\sum_i \sum_j \sum_h \left\{ \sum_{f=1}^{n_{ijh}} Y_{ijhf}^2 - Y_{ijh}^2/n_{ijh} \right\}$

Table II. Analysis of variance for A and C combined and for D in a two-way classification

Source of variation	df	Sum of Squares
Correction for mean	1	$Y_{...}^2/n...$
A and C together (elim. mean; ign. all else)	ac-1	$\sum_{i=1}^a \sum_{j=1}^c Y_{ij.}^2/n_{ij.} - Y_{...}^2/n...$
D (elim. mean, A, C, AxC; ignoring all else)	d-1	Case I anova
Interaction = AxD+CxD +AxCxD (elim. all else)	(ac-1)(d-1)	by subtraction
Within subclasses	n...-acd	$\sum_{i=1}^a \sum_{j=1}^c \sum_{h=1}^d \left\{ \sum_{f=1}^{n_{ijh}} Y_{ijhf}^2 - Y_{ijh.}^2/n_{ijh} \right\}$
A+C+AxC (elim. all else)	ac-1	$\sum_i \sum_j w_{ij} \cdot \bar{y}_{ij.}^2$ $-\left(\sum_i \sum_j w_{ij} \cdot \bar{y}_{ij.} \right)^2 / \sum_i \sum_j w_{ij}$
D (elim. all else)	d-1	$\sum_h w_{...h} \bar{y}_{...h}^2$ $-\left(\sum_h w_{...h} \bar{y}_{...h} \right)^2 / \sum_h w_{...h}$

$$w_{...h} = \left(\frac{1}{a^2 c^2} \sum_i \sum_j \frac{1}{n_{ijh}} \right)^{-1} \quad w_{ij.} = \left(\frac{1}{d^2} \sum_h \frac{1}{n_{ijh}} \right)^{-1}$$

$$\bar{y}_{...h} = \frac{1}{ac} \sum_i \sum_j \bar{y}_{ijh} \cdot \quad \bar{y}_{ij.} = \frac{1}{d} \sum_h \bar{y}_{ijh} \cdot$$

[see pages 345 and 346 of reference cited for Case II anova]

Table III. Analysis of variance for A and D combined and for C in a two-way classification

Source of variation	df	Sum of Squares
Correction for mean	1	$Y^2 \dots / n \dots$
A and D together (elim. mean; ign. all else)	ad-1	$\sum_i \sum_h Y_{i \cdot h}^2 / n_{i \cdot h} - Y^2 \dots / n \dots$
C (elim. mean, A, D, AxD; ign. all else)	c-1	Case I anova
Interaction = AxC+CxD+ AxCxD (elim. all else)	(ad-1)(c-1)	by subtraction
Within subclasses	n...-acd	$\sum_i \sum_j \sum_h \left\{ \sum_f Y_{ijhf}^2 - Y_{ijh \cdot}^2 / n_{ijh} \right\}$
A+D+AxD (elim. all else)	ad-1	$\sum_i \sum_h w_{i \cdot h} \bar{y}_{i \cdot h}^2 - \left(\sum_i \sum_h w_{i \cdot h} \bar{y}_{i \cdot h} \right)^2 / \sum_i \sum_h w_{i \cdot h}$
C (elim. all else)	c-1	$\sum_j w_{\cdot j} \bar{y}_{\cdot j}^2 - \left(\sum_j w_{\cdot j} \bar{y}_{\cdot j} \right)^2 / \sum_j w_{\cdot j}$

$$w_{\cdot j} = \left(\frac{1}{a^2 d^2} \sum_i \sum_h \frac{1}{n_{ijh}} \right)^{-1} \quad w_{i \cdot h} = \left(\frac{1}{c^2} \sum_j \frac{1}{n_{ijh}} \right)^{-1}$$

$$\bar{y}_{\cdot j} = \frac{1}{ad} \sum_i \sum_h \bar{y}_{ijh} \quad \bar{y}_{i \cdot h} = \frac{1}{c} \sum_j \bar{y}_{ijh}$$

[see reference cited for Case II anova]

Table IV. Analysis of variance for C and D combined and for A in a two-way classification

Source of variation	df	Sum of Squares
Correction for mean	1	$Y_{...}^2/n...$
C and D together (elim. mean; ign. all else)	cd-1	$\sum_j \sum_h Y_{jh}^2/n_{jh} - Y_{...}^2/n...$
A (elim. mean, C, D, CxD; ignoring all else)	a-1	Case I anova
Interaction = Ax C + Ax D + Ax CxD (elim. all else)	(a-1)(cd-1)	by subtraction
Within subclasses	n...-acd	$\sum_i \sum_j \sum_h \left\{ \sum_f Y_{ijhf}^2 - Y_{ijh}^2/n_{ijh} \right\}$
C+D+CxD (elim. all others)	cd-1	$\sum_j \sum_h w_{jh} \bar{y}_{jh}^2 - \left(\sum_j \sum_h w_{jh} \bar{y}_{jh} \right)^2 / \sum_j \sum_h w_{jh}$
A (elim. all others)	a-1	$\sum_i w_{i..} \bar{y}_i^2 - \left(\sum_i w_{i..} \bar{y}_i \right)^2 / \sum_i w_{i..}$

$$w_{i..} = \left(\frac{1}{c^2 d^2} \sum_j \sum_h \frac{1}{n_{ijh}} \right)^{-1} \quad w_{jh} = \left(\frac{1}{a^2} \sum_i \frac{1}{n_{ijh}} \right)^{-1}$$

$$\bar{y}_{i..} = \frac{1}{cd} \sum_j \sum_h \bar{y}_{ijh} \quad \bar{y}_{jh} = \frac{1}{a} \sum_i \bar{y}_{ijh}$$

[see reference cited for Case II anova]

Case I Anova on Three-way Classification With Interactions

A Case I analysis of variance for a three-way classification is given in Table V. The analysis given in Table 3 of the reference cited assumes no interactions, but if they are present, the "Source of variation" listing in Table V is the appropriate one. The sum of the two-factor and the three-factor interactions is obtained by subtracting the within subclass sum of squares from the residual, R_{yy} , in Table 3 of the reference cited, or by subtracting the sums of squares for A+C+D (ignoring interactions) from the among subclasses sum of squares.

Table V. Case I Anova for a three-way classification with interactions present

Source of variation	df	Sum of Squares
Correction for mean	1	$Y_{...}^2/n...$
Among subclasses	$acd-1$	$\sum_i \sum_j \sum_h Y_{ijh}^2 / n_{ijh} - Y_{...}^2/n...$
A (elim. mean; ign. all else)	$a-1$	Case I anova; see Table 3 of reference cited
C (elim. mean, A; ign. all else)	$c-1$	
D (elim. mean, A, C; ign. all else)	$d-1$	
Interaction = $A \times D + A \times C + C \times D + A \times C \times D$ (elim. all else)	$acd-a-c-d+2$	by subtraction
Within subclasses	$n...-acd$	$\sum_i \sum_j \sum_h \left\{ \sum_{f=1}^{n_{ijh}} Y_{ijhf}^2 - Y_{ijh}^2 / n_{ijh} \right\}$

From Tables II to V there are four "Interaction" sums of squares containing some or all of the interactions. Since there are four equations and four unknowns one could obtain a solution for each of the interaction terms, thus:

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \text{AxC} \\ \text{AxD} \\ \text{Cx D} \\ \text{AxCxD} \end{pmatrix} = \begin{pmatrix} \text{Interaction sum of squares in Table II} = I_2 \\ " & " & " & " & " & " & \text{III} = I_3 \\ " & " & " & " & " & " & \text{IV} = I_4 \\ " & " & " & " & " & " & \text{V} = I_5 \end{pmatrix}$$

or,

$$\begin{pmatrix} \text{AxC} \\ \text{AxD} \\ \text{Cx D} \\ \text{AxCxD} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} I_2 \\ I_3 \\ I_4 \\ I_5 \end{pmatrix}$$

But, these solutions for the two factor interactions are the following sums of squares:

$$\begin{aligned} &\text{AxC (eliminating mean, A, C; ignoring all else)} \\ &\text{AxD (" " " A, D; " " ")} \\ &\text{Cx D (" " " C, D; " " ")} \end{aligned}$$

which could have been obtained by conducting a two-way classification analysis on the resulting data after summing over levels of the third classification. Likewise, the three factor interaction does not emerge unentangled from other effects. Hence, this approach yields no new information though these sums of squares are obtainable without the additional analyses.

Two-way Classification Analyses at Each Level of the Third Classification

Two-way classification analyses of variance could be performed at each level of the third category as described in Table VI for the A and C classifications at each level of the D classification. Three such tables could be computed. In each table the interaction sum of squares could be summed over the levels of the third classification to obtain these sums of squares:

$$\begin{array}{l} \text{Ax}C + \text{Ax}CxD \text{ (eliminating all other effects)} \\ \text{Ax}D + \text{Ax}CxD \left(\begin{array}{cccc} " & " & " & " \end{array} \right) \\ \text{Cx}D + \text{Ax}CxD \left(\begin{array}{cccc} " & " & " & " \end{array} \right) \end{array}$$

Differences between the above sums of squares and those in Tables II to IV yield quantities which are usually not of the desired form. For example, the difference between the Interaction sum of squares in Table IV and the first one above results in:

$$\begin{aligned} & \text{Ax}C + \text{Ax}D + \text{Ax}CxD \text{ (eliminating all other effects)} - \left\{ \sum_{h=1}^d I_h \text{ in Table VI} = \right. \\ & \left. \text{Ax}C + \text{Ax}CxD \text{ (eliminating all other effects)} \right\} \\ & = \text{Ax}D \text{ (eliminating } \mu, A, C, D, \text{Cx}D; \text{ ignoring Ax}C, \text{Ax}CxD) \end{aligned}$$

The usefulness here would be in by-passing the computation of sums of squares of the above form should they be desired.

However, it may be instructive to compute tables similar to Table VI in order to pin-point differences from level to level of the third factor. These tables may be desirable even though a complete combined analysis of the form of Table 5 in the reference cited is available.

Two-way Classification Analyses on Means of Subclass Means

If the subclass means, $\bar{y}_{1jh..}$, from the three-way classification are summed over one of the subscripts, say h , and divided by the number of levels, say d , one obtains a mean of subclass means, say $\frac{1}{d} \sum_{h=1}^d \bar{y}_{1jh..} = \bar{\bar{y}}_{1j..}$. Consider a two-way classification, say A and C , in which the corresponding subclass totals are the mean of subclass means, say $\bar{\bar{y}}_{1j..}$. The n_{1j} for this table will be $n_{1j} = w_{1j} = \left(\frac{1}{d^2} \sum_{h=1}^d \frac{1}{n_{1jh}} \right)^{-1}$. Thus, one now has quantities corresponding to subclass totals and number of observations for a two-way classification. One may now proceed to run an analysis of variance on these quantities in the same manner as Table 4 of the reference cited or as given in Table VII.

The quantities $\bar{\bar{y}}_{1j..}$ do not contain any of the remaining effects, i.e., D , $\text{Ax}D$, $\text{Cx}D$, or $\text{Ax}CxD$. Thus, these results contain only the A , C , and $\text{Ax}C$ effects, and an analysis of variance in the form of Table VII yields the $\text{Ax}C$ interaction sum of squares free of all other effects. The A and C

Table VI. Two-way classification analyses for A and C within each level of classification D

Source of variation	df	Sum of Squares for levels of category D				Sum
		1	2	...	d	
Correction for mean	1	$\frac{Y_{..1}^2}{n_{..1}}$	$\frac{Y_{..2}^2}{n_{..2}}$	...	$\frac{Y_{..d}^2}{n_{..d}}$	—
A (ign. C, AxC) (elim. mean)	a-1	$\sum_i \frac{Y_{i.1}^2}{n_{i.1}} - \frac{Y_{..1}^2}{n_{..1}}$	$\sum_i \frac{Y_{i.2}^2}{n_{i.2}} - \frac{Y_{..2}^2}{n_{..2}}$	...	$\sum_i \frac{Y_{i.d}^2}{n_{i.d}} - \frac{Y_{..d}^2}{n_{..d}}$	$\sum_h \left(\sum_i \frac{Y_{i.h}^2}{n_{i.h}} - \frac{Y_{..h}^2}{n_{..h}} \right)$
C (elim. mean, A; ign. AxC)	c-1	Case I	Case I	...	Case I	—
AxC (elim. mean, A, C)	(a-1)(c-1)	I_1 by subtraction	I_2 by subtraction	...	I_d by subtraction	$\sum_{h=1} I_h$
Within subclasses	$n_{..h} - ac$	$\sum_{i,j} \left\{ \sum_f Y_{ij1f}^2 - \frac{Y_{ij1.}^2}{n_{ij1.}} \right\}$	$\sum_{i,j} \left\{ \sum_f Y_{ij2f}^2 - \frac{Y_{ij2.}^2}{n_{ij2.}} \right\}$	...	$\sum_{i,j} \left\{ \sum_f Y_{ijd f}^2 - \frac{Y_{ij d.}^2}{n_{ij d.}} \right\}$	S_{yy}
Total	$n_{..h}$	$\sum_i \sum_j \sum_f Y_{ij1f}^2$	$\sum_i \sum_j \sum_f Y_{ij2f}^2$	...	$\sum_i \sum_j \sum_f Y_{ij d f}^2$	$\sum_h \sum_i \sum_j \sum_f Y_{ij h f}^2$

effects sum of squares free of all other effects are given in Tables IV and III, respectively.

Similarly, analysis of variance tables for A and D and for C and D classifications may be constructed in this manner if desired.

Table VII. Analysis of variance on $\bar{y}_{1j..}$ and for $w_{1j.} = n_{1j.}$ for the A and C classifications eliminating all other effects

Source of variation	df	Sum of Squares
Among subclasses	ac-1	$\sum_i \sum_j \bar{y}_{1j..}^2 / w_{1j.}$ $- \left(\sum_i \sum_j \bar{y}_{1j..} \right)^2 / \sum_i \sum_j w_{1j.}$
A (ignoring C and AxC)	a-1	$\sum_i \frac{\left(\sum_j \bar{y}_{1j..} \right)^2}{\sum_j w_{1j.}} - \left(\sum_i \sum_j \bar{y}_{1j..} \right)^2 / \sum_i \sum_j w_{1j.}$
C (elim. A, ign. AxC)	c-1	Case I anova
AxC (elim. A, C)	(a-1)(c-1)	subtraction
Within subclasses	n...-acd	$\sum_i \sum_j \sum_h \left\{ \sum_f Y_{1jhf}^2 - Y_{1jhf}^2 / n_{1jhf} \right\} = S_{yy}$
$\bar{y}_{1j..} = \frac{1}{d} \sum_{h=1}^d \bar{y}_{1jh.} \quad w_{1j.} = \left(\frac{1}{a} \sum_{h=1}^d \frac{1}{n_{1jh.}} \right)^{-1}$		

Discussion

Although much of the information for analysis of variance for three-way and higher-way classifications may be extracted from nested and two-way classifications analyses of variance, it may be more prudent to proceed directly to analyses of variance and estimation procedures for three-way or higher-way classifications. In some situations the arithmetic and algebraic procedures are not too difficult. In other situations the arithmetical and algebraic complications may become quite tedious. It is

suggested that the higher-way analyses be studied prior to the start of computational programming.

In the foregoing it has been assumed that the number of observations in any subclass is at least one. In the event that there are zero observations in some of the subclasses some assumptions must be made about parameters with zero observations, or a chance in the model must be made. Some discussion of this is given in the above reference and in some of the references cited therein.