# COVARIANCE ANALYSIS IN A TWO-WAY CLASSIFICATION WITH UNEQUAL NUMBERS IN THE SUBCLASSES

BU-51-M W. T. Federer April, 1954

Covariance analyses for a two-way classification with unequal numbers in the subclasses present computational difficulties as well as some theoretical complications. The purpose of this note is to set forth some of the considerations involved in various covariance analyses and to illustrate the procedures with a numerical example.

In the following covariance analyses it will be assumed that the usual assumptions in a covariance analysis are satisfied; i.e.,

(i) the independent variate (X) is measured without error,

(ii) the differences among treatments for the covariate represent random sampling fluctuations,

(iii) the regressions for the various treatments ~~must be~~ are estimates of a common regression, $\beta$, and

(iv) the treatment errors after adjusting for covariation are estimates of a common error variance, $\sigma_\epsilon^2$.

If conditions (ii), and/or (iii) do ~~does~~ not hold covariance may still be applicable but the interpretation of the results must take this into account. Whether or not meaningful conclusions can be reached will depend upon the nature of the experimental treatments and conditions.

Three cases will be considered in the present manuscript. These are:

Case I - The regression from the residual line is used to adjust the category means, the adjusted residual mean square is used

Table XVI-24. Yield and number of observations for the variate Y and for the covariate X from a randomized complete block design (totals per $n_{ij}$ observations).

| Treatment | Replicate 1 | 2 | ... | j | ... | r | Total |
|---|---|---|---|---|---|---|---|
| 1 | $n_{11}$ $Y_{11\cdot}$ $X_{11\cdot}$ | $n_{12}$ $Y_{12\cdot}$ $X_{12\cdot}$ | | $n_{1j}$ $Y_{1j\cdot}$ $X_{1j\cdot}$ | | $n_{1r}$ $Y_{1r\cdot}$ $X_{1r\cdot}$ | $n_{1\cdot}$ $Y_{1\cdot\cdot}$ $X_{1\cdot\cdot}$ |
| 2 | $n_{21}$ $Y_{21\cdot}$ $X_{21\cdot}$ | $n_{22}$ $Y_{22\cdot}$ $X_{22\cdot}$ | | $n_{2j}$ $Y_{2j\cdot}$ $X_{2j\cdot}$ | | $n_{2r}$ $Y_{2r\cdot}$ $X_{2r\cdot}$ | $n_{2\cdot}$ $Y_{2\cdot\cdot}$ $X_{2\cdot\cdot}$ |
| ⋮ | | | | | | | |
| i | $n_{i1}$ $Y_{i1\cdot}$ $X_{i1\cdot}$ | $n_{i2}$ $Y_{i2\cdot}$ $X_{i2\cdot}$ | | $n_{ij}$ $Y_{ij\cdot}$ $X_{ij\cdot}$ | | $n_{ir}$ $Y_{ir\cdot}$ $X_{ir\cdot}$ | $n_{i\cdot}$ $Y_{i\cdot\cdot}$ $X_{i\cdot\cdot}$ |
| ⋮ | | | | | | | |
| v | $n_{v1}$ $Y_{v1\cdot}$ $X_{v1\cdot}$ | $n_{v2}$ $Y_{v2\cdot}$ $X_{v2\cdot}$ | | $n_{vj}$ $Y_{vj\cdot}$ $X_{vj\cdot}$ | | $n_{vr}$ $Y_{vr\cdot}$ $X_{vr\cdot}$ | $n_{v\cdot}$ $Y_{v\cdot\cdot}$ $X_{v\cdot\cdot}$ |
| Total | $n_{\cdot 1}$ $Y_{\cdot 1\cdot}$ $X_{\cdot 1\cdot}$ | $n_{\cdot 2}$ $Y_{\cdot 2\cdot}$ $X_{\cdot 2\cdot}$ | | $n_{\cdot j}$ $Y_{\cdot j\cdot}$ $X_{\cdot j\cdot}$ | | $n_{\cdot r}$ $Y_{\cdot r\cdot}$ $X_{\cdot r\cdot}$ | $n_{\cdot\cdot}$ $Y_{\cdot\cdot\cdot}$ $X_{\cdot\cdot\cdot}$ |

to test hypotheses about category means, and interaction is assumed non-existent.

Case II - Case I with interaction present and for fixed effects.

Case III - The interaction regression is used to adjust the category means, and the adjusted interaction error is used to test hypotheses about category means.

Case I - Categories adjusted for residual regression, no interaction

If the items in the two classifications are considered to be fixed effects and if interaction is non-existent, the linear model in a two-way classification with a covariate is:

$$Y_{ijh} = \mu + \rho_j + \tau_i + \delta_{ijh} + \beta X_{ijh}, \tag{1}$$

where $\mu$, $\rho_j$, $\tau_i$, $\beta$, $X_{ijh}$, and $\delta_{ijh}$ represent the mean effect, the effect of the jth level of the first factor, the ith level of the second factor, the average regression coefficient from the deviations, the ijhth value of the independent variate corresponding to the ijhth value of the variable of interest Y, and a random error component, respectively; $i = 1, 2, \ldots, v$; $j = 1, 2, \ldots, r_i$; $h = 1, 2, \ldots, n_{ij}$; $\sum_{ij}\sum n_{ij} = n_{..}$ = total number of observations; $\sum_i r_i = r_.$ = total number of subclasses.

The number of individuals per subclass, $n_{ij}$, and the totals per subclass for the two variables, Y and X, are given in Table 1. The various row and column totals and the grand totals are defined in the margins. If any $n_{ij}$ = zero then the corresponding $X_{ij.}$ and $Y_{ij.}$ will also be zero, and the number of subclasses in the ith row will then be $r_i$. Likewise, if the second category* is considered there will be $v_j$ subclasses in the jth column.

* The first category in Table 1 is denoted as the treatment effect while the second category is called the replicate effect. For the present example it is more realistic to think of replicate as a second category rather than as a replicate in a randomized complete block design.

The estimates of the effects $\mu$, $\rho_j$, $\tau_i$, and $\beta$ are obtained by partial differentiation of the residual sum of squares with respect to the effects, by setting the resulting equations equal to zero, by imposing the restrictions $\sum_j \hat{\rho}_j = 0$ and $\sum_i \hat{\tau}_i = 0$, and by obtaining the estimates $\hat{\mu}$, $\hat{\rho}_j$, $\hat{\tau}_i$, and $\hat{\beta}$, which satisfy the resulting equations. The residual sum of squares is:

$$\text{Res} = \sum_{i=1}^{v} \sum_{j=1}^{r_i} \sum_{h=1}^{n_{ij}} (Y_{ijh} - \mu - \rho_j - \tau_i - \beta X_{ijh})^2. \tag{2}$$

The partial derivatives of (2) with respect to the effects are:

$$\frac{d\ \text{Res}^*}{d\ \mu} = -2\sum\sum\sum_{ijh} (Y_{ijh} - \mu - \rho_j - \tau_i - \beta X_{ijh}). \tag{3}$$

$$\frac{d\ \text{Res}}{d\rho_j} = -2\sum\sum_{ih} (Y_{ijh} - \mu - \rho_j - \tau_i - \beta X_{ijh}). \tag{4}$$

$$\frac{d\ \text{Res}}{d\tau_i} = -2\sum\sum_{jh} (Y_{ijh} - \mu - \rho_j - \tau_i - \beta X_{ijh}). \tag{5}$$

$$\frac{d\ \text{Res}}{d\ \beta} = -2\sum\sum\sum_{ijh} X_{ijh}(Y_{ijh} - \mu - \rho_j - \tau_i - \beta X_{ijh}). \tag{6}$$

Setting the above equations equal to zero we obtain the following:

<u>equation for $\hat{\mu}$:</u>

$$n_{..}\hat{\mu} + \sum_j n_{.j}\hat{\rho}_j + \sum_i n_{i.}\hat{\tau}_i + \hat{\beta} X_{...} = Y_{...} \tag{7}$$

<u>equations for $\hat{\rho}_j$:</u>

$$n_{.1}(\hat{\mu} + \hat{\rho}_1) + \sum n_{i1}\hat{\tau}_i + \hat{\beta} X_{.1.} = Y_{.1.} \tag{8}$$

$$n_{.2}(\hat{\mu} + \hat{\rho}_2) + \sum n_{i2}\hat{\tau}_i + \hat{\beta} X_{.2.} = Y_{.2.} \tag{9}$$

$\vdots$

---

* $\frac{d}{dX}$ used instead of $\frac{\partial}{\partial X}$ .

$$n_{.r}(\hat{\mu} + \hat{\rho}_r) + \Sigma n_{ir}\hat{\tau}_i + \hat{\beta}X_{.r.} = Y_{.r.} \tag{10}$$

equations for $\hat{\tau}_i$:

$$n_{1.}(\hat{\mu} + \hat{\tau}_1) + \Sigma n_{1j}\hat{\rho}_j + \hat{\beta}X_{1..} = Y_{1..} \tag{11}$$

$$n_{2.}(\hat{\mu} + \hat{\tau}_2) + \Sigma n_{2j}\hat{\rho}_j + \hat{\beta}X_{2..} = Y_{2..} \tag{12}$$

$$\vdots$$

$$n_{v.}(\hat{\mu} + \hat{\tau}_v) + \Sigma n_{vj}\hat{\rho}_j + \hat{\beta}X_{v..} = Y_{v..} \tag{13}$$

equation for $\hat{\beta}$:

$$X_{...}\hat{\mu} + \Sigma X_{.j.}\hat{\rho}_j + \Sigma X_{i..}\hat{\tau}_i + \hat{\beta}\Sigma\Sigma\Sigma X_{ijh}^2 = \sum_{ijh}\Sigma\Sigma X_{ijh}Y_{ijh} \tag{14}$$

The quantities $\Sigma\hat{\rho}_j$ and $\Sigma\hat{\tau}_i$ do not appear as such in equations (7) to (14). Therefore, it is necessary to consider the whole set of r + v + 1 + 1 equations in solving for the effects. If the $n_{ij}$'s are all equal the equations become considerably simplified. Since this is not the case we obtain the value of $\hat{\rho}_j$ and $\hat{\mu}$ from equations (8) to (10) and from the equation $\Sigma\hat{\rho}_j = 0$ ~~and substitute these values in equations (11) to (14)~~; thus,

$$\hat{\rho}_j = \frac{1}{n_{.j}}\left\{ Y_{.j.} - n_{.j}\hat{\mu} - \Sigma n_{ij}\hat{\tau}_i - \hat{\beta}X_{.j.} \right\} . \tag{15}$$

$$\mu = \frac{1}{r}\left\{ \Sigma\bar{y}_{.j.} - \sum_j \frac{1}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i - \hat{\beta}\Sigma\bar{x}_{.j.} \right\} . \tag{16}$$

Substituting for $\mu$ in equation (15) we obtain

$$\rho_j = \bar{y}_{.j.} - \frac{1}{r}\left\{ \Sigma\bar{y}_{.j.} - \sum_j \frac{1}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i - \hat{\beta}\sum_j\bar{x}_{.j.} \right\}$$

$$- \frac{1}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i - \hat{\beta}\bar{x}_{.j.} \quad . \tag{17}$$

Substituting these values in equation (14) we obtain

$$\hat{\beta} = \frac{\sum\sum\sum X_{ijh} Y_{ijh} - \sum X_{.j.} \bar{y}_{.j.} - \sum X_{i..} \hat{\tau}_i + \sum \bar{x}_{.j.} \sum_i n_{ij} \hat{\tau}_i}{\sum_{ijh}\sum\sum X_{ijh}^2 - \sum_j X_{.j.} \bar{x}_{.j.}} \qquad (18)$$

Substituting the values for $\hat{\mu}$, $\hat{\rho}_j$, and $\hat{\beta}$ obtained in equations (16), (17), and (18) in equations (11) to (13) we obtain the following

$$n_{1.}\hat{\tau}_1 - \sum_j \frac{n_{1j}}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i + (X_{1..} - \sum_j n_{1j}\bar{x}_{.j.}) \left\{ \frac{\sum_j \bar{x}_{.j.} \sum_i n_{ij}\hat{\tau}_i - \sum X_{i..}\hat{\tau}_i}{(\sum_{ijh}\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.})} \right\}$$

$$= Y_{1..} - \sum_j n_{1j}\bar{y}_{.j.} - (X_{1..} - \sum_j n_{1j}\bar{x}_{.j.}) \left( \frac{\sum_{ijh} X_{ijh} Y_{ijh} - \sum X_{.j.}\bar{y}_{.j.}}{\sum\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.}} \right) \qquad (19)$$

$$n_{2.}\hat{\tau}_2 - \sum_j \frac{n_{2j}}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i + (X_{2..} - \sum_j n_{2j}\bar{x}_{.j.}) \left( \frac{\sum_j \bar{x}_{.j.} \sum_i n_{ij}\hat{\tau}_i - \sum X_{i..}\hat{\tau}_i}{\sum\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.}} \right)$$

$$= Y_{2..} - \sum_j n_{2j}\bar{y}_{.j.} - (X_{2..} - \sum_j n_{2j}\bar{x}_{.j.}) \left( \frac{\sum\sum\sum X_{ijh} Y_{ijh} - \sum X_{.j.}\bar{y}_{.j.}}{\sum\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.}} \right) \qquad (20)$$

$$\vdots$$

$$n_{v.}\hat{\tau}_v - \sum_j \frac{n_{vj}}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i + (X_{v..} - \sum_j n_{vj}\bar{x}_{.j.}) \left( \frac{\sum_j \bar{x}_{.j.} \sum_i n_{ij}\hat{\tau}_i - \sum X_{i..}\hat{\tau}_i}{\sum\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.}} \right)$$

$$= Y_{v..} - \sum_j n_{vj}\bar{y}_{.j.} - \left(X_{v..} - \sum_j n_{vj}\bar{x}_{.j.}\right)\left(\frac{\sum\sum\sum X_{ijh}Y_{ijh} - \sum X_{.j.}\bar{y}_{.j.}}{\sum\sum\sum X_{ijh}^2 - \sum X_{.j.}\bar{x}_{.j.}}\right) \qquad (21)$$

The estimates, $\hat{\tau}_i$, of the $\tau_i$ which satisfy equations (19) to (21) and the equation $\Sigma\hat{\tau}_i = 0$ are the least squares estimates of the treatment effects adjusted for linear regression and assuming no interaction. The estimates of treatments effects not adjusted for linear regression may be obtained from equations (7) to (13) with the $\beta$ set equal to zero; these are (see Snedecor, 1946, chapter 11):

$$\hat{\tau}_1\left(n_{1.} - \sum_j \frac{n_{1j}^2}{n_{.j}}\right) - \sum_{i=2}^{v} \hat{\tau}_i \sum_j \frac{n_{1j}n_{ij}}{n_{.j}} = Y_{1..} - \sum_j n_{1j}\bar{y}_{.j.} \qquad (22)$$

$$-\hat{\tau}_1 \sum_j \frac{n_{2j}n_{1j}}{n_{.j}} + \hat{\tau}_2\left(n_{2.} - \sum_j \frac{n_{2j}^2}{n_{.j}}\right) - \sum_{i=3}^{v} \hat{\tau}_i \sum_j \frac{n_{2j}n_{ij}}{n_{.j}} = Y_{2..} - \sum_j n_{2j}\bar{y}_{.j.} \qquad (23)$$

$$\vdots$$

$$-\sum_{i=1}^{v-1} \hat{\tau}_i \sum_j n_{vj} \frac{n_{ij}}{n_{.j}} + \hat{\tau}_v\left(n_{v.} - \sum_j \frac{n_{vj}^2}{n_{.j}}\right) = Y_{v..} - \sum_j n_{vj}\bar{y}_{.j.} \qquad (24)$$

If the above equations are multiplied by a minus one they are the equations used by Snedecor (1946, section 11.12). Also, if $\beta$ is set equal to zero (i.e., all terms involving X are deleted) in equations (19) to (21) we obtain equations (22) to (24). In order to obtain unique solutions for the $\hat{\tau}_i$ we ~~must~~ impose ~~another~~ the restriction ~~. In this case~~ that $\Sigma\hat{\tau}_i = 0$ ~~is used~~.

~~Also,~~ From equations (11) to (13) it is possible to obtain equations for the $\hat{\tau}_i$ and for $\hat{\mu}$; thus,

$$\hat{\mu} = \frac{1}{v}\left\{\sum_i \bar{y}_{i..} - \sum_i \frac{1}{n_{i.}} \sum_j n_{ij}\hat{\rho}_j - \hat{\beta}\sum \bar{x}_{i..}\right\} \tag{25}$$

and

$$\hat{\tau} = \bar{y}_{i..} - \mu - \frac{1}{n_{i.}} \sum_j n_{ij}\rho_j - \beta\bar{x}_{i..}$$

$$= \bar{y}_{i..} - \frac{1}{v}\left\{\sum \bar{y}_{i..} - \sum \frac{1}{n_{i.}} \sum n_{ij}\hat{\rho}_j - \hat{\beta}\sum \bar{x}_{i..}\right\}$$

$$- \frac{1}{n_{i.}} \sum_j n_{ij}\hat{\rho}_j - \hat{\beta}\bar{x}_{i..} \; . \tag{26}$$

It should be noted that equations (16) and (25) are both expressions for $\hat{\mu}$. Making use of equations (25) and (26) we could obtain r equations involving only the $\hat{\rho}_j$'s. These equations would be similar to equations (19) to (21). The resulting estimates, $\hat{\rho}_j$, of $\rho_j$ would be the replicate effect adjusted for linear regression and assuming zero interaction. The equations are:

$$n_{.1}\hat{\rho}_1 - \sum_i \frac{n_{i1}}{n_{i.}} \sum_j n_{ij}\hat{\rho}_j + \left(X_{.1.} - \sum_i n_{i1}\bar{x}_{i..}\right)\left(\frac{\sum_i \bar{x}_{i..}\sum_j n_{ij}\hat{\rho}_j - \sum X_{.j.}\hat{\rho}_j}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}}\right)$$

$$= Y_{.1.} - \sum_i n_{i1}\bar{y}_{i..} - \left(X_{.1.} - \sum_i n_{i1}\bar{x}_{i..}\right)\left(\frac{\sum\sum\sum X_{ijh}Y_{ijh} - \sum X_{i..}\bar{y}_{i..}}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}}\right) \tag{27}$$

$$n_{.2}\hat{\rho}_2 - \sum_i \frac{n_{i2}}{n_{i.}} \sum_j n_{ij}\hat{\rho}_j + (X_{.2.} - \sum_i n_{i2}\bar{x}_{i..}) \left( \frac{\sum \bar{x}_i \sum n_{ij}\hat{\rho}_j - \sum X_{.j.}\hat{\rho}_j}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}} \right)$$

$$= Y_{.2.} - \sum n_{i2}\bar{y}_{i..} - (X_{.2.} - \sum_i n_{i2}\bar{x}_{i..}) \left( \frac{\sum\sum\sum X_{ijh}Y_{ijh} - \sum X_{i..}\bar{y}_{i..}}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}} \right) \tag{28}$$

$$\vdots$$

$$n_{.r}\hat{\rho}_r - \sum_i \frac{n_{ir}}{n_{i.}} \sum_j n_{ij}\hat{\rho}_j + (X_{.r.} - \sum_i n_{ir}\bar{x}_{i..}) \left( \frac{\sum \bar{x}_{i..} \sum_j n_{ij}\hat{\rho}_j - \sum X_{.j.}\hat{\rho}_j}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}} \right)$$

$$= Y_{.r.} - \sum n_{ir}\bar{y}_{i..} - (X_{.r.} - \sum_i n_{ir}\bar{x}_{i..}) \left( \frac{\sum\sum\sum X_{ijh}Y_{ijh} - \sum X_{i..}\bar{y}_{i..}}{\sum\sum\sum X_{ijh}^2 - \sum X_{i..}\bar{x}_{i..}} \right) \tag{29}$$

Adding the restriction $\sum\hat{\rho}_j = 0$ results in unique solutions for the $\hat{\rho}_j$'s.

If the terms involving the X's are omitted we obtain equations for the $\rho_j'$'s which are of the form used by Snedecor (1946, chapter 11); thus,

$$\hat{\rho}_1 (n_{.1} - \sum_i \frac{n_{i1}^2}{n_{i.}}) - \sum_{j=2}^{r} \hat{\rho}_j \sum_i \frac{n_{i1}n_{ij}}{n_{i.}} = Y_{.1.} - \sum_i n_{i1}\bar{y}_{i..} \tag{30}$$

$$-\hat{\rho}_1 \sum_i \frac{n_{i1}n_{i2}}{n_{i.}} + \hat{\rho}_2 (n_{.2} - \sum_i \frac{n_{i2}^2}{n_{i.}}) - \sum_{j=3}^{r} \hat{\rho}_j \sum_i \frac{n_{i2}n_{ij}}{n_{i.}} = Y_{.2.} - \sum_i n_{i2}\bar{y}_{i..} \, . \tag{31}$$

$$\vdots$$

$$- \sum_{j=1}^{r} \hat{\rho}_j \sum_i \frac{n_{ir} n_{ij}}{n_{i.}} + \hat{\rho}_r \sum_i \frac{n_{ir}^2}{n_{i.}} \Big) = Y_{.r.} - \sum n_{ir} \bar{y}_{i..} \; . \tag{32}$$

The sums of squares for the various sources of variation are given in table 2, and are obtained as follows:

$$T_{yy} = \sum_{ijh}\sum\sum Y_{ijh}^2 - \frac{Y_{...}^2}{n_{..}} \tag{33}$$

$$T_{xy} = \sum_{ijh}\sum\sum Y_{ijh} X_{ijh} - \frac{X_{...} Y_{...}}{n_{..}} \tag{34}$$

$$T_{xx} = \sum_{ijh}\sum\sum X_{ijh}^2 - \frac{X_{...}^2}{n_{..}} \tag{35}$$

$$R_{yy} = \sum_{j=1}^{r} \frac{Y_{.j.}^2}{n_{.j}} - \frac{Y_{...}^2}{n_{..}} \tag{36}$$

$$R_{xy} = \sum_{j=1}^{r} \frac{X_{.j.} Y_{.j.}}{n_{.j}} - \frac{X_{...} Y_{...}}{n_{..}} = \sum_{j=1}^{r} X_{.j.} \bar{y}_{.j.} - \frac{X_{...} Y_{...}}{n_{..}} \tag{37}$$

$$R_{xx} = \sum_{j=1}^{r} \frac{X_{.j.}^2}{n_{.j}} - \frac{X_{...}^2}{n_{..}} \tag{38}$$

$$W_{yy} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{r_j} Y_{ijh}^2 - \frac{Y_{.j.}^2}{n_{.j}} \right\} \tag{39}$$

$$W_{xy} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{r_j} \sum_{h=1}^{n_{ij}} X_{ijh} Y_{ijh} - \frac{X_{.j.} Y_{.j.}}{n_{.j}} \right\} . \tag{40}$$

$$W_{xx} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{r_j} \sum_{h=1}^{n_{ij}} X_{ijh}^2 - \frac{X_{.j.}^2}{n_{.j}} \right\} \tag{41}$$

The remaining sums of squares are much more difficult to obtain. The sum of squares for treatment (eliminating replicate effect only) is the sum of squares due to $\hat{\mu}$, $\hat{\rho}_j$ and $\hat{\tau}_i$ - sum of squares due to $\hat{\mu}$ and $\hat{\rho}_j$ assuming no treatment effect and no regression effect; thus from the estimates of the effects obtained in equations (22) to (25) and (27) to (29) the following sum of squares results in $V_{yy}$:

$$\begin{aligned} V_{yy} &= \hat{\mu}Y_{...} + \Sigma\hat{\rho}_j Y_{.j.} + \Sigma\hat{\tau}_i Y_{i..} - \Sigma \frac{Y_{.j.}^2}{n_{.j}} \\ &= \sum_j (\hat{\mu} + \hat{\rho}_j) Y_{.j.} + \Sigma\hat{\tau}_i Y_{i..} - \Sigma\frac{Y_{.j.}^2}{n_{.j}} \\ &= \sum_{i=1}^{v} \hat{\tau}_i \left(Y_{i..} - \sum_{j=1}^{r} n_{ij}\bar{y}_{.j.}\right) , \end{aligned} \tag{42}$$

where the various totals represent the corresponding right hand members of the equations (7) to (13) and equations (62) to (64). Likewise, by obtaining estimates $\hat{\mu}_X$, $\hat{\rho}_{jX}$, and $\hat{\tau}_{iX}$ for the X values

$$V_{XX} = \hat{\mu}_X X_{...} + \Sigma\hat{\rho}_{jX} X_{.j.} + \Sigma\hat{\tau}_{iX} X_{i..} - \Sigma \frac{X_{.j.}^2}{n_{.j}} , \tag{43}$$

and

$$V_{xy} = \hat{\mu}X_{...} + \Sigma\hat{\rho}_j X_{.j.} + \Sigma\hat{\tau}_i X_{i..} - \Sigma \frac{X_{.j.}Y_{.j.}}{n_{.j}}$$

$$= \hat{\mu}_X Y_{...} + \Sigma\hat{\rho}_{jX} Y_{.j.} + \Sigma\hat{\tau}_{iX} Y_{i..} - \Sigma \frac{X_{.j.} Y_{.j.}}{n_{.j}} \tag{44}$$

The residual sums of products are obtained as

$$D_{yy} = \Sigma\Sigma\Sigma Y_{ijh}^2 - (\hat{\mu} Y_{...} + \Sigma\hat{\rho} Y_{.j.} + \Sigma\hat{\tau}_i Y_{i..}) \tag{45}$$

$$D_{xy} = \Sigma\Sigma\Sigma Y_{ijh} X_{ijh} - (\hat{\mu} X_{...} + \Sigma\hat{\rho}_j X_{.j.} + \Sigma\hat{\tau}_i X_{i..}) \tag{46}$$

$$D_{xx} = \Sigma\Sigma\Sigma X_{ijh}^2 - (\hat{\mu}_X X_{...} + \Sigma\hat{\rho}_{jX} X_{.j.} + \Sigma\hat{\tau}_{iX} X_{i..}) \tag{47}$$

The errors of estimate may be computed as indicated in Table 2, or as follows:

$$D_{yy}' = \Sigma\Sigma\Sigma Y_{ijh}^2 - \left\{ \hat{\mu} Y_{...} + \Sigma\hat{\rho}_j Y_{.j.} + \Sigma\hat{\tau}_i Y_{i..} + \hat{\beta}\Sigma\Sigma\Sigma X_{ijh} Y_{ijh} \right\} , \tag{48}$$

where the estimates $\hat{\mu}$, $\hat{\rho}_j$, $\hat{\tau}_i$, and $\hat{\beta}$ are obtained from equations (7) to (13), 14

$$V_{yy}' = \hat{\mu} Y_{...} + \sum_j \hat{\rho}_j Y_{.j.} + \sum_i \hat{\tau}_i Y_{i..} + \hat{\beta}\Sigma\Sigma\Sigma X_{ijh} Y_{ijh} \tag{49}$$

$$- \sum_j \frac{Y_{.j.}^2}{n_{.j}} - \frac{W_{xy}^2}{W_{xx}} .$$

In the last equation the sum of squares due to the mean effect, the replicate effects, and the linear regression on X ~~with~~ assuming the treatment effects equal to zero is given by the last two terms. ~~This is obvious~~ that this is so may be seen from the fact that since we have only a replicates and within replicates covariance analysis.

Table 2. Analysis of covariance for Case I.

| Source of variation | df | Sums of products $y^2$ | $xy$ | $x^2$ |
|---|---|---|---|---|
| Total | $n_{..} - 1$ | $T_{yy}$ | $T_{xy}$ | $T_{xx}$ |
| Replicate (ignoring treatment) | $r - 1$ | $R_{yy}$ | $R_{xy}$ | $R_{xx}$ |
| Treatments (eliminating replicate) | $v - 1$ | $V_{yy}$ | $V_{xy}$ | $V_{xx}$ |
| Residual | $n_{..} - r - v + 1 = f_r$ | $D_{yy}$ | $D_{xy}$ | $D_{xx}$ |
| Treatments + residual | $n_{..} - r$ | $W_{yy}$ | $W_{xy}$ | $W_{xx}$ |

adjusted y's

| Source of variation | ~~Errors of estimate~~ df | ss |
|---|---|---|
| Residual | $f_r - 1$ | $D_{yy}' = D_{yy} - D_{xy}^2/D_{xx}$ |
| Treatment + residual | $n_{..} - r - 1$ | $W_{yy}' = W_{yy} - W_{xy}^2/W_{xx}$ |
| Treatment (eliminating replicate and linear regression on X) | $v - 1$ | $V_{yy}' = W_{yy}' - E_{yy}'$ D |

Example 1. The data in Table 3 were selected to illustrate the analysis of covariance for a two-way classification with unequal numbers of observations in the subclasses. Strictly speaking the linear model in Case I does not apply to an example of this type, but the example will serve to illustrate the computational procedure.

The first step in the analysis is to use equations (22) to (25) and (30) to (32) to estimate the values of $\hat{\mu}$, $\hat{\rho}_j$, and $\hat{\tau}_i$ for the Y variate and the $\hat{\mu}_x$, $\hat{\rho}_{jx}$, and $\hat{\tau}_{ix}$ for the X variate. Using equations (22) to (24) the following equations for the $\hat{\tau}_i$ are obtained (see Table 5):

equation for $\hat{\tau}_1$:

$$\hat{\tau}_1 \left[13 - \{1(.50000) + 1(.20000) + \ldots + 1(.25000) + 2(.33333)\}\right]$$

$$- \hat{\tau}_2 \left[1(.50000) + 1(.40000) + \ldots + 1(.25000) + 2(.33333)\right]$$

$$- \hat{\tau}_3 \left[1(.00000) + 1(.40000) + \ldots + 1(.50000) + 2(.33333)\right]$$

$$= 7.71667\ \hat{\tau}_1 - 4.64999\ \hat{\tau}_2 - 3.06666\ \hat{\tau}_3$$

$$= 85.68 - \left[1(6.52500) + 1(5.35200) + \ldots + 2(5.83333)\right]$$

$$= 3.50635.$$

equation for $\hat{\tau}_2$:

$$- \hat{\tau}_1 \left[1(.50000) + 2(.20000) + \ldots + 1(.25000) + 2(.33333)\right]$$

$$+ \hat{\tau}_2 \left[14 - \{1(.50000) + 2(.40000) + \ldots + 1(.25000) + 2(.33333)\}\right]$$

$$- \hat{\tau}_3 \left[1(.00000) + 2(.40000) + \ldots + 1(.50000) + 2(.33333)\right]$$

$$= -4.64999\ \hat{\tau}_1 + 7.86667\ \hat{\tau}_2 - 3.21666\ \hat{\tau}_3$$

$$= 88.58 - \left[1(6.52500) + 2(5.35200) + \ldots + 2(5.83333)\right]$$

$$= 0.26685.$$

<u>equation for $\hat{\tau}_3$:</u>

$$- \hat{\tau}_1 [0(.50000) + 2(.20000) + \dots + 2(.25000) + 2(.33333)]$$

$$- \hat{\tau}_2 [0(.50000) + 2(.40000) + \dots + 2(.25000) + 2(.33333)]$$

$$+ \hat{\tau}_3 [10 - \{0(.00000) + 2(.40000) + \dots + 2(.50000) + 2(.33333)\}]$$

$$= -3.06666\ \hat{\tau}_1 - 3.21666\ \hat{\tau}_2 + 6.28334\ \hat{\tau}_3$$

$$= 57.13 - [0(6.52500) + 2(5.35200) + \dots + 2(5.83333)]$$

$$= -3.77316.^{*}$$

Remembering that $\Sigma\hat{\tau}_i = 0$, we substitute $\hat{\tau}_3 = -\hat{\tau}_1 - \hat{\tau}_2$ in the above equations. Then we subtract the last equation from the remaining equations to obtain two equations involving $\hat{\tau}_1$ and $\hat{\tau}_2$. After eliminating $\hat{\tau}_1$ ($\hat{\tau}_2$) we find the solution for $\hat{\tau}_2$. Then, the solutions for $\hat{\tau}_i$ are found to be:

$$\hat{\tau}_1 = 0.33507;$$

$$\hat{\tau}_2 = 0.06741;$$

$$\hat{\tau}_3 = -0.40248.$$

Following a similar procedure for the X's we obtain the following three equations for the $\hat{\tau}_{ix}$'s:

$$7.71667\ \hat{\tau}_{1x} - 4.64999\ \hat{\tau}_{2x} - 3.06666\ \hat{\tau}_{3x} = 52.88334;$$

$$-4.64999\ \hat{\tau}_{1x} + 7.86667\ \hat{\tau}_{2x} - 3.21666\ \hat{\tau}_{3x} = 19.18334;$$

$$-3.06666\ \hat{\tau}_{1x} - 3.21666\ \hat{\tau}_{2x} + 6.28334\ \hat{\tau}_{3x} = -72.06666.$$

From the above three equations and from the equation $\Sigma\hat{\tau}_{ix} = 0$ we obtain the following solutions for the $\hat{\tau}_{ix}$:

* Note that the sum of coefficients for any $\hat{\tau}_i = 0$ and that the right hand side of the equations also sums to zero.

$\hat{\tau}_{1x} = 5.25816;$

$\hat{\tau}_{2x} = 2.41083;$

$\hat{\tau}_{3x} = -7.66899.$

If $\beta$ is set equal to zero in equations (8) to (10), and if the equations are divided by $n_{.j}$ we obtain the following equation for the jth replicate:

$$\hat{\mu} + \hat{\rho}_j = \bar{y}_{.j.} - \left(\sum_i \frac{n_{ij}}{n_{.j}} \hat{\tau}_i\right). \tag{50}$$

Making use of equation (50) we compute the various $\hat{\mu} + \hat{\rho}_j$ values; thus,

$$\hat{\mu} + \hat{\rho}_1 = 6.52500 - [.50000(.33507) + .50000(.06741) + .00000(-.40248)]$$

$$= 6.32376$$

$$\hat{\mu} + \hat{\rho}_2 = 5.35200 - [.20000(.33507) + .40000(.06741) + .40000(-.40248)]$$

$$= 5.41901.$$

$$\vdots$$

$$\hat{\mu} + \hat{\rho}_{10} = 5.83333 - [.33333(.33507) + .33333(.06741) + .33333(-.40248)]$$

$$= 5.83333.$$

Since $\sum_i \hat{\rho}_j = 0$ the sum of the above 10 equations is equal to $10\hat{\mu}$, which is 62.16463. Since $\hat{\mu}$ is known it is possible to evaluate each of the $\hat{\rho}_j$'s; for example, $\hat{\rho}_1 = 6.32376 - 6.21646 = .10730.$

With the values for $\hat{\mu}$, $\hat{\rho}_j$, and $\hat{\tau}_i$ it is now possible to compute the various sums of products in Tables 2 and 4. The sums of products for the total, replicate (ignoring treatment), and within replicate are straight-forward. ~~The sums of products $V_{yy}$, $V_{xx}$, and $V_{xy}$ are computed from equations (42) to (44) (see Table 5); thus,~~

Thus,

<u>correction for the mean (1 df):</u>

$Y$: $231.39^2/37 = 1447.06306.$

$XY$: $231.39(4457)/37 = 27873.11432.$

$X$: $4457^2/37 = 536887.81081.$

<u>total sum of squares ($n_{..} - 1$ df):</u>

$T_{yy} = 1550.27570 - 1447.06306 = 103.21267.$

$T_{xy} = 29516.20000 - 27873.11432 = 1643.08568.$

$T_{xx} = 569181.00000 - 536887.81081 = 32293.18919.$

<u>replicates (ignoring treatments)(r - 1 df):</u>

$R_{yy} = 1488.21038 - 1447.06303 = 41.14735.$

$R_{xy} = 28492.01367 - 27873.11432 = 618.89935.$

$R_{xx} = 546960.61667 - 536887.81081 = 10072.80586.$

<u>within replicates ($n_{..}$ - r df):</u>

$W_{yy} = 1550.27570 - 1488.21038 = 62.06532.$

$W_{xy} = 29516.20000 - 28492.01367 = 1024.18633.$

$W_{xx} = 569181.00000 - 546960.61667 = 22220.38333.$

The sums of products $V_{yy}$, $V_{xy}$, and $V_{xx}$ are computed from equations (42) to (44) (see Table 5); thus,

$$V_{yy} = 13.05(6.32376) + 26.76(5.41901) + \ldots + 35.00(5.83333)$$

$$+ 85.68(.33507) + 88.58(.06471) + 57.13(-.40248)$$

$$- \left( \frac{13.05^2}{2} + \frac{26.76^2}{5} + \ldots + \frac{35.00^2}{6} \right) = 1490.92150 - 1488.21038$$

$$= 2.71112.$$

Also, $V_{yy} = .33507(3.50635) + .06741(.26685) - .40248(-3.77316) = 2.71148$,

which agrees with the above within rounding errors.

$$V_{xx} = 245(118.66550) + 526(106.25163) + \ldots + 698(116.33333)$$

$$+ 1634(5.25816) + 1713(2.41083) + 1110(-7.66899)$$

$$- \left( \frac{245^2}{2} + \frac{526^2}{5} + \ldots + \frac{698^2}{6} \right)$$

$$= 547837.60759 - 546960.61667 = 876.99092.$$

Also, $V_{xx} = 5.25816(52.88334) + 2.41083(19.18334) - 7.66899(-72.06666)$

$= 876.99533$, which agrees with the above within rounding errors.

$$V_{xy} = 245(6.32376) + 526(5.41901) + \ldots + 698(5.83333)$$

$$+ 1634(.33507) + 1713(.06741) + 1110(-.40248)$$

$$- \left( \frac{13.05(245)}{2} + \frac{26.76(526)}{5} + \ldots + \frac{35.00(698)}{6} \right)$$

$$= 28540.02481 - 28492.01367 = 48.01114.$$

Also, $V_{xy} = .33507(52.88334) + .06741(19.18334) - .40248(-72.06666) = 48.01816$,

and $5.25816(3.50635) + 2.41083(.26685) - 7.66899(-3.77316) = 48.01661$,

which equals the above within rounding errors.

The sums of products $D_{yy}$, $D_{xy}$, and $D_{xx}$ are obtained by subtraction or from equations (45) to (47); thus,

$$D_{yy} = 1550.27570 - 1490.92150 = 59.35420;$$

$$D_{xy} = 29516.20000 - 28540.02481 = 976.17519;$$

$$D_{xx} = 569181.00000 - 547837.60759 = 21343.39241.$$

The sums of squares adjusted for regression are computed in the usual manner ( see Tables 2 and 4).

As a check on the adjusted sums of squares obtained in the above manner, solutions for the $\hat{\mu}$, $\hat{\rho}_j$, $\hat{\beta}$, and $\hat{\tau}_i$ values may be obtained from equations (7) to (14). These are given by equations (15) to (23) and (25) to (29). Then, $D_{yy}'$ and $V_{yy}'$ are obtained from equations (48) and (49). The results for this example follow:

<u>equations for $\hat{\tau}_i$ (equations (19) to (21))</u>:

$$\hat{\tau}_1 \Big\{ 13 - [1(.50000) + 1(.20000) + \ldots + 1(.250000) + 2(.33333)]$$

$$- \frac{[1634 - 1(122.50000) - 1(105.20000) - \ldots - 1(133.75000) - 2(116.33333)]^2}{22220.38333} \Big\}$$

$$- \hat{\tau}_2 \Big\{ 1(.50000) + 1(.40000) + \ldots + 1(.25000) + 2(.33333)$$

$$+ [1634 - 1(122.50000) - 1(105.20000) - \ldots - 1(133.750000) - 2(116.33333)]$$

$$\left[ \frac{1713 - 1(122.50000) - 2(105.20000) - \ldots - 2(116.33333)}{22220.38333} \right] \Big\}$$

$$- \hat{\tau}_3 \Big\{ 1(.00000) + 1(.40000) + \ldots + 1(.50000) + 2(.33333)$$

$$+ [1634 - 1(122.50000) - 1(105.20000) - \ldots - 2(116.33333)]$$

$$\left[ \frac{1110 - 0(122.50000) - 2(105.20000) - \ldots - 2(133.75000) - 2(116.33333)}{22220.38333} \right] \Big\}$$

$$= \hat{\tau}_1 \left\{ 7.71667 - \frac{(52.88334)^2}{22220.38333} \right\} - \hat{\tau}_2 \left\{ 4.64999 + \frac{(52.88334)(19.18334)}{22220.38333} \right\}$$

$$- \hat{\tau}_3 \left\{ 3.06666 + \frac{52.88334(-72.06666)}{22220.38333} \right\} = 7.59081\ \hat{\tau}_1 - 4.69565\ \hat{\tau}_2$$

$$- 2.89515\ \hat{\tau}_3 = 3.50635 - \frac{52.88334(1024.18633)}{22220.38333}$$

$$= 1.06884.$$

$$- \hat{\tau}_1 \left\{ 4.64999 + \frac{19.18334(52.88334)}{22220.38333} \right\} + \hat{\tau}_2 \left\{ 7.86667 - \frac{(19.18334)^2}{22220.38333} \right\}$$

$$- \hat{\tau}_3 \left\{ 3.21666 + \frac{19.18334(-72.06666)}{22220.38333} \right\} = -4.69565\ \hat{\tau}_1 + 7.85011\ \hat{\tau}_2 - 3.15444\ \hat{\tau}_3$$

$$= .26685 - \frac{19.18334(1024.18633)}{22220.38333} = -.61735.$$

$$- \hat{\tau}_1 \left\{ 3.06666 + \frac{52.88334(-72.06666)}{22220.38333} \right\} - \hat{\tau}_2 \left\{ 3.21666 + \frac{19.18334(-72.06666)}{22220.38333} \right\}$$

$$+ \hat{\tau}_3 \left\{ 6.28334 - \frac{(-72.06666)^2}{22220.38333} \right\} = -2.89515\ \hat{\tau}_1 - 3.15444\ \hat{\tau}_2 + 6.04961\ \hat{\tau}_3$$

$$= -3.77316 - \frac{(-72.06666)(1024.18633)}{22220.38333} = -.45145.$$

The equality of the various terms in the above equations and of those used in solving for the $\hat{\tau}_i$ and $\hat{\tau}_{ix}$ assuming no regression should be noted.

The estimates $\hat{\tau}_i$ which satisfy the above three equations and the equation $\Sigma\hat{\tau}_i = 0$ are:

$$\hat{\tau}_1' = .09457;\quad \hat{\tau}_2' = -.04285;\quad \hat{\tau}_3' = -.05172.^*$$

* The primes are used to denote that these are different from the estimated treatment effects prior to adjustment for linear regression.

The sum of squares for treatments adjusted for the residual linear regression is (see equation (49)):

$$.09457(1.06884) - .04285(-.61735) - .05172(-\ .45145) \ = \ .151$$

which agrees, within rounding errors, with the sum of squares, .152, obtained in Table 4.

The estimate, $\hat{\beta}$, of $\beta$ may be obtained from equation (18); thus,

$$\hat{\beta} = \frac{29516.20000 - 28492.01367 - 23.71613 + 15.80967}{569181 - 546960.61667 = 22220.38333}$$

$$= \frac{1016.27987}{22220.38333} = .0457363 79 = \frac{D_{xy}}{D_{xx}}$$

Knowing $\hat{\beta}$ and $D_{xx}$ we may compute $D_{xy}$; thus, $\cancel{D_{xy} = \hat{\beta} D_{xx}} =$

$$D_{xy} = \hat{\beta} D_{xx} = .0457363 79(21343.392) = 976.16947$$

which is equal, except for rounding errors, to the sum of squares obtained in Table 4. If one proceeds directly to the computation of $V_{yy}'$ the above method of computing $D_{xy}$ may be used. Consequently, $V_{xy} = W_{xy} - \hat{D}_{xy}$. (the [illegible])

The following formula may be used to estimate the $\hat{\mu} + \rho_j$ quantities:

$$\hat{\mu} + \hat{\rho}_j = \bar{y}_{.j.} - \hat{\beta}\bar{x}_{.j.} - \frac{1}{n_{.j}} \sum_i n_{ij} \hat{\tau}_i, \tag{51}$$

where the $\hat{\tau}_i$ and the $\hat{\beta}$ are obtained from equations (18) to (22). The resulting $\hat{\rho}_j$ values are the same as those obtained from equations (27) to (29).

Table 3. Number of plants per replicate (n), yield of rubber in grams (Y), and dry weight of shrub in grams (X) for the normal plants ~~to~~ of three guayule varieties.

| Rep. no. | Variety 405 | | | | | Variety 407 | | | | | Variety 416 | | | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | Y | X | Y | X | n | Y | X | Y | X | n | Y | X | Y | X | n | Y | X |
| 1 | 1 | 6.93 | 143 | --- | --- | 1 | 6.12 | 102 | --- | --- | 0 | --- | --- | --- | --- | 2 | 13.05 | 245 |
| 2 | 1 | 6.44 | 146 | --- | --- | 2 | 5.00 | 88 | 5.12 | 92 | 2 | 4.35 | 96 | 5.85 | 104 | 5 | 26.76 | 526 |
| 3 | 1 | 5.22 | 111 | --- | --- | 2 | 2.59 | 61 | 4.79 | 93 | 1 | 5.08 | 91 | --- | --- | 4 | 17.66 | 356 |
| 4 | 1 | 8.55 | 170 | --- | --- | 2 | 6.46 | 127 | 10.70 | 194 | 1 | 5.41 | 101 | --- | --- | 4 | 31.12 | 592 |
| 5 | 2 | 7.30 | 132 | 4.19 | 83 | 1 | 7.66 | 169 | --- | --- | 0 | --- | --- | --- | --- | 3 | 19.15 | 384 |
| 6 | 2 | 6.77 | 115 | 9.03 | 120 | 1 | 6.81 | 139 | --- | --- | 1 | 6.48 | 132 | --- | --- | 4 | 29.09 | 506 |
| 7 | 1 | 7.10 | 137 | --- | --- | 2 | 6.25 | 116 | 8.08 | 146 | 0 | --- | --- | --- | --- | 3 | 21.43 | 399 |
| 8 | 1 | 5.57 | 129 | --- | --- | 0 | --- | --- | --- | --- | 1 | 4.22 | 87 | --- | --- | 2 | 9.79 | 216 |
| 9 | 1 | 9.17 | 151 | --- | --- | 1 | 5.44 | 116 | --- | --- | 2 | 5.04 | 90 | 8.69 | 178 | 4 | 28.34 | 535 |
| 10 | 2 | 5.41 | 106 | 4.02 | 91 | 2 | 6.56 | 130 | 7.00 | 140 | 2 | 4.45 | 91 | 7.56 | 140 | 6 | 35.00 | 698 |
| Sum | 13 | --- | --- | 85.68 | 1634 | 14 | --- | --- | 88.58 | 1713 | 10 | --- | --- | 57.13 | 1110 | 37 | 231.39 | 4457 |

Table 4. Analysis of covariance for the data of Table 3 assuming the following linear model is appropriate:

$$Y_{ijh} = \mu + \rho_j + \tau_i + \delta_{ijh} + \beta X_{ijh}.$$

| Source of variation | df | Sums of products $y^2$ | $xy$ | $x^2$ |
|---|---|---|---|---|
| Total | 36 | 103.213 | 1643.086 | 32293.189 |
| Replicate (ignoring treatment) | 9 | 41.147 | 618.899 | 10072.806 |
| Treatment (eliminating replicate) | 2 | 2.712 | 48.017 | 876.991 |
| Residual | 25 | 59.354 | 976.170 | 21343.392 |
| Treatment + residual | 27 | 62.066 | 1024.187 | 22220.383 |

| Source of variation | df | ss | ms |
|---|---|---|---|
| Residual adjusted for regression | 24 | 14.707 | 0.6128 |
| Treatment + residual adjusted | 26 | 14.859 | --- |
| Treatment (eliminating replicate and regression) | 2 | 0.152 | 0.0760 |

Table 5. Values of $n_{ij}$, $n_{ij}/n_{.j}$, totals, means, and effects.

| Rep. | Variety 405 | 407 | 416 | $n_{.j}$ | $Y_{.j.}$ | $Y_{.j.}/n_{.j}$ | $\hat{\rho}_j + \hat{\mu}$ | $X_{.j.}$ | $X_{.j.}/n_{.j}$ | $\hat{\rho}_{jx} + \hat{\mu}_x$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1<br>.50000 | 1<br>.50000 | 0<br>.00000 | 2 | 13.05 | 6.52500 | 6.32376 | 245 | 122.50 | 118.66550 |
| 2 | 1<br>.20000 | 2<br>.40000 | 2<br>.40000 | 5 | 26.76 | 5.35200 | 5.41901 | 526 | 105.20 | 106.25163 |
| 3 | 1<br>.25000 | 2<br>.50000 | 1<br>.25000 | 4 | 17.66 | 4.41500 | 4.39815 | 356 | 89.00 | 88.39729 |
| 4 | 1<br>.25000 | 2<br>.50000 | 1<br>.25000 | 4 | 31.12 | 7.78000 | 7.76315 | 592 | 148.00 | 147.39729 |
| 5 | 2<br>.66667 | 1<br>.33333 | 0<br>.00000 | 3 | 19.15 | 6.38333 | 6.13748 | 384 | 128.00 | 123.69094 |
| 6 | 2<br>.50000 | 1<br>.25000 | 1<br>.25000 | 4 | 29.09 | 7.27250 | 7.18873 | 506 | 126.50 | 125.18546 |
| 7 | 1<br>.33333 | 2<br>.66667 | 0<br>.00000 | 3 | 21.43 | 7.14333 | 6.98670 | 399 | 133.00 | 129.64007 |
| 8 | 1<br>.50000 | 0<br>.00000 | 1<br>.50000 | 2 | 9.79 | 4.89500 | 4.92870 | 216 | 108.00 | 109.20542 |
| 9 | 1<br>.25000 | 1<br>.25000 | 2<br>.50000 | 4 | 28.34 | 7.08500 | 7.18562 | 535 | 133.75 | 135.66725 |
| 10 | 2<br>.33333 | 2<br>.33333 | 2<br>.33333 | 6 | 35.00 | 5.83333 | 5.83333 | 698 | 116.33 | 116.33333 |
| $n_{i.}$ | 13 | 14 | 10 | 37 | | | 62.16463 | | | 1200.43418 |
| $Y_{i..}$ | 85.68 | 88.58 | 57.13 | | 231.39 | | | | | |
| $Y_{i..}/n_{i.}$ | 6.5908 | 6.3271 | 5.7130 | | | | $\hat{\mu}$ = 6.21646 | | | |
| $\hat{\tau}_i$ | .33507 | .06741 | -.40248 | | | | | | | |
| $X_{i..}$ | 1634 | 1713 | 1110 | | | | | 4457 | $\hat{\mu}_x$ = 120.04342 | |
| $X_{i..}/n_{i.}$ | 125.6923 | 122.3571 | 111.0 | | | | | | | |
| $\hat{\tau}_{ix}$ | 5.25816 | 2.41083 | -7.66899 | | | | | | | |

Case II - Interaction present and for fixed effects

If interaction is present there may be no reason for estimating the category means. Instead one may wish to consider the subclass means, and, in such a case, the analysis of covariance is simply a within subclasses and among subclasses analysis (Snedecor, 1946; Day and Fisher, 1937). In the event that the category means are to be compared and the interaction is to be tested the following analysis is appropriate (Yates, 1934).

For this case the linear model is assumed to be:

$$Y_{ijh} = \mu + \rho_j + \tau_i + \rho\tau_{ij} + \beta X_{ijh} + \delta_{ijh}, \tag{52}$$

where $(\rho\tau)_{ij}$ represents an effect peculiar to the ijth subclass and where the other effects are defined in equation (1); $i = 1, 2, \ldots, v$; $j = 1, 2, \ldots, r$; $h = 1, 2, \ldots, n_{ij}$. The sum of squares to be minimized is

$$\text{Res} = \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh})^2. \tag{53}$$

The partial derivatives of the above sums of squares with respect to the various effects are:

$$\frac{\delta \text{ Res}}{\delta \mu} = -2 \sum_{ijh}\sum\sum (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh}) ; \tag{54}$$

$$\frac{\delta \text{ Res}}{\delta \rho_j} = -2 \sum_{ih}\sum (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh}); \tag{55}$$

$$\frac{\delta \text{ Res}}{\delta \tau_i} = -2 \sum_{jh}\sum (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh}); \tag{56}$$

$$\frac{\delta \text{ Res}}{\delta \rho\tau_{ij}} = -2 \sum_{h} (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh}); \tag{57}$$

$$\frac{\delta \text{ Res}}{\delta \beta} = -2 \sum_{ijh}\sum\sum (Y_{ijh} - \mu - \rho_j - \tau_i - \rho\tau_{ij} - \beta X_{ijh}). \tag{58}$$

Setting the above equations equal to zero, we obtain the following:

equation for $\mu$:

$$n_{..}\hat{\mu} + \sum_j n_{.j}\hat{\rho}_j + \sum_i n_{i.}\hat{\tau}_i + \sum_{i,j} n_{ij}\widehat{\rho\tau}_{ij} + \hat{\beta}X_{...} = Y_{...} \tag{59}$$

equations for $\rho_j$:

$$n_{.1}(\hat{\mu} + \hat{\rho}_1) + \sum_i n_{i1}(\hat{\tau}_i + \widehat{\rho\tau}_{i1}) + \hat{\beta}X_{.1.} = Y_{.1.} \tag{60}$$

$$n_{.2}(\hat{\mu} + \hat{\rho}_2) + \sum_i n_{i2}(\hat{\tau}_i + \widehat{\rho\tau}_{i2}) + \hat{\beta}X_{.2.} = Y_{.2.} \tag{61}$$

$$\vdots$$

$$n_{.r}(\hat{\mu} + \hat{\rho}_r) + \sum_i n_{ir}(\hat{\tau}_i + \widehat{\rho\tau}_{ir}) + \beta X_{.r.} = Y_{.r.} \tag{62}$$

equations for $\tau_i$:

$$n_{1.}(\hat{\mu} + \hat{\tau}_1) + \sum_j n_{1j}(\hat{\rho}_j + \widehat{\rho\tau}_{1j}) + \hat{\beta}X_{1..} = Y_{1..} \tag{63}$$

$$n_{2.}(\hat{\mu} + \hat{\tau}_2) + \sum_j n_{2j}(\hat{\rho}_j + \widehat{\rho\tau}_{2j}) + \hat{\beta}X_{2..} = Y_{2..} \tag{64}$$

$$\vdots$$

$$n_{v.}(\hat{\mu} + \hat{\tau}_v) + \sum_j n_{vj}(\hat{\rho}_j + \widehat{\rho\tau}_{vj}) + \hat{\beta}X_{v..} = Y_{v..} \tag{65}$$

equations for $\rho\tau_{ij}$:

$$n_{11}(\hat{\mu} + \hat{\rho}_1 + \hat{\tau}_1 + \widehat{\rho\tau}_{11}) + \hat{\beta}X_{11.} = Y_{11.} \tag{66}$$

$$\vdots$$

$$n_{1r}(\hat{\mu} + \hat{\rho}_r + \hat{\tau}_1 + \widehat{\rho\tau}_{1r}) + \hat{\beta}X_{1r.} = Y_{1r.} \tag{67}$$

$$n_{21}(\hat{\mu} + \hat{\rho}_1 + \hat{\tau}_2 + \widehat{\rho\tau}_{21}) + \hat{\beta}X_{21.} = Y_{21.} \tag{68}$$

$$\vdots$$

$$n_{2r}(\hat{\mu} + \hat{\rho}_r + \hat{\tau}_2 + \widehat{\rho\tau}_{2r}) + \hat{\beta}X_{2r.} = Y_{2r.} \tag{69}$$

$$\vdots$$

$$n_{vr}(\hat{\mu} + \hat{\rho}_r + \hat{\tau}_v + \widehat{\rho\tau}_{vr}) + \hat{\beta}X_{vr.} = Y_{vr.} \tag{70}$$

<u>equation for $\beta$:</u>

$$X_{...}\hat{\mu} + \sum_j X_{.j.}\hat{\rho}_j + \sum_i X_{i..}\hat{\tau}_i + \sum\sum_{ij} X_{ij.}\widehat{\rho\tau}_{ij} + \hat{\beta}\sum\sum\sum_{ijh} X_{ijh}^2 = \sum\sum\sum_{ijh} X_{ijh}Y_{ijh} \tag{71}$$

In order to obtain unique solutions for the effects in the above equations, the following restrictions are imposed:

$$\sum_{j=1}^{v} \hat{\rho}_j = 0; \tag{72}$$

$$\sum_{i=1}^{v} \hat{\tau}_i = 0; \tag{73}$$

$$\sum_{i=1}^{v} \widehat{\rho\tau}_{ij} = \sum_{j=1}^{r} \widehat{\rho\tau}_{ij} = 0. \tag{74}$$

In the $\rho\tau_{ij}$ equations if we divide each equation by the corresponding $n_{ij}$ and sum over j we obtain the following set of equations:*

* It is assumed that no $n_{ij}$ is equal to zero. If one or more $n_{ij} = 0$ no solution is possible unless equation (74) is altered to the following:

$$\sum_{i=1}^{v_j} \widehat{\rho\tau}_{ij} = 0 = \sum_{j=1}^{r_i} \widehat{\rho\tau}_{ij}. \tag{74a}$$

The above restrictions assume that the missing $\widehat{\rho\tau}_{ij}$ are set equal to zero (see Henderson, 1948).

$$\hat{\mu} + \hat{\tau}_1 = \frac{1}{r} \sum_j (\bar{y}_{1j.} - \hat{\beta}\bar{x}_{1j.}), \tag{75}$$

$$\hat{\mu} + \hat{\tau}_2 = \frac{1}{r} \sum_j (\bar{y}_{2j.} - \hat{\beta}\bar{x}_{2j.}), \tag{76}$$

$$\vdots$$

$$\hat{\mu} + \hat{\tau}_v = \frac{1}{r} \sum_j (\bar{y}_{vj.} - \beta\bar{x}_{vj}). \tag{77}$$

Summing equations (75) to (77) we obtain:

$$\hat{\mu} = \frac{1}{rv} \sum_{i=1}^{v} \sum_{j=1}^{r} (\bar{y}_{ij} - \hat{\beta}\bar{x}_{ij.}). \tag{78}$$

Since $\hat{\beta}$ in the above equations is a within subclasses regression it may be estimated from the formula (see Table 6):

$$\hat{\beta} = \frac{\sum_{i=1}^{v} \sum_{j=1}^{r_i} \left\{ \sum_{h=1}^{n_{ij}} X_{ijh} Y_{ijh} - X_{ij.} Y_{ij.} / n_{ij} \right\} = S_{xy}}{\sum_{i=1}^{v} \sum_{j=1}^{r_i} \left\{ \sum_{h=1}^{n_{ij}} X_{ijh}^2 - X_{ij.}^2 / n_{ij} \right\} = S_{xx}} . \tag{79}$$

Verification that the above solution for $\beta$ is the same as that obtained from solving ~~from~~ equations (59) to (71) is made by substituting the values of $\hat{\mu}$ from equation (78), $\hat{\tau}_i$ from equations (75) to (77), and the following estimates for $\hat{\rho}_j$ and $\widehat{\rho\tau}_{ij}$ in equation (71):

$$\hat{\rho}_1 + \hat{\mu} = \frac{1}{v} \sum_i \left\{ \bar{y}_{i1.} - \hat{\beta}\bar{x}_{i1.} \right\} , \tag{80}$$

$$\vdots$$

$$\hat{\rho}_j + \hat{\mu} = \frac{1}{v} \sum_i \left\{ \bar{y}_{ij.} - \hat{\beta}\bar{x}_{ij.} \right\} , \tag{81}$$

$$\vdots$$

This estimate of $\hat{\beta}$ holds whether or not any $n_{ij}$ is equal to zero.

$$\hat{\rho}_r + \hat{\mu} = \frac{1}{v} \sum_i \left\{ \bar{y}_{ir.} - \hat{\beta}\bar{x}_{ir.} \right\} ; \tag{82}$$

$$\widehat{\rho\tau}_{11} = \bar{y}_{11.} - \hat{\beta}\bar{x}_{11.} - \hat{\mu} - \hat{\rho}_1 - \hat{\tau}_1 , \tag{83}$$

$$\widehat{\rho\tau}_{12} = \bar{y}_{12.} - \hat{\beta}\bar{x}_{12.} - \hat{\mu} - \hat{\rho}_2 - \hat{\tau}_1 , \tag{84}$$

$$\vdots$$

$$\widehat{\rho\tau}_{ij} = \bar{y}_{ij.} - \hat{\beta}\bar{x}_{ij.} - \hat{\mu} - \hat{\rho}_j - \hat{\tau}_i , \tag{85}$$

$$\vdots$$

$$\widehat{\rho\tau}_{vr} = \bar{y}_{vr.} - \hat{\beta}\bar{x}_{vr.} - \hat{\mu} - \hat{\rho}_r - \hat{\tau}_v . \tag{86}$$

In fact, the simplest form for the solution for $\hat{\beta}$ is obtained by substituting only for the $\widehat{\rho\tau}_{ij}$'s. The terms involving $\hat{\mu}$, $\hat{\rho}_j$, and $\hat{\tau}_i$ sum to zero.

In order to simplify the procedure we shall consider the variance analysis first. (The covariance analysis (Table 6) will be considered later.) The preceding equations hold for the variance analysis if $\hat{\beta}$ is set equal to zero in all equations.

If the variance of a single observation is $\sigma_\delta^2$ then the variance of any $\bar{y}_{ij.}$ is

$$V(\bar{y}_{ij.}) = \frac{\sigma_\delta^2}{n_{ij}} . \tag{87}$$

The variance of $\hat{\mu} + \hat{\tau}_i$ is:

$$V(\hat{\mu} + \hat{\tau}_i) = V(\frac{1}{r} \sum_j \bar{y}_{ij}) = \frac{\sigma_\delta^2}{r^2} \sum_j \frac{1}{n_{ij}} = \frac{\sigma_\delta^2}{w_i} , \tag{88}$$

where

$$\frac{1}{w_i} = \frac{1}{r^2} \sum_{j=1}^{r} \frac{1}{n_{ij}} . \tag{89}$$

From this it follows that

$$\sum_{i=1}^{v} w_i (\bar{\bar{y}}_{i..} - \bar{y}_w)^2 = \Sigma w_i \bar{\bar{y}}_{i..}^2 - \frac{(\Sigma w_i \bar{\bar{y}}_{i..})^2}{\Sigma w_i} \tag{90}$$

is distributed as $\chi^2\sigma^2$ with $v - 1$ degrees of freedom if there are no true differences between the treatments. In the above

$$\bar{\bar{y}}_{i..} = \frac{1}{r} \sum_{j=1}^{r} \bar{y}_{ij} \tag{91}$$

and

$$\bar{y}_w = \frac{\sum_{i=1}^{v} w_i \bar{\bar{y}}_{i..}}{\sum_{i=1}^{v} w_i} . \tag{92}$$

The above all follows from a theorem given by Yates (1934):

<u>Theorem</u>:

"If the numbers are $u_1$, $u_2$, ..., and their variances are $\frac{1}{w_1}$, $\frac{1}{w_2}$, ... of the general variance $\sigma^2$, the estimate $s^2$ of $\sigma^2$ is given by

$$Q = (v - 1)s^2 = \Sigma w_i(u_i - \bar{u})^2$$

$$= \Sigma w_i u_i^2 - \frac{(\Sigma w_i u_i)^2}{\Sigma w_i} ,$$

where $\bar{u} = \Sigma w_i u_i / \Sigma w_i$. From this then

$$\frac{1}{v - 1} \Sigma w_i(\bar{\bar{y}}_{i..} - \bar{y}_w)^2 = \frac{1}{v - 1}\left\{\Sigma w_i \bar{\bar{y}}_{i..}^2 - \frac{(\Sigma w_i \bar{\bar{y}}_{i..})^2}{\Sigma w_i}\right\}$$

is an estimate of $\sigma_\delta^2$ obtained from the treatment means. This is compared in an F test with the within subclasses mean square which is also an independent estimate of $\sigma_\delta^2$."

Thus, when interaction is present the treatment sum of squares is obtained from equation (90). The replicate (the second category) sum of squares will be derived from a similar sum of squares; thus,

$$\Sigma w_j \bar{\bar{y}}_{.j.}^2 - \frac{(\Sigma w_j \bar{\bar{y}}_{.j.})^2}{\Sigma w_j} , \qquad (93)$$

where

$$\bar{\bar{y}}_{.j.} = \frac{1}{v} \sum_{i=1}^{v} \bar{y}_{ij.} , \qquad (94)$$

and

$$\frac{1}{w_j} = \frac{1}{v} \sum_{i=1}^{v} \frac{1}{n_{ij}} . \qquad (95)$$

The sum of squares in (93) is divided by r - 1 to obtain the mean square for the second category which is compared with the within classes mean square.

The sum of squares for interaction is:

$$\sum_{i=1}^{v} \sum_{j=1}^{r} \frac{Y_{ij.}^2}{n_{ij}} - \text{sum of squares due to } \hat{\mu},\ \hat{\rho}_j,\ \hat{\tau}_i,\ \text{assuming no interaction}$$

$$= \sum_{i=1}^{v} \sum_{j=1}^{r} \frac{Y_{ij.}^2}{n_{ij}} - \hat{\mu}Y_{...} - \Sigma\hat{\rho}_j Y_{.j.} - \Sigma\hat{\tau}_i Y_{i..}$$

$$= D_{yy} - \sum_{ij}\sum \left\{ \sum_h Y_{ijh}^2 - Y_{ij.}^2/n_{ij} \right\}, \tag{96}$$

where the quantities are obtained from equations (22) to (24), (30) to (32), and (50) (also, see (42) and (45)). Thus, it is necessary to go through the case I analysis in order to obtain the correct interaction sum of squares. To illustrate that this is correct consider the case where v = 2 and r = r. Then the interaction is obtained as the sum of squares of the differences. If these differences turn out to be weighted inversely to their variances then the interaction sum of squares follows the $\chi^2$ distribution with (2 - 1)(r - 1) = r - 1 degrees of freedom. Now, the two equations involving only $\hat{\tau}_1$ and $\hat{\tau}_2$ are (see equations (22) and (23)):

$$\hat{\tau}_1\left(n_{1.} - \sum_{j=1}^{r} \frac{n_{1j}^2}{n_{.j}}\right) - \hat{\tau}_2 \sum_j \frac{n_{1j}n_{2j}}{n_{.j}} = Y_{1..} - \Sigma n_{1j}\bar{y}_{.j.}\ ; \tag{97}$$

$$-\hat{\tau}_1 \Sigma \frac{n_{1j}n_{2j}}{n_{.j}} + \hat{\tau}_2\left(n_{2.} - \Sigma \frac{n_{2j}^2}{n_{.j}}\right) = Y_{2..} - \Sigma n_{2j}\bar{y}_{.j.}\ . \tag{98}$$

Remembering that the sum of the coefficients for any $\hat{\tau}_i$ is equal to zero and that the sum of the right hand sides of equations (97) and (98) is zero,

$$n_{1.} - \Sigma \frac{n_{1j}^{\;2}}{n_{.j}} = \Sigma \frac{n_{1j}n_{2j}}{n_{.j}} = n_{2.} - \Sigma \frac{n_{2j}^{\;2}}{n_{.j}}$$

$$= \Sigma n_{1j} - \Sigma \frac{n_{1j}}{n_{1j} + n_{2j}} = \Sigma_j n_{1j} \frac{(n_{1j} + n_{2j} - n_{1j})}{n_{1j} + n_{2j}}$$

$$= \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} \tag{99}$$

and

$$Y_{1..} - \Sigma n_{1j}\bar{y}_{.j.} = \Sigma n_{1j}\bar{y}_{1j.} - \Sigma n_{1j} \frac{(n_{1j}\bar{y}_{1j.} + n_{2j}\bar{y}_{2j.})}{n_{1j} - n_{2j}}$$

$$= \Sigma \frac{n_{1j}}{n_{1j} + n_{2j}} (n_{1j}\bar{y}_{1j.} + n_{2j}\bar{y}_{1j.} - n_{1j}\bar{y}_{1j.} - n_{2j}\bar{y}_{2j})$$

$$= \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.}) \tag{100}$$

Substituting in the results in equations (99) and (100) and $\hat{\tau}_1 = -\hat{\tau}_2$ in equation (97) we obtain

$$2\hat{\tau}_1 = \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.}) \Bigg/ \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} \, . \tag{101}$$

From equation (42) the sum of squares for treatments eliminating replicates is

$$\hat{\tau}_1(Y_{1..} - \Sigma n_{1j}\bar{y}_{.j.}) + \hat{\tau}_2(Y_{2..} - \Sigma n_{2j}\bar{y}_{.j.})$$

$$= \left[\Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.})\right]^2 \Bigg/ \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} \quad . \qquad (102)$$

Now, the variance of

$$\bar{y}_{1j.} - \bar{y}_{2j.} = \sigma_\delta^2\left( \frac{1}{n_{1j}} + \frac{1}{n_{2j}} \right) \quad . \qquad (103)$$

Therefore, from the above theorem the sum of the squares of the differences weighted inversely to the variance with which they are estimated is distributed as chi-square. This sum of squares is

$$\sum_{j=1}^{r} w_j d_j^2 - \frac{(\Sigma w_j d_j)^2}{\Sigma w_j}$$

$$= \Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.})^2 - \frac{\left[\Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.})\right]^2}{\Sigma \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}}}$$

$$= \sum_{j=1}^{r} \left\{ \frac{Y_{1j.}^2}{n_{1j}} + \frac{Y_{2j.}^2}{n_{2j}} - \frac{Y_{.j.}^2}{n_{1j} + n_{2j}} \right\} - \text{equation (102)}$$

= among subclasses within replicates - treatments eliminating replicates.
(104)

In the above,

$$\Sigma \frac{n_{1j}n_{2j}}{n_{1j}+n_{2j}} (\bar{y}_{1j.} - \bar{y}_{2j.})^2 = \Sigma \frac{n_{1j}n_{2j}}{n_{1j}+n_{2j}} \left\{ \left(\frac{Y_{1j.}}{n_{1j}}\right)^2 + \left(\frac{Y_{2j.}}{n_{2j}}\right)^2 - 2\,\frac{Y_{1j.}Y_{2j.}}{n_{1j}n_{2j}} \right\}$$

$$= \Sigma \left\{ \frac{Y_{1j}^{\ 2}}{n_{1j}} \left(1 - \frac{n_{1j}}{n_{1j}+n_{2j}}\right) + \frac{Y_{2j.}^{\ 2}}{n_{2j.}} \left(1 - \frac{n_{2j}}{n_{1j}+n_{2j}}\right) - 2\,\frac{Y_{1j.}Y_{2j.}}{n_{1j}+n_{2j}} \right\}$$

$$= \sum_{j} \left\{ \frac{Y_{1j.}^{\ 2}}{n_{1j}} + \frac{Y_{2j.}^{\ 2}}{n_{2j}} - \frac{(Y_{1j.}+Y_{2j.})^2}{n_{1j}+n_{2j}} \right\}, \qquad (105)$$

which is the within replicate sum of squares of the subclass totals. Thus, the interaction sum of squares obtained by subtraction is the correct sum of squares.

Proceeding now to a covariance analysis the various sums of products given in Table 6 are computed as follows:

$$T_{yy} = \Sigma\Sigma\Sigma Y_{ijh}^{\ 2} - Y_{...}^{\ 2}/n_{..} \qquad (106)$$

$$T_{xy} = \Sigma\Sigma\Sigma Y_{ijh}X_{ijh} - Y_{...}X_{...}/n_{..} \qquad (107)$$

$$T_{xx} = \Sigma\Sigma\Sigma X_{ijh}^{\ 2} - X_{...}^{\ 2}/n_{..} \qquad (108)$$

$$R_{yy} = \sum_{j=1}^{r} Y_{.j.}^{\ 2}/n_{.j} - Y_{...}^{\ 2}/n_{..} \qquad (109)$$

$$R_{xy} = \sum_{j=1}^{r} Y_{.j.}X_{.j.}/n_{.j} - Y_{...}X_{...}/n_{..} \qquad (110)$$

$$R_{xx} = \sum_{j=1}^{r} X_{.j.}^{\ 2}/n_{.j} - X_{...}^{\ 2}/n_{..} \qquad (111)$$

$$W_{yy} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{v} Y_{ij.}^{2}/n_{ij} - Y_{.j.}^{2}/n_{.j} \right\} \tag{112}$$

$$W_{xy} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{v} Y_{ij.}X_{ij.}/n_{ij} - Y_{.j.}X_{.j.}/n_{.j} \right\} \tag{113}$$

$$W_{xx} = \sum_{j=1}^{r} \left\{ \sum_{i=1}^{v} X_{ij.}^{2}/n_{ij} - X_{.j.}^{2}/n_{.j} \right\} \tag{114}$$

$$S_{yy} = \sum_{ij}\sum \left\{ \sum_{h=1}^{n_{ij}} Y_{ijh}^{2} - Y_{ij.}^{2}/n_{ij} \right\} \tag{115}$$

$$S_{xy} = \sum_{ij}\sum \left\{ \sum_{h} Y_{ijh}X_{ijh} - Y_{ij.}X_{ij.}/n_{ij} \right\} \tag{116}$$

$$S_{xx} = \sum_{ij}\sum \left\{ \sum_{h} X_{ijh}^{2} - X_{ij.}^{2}/n_{ij} \right\} \tag{117}$$

$$V_{yy} = \sum_{i=1}^{v} w_i \bar{\bar{y}}_{i..}^{2} - (\Sigma w_i \bar{\bar{y}}_{i..})^2/\Sigma w_i \tag{118}$$

$$V_{xy} = \sum_{i=1}^{v} w_i \bar{\bar{y}}_{i..}\bar{\bar{x}}_{i..} - (\Sigma w_i \bar{\bar{y}}_{i..})(\Sigma w_i \bar{\bar{x}}_{i..})/\Sigma w_i \tag{119}$$

$$V_{xx} = \sum_{i=1}^{v} w_i \bar{\bar{x}}_{i..}^{2} - (\Sigma w_i \bar{\bar{x}}_{i..})^2/\Sigma w_i \tag{120}$$

where $w_i$ is defined in equation (89) and $\bar{\bar{y}}_{i..}$ is defined in equation (91). $\bar{\bar{x}}_{i..}$ is obtained from an equation similar to equation (91).

$$I_{yy} = \sum_{ij}\sum \frac{Y_{ij.}^{2}}{n_{ij}} - \left\{ \hat{\mu}Y_{...} + \sum_{j} \hat{\rho}_j Y_{.j.} + \Sigma \hat{\tau}_i Y_{i..} \right\} = D_{yy} - S_{yy} \tag{121}$$

$$I_{xy} = \sum_{ij}\sum \frac{Y_{ij.}X_{ij.}}{n_{ij}} - \left\{\hat{\mu}X_{...} + \Sigma\hat{\rho}_j X_{.j.} + \Sigma\hat{\tau}_i X_{i..}\right\} = D_{xy} - S_{xy} \qquad (122)$$

$$I_{xx} = \sum_{ij}\sum \frac{X_{ij.}^{2}}{n_{ij}} - \left\{\hat{\mu}_x X_{...} + \Sigma\hat{\rho}_{jx} X_{.j.} + \Sigma\hat{\tau}_{ix} X_{i..}\right\} = D_{xx} - S_{xx} \qquad (123)$$

where the quantities in parentheses are defined in equations (45) to (47). The adjusted sums of squares are obtained as indicated in Table 6. The means are adjusted as indicated in equations (75) to (77) and (80) to (82). If desired the interaction effects may also be adjusted for regression as indicated in equations (83) to (86).

Example 2. The first example used to illustrate the computational procedure is synthesized from the data in Table 3. All replicates for which no plants were obtained for any variety are omitted. The result is that replicates 1, 5, 7, and 8 were omitted (see Table 7) in order that any plot would contain one or more plants. The various ratios required to compute the interaction sums of products are presented in Table 8. The various sums of products are:

<u>Example 2</u>. The first example used to illustrate the computational procedure for Case II is synthesized from the data in Table 3. All replicates for which no plants were obtained for any variety were omitted. The result is that replicates 1, 5, 7, and 8 were omitted (see Table 7) in order that all plots contain one or more plants. The various ratios required to compute the interaction sums of products are presented in Table 8. The various sums of products are:

<u>correction factor:</u>

for Y: $167.97^2/27 = 1044.96003.$

for XY: $167.97(3213)/27 = 19988.43000.$

for X: $3213^2/27 = 382347.00000.$

<u>total (formulae (106) to (108)); $n_{..} - 1$ df:</u>

$T_{yy} = 1131.68250 - 1044.96003 = 86.72247.$

$T_{xy} = 21332.01000 - 19988.43000 = 1343.58000.$

$T_{xx} = 407703.00000 - 382347.00000 = 25356.00000.$

<u>replicate (ignoring treatment)(formulae (109) to (111)); $r - 1$ df:</u>

$R_{yy} = 1079.81461 - 1044.96003 = 34.85458.$

$R_{xy} = 20534.67867 - 19988.43000 = 546.24867.$

$R_{xx} = 391401.11667 - 382347.00000 = 9054.11667.$

<u>among plots within replicates (formulae (112) to (114)); $r(v - 1)$ df:</u>

$W_{yy} = 1104.02755 - 1079.81461 = 24.21294.$

$W_{xy} = 20893.46000 - 20534.67867 = 358.78133.$

$W_{xx} = 399659.00000 - 391401.11667 = 8257.88333.$

among plants within plots (formulae (115) to (116)); $n_{..} - rv$ df;

$S_{yy} = 1131.68250 - 1104.02755 = 27.65495.$

$S_{xy} = 21332.01000 - 20893.46000 = 438.55000.$

$S_{xx} = 407703.00000 - 399659.00000 = 8044.00000.$

In order to compute the last three sums of products it is necessary to solve for $\hat{\tau}_i$, $\hat{\tau}_{ix}$, $\hat{\mu} + \hat{\rho}_j$, and $\hat{\mu}_x + \hat{\rho}_{jx}$. The three equations for the $\hat{\tau}_i$ are:

$$5.38334\ \hat{\tau}_1 - 2.81667\ \hat{\tau}_2 - 2.56667\ \hat{\tau}_3 = 3.74634.$$

$$-2.81667\ \hat{\tau}_1 + 6.03334\ \hat{\tau}_2 - 3.21667\ \hat{\tau}_3 = -.64816.$$

$$-2.56667\ \hat{\tau}_1 - 3.21667\ \hat{\tau}_2 + 5.78334\ \hat{\tau}_3 = -3.09816.$$

The solutions for the $\hat{\tau}_i$ are:

$$\hat{\tau}_1 = .46840; \quad \hat{\tau}_2 = -.09033; \quad \hat{\tau}_3 = -.37807.$$

The three equations for the $\hat{\tau}_{ix}$ are:

$$5.38334\ \hat{\tau}_{1x} - 2.81667\ \hat{\tau}_{2x} - 2.56667\ \hat{\tau}_{3x} = 48.38334.$$

$$-2.81667\ \hat{\tau}_{1x} + 6.03334\ \hat{\tau}_{2x} - 3.21667\ \hat{\tau}_{3x} = 2.68334.$$

$$-2.56667\ \hat{\tau}_{1x} - 3.21667\ \hat{\tau}_{2x} + 5.78334\ \hat{\tau}_{3x} = -51.06667.$$

The solutions for the $\hat{\tau}_{ix}$ are:

$$\hat{\tau}_{1x} = 6.08679; \quad \hat{\tau}_{2x} = 0.02688; \quad \hat{\tau}_{3x} = -6.11367.$$

The $\hat{\mu} + \hat{\rho}_j$ and the $\hat{\mu}_x + \hat{\rho}_{jx}$ values are computed from equation (50) and are given in Table 8. For example,

$$\hat{\mu} + \hat{\rho}_1 = 5.35200 - \left[.46840(.20000) - .09033(.40000) - .37807(.40000)\right]$$

$$= 5.44568.$$

<u>interaction (eliminating variety and replicate) (formulae (122) to (124));</u>

<u>$(r - 1)(v - 1)$ df:</u>

$$I_{yy} = 1104.02755 - \{26.76(5.44568) + 17.66(4.43758) + \dots + 35.00(5.83333)$$
$$+ .46840(54.59) - .09033(60.47) - .37807(52.91)\}$$
$$= 1104.02755 - 1082.79910 = 21.22845 = \text{(within rounding error)}$$
$$24.21294 - \{.46840(3.74634) - .09033(-.64816) - .37807(-3.09816)\}$$
$$= 24.21294 - 2.98466 = 21.22828.$$

$$I_{xy} = 20893.46000 - \{5.44568(526) + \dots + 5.83333(698) + .46840(1010)$$
$$- .09033(1180) - .37807(1023)\} = 20893.46000 - 20576.40245$$
$$= 317.05755 = \text{(within rounding errors)}\ 358.78133$$
$$- \{.46840(48.38334) - .09033(2.68334) - .37807(-51.06667)\}$$
$$= 358.78133 - 41.72715 = 317.05418.$$

$$I_{xx} = 399659 - \{526(106.41736) + 356(88.99328) + \dots + 698(116.33333)$$
$$+ 1010(6.08679) + 1180(.02688) + 1023(-6.11367)\} = 399659$$
$$- 392007.89153 = 7651.10847 = \text{(within rounding errors)}\ 8257.88333$$
$$- \{6.08679(48.38334) + .02688(2.68334) - 6.11367(-51.06667)\}$$
$$= 8257.88333 - 606.77613 = 7651.10720.$$

In order to compute the variety sums of products weighted inversely to the variance with which they are estimated it is convenient to construct a table of weights (Table 10). This table is similar to the one given by Snedecor (1946, table 11.30). Each cell in the table contains the number of individuals in the subclass, $n_{ij}$, the reciprocal of this number, $1/n_{ij}$, and the subclass means, $\bar{y}_{ij.}$ and $\bar{x}_{ij.}$, for the two variates. The next step is to sum the reciprocals of the numbers in each row and in each column of the table. The reciprocal of the sum of the reciprocals (i.e., $1/\sum_j \frac{1}{n_{ij}}$ and $1/\sum_i \frac{1}{n_{ij}}$) result in the weights $w_j'$ and $w_i'$. These weights are proportional to the weights $w_j$ and $w_i$ given by equations (95) and (89), respectively. In the former case the factor $v^2 w_j' = w_j$, and in the latter case $w_i' r^2 = w_i$. The proportionality factors are taken care of in the final sums of squares. The sums of products for variety are (equation (90)):

$$V_{yy} = 6^2 \left\{ .20000(6.99667)^2 + .25000(6.06000)^2 + .22222(5.82167)^2 - \frac{(4.20802)^2}{.67222} \right\}$$

$$= 6^2 \left\{ 26.50302 - 26.34172 \right\} = 5.80680$$

$$V_{xy} = 6^2 \left\{ .20000(6.99667)(132.33333) + .25000(6.06000)(119.58333) \right.$$

$$\left. + .22222(5.82167)(112.50000) - (4.20802)(81.30670)/.67222 \right\}$$

$$= 6^2 \left\{ 511.56414 - 508.97060 \right\} = 93.36744,$$

$$V_{xx} = 6^2 \left\{ .20000(132.33333)^2 + .25000(119.58333)^2 + .22222(112.50000)^2 \right.$$

$$\left. - 81.30670^2/.67222 \right\} = 1555.24536 ,$$

where the various quantities are given in the last 5 rows of Table 10. The $\bar{\bar{y}}_{i..}$ and $\bar{\bar{x}}_{i..}$ in Table 10 are the means of the subclass means (see equation (91)).

The sums of products for replicates is obtained from the last 5 columns of Table 10 is (see equation (93)):

$$R_{yy} = 3^2 \left\{ 5.53333(2.76666) + \dots + (5.83333)(3.88891) - 17.21156^2/2.76666 \right\}$$

$$= 3^2 \left\{ 109.69144 - 107.07416 \right\} = 23.55552$$

$$R_{xy} = 3^2 \left\{ 112.00000(2.76666) + \dots + 116.33333(3.88891) \right.$$

$$\left. - 17.21156(333.55594)/2.76666 \right\}$$

$$= 3^2 \left\{ 2116.29450 - 2075.07177 \right\} = 371.00457$$

$$R_{xx} = 3^2 \left\{ 112.00000(56.00000) + \dots + 116.33333(77.55594) \right.$$

$$\left. - 333.55594^2/2.76666 \right\}$$

$$= 3^2 \left\{ 40883.96294 - 40214.39750 \right\} = 6026.08896$$

The errors of estimate are computed as described in Tables 6 and 8. For example the error of estimate sum of squares for variety plus subclasses is

$$27.65495 + 5.80680 - \frac{(438.55000 + 93.36744)^2}{8044.00000 + 1555.24536}$$

$$= 33.46175 - 29.47483 = 3.98692,$$

and the sum of squares for varieties adjusted for replicates, interaction, and linear regression is

$$27.65495 + 5.80680 - \frac{(531.91744)^2}{9599.24536}$$

$$= 33.46175 - \frac{282936.16298}{9599.24536}$$

$$= 33.46175 - 29.47483$$

$$= 3.98692.$$

The variety, $\bar{\bar{y}}_{i..}$, and replicate, $\bar{\bar{y}}_{.j.}$, means given in Table 10 are adjusted for regression by equations (75) to (77) and equations (80) to (82), respectively. Thus, the adjusted treatment mean for variety 405 is equal to

$$6.99667 - \frac{438.55000}{8044.00000}\left(132.33333 - \frac{(146 + 90 + \ldots + 115.5)}{3 \times 6}\right)$$

$$= 6.99667 - (0.59969) = 6.39998.$$

Table 6. Analysis of covariance for a two-way classification with unequal numbers in the subclasses - Case II.

| Source of variation | df | Sum of products $y^2$ | $xy$ | $x^2$ |
|---|---|---|---|---|
| Total = T | $n_{..} - 1$ | $T_{yy}$ | $T_{xy}$ | $T_{xx}$ |
| Among subclasses = A | $rv - 1$ | $A_{yy}$ | $A_{xy}$ | $A_{xx}$ |
| Replicate (ignoring treatment) | $r - 1$ | $R_{yy}$ | $R_{xy}$ | $R_{xx}$ |
| Within replicate | $rv - r$ | $W_{yy}$ | $W_{xy}$ | $W_{xx}$ |
| Treatments (elim. replicate) | $v - 1$ | $V_{yy}$ | $V_{xy}$ | $V_{xx}$ |
| * Interaction (elim. treatment and replicate) = I | $f_e$ | $I_{yy}$ | $I_{xy}$ | $I_{xx}$ |
| Within subclasses = S = error | $n_{..} - rv$ | $S_{yy}$ | $S_{xy}$ | $S_{xx}$ |

| Source of variation | Regression df | Regression ss | Errors of estimate df | Errors of estimate ss |
|---|---|---|---|---|
| S | 1 | $S_{xy}{}^2/S_{xx}$ | $n_{..}-rv-1$ | $S_{yy}' = S_{yy} - S_{xy}{}^2/S_{xx}$ |
| S + I | 1 | $\frac{(S_{xy} + I_{xy})^2}{(S_{xx} + I_{xx})}$ | $n_{..}-r-v$ | $B_{yy}' = S_{yy} + I_{yy} - \frac{(S_{xy} + I_{xy})^2}{(S_{xx} + I_{xx})}$ |
| S + V | 1 | $\frac{(S_{xy} + V_{xy})^2}{(S_{xx} + V_{xx})}$ | $n_{..}-rv+v-2$ | $C_{yy}' = S_{yy} + V_{yy} - \frac{(S_{xy} + V_{xy})^2}{(S_{xx} + V_{xx})}$ |
| Residual (elim. treatment and replicate) adjusted for regression | | | $f_e$ | $I_{yy}' = B_{yy}' - S_{yy}'$ |
| Treatment (elim. replicate) adjusted for regression | | | $v-1$ | $V_{yy}' = C_{yy}' - S_{yy}'$ |

* $f_e = rv - r - v + 1.$

Table 7. Number of plants per replicate (n), yield of rubber in grams (Y), and dry weight of shrub in grams (X) for the normal plants of three guayule varieties in a selected set of replicates.

| | Variety 405 | | | | | Variety 407 | | | | | Variety 416 | | | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rep. | n | Y | X | Y | X | n | Y | X | Y | X | n | Y | X | Y | X | N | Y | X |
| 2 | 1 | 6.44 | 146 | - | - | 2 | 5.00 | 88 | 5.12 | 92 | 2 | 4.35 | 96 | 5.85 | 104 | 5 | 26.76 | 526 |
| 3 | 1 | 5.20 | 111 | - | - | 2 | 2.59 | 61 | 4.79 | 93 | 1 | 5.08 | 91 | - | - | 4 | 17.66 | 356 |
| 4 | 1 | 8.55 | 170 | - | - | 2 | 6.46 | 127 | 10.70 | 194 | 1 | 5.41 | 101 | - | - | 4 | 31.12 | 592 |
| 6 | 2 | 6.77 | 115 | 9.03 | 120 | 1 | 6.81 | 139 | - | - | 1 | 6.48 | 132 | - | - | 4 | 29.09 | 506 |
| 9 | 1 | 9.17 | 151 | - | - | 1 | 5.44 | 116 | - | - | 2 | 5.04 | 90 | 8.69 | 178 | 4 | 28.34 | 535 |
| 10 | 2 | 5.41 | 106 | 4.02 | 91 | 2 | 6.56 | 130 | 7.00 | 140 | 2 | 4.45 | 91 | 7.56 | 140 | 6 | 35.00 | 698 |
| | 8 | | | 54.59 | 1010 | 10 | | | 60.47 | 1180 | 9 | | | 52.91 | 1023 | 27 | 167.97 | 3213 |

Table 8. Analysis of covariance for the data of Table 7 assuming the following linear model:

$$Y_{ijh} = \rho + \rho_j + \tau_i + \rho\tau_{ij} + \beta X_{ijh} + \delta_{ijh} .$$

| Source of variation | df | Sums of products $y^2$ | $xy$ | $x^2$ |
|---|---|---|---|---|
| Total | 26 | 86.72247 | 1343.58000 | 25356.00000 |
| Among subclasses | 17 | 59.06752 | 905.03000 | 17312.00000 |
| Replicates | 5 | 34.85458 | 546.24867 | 9054.11667 |
| Within replicates | 12 | 24.21294 | 358.78133 | 8257.88333 |
| Variety (elim. replicate) | 2 | 2.985 | 41.727 | 606.776 |
| Interaction (I) | 10 | 21.228 | 317.054 | 7651.107 |
| Within subclasses | 9 | 27.65495 | 438.55000 | 8044.00000 |
| Variety (equation (90))(V) | 2 | 5.80680 | 93.36744 | 1555.24536 |
| Replicate (equation (93))(R) | 5 | 23.55552 | 371.00457 | 6026.08896 |

| | Regression df | Regression ss | Errors of estimate df | Errors of estimate ss | Errors of estimate ms |
|---|---|---|---|---|---|
| Within subclasses (S) | 1 | 23.90926 | 8 | 3.74569 | .46821 |
| S + I | 1 | 36.37678 | 18 | 12.50617 | -- |
| S + V | 1 | 29.47483 | 10 | 3.98692 | -- |
| S + R | 1 | 46.57956 | 13 | 4.63091 | -- |
| Interaction (adjusted for variety, replicate, regression) | | | 10 | 8.76048 | .87605 |
| Variety (adjusted for replicate, interaction, regression) | | | 2 | 0.24123 | .12062 |
| Replicate (adjusted for variety, interaction, regression) | | | 5 | 0.88522 | .17704 |

Table 9. Values of $n_{ij}$, $n_{ij}/n_{\cdot j}$, totals, means, and effects for the data of Table 7.

| Rep. | Variety 405 | 407 | 416 | $n_{\cdot j}$ | $Y_{\cdot j \cdot}$ | $Y_{\cdot j \cdot}/n_{\cdot j}$ | $\hat{\rho}_j+\hat{\mu}_y$ | $X_{\cdot j \cdot}$ | $X_{\cdot j \cdot}/n_{\cdot j}$ | $\hat{\rho}_j+\hat{\mu}_x$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 2 | 2 | | | | | | | |
| | .20000 | .40000 | .40000 | 5 | 26.76 | 5.35200 | 5.44568 | 526 | 105.20000 | 106.41736 |
| 3 | 1 | 2 | 1 | | | | | | | |
| | .25000 | .50000 | .25000 | 4 | 17.66 | 4.41500 | 4.43758 | 356 | 89.00000 | 88.99328 |
| 4 | 1 | 2 | 1 | | | | | | | |
| | .25000 | .50000 | .25000 | 4 | 31.12 | 7.78000 | 7.80258 | 592 | 148.00000 | 147.99328 |
| 6 | 2 | 1 | 1 | | | | | | | |
| | .50000 | .25000 | .25000 | 4 | 29.09 | 7.27250 | 7.15540 | 506 | 126.50000 | 124.97830 |
| 9 | 1 | 1 | 2 | | | | | | | |
| | .25000 | .25000 | .50000 | 4 | 28.34 | 7.08500 | 7.17952 | 535 | 133.75000 | 135.27842 |
| 10 | 2 | 2 | 2 | | | | | | | |
| | .33333 | .33333 | .33333 | 6 | 35.00 | 5.83333 | 5.83333 | 698 | 116.33333 | 116.33333 |
| | | | | | | | 37.85409 | | | 719.99397 |
| $n_{i\cdot}$ | 8 | 10 | 9 | 27 | | | | | | |
| $Y_{i\cdot}$ | 54.59 | 60.47 | 52.91 | | 167.97 | | | | | |
| $Y_{i\cdot}/n_{i\cdot}$ | 6.82375 | 6.04700 | 5.87889 | | | | | | | |
| $\hat{\tau}_i$ | 0.46840 | -0.09033 | -0.37807 | | | | $\hat{\mu}_y = 6.30902$ | | | $\hat{\mu}_x = 119.99900$ |
| $X_{i\cdot\cdot}$ | 1010 | 1180 | 1023 | | | | | 3213 | | |
| $X_{i\cdot\cdot}/n_i$ | 126.25000 | 118.00000 | 113.66667 | | | | | | | |
| $\hat{\tau}_{ix}$ | 6.08679 | 0.02688 | -6.11367 | | | | | | | |

Table 10. Computations in an rxv table when interaction is present.

| Rep. | Variety 405 | 407 | 416 | $\sum_i \frac{1}{n_{ij}}$ | $w_j'$ | $\bar{\bar{y}}_{.j.}$ | $\bar{\bar{x}}_{.j.}$ | $w_j'\bar{\bar{y}}_{.j.}$ | $w_j'\bar{\bar{x}}_{.j.}$ |
|---|---|---|---|---|---|---|---|---|---|
| 2 | $n_{11}=1$ | $n_{21}=2$ | $n_{31}=2$ | - | | | | | |
| | $1/n_{11}=1.0$ | $1/n_{21}=.5$ | $1/n_{31}=.5$ | 2.0 | .50000 | | | | |
| | $\bar{y}_{11.}=6.44$ | $\bar{y}_{21.}=5.06$ | $\bar{y}_{31.}=5.10$ | | | 5.53333 | | 2.76666 | |
| | $\bar{x}_{11.}=146.0$ | $\bar{x}_{21.}=90.0$ | $\bar{x}_{31.}=100.0$ | | | | 112.00000 | | 56.00000 |
| 3 | $n_{12}=1$ | $n_{22}=2$ | $n_{32}=1$ | - | | | | | |
| | $1/n_{12}=1.0$ | $1/n_{22}=.5$ | $1/n_{32}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{12.}=5.20$ | $\bar{y}_{22.}=3.69$ | $\bar{y}_{32.}=5.08$ | | | 4.65667 | | 1.86267 | |
| | $\bar{x}_{12.}=111.0$ | $\bar{x}_{22.}=77.0$ | $\bar{x}_{32.}=91.0$ | | | | 93.00000 | | 37.20000 |
| 4 | $n_{13}=1$ | $n_{23}=2$ | $n_{33}=1$ | - | | | | | |
| | $1/n_{13}=1.0$ | $1/n_{23}=.5$ | $1/n_{33}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{13.}=8.55$ | $\bar{y}_{23.}=8.58$ | $\bar{y}_{33.}=5.41$ | | | 7.51333 | | 3.00533 | |
| | $\bar{x}_{13.}=170.0$ | $\bar{x}_{23.}=160.5$ | $\bar{x}_{33.}=101.0$ | | | | 143.83333 | | 57.53333 |
| 6 | $n_{14}=2$ | $n_{24}=1$ | $n_{34}=1$ | | | | | | |
| | $1/n_{14}=.5$ | $1/n_{24}=1.0$ | $1/n_{34}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{14.}=7.90$ | $\bar{y}_{24.}=6.81$ | $\bar{y}_{34.}=6.48$ | | | 7.06333 | | 2.82533 | |
| | $\bar{x}_{14.}=117.5$ | $\bar{x}_{24.}=139.0$ | $\bar{x}_{34.}=132.0$ | | | | 129.50000 | | 51.80000 |

Table 10, continued

| Rep. | Variety 405 | 407 | 416 | $\sum_i \frac{1}{n_{ij}}$ | $w_j'$ | $\bar{\bar{y}}_{.j.}$ | $\bar{\bar{x}}_{.j.}$ | $w_j'\bar{\bar{y}}_{.j.}$ | $w_j'\bar{\bar{x}}_{.j.}$ |
|---|---|---|---|---|---|---|---|---|---|
| 9 | $n_{15}=1$ | $n_{25}=1$ | $n_{35}=2$ | | | | | | |
| | $1/n_{15}=1.0$ | $1/n_{25}=1.0$ | $1/n_{35}=.5$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{15.}=9.17$ | $\bar{y}_{25.}=5.44$ | $\bar{y}_{35.}=6.86$ | | | 7.15667 | | 2.86267 | |
| | $\bar{x}_{15.}=151.0$ | $\bar{x}_{25.}=116.0$ | $\bar{x}_{35.}=134.0$ | | | | 133.66667 | | 53.46667 |
| 10 | $n_{16}=2$ | $n_{26}=2$ | $n_{36}=2$ | | | | | | |
| | $1/n_{16}=.5$ | $1/n_{26}=.5$ | $1/n_{36}=.5$ | 1.5 | .66667 | | | | |
| | $\bar{y}_{16.}=4.72$ | $\bar{y}_{26.}=6.78$ | $\bar{y}_{36.}=6.00$ | | | 5.83333 | | 3.88891 | |
| | $\bar{x}_{16.}=98.5$ | $\bar{x}_{26.}=135.0$ | $\bar{x}_{36.}=115.5$ | | | | 116.33333 | | 77.55594 |
| $\sum_j \frac{1}{n_{ij}}$ | 5.0 | 4.0 | 4.5 | | 2.76667 | 37.75666 | 728.33333 | 17.21156 | 333.55594 |
| $w_i'$ | .20000 | .25000 | .22222 | .67222 | | | | | |
| $\bar{\bar{y}}_{i..}$ | 6.99667 | 6.06000 | 5.82167 | 18.87834 | | | | | |
| $\bar{\bar{x}}_{i..}$ | 132.33333 | 119.58333 | 112.25000 | 364.16667 | | | | | |
| $w_i'\bar{\bar{y}}_{i..}$ | 1.39933 | 1.51500 | 1.29369 | 4.20802 | | | | | |
| $w_i'\bar{\bar{x}}_{i..}$ | 26.46667 | 29.89583 | 24.94420 | 81.30670 | | | | | |

Example 3. This example is included to illustrate the calculational procedure when one or more of the $n_{ij}$ = zero. It should be emphasized that the procedure is approximate and that there appears to be little else that can be done if an analysis is desired. (see Henderson, 1948) In order to obtain an approximate analysis it is assumed that

$$\sum_{i=1}^{v_j} \widehat{\rho\tau}_{ij} = \sum_{j=1}^{r_i} \widehat{\rho\tau}_{ij} = 0.$$

These assumptions are probably not too bad for the infinite model since the $E(\widehat{\rho\tau}_{ij})$ = zero. However, in the finite model the $E(\widehat{\rho\tau}_{ij}) = \widehat{\rho\tau}_{ij}$. Thus, if a zero is substituted for $\widehat{\rho\tau}_{ij}$ for a cell in which the $n_{ij} = 0$, a bias results. If the number of subclasses in which there are no observations is relatively small the amount of the bias probably will not seriously affect the results. However, with several missing subclasses the appropriateness of the resulting analysis is open to question.

The data of Table 3 are used to illustrate the computational procedure for a covariance analysis for data from a two-way classification with unequal numbers in the subclasses and with interaction assumed present. The various ratios, total, and means required for the case I analysis are given in Table 5, and the sums of squares are presented in Table 12. (A number of these sums of squares are obtained directly from Table 4 and from the calculations associated with Table 4.) Thus,

<u>total with 36 df:</u>

$$T_{yy} = 1550.27570 - 1447.06303 = 103.21267.$$
$$T_{xy} = 29516.20000 - 27873.11432 = 1643.08568.$$
$$T_{xx} = 569181.00000 - 536887.81081 = 32293.18919.$$

<u>among subclasses with 25 df:</u>

$$A_{yy} = 1516.11025 - 1447.06303 = 69.04722.$$
$$A_{xy} = 28974.00500 - 27873.11432 = 1100.89068.$$
$$A_{xx} = 559486.50000 - 536887.81081 = 22598.68919.$$

<u>within subclasses with 11 df:</u>

$$S_{yy} = 1550.27570 - 1516.11025 = 34.16545.$$
$$S_{xy} = 29516.20000 - 28974.00500 = 542.19500.$$
$$S_{xx} = 569181.00000 - 559486.50000 = 9694.50000.$$

<u>interaction with 14 df (formulae (121) to (123))(see example 1):</u>

$$I_{yy} = 1516.11025 - 1490.92150 = 25.18875.$$
$$I_{xy} = 28974.00500 - 28540.02481 = 433.98019.$$
$$I_{xx} = 559486.50000 - 547837.60759 = 11648.89241.$$

<u>variety with 2 df (formula (90) and Table 11):</u>

$$V_{yy} = 10^2 \left\{ 6.73200(.79202) + 6.36667(.97951) + 5.59286(1.01689) - \frac{2.78842^2}{.45332} \right\}$$
$$= 100(17.2554190 - 17.1518709) = 10.35481.$$

$$V_{xy} = 10^2 \left\{ 6.73200(15.41803) + 6.36667(19.13723) \right.$$
$$\left. + 5.59286(19.75344) - \frac{2.78842(54.30870)}{.45332} \right\}$$
$$= 100(336.1128305 - 334.0586457) = 205.41848.$$

$$V_{xx} = 10^2 \left\{ 131.05000(15.41803) + 124.38889(19.13723) \right.$$

$$\left. + 108.64286(19.75344) - \frac{54.30870^2}{.45332} \right\}$$

$$= 100(6547.0618453 - 6506.2977492) = 4076.40961.$$

replicate with 9 df (formula (93) and Table 11):

$$R_{yy} = 3^2 \left\{ 6.52500(3.26250) + \ldots + 5.83333(3.88891) - \frac{32.14162^2}{5.10001} \right\}$$

$$= \cancel{37.94613}.\ 37.91670.$$

$$R_{xy} = 3^2 \left\{ 122.50000(3.26250) + \ldots + 116.33333(3.88891) \right.$$

$$\left. - \frac{32.14162(630.30685)}{5.10001} \right\}$$

$$= 561.50469.$$

$$R_{xx} = 3^2 \left\{ 122.50000(61.25000) + \ldots + 116.33333(77.55594) - \frac{630.30685^2}{5.10001} \right\}$$

$$= 9294.435\cancel{72}^{45}.$$

The error of estimate sums of squares are obtained in the usual manner. The variety means adjusted for linear regression ~~be~~ are computed from formulas (75) to (77) with $r_1$ replacing r; thus, the adjusted mean for variety 416 is

$$\frac{1}{7}(5.10 + \ldots + 6.00) - \frac{542.19500}{9694.50000} \left\{ \frac{100.0 + \ldots + 115.5}{7} - \frac{143.0 + \ldots + 115.5}{26} \right\}$$

$$= 5.59286 - .055928 \left\{ 108.64286 - 122.71154 \right\}$$

$$= 6.37969.$$

Table 11. Computations in an rxv table when interaction is present and when some of the $n_{ij} = 0$.

| Rep. | Variety 405 | 407 | 416 | $\sum_{j=1}^{v_j} \frac{1}{n_{ij}}$ | $w_j'$ | $\sum_{i=1}^{v_j} \bar{y}_{ij.}/v_j = \bar{\bar{y}}_{.j.}$ | $\sum_{i=1}^{v_j} \bar{x}_{ij.}/v_j = \bar{\bar{x}}_{.j.}$ | $w_j' \bar{\bar{y}}_{.j.}$ | $w_j' \bar{\bar{x}}_{.j.}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | $n_{11}=1$ | $n_{21}=1$ | 0 | | | | | | |
| | $1/n_{11}=1.0$ | $1/n_{21}=1.0$ | - | 2.0 | .50000 | | | | |
| | $\bar{y}_{11.}=6.93$ | $\bar{y}_{21.}=61.2$ | - | | | 6.52500 | | 3.26250 | |
| | $\bar{x}_{11.}=143.0$ | $\bar{x}_{21.}=102.0$ | - | | | | 122.50000 | | 61.25000 |
| 2 | $n_{12}=1$ | $n_{22}=2$ | $n_{32}=2$ | | | | | | |
| | $1/n_{12}=1.0$ | $1/n_{22}=.5$ | $1/n_{32}=.5$ | 2.0 | .50000 | | | | |
| | $\bar{y}_{12.}=6.44$ | $\bar{y}_{22.}=5.06$ | $\bar{y}_{32.}=5.10$ | | | 5.53333 | | 2.76666 | |
| | $\bar{x}_{12.}=146.0$ | $\bar{x}_{22.}=90.0$ | $\bar{x}_{32.}=100.0$ | | | | 112.00000 | | 56.00000 |
| 3 | $n_{13}=1$ | $n_{23}=2$ | $n_{33}=1$ | | | | | | |
| | $1/n_{13}=1.0$ | $1/n_{23}=.5$ | $1/n_{33}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{13.}=5.20$ | $\bar{y}_{23.}=3.69$ | $\bar{y}_{33.}=5.08$ | | | 4.65667 | | 1.86267 | |
| | $\bar{x}_{13.}=111.0$ | $\bar{x}_{23.}=77.0$ | $\bar{x}_{33.}=91.0$ | | | | 93.00000 | | 37.20000 |
| 4 | $n_{14}=1$ | $n_{24}=1$ | $n_{34}=1$ | | | | | | |
| | $1/n_{14}=1.0$ | $1/n_{24}=.5$ | $1/n_{34}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{14.}=8.55$ | $\bar{y}_{24.}=8.58$ | $\bar{y}_{34.}=5.41$ | | | 7.51333 | | 3.00533 | |
| | $\bar{x}_{14.}=170.0$ | $\bar{x}_{24.}=160.5$ | $\bar{x}_{34.}=101.0$ | | | | 143.83333 | | 57.53333 |

Table 11, continued

| Rep. | Variety 405 | 407 | 416 | $\sum_{j=1}^{v_j} \frac{1}{n_{ij}}$ | $w_j'$ | $\sum_{i=1}^{v_j} \bar{y}_{ij.}/v_j = \bar{\bar{y}}_{.j.}$ | $\sum_{i=1}^{v_j} \bar{x}_{ij.}/v_j = \bar{\bar{x}}_{.j.}$ | $w_j' \bar{\bar{y}}_{.j.}$ | $w_j' \bar{\bar{x}}_{.j.}$ |
|---|---|---|---|---|---|---|---|---|---|
| 5 | $n_{15}=2$ | $n_{25}=1$ | 0 | | | | | | |
| | $1/n_{15}=.5$ | $1/n_{25}=1.0$ | - | 1.5 | .66667 | | | | |
| | $\bar{y}_{15.}=5.74$ | $\bar{y}_{25.}=7.66$ | - | | | 6.70000 | | 4.46669 | |
| | $\bar{x}_{15.}=107.5$ | $\bar{x}_{25.}=169.0$ | - | | | | 138.25000 | | 92.16713 |
| 6 | $n_{16}=2$ | $n_{26}=1$ | $n_{36}=1$ | | | | | | |
| | $1/n_{16}=.5$ | $1/n_{26}=1.0$ | $1/n_{36}=1.0$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{16.}=7.90$ | $\bar{y}_{26.}=6.81$ | $\bar{y}_{36.}=6.48$ | | | 7.06333 | | 2.82533 | |
| | $\bar{x}_{16.}=117.5$ | $\bar{x}_{26.}=139.0$ | $\bar{x}_{36.}=132.0$ | | | | 129.50000 | | 51.80000 |
| 7 | $n_{17}=1$ | $n_{27}=2$ | 0 | | | | | | |
| | $1/n_{17}=1.0$ | $1/n_{27}=.5$ | - | 1.5 | .66667 | | | | |
| | $\bar{y}_{17.}=7.10$ | $\bar{y}_{27.}=7.16$ | - | | | 7.13000 | | 4.75336 | |
| | $\bar{x}_{17.}=137.0$ | $\bar{x}_{27.}=131.0$ | - | | | | 134.00000 | | 89.33378 |
| 8 | $n_{18}=1$ | $n_{28}=0$ | $n_{38}=1$ | | | | | | |
| | $1/n_{18}=1.0$ | - | $1/n_{38}=1.0$ | 2.0 | .50000 | | | | |
| | $\bar{y}_{18.}=5.57$ | - | $\bar{y}_{38.}=4.22$ | | | 4.89500 | | 2.44750 | |
| | $\bar{x}_{18.}=129.0$ | - | $\bar{x}_{38.}=87.0$ | | | | 108.00000 | | 54.00000 |

Table 11, continued

| Rep. | Variety | | | $\sum_{j=1}^{v_j} \frac{1}{n_{ij}}$ | $w_j'$ | $\sum_{i=1}^{v_j} \bar{y}_{ij.}/v_j = \bar{\bar{y}}_{.j.}$ | $\sum_{i=1}^{v_j} \bar{x}_{ij.}/v_j = \bar{\bar{x}}_{.j.}$ | $w_j' \bar{\bar{y}}_{.j.}$ | $w_j' \bar{\bar{x}}_{.j.}$ |
|---|---|---|---|---|---|---|---|---|---|
| | 405 | 407 | 416 | | | | | | |
| 9 | $n_{19}=1$ | $n_{29}=1$ | $n_{39}=2$ | | | | | | |
| | $1/n_{19}=1.0$ | $1/n_{29}=1.0$ | $1/n_{39}=.5$ | 2.5 | .40000 | | | | |
| | $\bar{y}_{19.}=9.17$ | $\bar{y}_{29.}=5.44$ | $\bar{y}_{39.}=6.86$ | | | 7.15667 | | 2.86267 | |
| | $\bar{x}_{19.}=151.0$ | $\bar{x}_{29.}=116.0$ | $\bar{x}_{39.}=134.0$ | | | | 133.66667 | | 53.46667 |
| 10 | $n_{10}=2$ | $n_{20}=2$ | $n_{30}=2$ | | | | | | |
| | $1/n_{10}=.5$ | $1/n_{20}=.5$ | $1/n_{30}=.5$ | 1.5 | .66667 | | | | |
| | $\bar{y}_{10.}=4.72$ | $\bar{y}_{20.}=6.78$ | $\bar{y}_{30.}=6.00$ | | | 5.83333 | | 3.88891 | |
| | $\bar{x}_{10.}=98.5$ | $\bar{x}_{20.}=135.0$ | $\bar{x}_{30.}=115.5$ | | | | 116.33333 | | 77.55594 |
| $\sum_i 1/n_{ij}$ | 8.5 | 6.5 | 5.5 | | 5.10001 | 63.00666 | 1231.08333 | 32.14162 | 630.30685 |
| $w_i'$ | .11765 | .15385 | .18182 | .45332 | | | | | |
| $\bar{\bar{y}}_{i..}$ | 6.73200 | 6.36667 | 5.59286 | 18.69153 | | | | | |
| $\bar{\bar{x}}_{i..}$ | 131.05000 | 124.38889 | 108.64286 | 364.08175 | | | | | |
| $w_i' \bar{\bar{y}}_{i..}$ | .79202 | .97951 | 1.01689 | 2.78842 | | | | | |
| $w_i' \bar{\bar{x}}_{i..}$ | 15.41803 | 19.13723 | 19.75344 | 54.30870 | | | | | |

Table 12. Analysis of covariance for data of Table 3 assuming the following linear model:

$$Y_{ijh} = \mu + \rho_j + \tau_i + \rho\tau_{ij} + \beta X_{ijh} + \delta_{ijh} .$$

| | | Sums of products | | |
|---|---|---|---|---|
| Source of variation | df | $y^2$ | xy | $x^2$ |
| Total | 36 | 103.21267 | 1643.08568 | 32293.18919 |
| Among subclasses | 25 | 69.04722 | 1100.89068 | 22598.68919 |
| Replicate and variety | 11 | 43.85847 | 666.91049 | 10949.79678 |
| Interaction (I) | 14 | 25.18875 | 433.98019 | 11648.89241 |
| Within subclasses (S) | 11 | 34.16545 | 542.19500 | 9694.50000 |
| Variety (equation (90))(V) | 2 | 10.35481 | 205.41848 | 4076.40961 |
| Replicate (equation (93))(R) | 9 | 37.91670 | 561.50469 | 9294.43545 |

| | Regression | | | Errors of estimate | |
|---|---|---|---|---|---|
| | df | ss | df | ss | ms |
| Within subclasses | 1 | 30.32394 | 10 | 3.84151 | .38415 |
| S + V | 1 | 40.58744 | 12 | 3.93282 | -- |
| S + R | 1 | 64.15067 | 19 | 7.93148 | -- |
| S + I | 1 | 44.64698 | 24 | 14.70722 | -- |
| Variety (adjusted for replicate, interaction, regression) | | | 2 | 0.09131 | .04566 |
| Replicate (adjusted for variety, interaction, regression) | | | 9 | 4.08997 | .45444 |
| Interaction (adjusted for variety, replicate, regression) | | | 14 | 10.86571 | .77612 |

Case III. Interaction regression used to adjust treatment means and interaction error mean square used to test hypotheses about means (infinite model for one or both categories).

If one or more of the categories is considered to be a sample of categories from a large population of categories the interaction mean square is used to test hypotheses about equality of means. In such a case, there may not be much interest in estimating an interaction effect. Instead, one may wish to estimate the variance component associated with the interaction effects. The problem of estimating effects is complicated by the fact that the weights are usually unknown. If the interaction regression and mean square are used the sums of squares to be minimized is

$$\text{Res} = \sum_{i=1}^{v} \sum_{j=1}^{r} w_{ij}(\bar{y}_{ij.} - \mu - \tau_i - \rho_j - \beta_1 \bar{x}_{ij.})^2, \tag{124}$$

where

$$Y_{ijh} = \mu + \tau_i + \rho_j + \epsilon_{ij} + \beta_1 \bar{x}_{ij} + \delta_{ijh} + \beta(X_{ijh} - \bar{x}_{ij.}), \tag{125}$$

where $\mu$, $\tau_i$, $\rho_j$, $\beta$, and $\delta_{ijh}$ are defined in formula (52), where $\epsilon_{ij}$ represents a random component associated with the ijth subclass, and where $\beta_1$ equals the interaction regression. One could minimize the sum of squares:

$$\sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} (Y_{ijh} - \mu - \tau_i - \rho_j - \beta_1 \bar{x}_{ij.} - \beta(X_{ijh} - \bar{x}_{ij}))^2, \tag{126}$$

but the same estimates of $\mu$, $\tau_i$, $\rho_j$, and $\beta_1$ obtained from minimizing (124) will be the same as from minimizing (126) if $w_{ij}$ in (124) is set equal to $n_{ij}$.

Since this is true and since

$$\hat{\beta} = \frac{\sum\sum_{ij} \left\{ \sum_h X_{ijh} Y_{ijh} - X_{ij.} Y_{ij.}/n_{ij} \right\} = S_{xy}}{\sum\sum_{ij} \left\{ \sum_h X_{ijh}^2 - X_{ij.}^2/n_{ij} \right\} = S_{xx}} , \tag{127}$$

the sum of squares in (124) instead of the sum of squares in (126) is minimized. Also, $w_{ij}$ is not always equal to $n_{ij}$.

The problem of the correct weighting is simple theoretically. Thus, the variance of any $\bar{y}_{ij.}$ in the absence of regression is

$$w_{ij} = 1 \Big/ \left( \sigma_\epsilon^2 + \frac{\sigma_\delta^2}{n_{ij}} \right) = \frac{n_{ij}}{n_{ij}\sigma_\epsilon^2 + \sigma_\delta^2} . \tag{128}$$

In practice, however, $\sigma_\epsilon^2$ and $\sigma_\delta^2$ are usually unknown, and, ~~[illegible]~~ consequently $w_{ij}$ is unknown. Also, the middle term of (128) implies that no $n_{ij} = 0$. However, an analysis is still possible even though one or more $n_{ij}$ = zero. The method presented below is applicable for one or more $n_{ij} = 0$, although the solution of the equations is more difficult. In order to obtain a partial solution consider the following two cases:

(i) $\sigma_\epsilon^2$ is large relative to $\sigma_\delta^2/n_{ij}$;

(ii) $\sigma_\epsilon^2$ is small relative to $\sigma_\delta^2/n_{ij}$.

In case (i) $w_{ij}$ is near unity, and for all practical purposes one may use $w_{ij} = 1$. If the $n_{ij}$ are large and if $\sigma_\epsilon^2$ is not small compared to $\sigma_\delta^2$, $w_{ij}$ will essentially be one. This results in an analysis of covariance on the subclass means. If all subclasses contain one or more observations, the analysis of covariance proceeds in the same manner as for equal numbers per subclass. If the second situation prevails then $w_{ij}$ is set equal to

$n_{ij}$ and the analysis goes through as described below.

However, the true situation is more often in between the two limiting cases cited above, and in order to effect a solution to the problem it will be necessary to have reasonably good estimates of $\sigma_\epsilon^2$ and $\sigma_\delta^2$. In order to illustrate the general procedure the formulae obtained below will contain $w_{ij}$. The appropriate weight may be substituted in effecting the solution to a given problem.

The partial derivatives of (124) with respect to $\mu$, $\tau_i$, $\rho_j$, and $\beta_1$ are:

$$\frac{\delta \text{ Res}}{\delta \mu} = -2\sum_{ij}\sum w_{ij}(\bar{y}_{ij.} - \mu - \tau_i - \rho_j - \beta_1\bar{x}_{ij.}) . \tag{129}$$

$$\frac{\delta \text{ Res}}{\delta \tau_i} = -2\sum_{j} w_{ij}(\bar{y}_{ij.} - \mu - \tau_i - \rho_j - \beta_1\bar{x}_{ij.}) . \tag{130}$$

$$\frac{\delta \text{ Res}}{\delta \rho_j} = -2\sum_{i} w_{ij}(\bar{y}_{ij.} - \mu - \tau_i - \rho_j - \beta_1\bar{x}_{ij.}) . \tag{131}$$

$$\frac{\delta \text{ Res}}{\delta \beta_1} = -2\sum_{ij}\sum w_{ij}\bar{x}_{ij.}(\bar{y}_{ij} - \mu - \tau_i - \rho_j - \beta_1\bar{x}_{ij.}) . \tag{132}$$

If we set the above equations equal to zero the following normal equations are obtained:

equation for $\mu$:

$$w_{..}\hat{\mu} + \sum_{i} w_{i.}\hat{\tau}_i + \sum_{j} w_{.j}\hat{\rho}_j + \hat{\beta}_1\sum_{ij}\sum w_{ij}\bar{x}_{ij.} = \sum_{ij}\sum w_{ij}\bar{y}_{ij.} . \tag{133}$$

equations for $\rho_j$:

$$w_{.1}(\hat{\mu} + \hat{\rho}_1) + \Sigma w_{i1}\hat{\tau}_i + \hat{\beta}_1\Sigma w_{i1}\bar{x}_{i1.} = \Sigma w_{i1}\bar{y}_{i1.} . \tag{134}$$

$$w_{.2}(\hat{\mu} + \hat{\rho}_2) + \Sigma w_{i2}\hat{\tau}_i + \hat{\beta}_1 \Sigma w_{i2}\bar{\bar{x}}_{i2.} = \Sigma w_{i2}\bar{\bar{y}}_{i2.} \; . \tag{135}$$

$\vdots$

$$w_{.r}(\hat{\mu} + \hat{\rho}_r) + \Sigma w_{ir}\hat{\tau}_i + \hat{\beta}_1 \Sigma w_{ir}\bar{\bar{x}}_{ir.} = \Sigma w_{ir}\bar{\bar{y}}_{ir.} \; . \tag{136}$$

<u>equations for $\tau_i$:</u>

$$w_{1.}(\hat{\mu} + \hat{\tau}_1) + \Sigma w_{1j}\hat{\rho}_j + \hat{\beta}_1 \Sigma w_{1j}\bar{\bar{x}}_{1j.} = \Sigma w_{1j}\bar{\bar{y}}_{1j.} \; . \tag{137}$$

$$w_{2.}(\hat{\mu} + \hat{\tau}_2) + \Sigma w_{2j}\hat{\rho}_j + \hat{\beta}_1 \Sigma w_{2j}\bar{\bar{x}}_{2j.} = \Sigma w_{2j}\bar{\bar{y}}_{2j.} \; . \tag{138}$$

$\vdots$

$$w_{r}(\hat{\mu} + \hat{\tau}_v) + \Sigma w_{vj}\hat{\rho}_j + \hat{\beta}_1 \Sigma w_{vj}\bar{\bar{x}}_{vj.} = \Sigma w_{vj}\bar{\bar{y}}_{vj.} \; . \tag{139}$$

<u>equation for $\beta_1$:</u>

$$\hat{\mu}\Sigma\Sigma w_{ij}\bar{\bar{x}}_{ij.} + \sum_i \hat{\tau}_i \sum_j w_{ij}\bar{\bar{x}}_{ij.} + \sum_j \hat{\rho}_j \sum_i w_{ij}\bar{\bar{x}}_{ij.} + \hat{\beta}_1 \Sigma w_{ij}\bar{\bar{x}}_{ij.}^2$$

$$= \Sigma\Sigma w_{ij}\bar{\bar{x}}_{ij.}\bar{\bar{y}}_{ij.} \tag{140}$$

If the jth equation in equations (134) to (136) is divided through by $w_{.j}$, then

$$\hat{\mu} + \hat{\rho}_j = \frac{1}{w_{.j}} \sum_i w_{ij}(\bar{\bar{y}}_{ij.} - \hat{\beta}_1\bar{\bar{x}}_{ij.} - \hat{\tau}_i). \tag{141}$$

Summing over j and imposing the restriction that

$$\sum_{j=1}^{r} \hat{\rho}_j = 0 , \tag{142}$$

the estimate of $\mu$ is:

$$\hat{\mu} = \frac{1}{r} \Sigma \frac{1}{w_{.j}} \sum_i w_{ij}(\bar{y}_{ij.} - \hat{\beta}_1 \bar{x}_{ij.} - \hat{\tau}_i) \tag{143}$$

Substituting the estimate for $\hat{\rho}_j$ in equation (140) we obtain:

$$\hat{\beta}_1 = \frac{\Sigma\Sigma w_{ij}\bar{x}_{ij.}\bar{y}_{ij.} - \sum_j \frac{1}{w_{.j}}(\sum_i w_{ij}\bar{y}_{ij})(\sum_i w_{ij}\bar{x}_{ij.}) - \sum_i \hat{\tau}_i \sum_j w_{ij}\bar{x}_{ij.} + \sum_j \frac{1}{w_{.j}} \sum_i w_{ij}\hat{\tau}_i \sum_i w_{ij}\bar{x}_{ij.}}{\sum_{ij}\Sigma w_{ij}\bar{x}_{ij.}^2 - \sum_j(\sum_i w_{ij}\bar{x}_{ij.})^2/w_{.j}} \tag{144}$$

Substituting the solution for $\hat{\rho}_j$ in equations (137) to (139) and setting $\hat{\beta}_1$ = zero, we obtain:

$$\hat{\tau}_1(w_{1.} - \sum_j w_{1j}^2/w_{.j}) - \sum_{i=2}^{v} \hat{\tau}_i \sum_{j=1}^{r} \frac{w_{1j}w_{ij}}{w_{.j}}$$

$$= \sum_j w_{1j}\bar{y}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{y}_{ij.} \; . \tag{145}$$

$$-\hat{\tau}_1 \sum_j \frac{w_{1j}w_{2j}}{w_{.j}} + \hat{\tau}_2(w_{2.} - \sum_j w_{2j}^2/w_{.j}) - \sum_{i=3}^{v} \hat{\tau}_i \Sigma \frac{w_{2j}w_{ij}}{w_{.j}}$$

$$= \sum_j w_{2j}\bar{y}_{2j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{y}_{ij.} \; . \tag{146}$$

$$\vdots$$

$$- \sum_{i=1}^{v-1} \hat{\tau}_i \sum_{j=1}^{r} \frac{w_{vj}w_{ij}}{w_{.j}} + \hat{\tau}_v(w_{v.} - \sum_{j=1}^{r} \frac{w_{vj}^2}{w_{.j}})$$

$$= \sum_j w_{vj}\bar{y}_{vj.} - \sum_{j=1}^{r} \frac{w_{vj}}{w_{.j}} \sum_i w_{ij}\bar{y}_{ij.} \; . \tag{147}$$

If $\hat{\beta}_1$ is not zero the equations involving only the $\hat{t}_i$ are:

$$\hat{t}_1 \left\{ w_{1.} - \sum_{j=1}^{r} \frac{w_{1j}^{2}}{w_{.j}} - \frac{1}{G_{xx}} \Big(\sum_j w_{1j}\bar{\bar{x}}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big)^2 \right\}$$

$$- \sum_{i=2}^{v} \hat{t}_i \left\{ \sum_j \frac{w_{1j}w_{ij}}{w_{.j}} + \frac{1}{G_{xx}} \Big(\sum_j w_{1j}\bar{\bar{x}}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big)\right.$$

$$\left.\Big(\sum_j \frac{1}{w_{.j}} \sum_{i=2}^{v} w_{ij}\hat{t}_i \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij} - \sum_{i=2}^{v} \hat{t}_i \sum_{j=1}^{r} w_{ij}\bar{\bar{x}}_{ij.}\Big) \right\}$$

$$= \sum_j w_{1j}\bar{\bar{y}}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{y}}_{ij.} - \frac{G_{xy}}{G_{xx}} \Big(\sum_j w_{1j}\bar{\bar{x}}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big). \qquad (148)$$

$$- \hat{t}_1 \left\{ \sum_{j=1}^{r} \frac{w_{1j}w_{2j}}{w_{.j}} + \frac{1}{G_{xx}} \Big(\sum_j w_{1j}\bar{\bar{x}}_{1j.} - \sum_j \frac{w_{1j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big)\Big(\sum_j w_{2j}\bar{\bar{x}}_{2j.} - \sum_j \frac{w_{2j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big) \right\}$$

$$+ \hat{t}_2 \left\{ w_{2.} - \sum_{j=1}^{r} \frac{w_{2j}^{2}}{w_{.j}} - \frac{1}{G_{xx}} \Big(\sum_j w_{2j}\bar{\bar{x}}_{2j.} - \sum_j \frac{w_{2j}}{w_{.j}} \sum_i w_{ij}\bar{\bar{x}}_{ij.}\Big)^2 \right\}$$

$$- \sum_{i=3}^{v} \hat{t}_i \left\{ \sum_{j=1}^{r} \frac{w_{ij}w_{2j}}{w_{.j}} + \frac{1}{G_{xx}} \Big(\sum_i w_{2j}\bar{\bar{x}}_{2j.} - \sum_j \frac{w_{2j}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\Big)\right.$$

$$\left.\Big(\sum_j w_{ij}\bar{\bar{x}}_{ij.} - \sum_j \frac{w_{ij}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\Big) \right\}$$

$$= \sum_j w_{2j}\bar{\bar{y}}_{2j.} - \sum \frac{w_{2j}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{y}}_{ij.} - \frac{G_{xy}}{G_{xx}} \left(\sum_j w_{2j}\bar{\bar{x}}_{2j.} - \sum_j \frac{w_{2j}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\right) \qquad (149)$$

$$\vdots$$

$$- \sum_{i=1}^{v-1} \hat{\tau}_i \left\{ \sum_{j=1}^{r} \frac{w_{ij}w_{vj}}{w_{.j}} + \frac{1}{G_{xx}} \left(\sum_j w_{2j}\bar{\bar{x}}_{2j.} - \sum_j \frac{w_{2j}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\right)\left(\sum_j w_{ij}\bar{\bar{x}}_{ij.}\right.\right.$$

$$\left.\left. - \sum_j \frac{w_{ij}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\right)\right\} + \hat{\tau}_v \left\{ w_{v.} - \sum \frac{w_{vj}^2}{w_{.j}} - \left(\sum_j w_{vj}\bar{\bar{x}}_{vj.} - \sum_j \frac{w_{vj}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\right)^2 \right\}$$

$$= \sum_j w_{vj}\bar{\bar{y}}_{vj.} - \sum \frac{w_{vj}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{y}}_{ij.} - \frac{G_{xy}}{G_{xx}} \left(\sum_j w_{vj}\bar{\bar{x}}_{vj.} - \sum_j \frac{w_{vj}}{w_{.j}} \sum_{i=1}^{v} w_{ij}\bar{\bar{x}}_{ij.}\right) \qquad (150)$$

where

$$G_{xy} = \sum\sum w_{ij}\bar{\bar{x}}_{ij.}\bar{\bar{y}}_{ij.} - \sum_j \frac{1}{w_{.j}} \left(\sum_i w_{ij}\bar{\bar{y}}_{ij.}\right)\left(\sum_i w_{ij}\bar{\bar{x}}_{ij.}\right) \qquad (151)$$

and

$$G_{xx} = \sum\sum w_{ij}\bar{\bar{x}}_{ij.}^2 - \sum_j \left(\sum_i w_{ij}\bar{\bar{x}}_{ij.}\right)^2 / w_{.j} \; . \qquad (152)$$

Solution of the above equations results in estimates of $\hat{\tau}_i$ which are [redacted] adjusted for replicate effect and for regression. The $\hat{\mu} + \hat{\tau}_i$ is the adjusted treatment mean.

With the estimates $\hat{\tau}_i$ obtained from equations (146) to (148) and with the various totals it is possible to compute the various sums of squares in Table 15. Thus, total sum of squares with $n_{..} - 1$ df:

$$T_{yy} = \Sigma\Sigma\Sigma Y_{ijh}^2 - Y_{...}^2/n_{..} \; . \tag{153}$$

$$T_{xy} = \Sigma\Sigma\Sigma X_{ijh}Y_{ijh} - X_{...}Y_{...}/n_{..} \; . \tag{154}$$

$$T_{xx} = \Sigma\Sigma\Sigma X_{ijh}^2 - X_{...}^2/n_{..} \; . \tag{155}$$

among subclasses with rv - 1 df:

$$A_{yy} = \Sigma\Sigma Y_{ij.}^2/n_{ij} - Y_{...}^2/n_{..} \; . \tag{156}$$

$$A_{xy} = \Sigma\Sigma Y_{ij.}X_{ij.}/n_{ij} - X_{...}Y_{...}/n_{..} \; . \tag{157}$$

$$A_{xx} = \Sigma\Sigma X_{ij.}^2/n_{ij} - X_{...}^2/n_{..} \; . \tag{158}$$

within subclasses with $n_{..} - rv$ df:

$$S_{yy} = \Sigma\Sigma \left\{ \sum_h Y_{ijh}^2 - Y_{ij.}^2/n_{ij} \right\} . \tag{159}$$

$$S_{xy} = \Sigma\Sigma \left\{ \sum_h Y_{ijh}X_{ijh} - Y_{ij.}X_{ij.}/n_{ij} \right\} . \tag{160}$$

$$S_{xx} = \Sigma\Sigma \left\{ \sum_h X_{ijh}^2 - X_{ij.}^2/n_{ij} \right\} . \tag{161}$$

replicate (ignoring treatment) with r - 1 df:

$$R_{yy} = \Sigma Y_{.j.}^2/n_{.j} - Y_{...}^2/n_{..} \; . \tag{162}$$

$$R_{xy} = \Sigma Y_{.j.}X_{.j.}/n_{.j} - X_{...}Y_{...}/n_{..} \; . \tag{163}$$

$$R_{xx} = \Sigma X_{.j.}^2/n_{.j} - X_{...}^2/n_{..} \; . \tag{164}$$

among subclasses within replicates with r(v - 1)df:

$$W_{yy} = A_{yy} - R_{yy} = \Sigma \left\{ \Sigma Y_{ij.}^2/n_{ij} - Y_{.j.}^2/n_{.j} \right\} . \quad (165)$$

$$W_{xy} = A_{xy} - R_{xy} = \Sigma \left\{ \Sigma Y_{ij.}X_{ij.}/n_{ij} - Y_{.j.}X_{.j.}/n_{.j} \right\} . \quad (166)$$

$$W_{xx} = A_{xx} - R_{xx} = \Sigma \left\{ \Sigma X_{ij.}^2/n_{ij} - X_{.j.}^2/n_{.j} \right\} . \quad (167)$$

treatment (eliminating replicate) with v - 1 df:

$$V_{yy} = \hat{\mu}\Sigma\Sigma w_{ij}\bar{y}_{ij.} + \sum_i \hat{\tau}_i \sum_j w_{ij}\bar{y}_{ij.} + \sum_j \hat{\rho}_j \sum_i w_{ij}\bar{y}_{ij.} - R_{yy} , \quad (168)$$

$$V_{xy} = \hat{\mu}\Sigma\Sigma w_{ij}\bar{x}_{ij.} + \sum_i \hat{\tau}_i \sum_j w_{ij}\bar{x}_{ij.} + \sum_j \hat{\rho}_j \sum_i w_{ij}\bar{x}_{ij.} - R_{xy}, \text{ and} \quad (169)$$

$$V_{xx} = \hat{\mu}_x\Sigma\Sigma w_{ij}\bar{x}_{ij.} + \sum_i \hat{\tau}_{ix} \sum_j w_{ij}\bar{x}_{ij.} + \sum_j \hat{\rho}_{jx} \sum_i w_{ij}\bar{x}_{ij.} - R_{xx} , \quad (170)$$

where the $\hat{\tau}_i$ are obtained from equations (145) to (147) and the $\hat{\tau}_{ix}$ are obtained from the same equations with $\bar{x}_{ij.}$ substituted for $\bar{y}_{ij.}$, and where the $\hat{\mu} + \hat{\rho}_j$ values are obtained from equation (141) with $\hat{\beta}_1$ set equal to zero, and the $\hat{\mu}_x + \hat{\rho}_x$ values are obtained from the same equation with $\bar{x}_{ij.}$ substituted for $\bar{y}_{ij.}$ .

experimental error with (r - 1)(v - 1) df:

$$E_{yy} = W_{yy} - V_{yy} . \quad (171)$$

$$E_{xy} = W_{xy} - V_{xy} . \quad (172)$$

$$E_{xx} = W_{xx} - V_{xx} . \quad (173)$$

The sum of squares for the errors of estimates may be obtained as:

$$E_{yy}' = W_{yy} - \Sigma(\hat{\mu} + \hat{\rho}_j) \sum_j w_{ij}\bar{y}_{ij.} - \Sigma\hat{\tau}_i \sum_j w_{ij}\bar{y}_{ij.}$$

$$- \hat{\beta}_1 \Sigma\Sigma w_{ij}\bar{x}_{ij.}\bar{y}_{ij.} , \tag{174}$$

and

$$V_{yy}' = \sum_j(\hat{\mu} + \hat{\rho}_j) \sum_i w_{ij}\bar{y}_{ij.} + \sum_i \hat{\tau}_i \sum_j w_{ij}\bar{y}_{ij.} + \hat{\beta}_1 \Sigma\Sigma w_{ij}\bar{x}_{ij.}\bar{y}_{ij.}$$

- sum of squares due to $\hat{\mu}'$, $\hat{\rho}_j'$, and $\hat{\beta}_1'$ assuming zero treatment effect. (175)

An alternative procedure for computing $E_{yy}'$ and $V_{yy}'$ is presented in Table 15.

The covariance analysis for $w_{ij} = n_{ij}$ is given in chapter XVI of "Experimental Design - Theory and Application", Federer, 1954.

Table 15. Analysis of covariance for a two-way classification with unequal numbers in the subclasses - Case III.

| Source of variation | df | Sum of products $y^2$ | $xy$ | $x^2$ |
|---|---|---|---|---|
| Total = T | $n_{..} - 1$ | $T_{yy}$ | $T_{xy}$ | $T_{xx}$ |
| Among subclasses = A | $rv - 1$ | $A_{yy}$ | $A_{xy}$ | $A_{xx}$ |
| Replicate (ignoring treatment) | $r - 1$ | $R_{yy}$ | $R_{xy}$ | $R_{xx}$ |
| Within replicate | $vr - r$ | $W_{yy}$ | $W_{xy}$ | $W_{xx}$ |
| Treatment (elim. replicate) = V | $v - 1$ | $V_{yy}$ | $V_{xy}$ | $V_{xx}$ |
| Error (elim. treatment and replicate) | $f_e$ | $E_{yy}$ | $E_{xy}$ | $E_{xx}$ |
| Within subclasses = S | $n_{..} - rv$ | $S_{yy}$ | $S_{xy}$ | $S_{xx}$ |

| Source of variation | Regression df | Regression ss | Errors of estimate df | Errors of estimate ss |
|---|---|---|---|---|
| Within subclasses | 1 | $S_{xy}^2/S_{xx}$ | $n_{..}-rv-1$ | $S_{yy}' = S_{yy} - S_{xy}^2/S_{xx}$ |
| Error | 1 | $E_{xy}^2/E_{xx}$ | $rv-r-v$ | $E_{yy}' = E_{yy} - E_{xy}^2/E_{xx}$ |
| Error + V | 1 | $W_{xy}^2/W_{xx}$ | $rv-r-1$ | $W_{yy}' = W_{yy} + W_{xy}^2/W_{xx}$ |
| Treatment (elim. replicate) adjusted for experimental error regression | | | $v - 1$ | $V_{yy}' = W_{yy}' - E_{yy}'$ |