Analysis of Variance for Multiple Factors

Two Factor Analysis

Consider two factors (treatments) A and B with A done at a levels and B done at b levels. Within a given treatment combination of A and B levels, labeled by i and j respectively, r repeated measurements are made of a response variable y. Since the number of measurements for each i, j level is the same, this kind of experimental design is called "balanced". The k'th score in the i, j treatment combination is designated as

$$y_{i,j,k}$$

i is the the treatment factor index for factor A ; $1 \le i \le a$
j is the the treatment factor index for factor B ; $1 \le j \le b$
k labels the score the within the i, j level ; $1 \le k \le r$.

The null hypothesis asserts that no treatment population differs from any other. Thus, the scores in the ab treatment levels should have the same mean and variance. For theoretical convenience, we will assume the scores are normally distributed with a common variance within each treatment level, but the tests which follow are fairly robust for departures from normality.

Null Hypothesis: $\mathrm{H}_0:\ \mu_{i,j} = \mu_0$ for $1 \le i \le a,\ 1 \le j \le b$, where μ_0 is the common population mean of the ab treatment populations. Thus, all of the variation seen in the measured scores is due to sample variations, i.e., randomness.

The total number of scores is $n = abr$. The average of all scores is the "grand mean" and is given by

$$\bar{y} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{i,j,k}}{abr} . \qquad \textbf{(1)}$$

The sample mean of the A B treatment combination i, j is given by

$$\bar{y}_{\mathrm{AB}\ i,j} = \frac{\sum_{k=1}^{r} y_{i,j,k}}{r} . \qquad \textbf{(2)}$$

The sample mean of treatment A at level i is

$$\bar{y}_{\mathrm{A}\ i} = \frac{\sum_{j=1}^{b}\sum_{k=1}^{r} y_{i,j,k}}{br} . \qquad \textbf{(3)}$$

Similarly, the sample mean of the B treatment at level j is

$$\bar{y}_{\mathrm{B}\ j} = \frac{\sum_{i=1}^{a}\sum_{k=1}^{r} y_{i,j,k}}{ar} . \qquad \textbf{(4)}$$

Now, the definitions of these means imply that the following sums of deviations vanish.

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\,i}-\bar{y}\right)=0 \qquad \textbf{(5)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\,j}-\bar{y}\right)=0 \qquad \textbf{(6)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{A}\,i}\right)=0 \qquad \textbf{(7)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{B}\,j}\right)=0 \qquad \textbf{(8)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{A}\,i}-\bar{y}_{\mathrm{B}\,j}+\bar{y}\right)=0 \qquad \textbf{(9)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{A}\,i}-\bar{y}_{\mathrm{B}\,j}+\bar{y}\right)=0 \qquad \textbf{(10)}$$

$$\sum_{k=1}^{r}\left(y_{i,j,k}-\bar{y}_{\mathrm{AB}\,i,j}\right)=0 \qquad \textbf{(11)}$$

Any effects due to treatments should manifest themselves in the deviations of the treatment sample means from the grand mean. The effects can be listed as follows:

The Effect due to A at level *i*: $\bar{y}_{\mathrm{A}\,i}-\bar{y}$

The Effect due to B at level *j*: $\bar{y}_{\mathrm{B}\,j}-\bar{y}$

The Interaction Effect of A at level *i* and B at level *j*:

$$\left(\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}\right)-\left(\bar{y}_{\mathrm{A}\,i}-\bar{y}\right)-\left(\bar{y}_{\mathrm{B}\,j}-\bar{y}\right)=\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{A}\,i}-\bar{y}_{\mathrm{B}\,j}+\bar{y}$$

The Residual or Error Effect: $y_{i,j,k}-\bar{y}_{\mathrm{AB}\,i,j}$

The AB Interaction Effect measures how A and B when both present can cause an effect different from their separate effects. The Effect due to Error measures the assumed random variation of the response variable within the treatment combination i, j .

Now, from equation (5), only $a-1$ of the deviations $\bar{y}_{\mathrm{A}\,i}-\bar{y}$ can be independently specified. Similarly, the deviations $\bar{y}_{\mathrm{B}\,j}-\bar{y}$ have $b-1$ degrees of freedom. The equation for the AB interaction effect, $\bar{y}_{\mathrm{AB}\,i,j}-\bar{y}_{\mathrm{A}\,i}-\bar{y}_{\mathrm{B}\,j}+\bar{y}$, defines the components of an $a\times b$ matrix. From equations (9) and (10) both the row and column sums of this matrix are constrained. As a consequence, in the first $a-1$ rows there are only $b-1$ independent entries. The entries in the last row are then completely determined by the column sum requirement. Thus, the AB interaction has $(a-1)\times(b-1)$ degrees of freedom. Finally, from equation (11), for each of the ab A and B factor combinations there are $r-1$ independent deviations from the sample mean $\bar{y}_{\mathrm{AB}\,i,j}$. The following table summarizes the degrees of freedom, ν , associated with each effect.

Effect	Degrees of Freedom
Grand Mean	$\nu_{\bar{y}} = 1$
Factor A	$\nu_{\text{A}} = a - 1$
Factor B	$\nu_{\text{B}} = b - 1$
Interaction of Factors AB	$\nu_{\text{AB}} = (a-1)(b-1)$
Residual or Error	$\nu_E = ab(r-1)$

Note that the sum of the degrees of freedom is the total number of scores $n = abr$.

Formally, we can write the following.

$$y_{i,j,k} = \bar{y} + \left(\bar{y}_{\text{A}\,i} - \bar{y}\right) + \left(\bar{y}_{\text{B}\,j} - \bar{y}\right) + \left(\bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y}\right) + \left(y_{i,j,k} - \bar{y}_{\text{AB}\,i,j}\right) \tag{12}$$

If the null hypothesis is true and all of the treatment combinations give the same mean result, then $\bar{y}$ should be a fair estimate of every score and the effects just listed all measure random variation.

Equation (12) can be thought of an orthogonal vector decomposition of the response variable *y*. Each of the *abr* measurements can be thought of as a component of an *abr* dimensional column vector.

$$\begin{pmatrix} y_{1,1,1} \\ y_{1,1,2} \\ \vdots \\ y_{1,1,r} \\ y_{1,2,1} \\ y_{1,2,2} \\ \vdots \\ y_{1,2,r} \\ \vdots \\ y_{i,j,1} \\ \vdots \\ y_{i,j,r} \\ \vdots \\ y_{a,b,r} \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \\ \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \\ \vdots \\ \bar{y} \\ \vdots \\ \bar{y} \\ \vdots \\ \bar{y} \end{pmatrix} + \begin{pmatrix} \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,a} - \bar{y} \end{pmatrix} + \begin{pmatrix} \bar{y}_{\text{B}\,1} - \bar{y} \\ \bar{y}_{\text{B}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,1} - \bar{y} \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,b} - \bar{y} \end{pmatrix} + \begin{pmatrix} \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,a,b} - \bar{y}_{\text{A}\,a} - \bar{y}_{\text{B}\,b} + \bar{y} \end{pmatrix} + \begin{pmatrix} y_{1,1,1} - \bar{y}_{\text{AB}\,1,1} \\ y_{1,1,2} - \bar{y}_{\text{AB}\,1,1} \\ \vdots \\ y_{1,1,r} - \bar{y}_{\text{AB}\,1,1} \\ y_{1,2,1} - \bar{y}_{\text{AB}\,1,2} \\ y_{1,2,2} - \bar{y}_{\text{AB}\,1,2} \\ \vdots \\ y_{1,2,r} - \bar{y}_{\text{AB}\,1,2} \\ \vdots \\ y_{i,j,1} - \bar{y}_{\text{AB}\,i,j} \\ \vdots \\ y_{i,j,r} - \bar{y}_{\text{AB}\,i,j} \\ \vdots \\ y_{a,b,r} - \bar{y}_{\text{AB}\,a,b} \end{pmatrix} \tag{13}$$

From the deviation constraints of equations (5) through (11) the dot product of any two different column vectors on the right hand side of equation (13) must be zero, i.e., the column vectors on the right side of equation (13) are mutually orthogonal. The column vector which gives the total deviation from the grand mean, $y_{i,j,k} - \bar{y}$, can be decomposed as the sum $V_{\text{A}} + V_{\text{B}} + V_{\text{AB}} + V_E$, where the mutually orthogonal column vectors are given by the following equations.

$$V_{\mathrm{A}} = \begin{pmatrix} \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{A}\,a} - \bar{y} \end{pmatrix};\; V_{\mathrm{B}} = \begin{pmatrix} \bar{y}_{\mathrm{B}\,1} - \bar{y} \\ \bar{y}_{\mathrm{B}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{B}\,1} - \bar{y} \\ \bar{y}_{\mathrm{B}\,2} - \bar{y} \\ \bar{y}_{\mathrm{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{B}\,b} - \bar{y} \end{pmatrix};\; V_{\mathrm{AB}} = \begin{pmatrix} \bar{y}_{\mathrm{AB}\,1,1} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,1} + \bar{y} \\ \bar{y}_{\mathrm{AB}\,1,1} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,1} + \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{AB}\,1,1} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,1} + \bar{y} \\ \bar{y}_{\mathrm{AB}\,1,2} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,2} + \bar{y} \\ \bar{y}_{\mathrm{AB}\,1,2} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{AB}\,1,2} - \bar{y}_{\mathrm{A}\,1} - \bar{y}_{\mathrm{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\mathrm{AB}\,a,b} - \bar{y}_{\mathrm{A}\,a} - \bar{y}_{\mathrm{B}\,b} + \bar{y} \end{pmatrix};\; V_E = \begin{pmatrix} y_{1,1,1} - \bar{y}_{\mathrm{AB}\,1,1} \\ y_{1,1,2} - \bar{y}_{\mathrm{AB}\,1,1} \\ \vdots \\ y_{1,1,r} - \bar{y}_{\mathrm{AB}\,1,1} \\ y_{1,2,1} - \bar{y}_{\mathrm{AB}\,1,2} \\ y_{1,2,2} - \bar{y}_{\mathrm{AB}\,1,2} \\ \vdots \\ y_{1,2,r} - \bar{y}_{\mathrm{AB}\,1,2} \\ \vdots \\ y_{i,j,1} - \bar{y}_{\mathrm{AB}\,i,j} \\ \vdots \\ y_{i,j,r} - \bar{y}_{\mathrm{AB}\,i,j} \\ \vdots \\ y_{a,b,r} - \bar{y}_{\mathrm{AB}\,a,b} \end{pmatrix}$$

Because of the orthogonality condition on the four V vectors, the total variation in y can be decomposed into contributions associated with effects A, B, AB and Error.

$$\begin{aligned} SS_y &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(y_{i,j,k} - \bar{y}\right)^2 = \left(V_{\mathrm{A}} + V_{\mathrm{B}} + V_{\mathrm{AB}} + V_E\right)\bullet\left(V_{\mathrm{A}} + V_{\mathrm{B}} + V_{\mathrm{AB}} + V_E\right) \\ &= SS_{\mathrm{A}} + SS_{\mathrm{B}} + SS_{\mathrm{AB}} + SSE \end{aligned} \quad \mathbf{(14)}$$

$$SS_{\mathrm{A}} = V_{\mathrm{A}}\bullet V_{\mathrm{A}} = br\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\,i} - \bar{y}\right)^2 \quad \mathbf{(15)}$$

$$SS_{\mathrm{B}} = V_{\mathrm{B}}\bullet V_{\mathrm{B}} = ar\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\,j} - \bar{y}\right)^2 \quad \mathbf{(16)}$$

$$SS_{\mathrm{AB}} = V_{\mathrm{AB}}\bullet V_{\mathrm{AB}} = r\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y}\right)^2 \quad \mathbf{(17)}$$

$$SSE = V_E\bullet V_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(y_{i,j,k} - \bar{y}_{\mathrm{AB}\,i,j}\right)^2 \quad \mathbf{(18)}$$

This decomposition allows us to test the null hypothesis. If the null hypothesis is true then the mean square of the Effects A, B and AB all estimate the same common population variance of each A B factor combination. The mean squares are computed as a variation divided by the associated degrees of freedom.

The Mean Square due to Factor A: $MS_{\mathrm{A}} = \dfrac{SS_{\mathrm{A}}}{a-1} = \dfrac{br\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\,i} - \bar{y}\right)^2}{a-1}$ **(19)**

The Mean Square due to Factor B: $MS_{\text{B}} = \dfrac{SS_{\text{B}}}{b-1} = \dfrac{ar\sum_{j=1}^{b}\left(\bar{y}_{\text{B}\,j} - \bar{y}\right)^2}{b-1}$ **(20)**

The Mean Square due to Interaction of A with B:

$$MS_{\text{AB}} = \frac{SS_{\text{AB}}}{(a-1)(b-1)} = \frac{r\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y}\right)^2}{(a-1)(b-1)} \qquad \textbf{(21)}$$

The Error or Residual Mean Square: $MSE = \dfrac{SSE}{ab(r-1)} = \dfrac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(y_{i,j,k} - \bar{y}_{\text{AB}\,i,j}\right)^2}{ab(r-1)}$ **(22)**

To test the null hypothesis that all effects measure the same common random variation associated with the Mean Square due to Error, we compute the observed Fisher F score.

$$F_{\text{obs}} = \frac{\text{Effect Mean Square}}{MSE} \qquad \textbf{(23)}$$

For a given level of significance, α, this value is compared against a critical score calculated from an F distribution used to compare two variances obtained from sampling variances from two normally distributed populations. The numerator degrees of freedom is appropriate to the effect in question and the denominator degrees of freedom is $ab(r-1)$. This information is easily summarized in a Two-Factor ANOVA table.

Source	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	$\mathbf{F_{obs}}$
Factor A	SS_{A}	$a-1$	$MS_{\text{A}} = \dfrac{SS_{\text{A}}}{a-1}$	$\dfrac{MS_{\text{A}}}{MSE}$
Factor B	SS_{B}	$b-1$	$MS_{\text{B}} = \dfrac{SS_{\text{B}}}{b-1}$	$\dfrac{MS_{\text{B}}}{MSE}$
Interaction AB	SS_{AB}	$(a-1)(b-1)$	$MS_{\text{AB}} = \dfrac{SS_{\text{AB}}}{(a-1)(b-1)}$	$\dfrac{MS_{\text{AB}}}{MSE}$
Error	SSE	$ab(r-1)$	$MSE = \dfrac{SSE}{ab(r-1)}$	
Total	SS_y	$n-1 = abr-1$	$s_y^2 = \dfrac{SS_y}{n-1}$	

Each computed $\mathbf{F_{obs}}$ is compared against $F_\alpha\left(\nu_{\text{effect}}, \nu_E\right)$ for a stated level of significance, α. If $F_{\text{obs}} < F_\alpha\left(\nu_{\text{effect}}, \nu_E\right)$, we fail to reject the null hypothesis that the effect is just an alias for random error (variation). If $F_{\text{obs}} > F_\alpha\left(\nu_{\text{effect}}, \nu_E\right)$, we reject H_0 and conclude the effect in question is real. **Note**: It is possible that the AB Interaction Effect can test as significant even when one or both of the separate A Factor and B Factor Effects are insignificant.

Two Factor Analysis with Replication

A problem that often occurs, particularly in industry or manufacturing, is that of Replication. This refers to seemingly uncontrolled variation in measurements made on what are nominally samples set up under identical treatments. This can sometimes be attributed to uncontrolled factors that vary at the time of unit construction such as barometric pressure, humidity, human assembly variations, etc. Replication can be thought of as including time as a factor in product performance. As a result, if two treatment factors A and B are already being considered, Replication can be treated as a third factor to be included in a three factor analysis (see the next section). A simpler, if less complete, approach is to ignore Replication interaction effects with the other treatment factors, but still introduce a main effect due to Replication.

Consider two factors (treatments) A and B with A done at a levels and B done at b levels. Within a given treatment combination of A and B levels labeled by i, j , there were c replications done in random order at different times of r repeated measurements of a response variable y. Once again, since the number of measurements for each i, j replication level is the same, the experimental design is balanced. The k'th score in the i, j treatment combination and g'th replication is designated as

$$y_{i,j,g,k}$$

i is the the treatment factor index for factor A ; $1 \le i \le a$
j is the the treatment factor index for factor B ; $1 \le j \le b$
g is the replication index ; $1 \le g \le c$
k labels the score the within the i, j, g level ; $1 \le k \le r$.

The null hypothesis asserts that no treatment population differs from any other. Thus, the scores in the abc treatment levels/replications should all have the same mean and variance. Again we will assume the scores are normally distributed with a common variance within each treatment/ replication level.

Null Hypothesis: $\mathrm{H}_0: \ \mu_{i,j,g} = \mu_0$ for $1 \le i \le a, \ 1 \le j \le b, \ 1 \le g \le c$, where μ_0 is the common population mean of the abc treatment populations. All of the variation seen in the measured scores is due to randomness.

The total number of scores is $n = abcr$. The "grand mean" is given by

$$\bar{y} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r} y_{i,j,g,k}}{abcr} . \qquad \textbf{(24)}$$

The sample mean of the A B treatment combination i, j and replication g is given by

$$\bar{y}_{\mathrm{AB}\ i,j,g} = \frac{\sum_{k=1}^{r} y_{i,j,g,k}}{r} . \ \textbf{(25)}$$

The sample mean of the A B treatment combination i, j over all replications is given by

$$\bar{y}_{\mathrm{AB}\ i,j} = \frac{\sum_{g=1}^{c}\sum_{k=1}^{r} y_{i,j,g,k}}{cr} . \qquad \textbf{(26)}$$

The sample mean of treatment A at level *i* is

$$\bar{y}_{\mathrm{A}\ i} = \frac{\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r} y_{i,j,g,k}}{bcr}. \qquad \textbf{(27)}$$

Similarly, the sample mean of the B treatment at level *j* is

$$\bar{y}_{\mathrm{B}\ j} = \frac{\sum_{i=1}^{a}\sum_{g=1}^{c}\sum_{k=1}^{r} y_{i,j,g,k}}{acr}. \qquad \textbf{(28)}$$

Once again these definitions imply deviation constraints.

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\ i} - \bar{y}\right) = 0 \qquad \textbf{(29)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\ j} - \bar{y}\right) = 0 \qquad \textbf{(30)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{A}\ i}\right) = 0 \qquad \textbf{(31)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{B}\ j}\right) = 0 \qquad \textbf{(32)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{A}\ i} - \bar{y}_{\mathrm{B}\ j} + \bar{y}\right) = 0 \qquad \textbf{(33)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{A}\ i} - \bar{y}_{\mathrm{B}\ j} + \bar{y}\right) = 0 \qquad \textbf{(34)}$$

$$\sum_{g=1}^{c}\left(\bar{y}_{\mathrm{AB}\ i,j,g} - \bar{y}_{\mathrm{AB}\ i,j}\right) = 0 \qquad \textbf{(35)}$$

$$\sum_{k=1}^{r}\left(y_{i,j,g,k} - \bar{y}_{\mathrm{AB}\ i,j,g}\right) = 0 \qquad \textbf{(36)}$$

The effects of the Treatments/Replications are listed as follows:

The Effect due to A at level *i*: $\bar{y}_{\mathrm{A}\ i} - \bar{y}$

The Effect due to B at level *j*: $\bar{y}_{\mathrm{B}\ j} - \bar{y}$

The Interaction Effect of A at level *i* and B at level *j*:

$$\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}\right) - \left(\bar{y}_{\mathrm{A}\ i} - \bar{y}\right) - \left(\bar{y}_{\mathrm{B}\ j} - \bar{y}\right) = \bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{A}\ i} - \bar{y}_{\mathrm{B}\ j} + \bar{y}$$

The Replication Effect: $\bar{y}_{\mathrm{AB}\ i,j,g} - \bar{y}_{\mathrm{AB}\ i,j}$

The Residual or Error Effect: $y_{i,j,g,k} - \bar{y}_{\mathrm{AB}\ i,j,g}$

The degrees of freedom associated with Effects A, B, and AB are the same as in the last section. From equation (35), for each of the *ab* combinations of A and B factors, there are $c-1$ independent deviations, $\bar{y}_{\text{AB}\ i,j,g} - \bar{y}_{\text{AB}\ i,j}$. From equation (36), for each of the *abc* A and B Factor/Replication combinations there are $r-1$ independent deviations from the sample mean $\bar{y}_{\text{AB}\ i,j,g}$. The following table summarizes the degrees of freedom, ν, associated with each effect.

Effect	Degrees of Freedom
Grand Mean	$\nu_{\bar{y}} = 1$
Factor A	$\nu_{\text{A}} = a-1$
Factor B	$\nu_{\text{B}} = b-1$
Interaction of Factors AB	$\nu_{\text{AB}} = (a-1)(b-1)$
Replication G	$\nu_{G} = ab(c-1)$
Residual or Error	$\nu_{E} = abc(r-1)$

Note that the sum of the degrees of freedom is still the total number of scores $n = abcr$.

Formally, we can write the following.

$$y_{i,j,g,k} = \bar{y} + (\bar{y}_{\text{A}\ i} - \bar{y}) + (\bar{y}_{\text{B}\ j} - \bar{y}) + (\bar{y}_{\text{AB}\ i,j} - \bar{y}_{\text{A}\ i} - \bar{y}_{\text{B}\ j} + \bar{y}) + (\bar{y}_{\text{AB}\ i,j,g} - \bar{y}_{\text{AB}\ i,j}) + (y_{i,j,g,k} - \bar{y}_{\text{AB}\ i,j,g}) \quad \textbf{(37)}$$

If the null hypothesis is true and all of the treatment combinations give the same mean result, then the effects just listed all measure the same random departure from the grand mean.

The constraint equations (29) through (36) verify that equation (37) describes an orthogonal vector decomposition of the response variable *y*. Each of the *abcr* measurements can be thought of as a component of an *abcr* dimensional column vector. The column vector which gives the total deviation from the grand mean, $y_{i,j,g,k} - \bar{y}$, can be decomposed as the sum of the five mutually orthogonal column vectors, $V_{\text{A}} + V_{\text{B}} + V_{\text{AB}} + V_{\text{G}} + V_{E}$.

$$V_{\text{A}} = \begin{pmatrix} \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,i} - \bar{y} \\ \vdots \\ \bar{y}_{\text{A}\,a} - \bar{y} \end{pmatrix}; \quad V_{\text{B}} = \begin{pmatrix} \bar{y}_{\text{B}\,1} - \bar{y} \\ \bar{y}_{\text{B}\,1} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,1} - \bar{y} \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,2} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,j} - \bar{y} \\ \vdots \\ \bar{y}_{\text{B}\,b} - \bar{y} \end{pmatrix}; \quad V_{\text{AB}} = \begin{pmatrix} \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,1,1} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,1} + \bar{y} \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,1,2} - \bar{y}_{\text{A}\,1} - \bar{y}_{\text{B}\,2} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j} - \bar{y}_{\text{A}\,i} - \bar{y}_{\text{B}\,j} + \bar{y} \\ \vdots \\ \bar{y}_{\text{AB}\,a,b} - \bar{y}_{\text{A}\,a} - \bar{y}_{\text{B}\,b} + \bar{y} \end{pmatrix}$$

$$V_{\text{G}} = \begin{pmatrix} \bar{y}_{\text{AB}\,1,1,1} - \bar{y}_{\text{AB}\,1,1} \\ \bar{y}_{\text{AB}\,1,1,1} - \bar{y}_{\text{AB}\,1,1} \\ \vdots \\ \bar{y}_{\text{AB}\,1,1,1} - \bar{y}_{\text{AB}\,1,1} \\ \bar{y}_{\text{AB}\,1,1,2} - \bar{y}_{\text{AB}\,1,1} \\ \bar{y}_{\text{AB}\,1,1,2} - \bar{y}_{\text{AB}\,1,1} \\ \vdots \\ \bar{y}_{\text{AB}\,1,1,2} - \bar{y}_{\text{AB}\,1,1} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j,g} - \bar{y}_{\text{AB}\,i,j} \\ \vdots \\ \bar{y}_{\text{AB}\,i,j,g} - \bar{y}_{\text{AB}\,i,j} \\ \vdots \\ \bar{y}_{\text{AB}\,a,b,c} - \bar{y}_{\text{AB}\,a,b} \end{pmatrix}; \quad V_E = \begin{pmatrix} y_{1,1,1,1} - \bar{y}_{\text{AB}\,1,1,1} \\ y_{1,1,1,2} - \bar{y}_{\text{AB}\,1,1,1} \\ \vdots \\ y_{1,1,1,r} - \bar{y}_{\text{AB}\,1,1,1} \\ y_{1,1,2,1} - \bar{y}_{\text{AB}\,1,1,2} \\ y_{1,1,2,2} - \bar{y}_{\text{AB}\,1,1,2} \\ \vdots \\ y_{1,1,2,r} - \bar{y}_{\text{AB}\,1,1,2} \\ \vdots \\ y_{i,j,g,1} - \bar{y}_{\text{AB}\,i,j,g} \\ \vdots \\ y_{i,j,g,r} - \bar{y}_{\text{AB}\,i,j,g} \\ \vdots \\ y_{a,b,c,r} - \bar{y}_{\text{AB}\,a,b,c} \end{pmatrix}$$

Because of the orthogonality condition on the five V vectors, the total variation in y can be decomposed into contributions associated with effects A, B, AB, G(Replication) and Error.

$$\begin{aligned} SS_y &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k} - \bar{y}\right)^2 = \left(V_{\text{A}} + V_{\text{B}} + V_{\text{AB}} + V_{\text{G}} + V_E\right)\bullet\left(V_{\text{A}} + V_{\text{B}} + V_{\text{AB}} + V_{\text{G}} + V_E\right) \\ &= SS_{\text{A}} + SS_{\text{B}} + SS_{\text{AB}} + SS_{\text{G}} + SSE \end{aligned} \tag{38}$$

$$SS_{\text{A}} = V_{\text{A}}\bullet V_{\text{A}} = bcr\sum_{i=1}^{a}\left(\bar{y}_{\text{A}\,i} - \bar{y}\right)^2 \tag{39}$$

$$SS_{\mathrm{B}} = V_{\mathrm{B}} \bullet V_{\mathrm{B}} = acr\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\,j} - \bar{y}\right)^2 \tag{40}$$

$$SS_{\mathrm{AB}} = V_{\mathrm{AB}} \bullet V_{\mathrm{AB}} = cr\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y}\right)^2 \tag{41}$$

$$SS_{\mathrm{G}} = V_{\mathrm{G}} \bullet V_{\mathrm{G}} = r\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(\bar{y}_{\mathrm{AB}\,i,j,g} - \bar{y}_{\mathrm{AB}\,i,j}\right)^2 \tag{42}$$

$$SSE = V_E \bullet V_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k} - \bar{y}_{\mathrm{AB}\,i,j,g}\right)^2 \tag{43}$$

If the null hypothesis is true then the mean square of the Effects A, B, AB and G all estimate the same common population variance of each A B factor combination. The mean squares are computed as a variation divided by the associated degrees of freedom.

The Mean Square due to Factor A:

$$MS_{\mathrm{A}} = \frac{SS_{\mathrm{A}}}{a-1} = \frac{bcr\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\,i} - \bar{y}\right)^2}{a-1} \tag{44}$$

The Mean Square due to Factor B:

$$MS_{\mathrm{B}} = \frac{SS_{\mathrm{B}}}{b-1} = \frac{acr\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\,j} - \bar{y}\right)^2}{b-1} \tag{45}$$

The Mean Square due to Interaction of A with B:

$$MS_{\mathrm{AB}} = \frac{SS_{\mathrm{AB}}}{(a-1)(b-1)} = \frac{cr\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y}\right)^2}{(a-1)(b-1)} \tag{46}$$

The Mean Square due to Replication:

$$MS_{\mathrm{G}} = \frac{SS_{\mathrm{G}}}{ab(c-1)} = \frac{r\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(\bar{y}_{\mathrm{AB}\,i,j,g} - \bar{y}_{\mathrm{AB}\,i,j}\right)^2}{ab(c-1)} \tag{47}$$

The Error or Residual Mean Square:

$$MSE = \frac{SSE}{abc(r-1)} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k} - \bar{y}_{\mathrm{AB}\,i,j,g}\right)^2}{abc(r-1)} \tag{48}$$

To test the null hypothesis that all effects measure the same common random variation associated with the Mean Square due to Error, we compute the observed Fisher F score.

$$F_{\text{obs}} = \frac{\text{Effect Mean Square}}{MSE} \qquad \textbf{(49)}$$

For a given level of significance, α, this value is compared against a critical score calculated from an F distribution used to compare two variances obtained from sampling variances from two normally distributed populations. The numerator degrees of freedom is appropriate to the effect in question and the denominator degrees of freedom is $abc(r-1)$. This information is summarized in a Two-Factor ANOVA table.

Source	Sum of Squares	Degrees of Freedom	Mean Square	$\mathbf{F_{obs}}$
Factor A	SS_{A}	$a-1$	$MS_{\text{A}} = \frac{SS_{\text{A}}}{a-1}$	$\frac{MS_{\text{A}}}{MSE}$
Factor B	SS_{B}	$b-1$	$MS_{\text{B}} = \frac{SS_{\text{B}}}{b-1}$	$\frac{MS_{\text{B}}}{MSE}$
Interaction AB	SS_{AB}	$(a-1)(b-1)$	$MS_{\text{AB}} = \frac{SS_{\text{AB}}}{(a-1)(b-1)}$	$\frac{MS_{\text{AB}}}{MSE}$
Replication	SS_{G}	$ab(c-1)$	$MS_{\text{G}} = \frac{SS_{\text{G}}}{ab(c-1)}$	$\frac{MS_{\text{G}}}{MSE}$
Error	SSE	$abc(r-1)$	$MSE = \frac{SSE}{abc(r-1)}$	
Total	SS_y	$n-1 = abcr-1$	$s_y^2 = \frac{SS_y}{n-1}$	

Each computed $\mathbf{F_{obs}}$ is compared against $F_\alpha(\nu_{\text{effect}}, \nu_E)$ for a stated level of significance, α. If $F_{\text{obs}} < F_\alpha(\nu_{\text{effect}}, \nu_E)$, we fail to reject the null hypothesis that the effect is just an alias for random error (variation). If $F_{\text{obs}} > F_\alpha(\nu_{\text{effect}}, \nu_E)$, we reject H_0 and conclude the effect in question is real. **Note**: It is possible that either the AB Interaction Effect or the Replication Effect can test as significant even when one or both of the separate A Factor and B Factor Effects are insignificant.

To facilitate the actual calculations, we define the following intermediate variables.

$$T_{i,j,g} = \sum_{k=1}^{r} y_{i,j,g,k} \qquad \textbf{(50)} \qquad\qquad Q_{i,j,g} = \sum_{k=1}^{r} \left(y_{i,j,g,k}\right)^2 \qquad \textbf{(51)}$$

$$T_{i,j} = \sum_{g=1}^{c} T_{i,j,g} \qquad \textbf{(52)} \qquad\qquad Q_{i,j} = \sum_{g=1}^{c} Q_{i,j,g} \qquad \textbf{(53)}$$

$$C_{\text{A}\,i} = \sum_{j=1}^{b}\sum_{g=1}^{c} \left(T_{i,j,g}\right)^2 \qquad \textbf{(54)} \qquad\qquad T_{\text{A}\,i} = \sum_{j=1}^{b} T_{i,j} \qquad \textbf{(55)}$$

$$T_{\mathrm{B}\,j}=\sum_{i=1}^{a}T_{i,j} \quad \textbf{(56)} \qquad T=\sum_{i=1}^{a}\sum_{j=1}^{b}T_{i,j} \quad \textbf{(57)}$$

$$W_{\mathrm{A}\,i}=\sum_{j=1}^{b}\left(T_{i,j}\right)^2 \quad \textbf{(58)} \qquad Q_{\mathrm{A}\,i}=\sum_{j=1}^{b}Q_{i,j} \quad \textbf{(59)}$$

$$\bar{y}_{i,j,g}=\frac{T_{i,j,g}}{r} \quad \textbf{(60)} \quad \bar{y}_{\mathrm{AB}\,i,j}=\frac{T_{i,j}}{cr} \quad \textbf{(61)} \quad \bar{y}_{\mathrm{A}\,i}=\frac{T_{\mathrm{A}\,i}}{bcr} \quad \textbf{(62)} \quad \bar{y}_{\mathrm{B}\,j}=\frac{T_{\mathrm{B}\,j}}{acr} \quad \textbf{(63)} \quad \bar{y}=\frac{T}{n} \quad \textbf{(64)}$$

Now,

$$SS_{\mathrm{A}}=bcr\sum_{i=1}^{a}\left(\frac{T_{\mathrm{A}\,i}}{bcr}-\frac{T}{abcr}\right)^2=\frac{1}{bcr}\sum_{i=1}^{a}\left(\left(T_{\mathrm{A}\,i}\right)^2-2\frac{T}{a}T_{\mathrm{A}\,i}+\frac{T^2}{a^2}\right)=\frac{1}{bcr}\left(\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2-\frac{2T^2}{a}+a\frac{T^2}{a^2}\right)$$

$$=\frac{1}{bcr}\left(\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2-\frac{T^2}{a}\right)$$

$$SS_{\mathrm{B}}=acr\sum_{j=1}^{b}\left(\frac{T_{\mathrm{B}\,j}}{acr}-\frac{T}{abcr}\right)^2=\frac{1}{acr}\sum_{j=1}^{b}\left(\left(T_{\mathrm{B}\,j}\right)^2-2\frac{T}{b}T_{\mathrm{B}\,j}+\frac{T^2}{b^2}\right)=\frac{1}{acr}\left(\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2-\frac{2T^2}{b}+b\frac{T^2}{b^2}\right)$$

$$=\frac{1}{acr}\left(\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2-\frac{T^2}{b}\right)$$

$$SS_{\mathrm{AB}}=cr\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\frac{T_{i,j}}{cr}-\frac{T_{\mathrm{A}\,i}}{bcr}-\frac{T_{\mathrm{B}\,j}}{acr}+\frac{T}{abcr}\right)^2$$

$$=\frac{1}{cr}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\left(T_{i,j}-\frac{T_{\mathrm{A}\,i}}{b}\right)^2-2\left(T_{i,j}-\frac{T_{\mathrm{A}\,i}}{b}\right)\left(\frac{T_{\mathrm{B}\,j}}{a}-\frac{T}{ab}\right)+\left(\frac{T_{\mathrm{B}\,j}}{a}-\frac{T}{ab}\right)^2\right)$$

$$=\frac{1}{cr}\left[\sum_{i=1}^{a}\sum_{j=1}^{b}\left(T_{i,j}-\frac{T_{\mathrm{A}\,i}}{b}\right)^2-\frac{2}{a}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}-\frac{T}{b}\right)\sum_{i=1}^{a}\left(T_{i,j}-\frac{T_{\mathrm{A}\,i}}{b}\right)+a\sum_{j=1}^{b}\left(\frac{T_{\mathrm{B}\,j}}{a}-\frac{T}{ab}\right)^2\right]$$

$$=\frac{1}{cr}\left[\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\left(T_{i,j}\right)^2-\frac{2T_{\mathrm{A}\,i}}{b}T_{i,j}+\frac{\left(T_{\mathrm{A}\,i}\right)^2}{b^2}\right)-\frac{2}{a}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}-\frac{T}{b}\right)^2+\frac{1}{a}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}-\frac{T}{b}\right)^2\right]$$

$$=\frac{1}{cr}\left[\sum_{i=1}^{a}\sum_{j=1}^{b}\left(T_{i,j}\right)^2-\frac{1}{b}\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2-\frac{1}{a}\sum_{j=1}^{b}\left(\left(T_{\mathrm{B}\,j}\right)^2-\frac{2T}{b}T_{\mathrm{B}\,j}+\frac{T^2}{b^2}\right)\right]$$

$$=\frac{1}{cr}\left[\sum_{i=1}^{a}\sum_{j=1}^{b}\left(T_{i,j}\right)^2-\frac{1}{b}\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2-\frac{1}{a}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2+\frac{T^2}{ab}\right]$$

$$=\frac{1}{cr}\left[\sum_{i=1}^{a}W_{\mathrm{A}\,i}-\frac{1}{b}\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2-\frac{1}{a}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2+\frac{T^2}{ab}\right]$$

$$SS_{\mathrm{G}} = r\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(\frac{T_{i,j,g}}{r}-\frac{T_{i,j}}{rc}\right)^2 = \frac{1}{r}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{g=1}^{c}\left(T_{i,j,g}\right)^2 - \frac{2T_{i,j}}{c}\sum_{g=1}^{c}T_{i,j,g} + \frac{\left(T_{i,j}\right)^2}{c}\right)$$

$$= \frac{1}{r}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{g=1}^{c}\left(T_{i,j,g}\right)^2 - \frac{\left(T_{i,j}\right)^2}{c}\right) = \frac{1}{r}\sum_{i=1}^{a}\left(C_{\mathrm{A}\,i} - \frac{W_{\mathrm{A}\,i}}{c}\right)$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k} - \frac{T_{i,j,g}}{r}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(\sum_{k=1}^{r}\left(y_{i,j,g,k}\right)^2 - \frac{2T_{i,j,g}}{r}\sum_{k=1}^{r}y_{i,j,g,k} + \frac{\left(T_{i,j,g}\right)^2}{r}\right)$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(\sum_{k=1}^{r}\left(y_{i,j,g,k}\right)^2 - \frac{\left(T_{i,j,g}\right)^2}{r}\right) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\left(Q_{i,j,g} - \frac{\left(T_{i,j,g}\right)^2}{r}\right) = \sum_{i=1}^{a}\left(Q_{\mathrm{A}\,i} - \frac{C_{\mathrm{A}\,i}}{r}\right)$$

Finally, as a check, the Total Variation must be the sum of the Effect Variations.

$$SS_y = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k} - \frac{T}{abcr}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{g=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,g,k}\right)^2 - \frac{T^2}{abcr} = \sum_{i=1}^{a}Q_{\mathrm{A}\,i} - \frac{T^2}{abcr}$$

$$SS_{\mathrm{A}} + SS_{\mathrm{B}} + SS_{\mathrm{AB}} + SS_{\mathrm{G}} + SSE$$
$$= \left(\frac{1}{bcr}\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2 - \frac{T^2}{abcr}\right) + \left(\frac{1}{acr}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2 - \frac{T^2}{abcr}\right) + \left(\frac{1}{cr}\sum_{i=1}^{a}W_{\mathrm{A}\,i} - \frac{1}{bcr}\sum_{i=1}^{a}\left(T_{\mathrm{A}\,i}\right)^2 - \frac{1}{acr}\sum_{j=1}^{b}\left(T_{\mathrm{B}\,j}\right)^2 + \frac{T^2}{abcr}\right)$$
$$+ \left(\frac{1}{r}\sum_{i=1}^{a}C_{\mathrm{A}\,i} - \frac{1}{cr}\sum_{i=1}^{a}W_{\mathrm{A}\,i}\right) + \left(\sum_{i=1}^{a}Q_{\mathrm{A}\,i} - \frac{1}{r}\sum_{i=1}^{a}C_{\mathrm{A}\,i}\right) = \sum_{i=1}^{a}Q_{\mathrm{A}\,i} - \frac{T^2}{abcr}$$

So a scheme to calculate a Two-Factor with Replication analysis of variance is to lay out the data in a spreadsheet, with adjacent rows in a given column representing the different A B Factor and Replication combinations (i, j, g) . For each such grouping of scores, calculate the sum of scores, $T_{i,j,g} = \sum_{k=1}^{r} y_{i,j,g,k}$, and the sum of squares of scores, $Q_{i,j,g} = \sum_{k=1}^{r}\left(y_{i,j,g,k}\right)^2$. Then calculate the sample mean, $\bar{y}_{\mathrm{AB}\,i,j,g} = \frac{T_{i,j,g}}{r}$ and the sample variance, $s^2_{\mathrm{AB}\,i,j,g} = \frac{Q_{i,j,g} - \left(T_{i,j,g}\right)^2 / r}{r-1}$. The latter is calculated for inspection purposes. Based on the values of the sample variances is it reasonable to assume that all of the treatment and replication combinations have the same population variance? Next calculate the sums given by equations (52) through (59). From these results construct the Two-Factor ANOVA table and test whether any of the Effects are significant.

Source	Sum of Squares	Mean Square	F_{obs}
A	$SS_A = \frac{1}{bcr}\left(\sum_{i=1}^{a}(T_{A\,i})^2 - \frac{T^2}{a}\right)$	$MS_A = \frac{SS_A}{a-1}$	$\frac{MS_A}{MSE}$
B	$SS_B = \frac{1}{acr}\left(\sum_{j=1}^{b}(T_{B\,j})^2 - \frac{T^2}{b}\right)$	$MS_B = \frac{SS_B}{b-1}$	$\frac{MS_B}{MSE}$
AB	$SS_{AB} = \frac{1}{cr}\left[\sum_{i=1}^{a}W_{A\,i} - \frac{1}{b}\sum_{i=1}^{a}(T_{A\,i})^2 - \frac{1}{a}\sum_{j=1}^{b}(T_{B\,j})^2 + \frac{T^2}{ab}\right]$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MSE}$
G	$SS_G = \frac{1}{r}\sum_{i=1}^{a}\left(C_{A\,i} - \frac{W_{A\,i}}{c}\right)$	$MS_G = \frac{SS_G}{ab(c-1)}$	$\frac{MS_G}{MSE}$
Error	$SSE = \sum_{i=1}^{a}\left(Q_{A\,i} - \frac{C_{A\,i}}{r}\right)$	$MSE = \frac{SSE}{abc(r-1)}$	
Total	$SS_y = SS_A + SS_B + SS_{AB} + SS_G + SSE$ $= \sum_{i=1}^{a}Q_{A\,i} - \frac{T^2}{abcr}$	$s_y^2 = \frac{SS_y}{abcr-1}$	

Three Factor Analysis

Now consider three factors (treatments) A, B and C with A done at *a* levels and B done at *b* levels, C done at *c* levels. Within a given treatment combination of A, B and C levels labeled by *i*, *j* , *m* there were *r* repeated measurements of a response variable *y*. Once again, since the number of measurements for each *i*, *j*, *m* level is the same, the experimental design is balanced. The *k*'th score in the *i*, *j*, *m* treatment combination is designated as

$$y_{i,j,m,k}$$

i is the treatment factor index for factor A ; $1 \le i \le a$
j is the treatment factor index for factor B ; $1 \le j \le b$
m is the treatment factor index for factor C; $1 \le m \le c$
k labels the score the within the *i*, *j*, *m* level ; $1 \le k \le r$.

The null hypothesis asserts that no treatment population differs from any other. Thus, the scores in the *abc* treatment levels should all have the same mean and variance. Again we will assume the scores are normally distributed with a common variance within each *i*, *j*, *m* treatment combination.

Null Hypothesis: $H_0: \ \mu_{i,j,m} = \mu_0$ for $1 \le i \le a, \ 1 \le j \le b, \ 1 \le m \le c$, where μ_0 is the common population mean of the *abc* treatment populations. All of the variation seen in the measured scores is due to randomness.

The total number of scores is $n = abcr$. The "grand mean" is given by

$$\bar{y} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\sum_{k=1}^{r} y_{i,j,m,k}}{abcr}. \qquad \textbf{(65)}$$

The sample mean of the A B C treatment combination *i*, *j*, *m* is given by

$$\bar{y}_{\mathrm{ABC}\ i,j,m} = \frac{\sum_{k=1}^{r} y_{i,j,m,k}}{r}. \qquad \textbf{(66)}$$

The three sample means of the two-way combinations A and B at levels *i* and *j*, A and C at levels *i* and *m*, and B and C at levels *j* and *m* are given by the following.

$$\bar{y}_{\mathrm{AB}\ i,j} = \frac{\sum_{m=1}^{c}\sum_{k=1}^{r} y_{i,j,m,k}}{cr} \qquad \textbf{(67)} \qquad \bar{y}_{\mathrm{AC}\ i,m} = \frac{\sum_{j=1}^{b}\sum_{k=1}^{r} y_{i,j,m,k}}{br} \qquad \textbf{(68)} \qquad \bar{y}_{\mathrm{BC}\ j,m} = \frac{\sum_{i=1}^{a}\sum_{k=1}^{r} y_{i,j,m,k}}{ar} \qquad \textbf{(69)}$$

Similar calculations give the sample means of treatment A at level *i*, treatment B at level *j*, and treatment C at level *m*.

$$\bar{y}_{\mathrm{A}\ i} = \frac{\sum_{j=1}^{b}\sum_{m=1}^{c}\sum_{k=1}^{r} y_{i,j,m,k}}{bcr} \qquad \textbf{(70)} \qquad \bar{y}_{\mathrm{B}\ j} = \frac{\sum_{i=1}^{a}\sum_{m=1}^{c}\sum_{k=1}^{r} y_{i,j,m,k}}{acr} \qquad \textbf{(71)} \qquad \bar{y}_{\mathrm{C}\ m} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{i,j,m,k}}{abr} \qquad \textbf{(72)}$$

Again these definitions imply constraints for the sums of deviations.

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{A}\ i} - \bar{y}\right) = 0 \qquad \textbf{(73)} \qquad \sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\ j} - \bar{y}\right) = 0 \qquad \textbf{(74)}$$

$$\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{C}\ m} - \bar{y}\right) = 0 \qquad \textbf{(75)} \qquad \sum_{j=1}^{b}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{A}\ i}\right) = 0 \qquad \textbf{(76)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AB}\ i,j} - \bar{y}_{\mathrm{B}\ j}\right) = 0 \qquad \textbf{(77)} \qquad \sum_{m=1}^{c}\left(\bar{y}_{\mathrm{AC}\ i,m} - \bar{y}_{\mathrm{A}\ i}\right) = 0 \qquad \textbf{(78)}$$

$$\sum_{i=1}^{a}\left(\bar{y}_{\mathrm{AC}\ i,m} - \bar{y}_{\mathrm{C}\ m}\right) = 0 \qquad \textbf{(79)} \qquad \sum_{m=1}^{c}\left(\bar{y}_{\mathrm{BC}\ j,m} - \bar{y}_{\mathrm{B}\ j}\right) = 0 \qquad \textbf{(80)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{BC}\ j,m} - \bar{y}_{\mathrm{C}\ m}\right) = 0 \qquad \textbf{(81)} \qquad \sum_{i=1}^{a}\left(\bar{y}_{\mathrm{ABC}\ i,j,m} - \bar{y}_{\mathrm{BC}\ j,m}\right) = 0 \qquad \textbf{(82)}$$

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{ABC}\ i,j,m} - \bar{y}_{\mathrm{AC}\ i,m}\right) = 0 \qquad \textbf{(83)} \qquad \sum_{m=1}^{c}\left(\bar{y}_{\mathrm{ABC}\ i,j,m} - \bar{y}_{\mathrm{AB}\ i,j}\right) = 0 \qquad \textbf{(84)}$$

$$\sum_{j=1}^{b}\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{ABC}\ i,j,m} - \bar{y}_{\mathrm{A}\ i}\right) = 0 \qquad \textbf{(85)} \qquad \sum_{i=1}^{a}\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{ABC}\ i,j,m} - \bar{y}_{\mathrm{B}\ j}\right) = 0 \qquad \textbf{(86)}$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{ABC}\ i,j,m}-\bar{y}_{\mathrm{C}\ m}\right)=0 \qquad \textbf{(87)} \qquad\qquad \sum_{k=1}^{r}\left(\bar{y}_{i,j,m,k}-\bar{y}_{\mathrm{ABC}\ i,j,m}\right)=0 \qquad \textbf{(88)}$$

The effects in the Three Factor analysis of are listed as follows:

The Effect due to Factor A at level *i*: $\bar{y}_{\mathrm{A}\ i}-\bar{y}$
The Effect due to Factor B at level *j*: $\bar{y}_{\mathrm{B}\ j}-\bar{y}$
The Effect due to Factor C at level *m*: $\bar{y}_{\mathrm{C}\ m}-\bar{y}$

The Two-Factor Interaction Effect of A at level *i* and B at level *j*:
$$\left(\bar{y}_{\mathrm{AB}\ i,j}-\bar{y}\right)-\left(\bar{y}_{\mathrm{A}\ i}-\bar{y}\right)-\left(\bar{y}_{\mathrm{B}\ j}-\bar{y}\right)=\bar{y}_{\mathrm{AB}\ i,j}-\bar{y}_{\mathrm{A}\ i}-\bar{y}_{\mathrm{B}\ j}+\bar{y}$$

The Two-Factor Interaction Effect of A at level *i* and C at level *m*:
$$\left(\bar{y}_{\mathrm{AC}\ i,m}-\bar{y}\right)-\left(\bar{y}_{\mathrm{A}\ i}-\bar{y}\right)-\left(\bar{y}_{\mathrm{C}\ m}-\bar{y}\right)=\bar{y}_{\mathrm{AC}\ i,m}-\bar{y}_{\mathrm{A}\ i}-\bar{y}_{\mathrm{C}\ m}+\bar{y}$$

The Two-Factor Interaction Effect of B at level *j* and C at level *m*:
$$\left(\bar{y}_{\mathrm{BC}\ j,m}-\bar{y}\right)-\left(\bar{y}_{\mathrm{B}\ j}-\bar{y}\right)-\left(\bar{y}_{\mathrm{C}\ m}-\bar{y}\right)=\bar{y}_{\mathrm{BC}\ j,m}-\bar{y}_{\mathrm{B}\ j}-\bar{y}_{\mathrm{C}\ m}+\bar{y}$$

The Three-Factor Interaction Effect of A at level *i*, B at level *j*, and C at level *m*:
$$\left(\bar{y}_{\mathrm{ABC}\ i,j,m}-\bar{y}\right)-\left(\bar{y}_{\mathrm{AB}\ i,j}-\bar{y}_{\mathrm{A}\ i}-\bar{y}_{\mathrm{B}\ j}+\bar{y}\right)-\left(\bar{y}_{\mathrm{AC}\ i,m}-\bar{y}_{\mathrm{A}\ i}-\bar{y}_{\mathrm{C}\ m}+\bar{y}\right)-\left(\bar{y}_{\mathrm{BC}\ j,m}-\bar{y}_{\mathrm{B}\ j}-\bar{y}_{\mathrm{C}\ m}+\bar{y}\right)$$
$$-\left(\bar{y}_{\mathrm{A}\ i}-\bar{y}\right)-\left(\bar{y}_{\mathrm{B}\ j}-\bar{y}\right)-\left(\bar{y}_{\mathrm{C}\ m}-\bar{y}\right)=\bar{y}_{\mathrm{ABC}\ i,j,m}-\bar{y}_{\mathrm{AB}\ i,j}-\bar{y}_{\mathrm{AC}\ i,m}-\bar{y}_{\mathrm{BC}\ j,m}+\bar{y}_{\mathrm{A}\ i}+\bar{y}_{\mathrm{B}\ j}+\bar{y}_{\mathrm{C}\ m}-\bar{y}$$

The Residual or Error Effect: $y_{i,j,m,k}-\bar{y}_{\mathrm{AB}\ i,j,m}$

Now since the sums of deviations is zero, i.e., from equations (73) through (87), the sum over any occurring index of the Factor effects, the Two-Factor interactions and the Three-Factor interaction vanishes. For example, the sum over *j* of the Three-Factor Interaction is zero by equations (83), (76), (81) and (74).

$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{ABC}\ i,j,m}-\bar{y}_{\mathrm{AB}\ i,j}-\bar{y}_{\mathrm{AC}\ i,m}-\bar{y}_{\mathrm{BC}\ j,m}+\bar{y}_{\mathrm{A}\ i}+\bar{y}_{\mathrm{B}\ j}+\bar{y}_{\mathrm{C}\ m}-\bar{y}\right)=$$
$$\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{ABC}\ i,j,m}-\bar{y}_{\mathrm{AC}\ i,m}\right)+\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{A}\ i}-\bar{y}_{\mathrm{AB}\ i,j}\right)+\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{C}\ m}-\bar{y}_{\mathrm{BC}\ j,m}\right)+\sum_{j=1}^{b}\left(\bar{y}_{\mathrm{B}\ j}-\bar{y}\right)=$$
$$0+0+0+0=0$$

Thus, for the Three-Factor interaction only the first $(b-1)$ of the *j* components are independent. Similar constraints apply to the *i* and *m* components, so the degrees of freedom of the Three-Factor interaction is $(a-1)(b-1)(c-1)$. The degrees of freedom associated with Effects A, B, C and the Two-Factor interactions are what we would expect from the Two Factor Analysis. From equation (88), for each of the *abc* combinations of A, B and C factors, there are $r-1$ independent deviations, $\bar{y}_{i,j,m,k}-\bar{y}_{\mathrm{ABC}\ i,j,m}$.

The following table summarizes the degrees of freedom, ν , associated with each effect.

Effect	Degrees of Freedom
Grand Mean	$\nu_{\bar{y}} = 1$
Factor A	$\nu_{\mathrm{A}} = a-1$
Factor B	$\nu_{\mathrm{B}} = b-1$
Factor C	$\nu_{\mathrm{C}} = c-1$
Two-Factor Interaction of Factors AB	$\nu_{\mathrm{AB}} = (a-1)(b-1)$
Two-Factor Interaction of Factors AC	$\nu_{\mathrm{AC}} = (a-1)(c-1)$
Two-Factor Interaction of Factors BC	$\nu_{\mathrm{BC}} = (b-1)(c-1)$
Three-Factor Interaction of Factors ABC	$\nu_{\mathrm{ABC}} = (a-1)(b-1)(c-1)$
Residual or Error	$\nu_E = abc(r-1)$

Note that the sum of the degrees of freedom is still the total number of scores $n = abcr$.

If the null hypothesis is true and all of the treatment combinations give the same mean result, and the eight effects just listed all measure the same random departure from the grand mean.

Formally, we can write the following.

$$\begin{aligned} y_{i,j,g,k} &= \bar{y} + (\bar{y}_{\mathrm{A}\,i} - \bar{y}) + (\bar{y}_{\mathrm{B}\,j} - \bar{y}) + (\bar{y}_{\mathrm{C}\,m} - \bar{y}) \\ &+ (\bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{B}\,j} + \bar{y}) + (\bar{y}_{\mathrm{AC}\,i,j} - \bar{y}_{\mathrm{A}\,i} - \bar{y}_{\mathrm{C}\,m} + \bar{y}) + (\bar{y}_{\mathrm{BC}\,i,m} - \bar{y}_{\mathrm{B}\,j} - \bar{y}_{\mathrm{C}\,m} + \bar{y}) \\ &+ (\bar{y}_{\mathrm{ABC}\,i,j,m} - \bar{y}_{\mathrm{AB}\,i,j} - \bar{y}_{\mathrm{AC}\,i,m} - \bar{y}_{\mathrm{BC}\,j,m} + \bar{y}_{\mathrm{A}\,i} + \bar{y}_{\mathrm{B}\,j} + \bar{y}_{\mathrm{C}\,m} - \bar{y}) \\ &+ (y_{i,j,m,k} - \bar{y}_{\mathrm{ABC}\,i,j,m}) \end{aligned} \tag{89}$$

The zero sum constraint of each effect insures that equation (89) describes an orthogonal vector decomposition of the response variable y. The column vector which gives the total deviation from the grand mean, $y_{i,j,m,k} - \bar{y}$, can be decomposed as the sum of the eight mutually orthogonal column vectors, $V_{\mathrm{A}} + V_{\mathrm{B}} + V_{\mathrm{C}} + V_{\mathrm{AB}} + V_{\mathrm{AC}} + V_{\mathrm{BC}} + V_{\mathrm{ABC}} + V_E$, with each vector associated with an effect. The orthogonality condition on the eight V vectors means that the total variation in y can be decomposed into contributions associated with the eight effects A, B, C, AB, AC, BC, ABC and Error.

$$\begin{aligned} SS_y &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\sum_{k=1}^{r}(y_{i,j,m,k} - \bar{y})^2 \\ &= (V_{\mathrm{A}} + V_{\mathrm{B}} + V_{\mathrm{C}} + V_{\mathrm{AB}} + V_{\mathrm{AC}} + V_{\mathrm{BC}} + V_{\mathrm{ABC}} + V_E) \bullet (V_{\mathrm{A}} + V_{\mathrm{B}} + V_{\mathrm{C}} + V_{\mathrm{AB}} + V_{\mathrm{AC}} + V_{\mathrm{BC}} + V_{\mathrm{ABC}} + V_E) \\ &= SS_{\mathrm{A}} + SS_{\mathrm{B}} + SS_{\mathrm{C}} + SS_{\mathrm{AB}} + SS_{\mathrm{AC}} + SS_{\mathrm{BC}} + SS_{\mathrm{ABC}} + SSE \end{aligned} \tag{90}$$

$$SS_{\mathrm{A}} = V_{\mathrm{A}} \bullet V_{\mathrm{A}} = bcr\sum_{i=1}^{a}(\bar{y}_{\mathrm{A}\,i} - \bar{y})^2 \tag{91}$$

$$SS_{\mathrm{B}} = V_{\mathrm{B}} \bullet V_{\mathrm{B}} = acr\sum_{j=1}^{b}\left(\overline{y}_{\mathrm{B}\,j} - \overline{y}\right)^2 \quad \textbf{(92)}$$

$$SS_{\mathrm{C}} = V_{\mathrm{C}} \bullet V_{\mathrm{C}} = abr\sum_{m=1}^{c}\left(\overline{y}_{\mathrm{C}\,m} - \overline{y}\right)^2 \quad \textbf{(93)}$$

$$SS_{\mathrm{AB}} = V_{\mathrm{AB}} \bullet V_{\mathrm{AB}} = cr\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{y}_{\mathrm{AB}\,i,j} - \overline{y}_{\mathrm{A}\,i} - \overline{y}_{\mathrm{B}\,j} + \overline{y}\right)^2 \quad \textbf{(94)}$$

$$SS_{\mathrm{AC}} = V_{\mathrm{AC}} \bullet V_{\mathrm{AC}} = br\sum_{i=1}^{a}\sum_{m=1}^{c}\left(\overline{y}_{\mathrm{AC}\,i,m} - \overline{y}_{\mathrm{A}\,i} - \overline{y}_{\mathrm{C}\,m} + \overline{y}\right)^2 \quad \textbf{(95)}$$

$$SS_{\mathrm{BC}} = V_{\mathrm{BC}} \bullet V_{\mathrm{BC}} = ar\sum_{j=1}^{b}\sum_{m=1}^{c}\left(\overline{y}_{\mathrm{BC}\,j,m} - \overline{y}_{\mathrm{B}\,j} - \overline{y}_{\mathrm{C}\,m} + \overline{y}\right)^2 \quad \textbf{(96)}$$

$$SS_{\mathrm{ABC}} = V_{\mathrm{ABC}} \bullet V_{\mathrm{ABC}} = r\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\left(\overline{y}_{\mathrm{ABC}\,i,j,m} - \overline{y}_{\mathrm{AB}\,i,j} - \overline{y}_{\mathrm{AC}\,i,m} - \overline{y}_{\mathrm{BC}\,j,m} + \overline{y}_{\mathrm{A}\,i} + \overline{y}_{\mathrm{B}\,j} + \overline{y}_{\mathrm{C}\,m} - \overline{y}\right)^2 \quad \textbf{(97)}$$

$$SSE = V_E \bullet V_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,m,k} - \overline{y}_{\mathrm{AB}\,i,j,m}\right)^2 \quad \textbf{(98)}$$

If the null hypothesis is true then the mean square of the Effects A, B, C, AB, AC, BC and ABC all estimate the same common population variance of each A B factor combination. The mean squares are computed as a variation divided by the associated degrees of freedom.

The Mean Square due to Factor A: $MS_{\mathrm{A}} = \dfrac{SS_{\mathrm{A}}}{a-1} = \dfrac{bcr\sum_{i=1}^{a}\left(\overline{y}_{\mathrm{A}\,i} - \overline{y}\right)^2}{a-1}$ **(99)**

The Mean Square due to Factor B: $MS_{\mathrm{B}} = \dfrac{SS_{\mathrm{B}}}{b-1} = \dfrac{acr\sum_{j=1}^{b}\left(\overline{y}_{\mathrm{B}\,j} - \overline{y}\right)^2}{b-1}$ **(100)**

The Mean Square due to Factor C: $MS_{\mathrm{C}} = \dfrac{SS_{\mathrm{C}}}{c-1} = \dfrac{abr\sum_{m=1}^{c}\left(\overline{y}_{\mathrm{C}\,m} - \overline{y}\right)^2}{c-1}$ **(101)**

The Mean Square due to theTwo-Factor Interaction of A with B:

$$MS_{\mathrm{AB}} = \frac{SS_{\mathrm{AB}}}{(a-1)(b-1)} = \frac{cr\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{y}_{\mathrm{AB}\,i,j} - \overline{y}_{\mathrm{A}\,i} - \overline{y}_{\mathrm{B}\,j} + \overline{y}\right)^2}{(a-1)(b-1)} \quad \textbf{(102)}$$

The Mean Square due to theTwo-Factor Interaction of A with C:

$$MS_{\mathrm{AC}} = \frac{SS_{\mathrm{AC}}}{(a-1)(c-1)} = \frac{br\sum_{i=1}^{a}\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{AC}\;i,m} - \bar{y}_{\mathrm{A}\;i} - \bar{y}_{\mathrm{C}\;m} + \bar{y}\right)^2}{(a-1)(c-1)} \quad \textbf{(103)}$$

The Mean Square due to theTwo-Factor Interaction of B with C:

$$MS_{\mathrm{BC}} = \frac{SS_{\mathrm{BC}}}{(b-1)(c-1)} = \frac{ar\sum_{j=1}^{b}\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{BC}\;j,m} - \bar{y}_{\mathrm{B}\;j} - \bar{y}_{\mathrm{C}\;m} + \bar{y}\right)^2}{(b-1)(c-1)} \quad \textbf{(104)}$$

The Mean Square due to theThree-Factor Interaction of A,B and C:

$$\begin{aligned} MS_{\mathrm{ABC}} &= \frac{SS_{\mathrm{ABC}}}{(a-1)(b-1)(c-1)} \\ &= \frac{r\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\left(\bar{y}_{\mathrm{ABC}\;i,j,m} - \bar{y}_{\mathrm{AB}\;i,j} - \bar{y}_{\mathrm{AC}\;i,m} - \bar{y}_{\mathrm{BC}\;j,m} + \bar{y}_{\mathrm{A}\;i} + \bar{y}_{\mathrm{B}\;j} + \bar{y}_{\mathrm{C}\;m} - \bar{y}\right)^2}{(a-1)(b-1)(c-1)} \end{aligned} \quad \textbf{(105)}$$

The Error or Residual Mean Square:

$$MSE = \frac{SSE}{abc(r-1)} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{m=1}^{c}\sum_{k=1}^{r}\left(y_{i,j,m,k} - \bar{y}_{\mathrm{ABC}\;i,j,m}\right)^2}{abc(r-1)} \quad \textbf{(106)}$$

To test the null hypothesis that all effects measure the same common random variation associated with the Mean Square due to Error, we compute the observed Fisher F score.

$$F_{\mathrm{obs}} = \frac{\text{Effect Mean Square}}{MSE} \quad \textbf{(107)}$$

For a given level of significance, α, this value is compared against a critical score calculated from an F distribution used to compare two variances obtained from sampling variances from two normally distributed populations. The numerator degrees of freedom is appropriate to the effect in question and the denominator degrees of freedom is $abc(r-1)$. This information is summarized in a Three-Factor ANOVA table.

Source	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	$\mathbf{F_{obs}}$
Factor A	SS_{A}	$a-1$	$MS_{\mathrm{A}} = \frac{SS_{\mathrm{A}}}{a-1}$	$\frac{MS_{\mathrm{A}}}{MSE}$
Factor B	SS_{B}	$b-1$	$MS_{\mathrm{B}} = \frac{SS_{\mathrm{B}}}{b-1}$	$\frac{MS_{\mathrm{B}}}{MSE}$
Factor C	SS_{C}	$c-1$	$MS_{\mathrm{C}} = \frac{SS_{\mathrm{C}}}{c-1}$	$\frac{MS_{\mathrm{C}}}{MSE}$

Interaction AB	SS_{AB}	$(a-1)(b-1)$	$MS_{\text{AB}} = \dfrac{SS_{\text{AB}}}{(a-1)(b-1)}$	$\dfrac{MS_{\text{AB}}}{MSE}$
Interaction AC	SS_{AC}	$(a-1)(c-1)$	$MS_{\text{AC}} = \dfrac{SS_{\text{AC}}}{(a-1)(c-1)}$	$\dfrac{MS_{\text{AC}}}{MSE}$
Interaction BC	SS_{BC}	$(b-1)(c-1)$	$MS_{\text{BC}} = \dfrac{SS_{\text{BC}}}{(b-1)(c-1)}$	$\dfrac{MS_{\text{BC}}}{MSE}$
Interaction ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	$MS_{\text{ABC}} = \dfrac{SS_{\text{ABC}}}{(a-1)(b-1)(c-1)}$	$\dfrac{MS_{\text{ABC}}}{MSE}$
Error	SSE	$abc(r-1)$	$MSE = \dfrac{SSE}{abc(r-1)}$	
Total	SS_y	$n-1 = abcr-1$	$s_y^2 = \dfrac{SS_y}{n-1}$	

Each computed $\mathbf{F_{obs}}$ is compared against $F_\alpha(\nu_{\text{effect}}, \nu_E)$ for a stated level of significance, α. If $F_{\text{obs}} < F_\alpha(\nu_{\text{effect}}, \nu_E)$, we fail to reject the null hypothesis that the effect is just an alias for random error (variation). If $F_{\text{obs}} > F_\alpha(\nu_{\text{effect}}, \nu_E)$, we reject H_0 and conclude the effect in question is real.

As this presentation demonstrates when there are many factors the analysis is quite complicated due to all of the interaction effects. In addition, the number of experimental runs increases as the product of the number of levels of each factor. To ease time, expense and analysis often only two levels are considered for each factor. Of course this results in a loss of detail, however, when there are many factors a “pilot experiment” done at two levels per factor can be used to eliminate insignificant factors. Such an experiment run with *n* factors is called a $\mathbf{2^n}$ **Factorial Experimental Design**. The use of only two levels considerably reduces the complexity of the analysis. Still 2^n is a big number for large *n*. For example, ten factors, which in an industrial process is not unreasonable, would require 1024 treatment groups ignoring replications! This is usually still too expensive. Fortunately, most higher order factor interactions turn out to be insignificant. Using the **Design Matrix** of the 2^n Factorial Design, one can “alias” factors with specific higher order interactions to reduce the number of treatment groups. This procedure is called a **Fractional Factorial Experimental Design**. The specific details of 2^n Factorial Experimental Designs and Fractional Factorial Experimental Designs are beyond the scope of these notes but are provided in the text.