Overview of the Two-way ANOVA

Factorial, With Replications, Balanced, Fixed Effect

Say there are two factors

Factor A has levels numbered i=1, ... a

Factor C has levels numbered j=1, ...c

and each combination has k=1, ... n replicatons.

The data could be laid out as follows:

Factor A ▼	Factor C j=1	 j=2	 ⋯	 j=c	Means for Factor A ▼
i=1	y_{111} y_{112} ⋮ y_{11n} $\bar{y}_{11\bullet}$	y_{121} y_{122} ⋮ y_{12n} $\bar{y}_{12\bullet}$	⋯	y_{1c1} y_{1c2} ⋮ y_{1cn} $\bar{y}_{1c\bullet}$	$\bar{y}_{1\bullet\bullet}$
i=2	y_{211} y_{212} ⋮ y_{21n} $\bar{y}_{21\bullet}$	y_{221} y_{222} ⋮ y_{22n} $\bar{y}_{22\bullet}$	⋯	y_{2c1} y_{2c2} ⋮ y_{2cn} $\bar{y}_{2c\bullet}$	$\bar{y}_{2\bullet\bullet}$
⋮	⋮	⋮	⋱	⋮	⋮
i=a	y_{a11} y_{a12} ⋮ x_{a1n} $\bar{y}_{a1\bullet}$	y_{a21} y_{a22} ⋮ y_{a2n} $\bar{y}_{a2\bullet}$	⋯	y_{ac1} y_{ac2} ⋮ y_{acn} $\bar{y}_{ac\bullet}$	$\bar{y}_{a\bullet\bullet}$
Means for Factor C ▶	$\bar{y}_{\bullet 1\bullet}$	$\bar{y}_{\bullet 2\bullet}$	⋯	$\bar{y}_{\bullet c\bullet}$	$\bar{y}_{\bullet\bullet\bullet}$

The model equation for this two way ANOVA could be written as:

$$y_{ijk}=\mu_{\text{baseline}}+\alpha_i+\gamma_j+(\alpha\gamma)_{ij}+\varepsilon_{ijk} \quad \text{for } i=1,\ldots a,\ j=1,\ldots c, \text{ and } k=1,\ldots n$$

where the y_{ijk} are the observations

μ_{baseline} is the baseline

$\alpha_1, \alpha_2, \ldots \alpha_a$ are the main effects for the levels of factor A

$\gamma_1, \gamma_2, \ldots \gamma_c$ are the main effects for the levels of factor C

$(\alpha\gamma)_{11}, (\alpha\gamma)_{12},\ldots (\alpha\gamma)_{21},\ldots (\alpha\gamma)_{ac}$ are the interactions for the combinations of A and C

and the ε_{ijk} are the errors that satisfy the conditions of mean equal to 0, equal variances, normality, and independence

The basic ANOVA table could then be written as:

Source	SS	df	MS	F
Between	$SSB=\sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(\bar{y}_{ij\bullet}-\bar{y}_{\bullet\bullet\bullet})^2$	*ac-1*	$MSB=\frac{SSB}{ac-1}$	$F=\frac{MSB}{MSW}$
▶Factor A	$SSA=\sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(\bar{y}_{i\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2$	*a-1*	$MSA=\frac{SSA}{a-1}$	$F=\frac{MSA}{MSW}$
▶Factor C	$SSC=\sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(\bar{y}_{\bullet j\bullet}-\bar{y}_{\bullet\bullet\bullet})^2$	*c-1*	$MSC=\frac{SSC}{c-1}$	$F=\frac{MSC}{MSW}$
▶AC Interaction	$SSAC=SSB-SSA-SSC$	*(ac-1)-(a-1)-(c-1)* *=(a-1)(c-1)*	$MSAC=\frac{SSAC}{(a-1)(c-1)}$	$F=\frac{MSAC}{MSW}$
Within	$SSW=\sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij\bullet})^2$	*acn-ac* *=ac(n-1)*	$MSW=\frac{SSW}{ac(n-1)}$	
Total	$TSS=\sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{\bullet\bullet\bullet})^2$	*acn-1*		

The formulas in this ANOVA table can be simplified to the following:

Source	SS	df	MS	F
Between	$SSB = n\sum_{i=1}^{a}\sum_{j=1}^{c}(\bar{y}_{ij\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	*ac-1*	$MSB = \frac{SSB}{ac-1}$	$F = \frac{MSB}{MSW}$
►Factor A	$SSA = nc\sum_{i=1}^{a}(\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	*a-1*	$MSA = \frac{SSA}{a-1}$	$F = \frac{MSA}{MSW}$
►Factor C	$SSC = na\sum_{j=1}^{a}(\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	*c-1*	$MSC = \frac{SSC}{c-1}$	$F = \frac{MSC}{MSW}$
►AC Interaction	$SSAC = n\sum_{i=1}^{a}\sum_{j=1}^{c}(\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} + \bar{y}_{\bullet\bullet\bullet})^2$	*ac-a-c+1* *=(a-1)(c-1)*	$MSAC = \frac{SSAC}{(a-1)(c-1)}$	$F = \frac{MSAC}{MSW}$
Within	$SSW = \sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij\bullet})^2$	*acn-ac* *=ac(n-1)*	$MSW = \frac{SSW}{ac(n-1)}$	
Total	$TSS = \sum_{i=1}^{a}\sum_{j=1}^{c}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{\bullet\bullet\bullet})^2$	*acn-1*		

Notice that there are four F statistics (each of which has its own p-value).
The null hypotheses tested by these F statistics are:

Between (Omnibus Test) H_0: $\alpha_1 = \ldots = \alpha_A$ and $\gamma_1 = \ldots = \gamma_c$ and $(\alpha\gamma)_{11} = (\alpha\gamma)_{12} = \ldots = (\alpha\gamma)_{21} = \ldots (\alpha\gamma)_{ac}$

Factor A (Factor A has no effect) H_0: $\alpha_1 = \alpha_2 = \ldots = \alpha_a$

Factor C (Factor C has no effect) H_0: $\gamma_1 = \gamma_2 = \ldots = \gamma_c$

Interaction AC (No Interactions) H_0: $(\alpha\gamma)_{11} = (\alpha\gamma)_{12} = \ldots = (\alpha\gamma)_{21} = \ldots (\alpha\gamma)_{ac}$

The reason that these various F-tests test the hypotheses we want can be seen by looking at the expected values of the mean squares. **If** we write the model so that the baseline sets the α, γ, and interactions to have mean zero, then the expected mean squares can be written as:

$$E(MSB) = \sigma_\varepsilon^2 + \frac{cn}{ac-1}\sum_{i=1}^{a}\alpha_i^2 + \frac{an}{ac-1}\sum_{j=1}^{c}\gamma_i^2 + \frac{n}{ac-1}\sum_{i=1}^{a}\sum_{j=1}^{c}(\alpha\gamma)_{ij}^2$$

$$E(MSA) = \sigma_\varepsilon^2 + \frac{cn}{a-1}\sum_{i=1}^{a}\alpha_i^2$$

$$E(MSC) = \sigma_\varepsilon^2 + \frac{an}{c-1}\sum_{j=1}^{c}\gamma_i^2$$

$$E(MSAC) = \sigma_\varepsilon^2 + \frac{n}{(a-1)(c-1)}\sum_{i=1}^{a}\sum_{j=1}^{c}(\alpha\gamma)_{ij}^2$$

$$E(MSW) = \sigma_\varepsilon^2$$

So, to test that there is no effect due to factor A, we would need to cancel out the σ_ε^2 in the E(MSA). We could do this by dividing the MS_A by the MSW, which is exactly what happens in the ANOVA table. The other tests are made similarly.

Note 1: If we use the baseline/control-case constraints (or any constraint besides the sum to zero ones) the above expected mean squares will be slightly different; replace α_i with $(\alpha_i - \overline{\alpha}_\bullet)$, γ_j with $(\gamma_j - \overline{\gamma}_\bullet)$, and $(\alpha\gamma)_{ij}$ with $((\alpha\gamma)_{ij} - \overline{(\alpha\gamma)}_{i\bullet} - \overline{(\alpha\gamma)}_{\bullet j} + \overline{(\alpha\gamma)}_{\bullet\bullet})$. The hypotheses being tested are the same. Notice what happens to these formulas if the sum to zero constrains are used.

Note 2: The basic rules (how to make the MS and F) for higher-way ANOVA tables are the same as for the One-way ANOVA. The only major difference is that we are splitting up the SSB.

Note 3: The construction of a three-way ANOVA table works similarly to that of the above two-way table. There are a variety of books on ANOVA and design of experiments that give the formulas if you need them.

Note 4: The formulas in the ANOVA table above, and the argument about the F tests, only work if the design is Factorial, With Replications, Balanced, and Fixed Effect