21.0 Two-Factor Designs

- Answer Questions
- RCBD
- Concrete Example
- Two-Way ANOVA
- Popcorn Example

21.4 RCBD

The **Randomized Complete Block Design** is also known as the two-way ANOVA without interaction. A key assumption in the analysis is that the effect of each level of the treatment factor is the same for each level of the blocking factor.

That assumption would be violated if, say, a particular fertilizer worked well for one stain but poorly for another; or if one cancer therapy were better for lung cancer but a different therapy were better for stomach cancer.

In RCBD, there is one observation for each combination of levels of the treatment and block factors.

RCBD notation:

- I is the number of treatments; J is the number of blocks
- X_{ij} is the measurement on the unit in block j that received treatment i.
- $n_{..} = I * J$ is the total number of experimental units.
- $n_{i.} = J$ and $n_{.j} = I$, the number of observations for a given treatment level or block level, respectively.
- $X_{i.}$ is the sum of all measurements for units receiving treatment i, and $X_{.j}$ is the sum of all measurements for units in the j block.
- $\bar{X}_{i.}$ is the average of all measurements for units receiving treatment i, or $\frac{1}{J}\sum_{j=1}^{J} X_{ij}$.
- $\bar{X}_{.j}$ is the average of all measurements for units in the jth block, or $\frac{1}{I}\sum_{i=1}^{I} X_{ij}$.
- $\bar{X}_{..}$ is the average of all measurements.

We continue to use the dot and bar conventions.

Some assumptions:

The model for an RCBD (or two-way ANOVA without interactions) is:

$$X_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

where μ is the overall mean of all experimental units, τ_i is the effect of treatment i, β_j is the effect of block j, and the ϵ_{ij} are random errors that are:

- normally distributed with mean zero and unknown standard deviation σ
- independent of the values for all other errors.

Note how this generalizes the one-way ANOVA model:

$$X_{ij} = \mu + \tau_i + \epsilon_{ij}.$$

There are two hypothesis tests in an RCBD, and they are always the same:

$\mathbf{H_0}$: The means of all treatments are equal or H_0: $\tau_1 = \cdots = \tau_I = 0$

versus

$\mathbf{H_A}$: At least one of the treatments has a different mean

and

$\mathbf{H_0}$: The means of all blocks are equal or H_0: $\beta_1 = \cdots = \beta_J = 0$.

versus

$\mathbf{H_A}$: At least one of the blocks has a different mean

In the case of an RCBD, we divide the total sum of squares (SS_{tot}) into the part attributable to differences between the treatment means (SS_{trt}), the part attributable to difference between the blocks (SS_{blk}) and the part attributable to differences within block-treatment groups, or the sum of squares due to pure error (SS_{err}).

The test statistics for an RCBD look at the standardized ratio of the between-group sum of squares to the within-group sum of squares for the treatment and block effects separately. If one or both are large, then it suggest that some of the treatment or block effects are not equal.

Formally, the test statistics are:

$$ts_{\mathbf{trt}} = \frac{\mathbf{SS_{trt}}/(I-1)}{\mathbf{SS_{err}}/(IJ-I-J+1)}$$

and

$$ts_{\mathbf{blk}} = \frac{\mathbf{SS_{blk}}/(J-1)}{\mathbf{SS_{err}}/(IJ-I-J+1)}.$$

These two test statistics are compared to an F distribution with $I-1$ or $J-1$ (respectively) degrees of freedom in the numerator and $IJ-I-J+1$ degrees of freedom in the denominator.

The numerator degrees of freedom for the test of treatments is $I-1$. That is because we make $I-1$ estimates to get it; one estimate for each of the I different groups, but the last one is not needed since the sum of the treatment effects is forced to add up to zero (since we have already found the overall mean, which cost us one 1 degree of freedom).

Similarly, the numerator degrees of freedom for the block term is $J-1$.

Since that the total degrees of freedom is $IJ-1$, then the degrees of freedom for the error term in the denominator must be $(IJ-1)-(I-1)-(J-1) = IJ-I-J+1 = (I-1)(J-1)$. As always, we have lost one degree of freedom due to estimating μ, the overall mean of the combined populations.

To understand this test statistic, consider the following definitional formulae:

$$
\begin{aligned}
\mathbf{SS_{tot}} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{..})^2 \\
\mathbf{SS_{trt}} &= \sum_{i=1}^{I} J(\bar{X}_{i.} - \bar{X}_{..})^2 \\
\mathbf{SS_{blk}} &= \sum_{j=1}^{J} I(\bar{X}_{.j} - \bar{X}_{..})^2 \\
\mathbf{SS_{err}} &= \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2
\end{aligned}
$$

In practice, if we aren't using a computer, then it is easier, faster, and more numerically stable to use the following computational formulae:

$$\begin{aligned}
\mathbf{SS_{tot}} &= \left(\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}^2\right) - \frac{X_{..}^2}{IJ} \\
\mathbf{SS_{blk}} &= \sum_{j=1}^{J} \frac{X_{.j}^2}{I} - \frac{X_{..}^2}{IJ} \\
\mathbf{SS_{trt}} &= \sum_{i=1}^{I} \frac{X_{i.}^2}{J} - \frac{X_{..}^2}{IJ} \\
\mathbf{SS_{err}} &= \left(\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}^2\right) - \sum_{i=1}^{I} \frac{X_{i.}^2}{J} - \sum_{j=1}^{J} \frac{X_{.j}^2}{I} + \frac{X_{..}^2}{IJ}
\end{aligned}$$

Note that $\mathrm{SS}_{\mathrm{tot}} = \mathrm{SS}_{\mathrm{trt}} + \mathrm{SS}_{\mathrm{blk}} + \mathrm{SS}_{\mathrm{err}}$. This is useful because it simplifies calculation—if you know three of the sum of square terms, you can find the fourth by subtraction.

Once you have calculated the test statistics, you refer them to F distributions. For the test of equal treatment effects, the numerator df is $I-1$. For the test of equal block effects, the numerator df is $J-1$. Both tests have the same denominator df, $(I-1)(J-1)$.

Note that both RCBD tests are one-sided—we reject if and only if we get large values of the test statistic. Also note that if we had only two treatments, then this would be informationally equivalent to a paired difference two-sample t-test and use an F with 1 df in the numerator and $J-1$ df in the denominator.

To simplify the organization of the calculations in an RCBD, it is customary to write things in a table.

Source	df	SS	MS	F
treatment	$I-1$	$\mathrm{SS}_{\mathrm{trt}}$	$\mathrm{SS}_{\mathrm{trt}}/(I-1)$	$\mathrm{MS}_{\mathrm{trt}}/\mathrm{MS}_{\mathrm{err}}$
block	$J-1$	$\mathrm{SS}_{\mathrm{blk}}$	$\mathrm{SS}_{\mathrm{blk}}/(J-1)$	$\mathrm{MS}_{\mathrm{blk}}/\mathrm{MS}_{\mathrm{err}}$
error	$(I-1)(J-1)$	$\mathrm{SS}_{\mathrm{err}}$	$\mathrm{SS}_{\mathrm{err}}/(I-1)(J-1)$	
total	$IJ-1$	$\mathrm{SS}_{\mathrm{tot}}$		

The MS column contains the **Mean Squares**, which are the average sum of squares attributable to each component in the partition. The F column contains the test statistic.

What happens if the block effect is not significant?

21.2 Concrete Example

Suppose you are manufacturing concrete cylinders for, say, bridge supports. There are three ways of drying green concrete (say A, B, and C), and you want to find the one that gives you the best compressive strength. The concrete is mixed in batches that are large enough to produce exactly three cylinders, and your production engineer believes that there is substantial variation in the quality of the concrete from batch to batch.

You have data from $J = 5$ batches on each of the $I = 3$ drying processes. Your measurements are the compressive strength of the cylinder in a destructive test. (So there is an economic incentive to learn as much as you can from a well-designed experiment.)

The data are:

	Batch					
Treatment	1	2	3	4	5	Trt Sum
A	52	47	44	51	42	236
B	60	55	49	52	43	259
C	56	48	45	44	38	231
Batch Mean	168	150	138	147	123	726

The primary null hypothesis is that all three drying techniques are equivalent, in terms of conferring compressive strength. The secondary null is that the batches are equivalent (but if they are, then we have wasted power by controlling for an effect that is small or non-existent).

We use the computational forms for the sum of squares calculation. Thus:

$$
\begin{aligned}
\mathbf{SS_{tot}} &= \left(\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}^2\right) - \frac{X_{..}^2}{IJ} \\
&= 499.6 \\
\mathbf{SS_{blk}} &= \sum_{j=1}^{J} \frac{X_{.j}^2}{I} - \frac{X_{..}^2}{IJ} \\
&= 363.6 \\
\mathbf{SS_{trt}} &= \sum_{i=1}^{I} \frac{X_{i.}^2}{J} - \frac{X_{..}^2}{IJ} \\
&= 89.2 \\
\mathbf{SS_{err}} &= \left(\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}^2\right) - \sum_{i=1}^{I} \frac{X_{i.}^2}{J} - \sum_{j=1}^{J} \frac{X_{.j}^2}{I} + \frac{X_{..}^2}{IJ} \\
&= 46.8
\end{aligned}
$$

We plug this into the ANOVA table:

Source	df	SS	MS	F
drying	2	89.2	44.6	7.62
batch	4	363.6	90.9	15.54
error	8	46.8	5.85	
total	14	499.6		

Our test statistics are 7.62 and 15.54. The test of the first uses an F with 2 df in the numerator, 8 in the denominator, so the 0.05 critical value is 4.46. We reject.

The secondary test has 4 df in the numerator, 8 in the denominator, and the 0.05 critical value is 3.84. The blocking was a good idea.

Suppose we had not blocked for batch. Then the data would be:

Treatment		Trt Sum
A	52, 47, 44, 51, 42	236
B	60, 55, 49, 52, 43	259
C	56, 48, 45, 44, 38	231

This is the same as before except now we ignore which batch the observation came from.

The one-way ANOVA table for this is:

Source	df	SS	MS	F
drying	2	89.2	44.6	1.30
error	$4 + 8$	46.8+363.6	34.2	
total	14	499.6		

Note that this gives a test statistic of 1.30, which is referred to an F-distribution with 2 df in the numerator, 12 in the denominator. The .05 critical value is 3.89. We fail to reject the null.

Using blocks gave us a more powerful test.

21.3 Two-Way ANOVA

The two-way Analysis of Variance is used when one suspects that there may be some kind of interaction between the levels of two experimental factors. This would occur, for example, if

- one kind of fertilizer were best for sunny fields, and a different kind for wet fields;
- one kind of chemotherapy was better for lung cancer, and a different kind for stomach cancer;
- female students learned better from reading assignments, but male students learned better from lecture.

The presence of interaction means that the simple RCBD model, with a block effect and a treatment effect that add, cannot explain the observations.

Two-Way ANOVA notation:

- I (J) is the number of levels of Factor A (B).
- X_{ijk} is the measurement on the kth experimental unit that receives level i of factor A and level j of factor B.
- the index k runs from 1 to $n_{ij.}$, the number of units that receive level i of factor and level j of factor B.
- $n_{...}$ is the total number of experimental units.
- $n_{i..}$ and $n_{.j.}$ are the number of units receiving the ith or j levels of factors A or B, respectively. And $n_{ij.}$ is the number of units that receive level i of factor A and level j of factor B.
- $X_{i...}$ is the sum of all measurements for units receiving level i of factor A; $X_{.j.}$ is the sum of all units receiving level j of factor B; $X_{ij.}$ is the sum for all units receiving both.

We continue to use the dot and bar conventions. So $\bar{X}_{i..}$ is the average of all measurements on units receiving level i of factor A, and so forth.

The model for the two-way ANOVA with interaction is:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}.$$

where γ_{ij} is the interaction between levels i and j of factors A and B, and the errors are:

- normally distributed with mean zero and unknown standard deviation σ
- independent of the values for all other errors.

As written, this model is not identifiable. To fit it, we add the additional constraints that:

$$\begin{aligned} \alpha_1 + \cdots + \alpha_I &= 0 \\ \beta_1 + \cdots + \beta_J &= 0 \\ \gamma_{11} + \cdots + \gamma_{IJ} &= 0 \end{aligned}$$

There are three hypothesis tests in a two-way ANOVA with interaction, and they are always the same:

H_0: The interaction effects are all zero or H_0: $\gamma_{11} = \cdots = \gamma_{IJ} = 0$

versus

H_A: At least one of the interactions is non-zero.

The next two tests concern factors A and B:

H_0: The means of all levels of Factor A are equal or
H_0: $\alpha_1 = \cdots = \alpha_I = 0$.

versus

H_A: At least one of the factor A levels has a non-zero mean

and similarly for the corresponding hypotheses for Factor B.

If the interaction term is statistically significant, we usually immediately decide that both main effects (Factors A and B) are also significant, no matter what the result of the test.

It can happen that both main effects tests are not significant but the interaction is. This situation is called **masking**.

As usual, we divide the total sum of squares (SS_{tot}) into the parts attributable to differences between the treatment means or main effects, written as SS_A and SS_B, the part attributable to the interaction, written as SS_{AB}, and the part attributable to random variation, or SS_{err}.

For the fixed effects model, the test statistics are:

$$ts_{\mathbf{A}} = \frac{\mathbf{SS_A}/(I-1)}{\mathbf{SS_{err}}/(n_{...} - IJ)}$$

$$ts_{\mathbf{B}} = \frac{\mathbf{SS_B}/(J-1)}{\mathbf{SS_{err}}/(n_{...} - IJ)}$$

$$ts_{\mathbf{AB}} = \frac{\mathbf{SS_{AB}}/(I-1)(J-1)}{\mathbf{SS_{err}}/(n_{...} - IJ)}$$

The test statistics for main effects A and B, and interaction AB, are compared to an F distribution with $I-1$ or $J-1$ or $(I-1)(J-1)$ degrees of freedom, respectively, in the numerator and $n_{\ldots} - IJ$ degrees of freedom in the denominator.

As usual, thoughtful counting will show that the degrees of freedom correspond to the number of estimates that one makes to find the effects of each of these quantities.

Since that the total degrees of freedom is $n_{\ldots} - 1$, then the degrees of freedom for the error term in the denominator must be $n_{\ldots} - (I-1) - (J-1) - (I-1)(J-1) - 1 = n_{\ldots} - IJ$. As always, we have lost one degree of freedom due to estimating μ, the overall mean of the combined populations.

The following definitional formulae for the sums of squares terms are:

$$\mathbf{SS_{tot}} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij.}} (X_{ijk} - \bar{X}_{...})^2$$

$$\mathbf{SS_{A}} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij.}} (\bar{X}_{i..} - \bar{X}_{...})^2$$

$$\mathbf{SS_{B}} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij.}} (\bar{X}_{.j.} - \bar{X}_{...})^2$$

$$\mathbf{SS_{AB}} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...})^2$$

One finds $\mathrm{SS_{err}}$ by subtraction.

To simplify the organization of the calculations in a fixed effects two-way ANOVA with interaction, it is customary to write things in a table.

Source	df	SS	MS	F
treatment A	$I-1$	SS_{A}	$\mathrm{SS}_{\mathrm{A}}/(I-1)$	$\mathrm{MS}_{\mathrm{A}}/\mathrm{MS}_{\mathrm{err}}$
treatment B	$J-1$	SS_{B}	$\mathrm{SS}_{\mathrm{B}}/(J-1)$	$\mathrm{MS}_{\mathrm{B}}/\mathrm{MS}_{\mathrm{err}}$
interaction	$(I-1)(J-1)$	$\mathrm{SS}_{\mathrm{AB}}$	$\mathrm{SS}_{\mathrm{AB}}/(I-1)(J-1)$	$\mathrm{MS}_{\mathrm{AB}}/\mathrm{MS}_{\mathrm{err}}$
error	$n_{\dots}-IJ$	$\mathrm{SS}_{\mathrm{err}}$	$\mathrm{SS}_{\mathrm{err}}/(n_{\dots}-IJ)$	
total	$n_{\dots}-1$	$\mathrm{SS}_{\mathrm{tot}}$		

The MS column contains the **Mean Squares**, which are the average sum of squares attributable to each component in the partition. The F column contains the test statistic.

21.4 Popcorn Example

Suppose you want to compare type of popcorn popper and brand of popcorn with respect to their yield (in terms of cups of popped corn). Factor A is the type of popper: oil-based versus air-based. Factor B is the brand of popcorn: gourmet versus national brand versus generic. For each combination of popper type and brand, you took three separate measurements.

	Corn			
Popper	Gourmet	Nat'l Brand	Generic	Row Sum
Oil	5.5, 5.5, 6	4.5, 4.5, 4	3.5, 4, 3	40.5
Air	6.5, 7, 7	5, 5.5, 5	4, 5, 4.5	49.5
col sum	37.5	28.5	24.0	90.0

We use the computational forms for the sum of squares calculation with n the number of observations for each combination of factor levels.

$$
\begin{aligned}
\mathbf{SS_{tot}} &= \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n_{ij.}} X_{ijk}^2 \right) - \frac{X_{...}^2}{n_{...}} = 22.0 \\
\mathbf{SS_{A}} &= \frac{1}{Jn} \left(\sum_{i=1}^{I} X_{i..}^2 \right) - \frac{X_{...}^2}{n_{...}} = 4.5 \\
\mathbf{SS_{B}} &= \frac{1}{In} \left(\sum_{j=1}^{J} X_{.j.}^2 \right) - \frac{X_{...}^2}{n_{...}} = 15.75 \\
\mathbf{SS_{AB}} &= \frac{1}{n} \left(\sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij.}^2 \right) - \frac{X_{...}^2}{n_{...}} - \mathbf{SS_{A}} - \mathbf{SS_{B}} = .083 \\
\mathbf{SS_{err}} &= \mathbf{SS_{tot}} - \mathbf{SS_{A}} - \mathbf{SS_{B}} - \mathbf{SS_{AB}} = 1.667
\end{aligned}
$$

We plug this into the ANOVA table:

Source	df	SS	MS	F
popper	1	4.5	4.5	32.4
corn	2	15.75	7.87	56.7
interaction	2	.083	.042	.3
error	12	1.667	.139	
total	17	22.0		

Our test statistics are .3 for the interaction, 56.7 for the corn, and 32.4 for the popper. The interaction test and the corn test both use an F with 2 df in the numerator, 12 in the denominator, so the 0.05 critical value is 3.89. We conclude that there is no interaction and that the brand of corn is significantly different. For the popper, there is 1 df in the numerator and the F critical value is 4.75. There is strong evidence that the popper is also significantly different.