COMPLETE BLOCK DESIGNS

DESIGN

Complete Block Designs

Fisher's basic principles of experimental design are often summarized as: randomization, blocking, and replication. The topic of this module is blocking.

The complete block design is a special case of the two-way (two-factor) layout. One of the factors identifies "blocks" of homogeneous material, within which levels of the other factor are compared.

Fisher's basic principles of experimental design are often summarized as: randomization, blocking, and replication. The topic of this module is blocking.

The complete block design is a special case of the two-way (two-factor) layout. One of the factors identifies "blocks" of homogeneous material, within which levels of the other factor are compared.

Example 1

Suppose r subjects have their blood pressure measured on each of t antihypertensive drugs. (This is a *crossover* design.)

Let y_{ij} be the blood pressure of subject j on drug i.

The goal is to determine whether the treatment means differ, after accounting for blocking by subject.

If $t = 2$, this is the paired design.

Let's first see some examples.

Suppose r subjects have their blood pressure measured on each of t antihypertensive drugs. This is a crossover design: each subject will receive all the treatments.

Let (y i j) be the blood pressure of subject j on drug i.

The goal is to determine whether the treatment means differ, after accounting for blocking by subject.

If t = 2, this is the paired design.

Example 1 (continued)

Treatment (drug)	Block (subject) 1	2	⋯	r
1	y_{11}	y_{12}	⋯	y_{1r}
2	y_{21}	y_{22}	⋯	y_{2r}
⋮	⋮	⋮		⋮
t	y_{t1}	y_{t2}	⋯	y_{tr}

The data structure in Example 1: each subject will try all t drugs in random order.

This table shows the data structure in Example 1.

In this so-called crossover design, each subject will try all t drugs in random order.

The main interest here is in comparing different drugs. Subjects serve as a blocking factor.

Discussion

In Example 1, the drugs are the treatments of interest. What is the experimental unit?

Hint: Recall that to correctly identify the experimental unit, we need to ask at which level the treatment is applied. What is the smallest unit that can receive different treatment?

In Example 1, the drugs are the treatments of interest. What is the experimental unit?

Recall that to correctly identify the experimental unit, we need to ask at which level the treatment is applied. What is the smallest unit that can receive different treatment?

(Think about this for one minute before moving on to the next slide.

Discussion (answer)

What is the experimental unit in Example 1?

Note that the same subject will receive all treatments. So the experiment unit is actually one treatment period of one subject. That is the smallest unit on which we can apply different treatments to.

Note that the same subject will receive all treatments. So the experiment unit is actually one treatment period of one subject. That is the smallest unit on which we can apply different treatments to.

A warning about crossover designs

In crossover designs, each subject will receive all treatments. That may lead to carryover effects, which are difficult to deal with.

In crossover designs, each subject will receive all treatments. That may lead to carryover effects, which are difficult to deal with.

Example 2

(Kuehl, p. 264) A study of the effect of fertilization method on the nitrogen content of wheat. Twenty-four plots in a field are available for testing six methods.

OPTION 1. Completely randomized experiment. For each of the 24 plots, randomly assign one of the 6 methods, so that each method is represented 4 times. (The randomization procedure is as discussed for the one-way layout.)

Problem: The plots differ in moisture content, due to an irrigation gradient. So the observed nitrogen values will have variability due to moisture, as well as random error.

Example 2

Suppose a researcher wants to study the effect of fertilization method on the nitrogen content of wheat. 24 plots in a field are available for testing 6 methods.

Option 1 will be to use a completely randomized design. For each of the 24 plots, randomly assign one of the six methods, so that each method is represented four times. The randomization procedure is as discussed for the one-way layout.

In fact, this is the design we studied in Module 3

The problem with this design is that the plots differ in moisture content, due to an irrigation gradient. So the observed nitrogen values will have variability due to moisture, as well as random error.

Valid, but not precise

In Fisher's 1926 paper (it's time to read it again), Fisher discussed "validity" and "precision".

If we use a completely randomized design in this scenario when there is moisture gradient in the field, our inference will still be valid, but it will not precise: after randomization, the random errors will be large.

One may say: but we can adjust for moisture afterwards using an ANCOVA model.

- That is true, but—this is subtle—the design will still be inefficient since the distribution of moisture levels among different treatment levels will unlikely to be balanced, so the covariate, moisture level, will not be **orthogonal** of the treatment factors.
- In a regression model, if two predictors are not orthogonal, the inference on their effects will be less precise.

In Fisher's 1926 paper (it's time to read it again), Fisher discussed "validity" and "precision".

If we use a completely randomized in this scenario when there is moisture gradient in the field, our inference will still be valid, but it will not precise: after randomization, the random errors will be large.

One may say: but we can adjust for moisture afterwards using an ANCOVA model.

That is true, but—this is subtle—the design will still be inefficient since the distribution of moisture levels among different treatments will unlikely to be balanced, so the covariate, moisture level, will not be orthogonal of the treatment factors.

Fully understanding orthogonal is an advanced topic. In this class, we just need the intuition that, in a regression model, if two predictors are not orthogonal, the inference on their effects will be less precise.

Example 2 (continued)

OPTION 2. ***Randomized complete block design*** (RCB design or RCBD).

- Divide the plots into four blocks in the direction that is perpendicular to the moisture gradient so that the moisture level within each block will be more similar.
- Test each method once in each of four blocks of plots, with randomization of treatments within blocks.

"Complete" means that each block includes all of the treatments.

The RCBD removes the block-to-block variability (related to the moisture gradient) from the error in the nitrogen data.

The figures on the next two slides show examples of these two designs for the wheat experiment, with letters indicating the six fertilization methods.

Another option for this study is to use a randomized complete block design (RCB design or RCDB).

We will divide the plots into four blocks in the direction that is perpendicular to the moisture gradient so that the moisture level within each block will be more similar.

We then test each method once in each of four blocks of plots, with randomization of treatments within blocks.

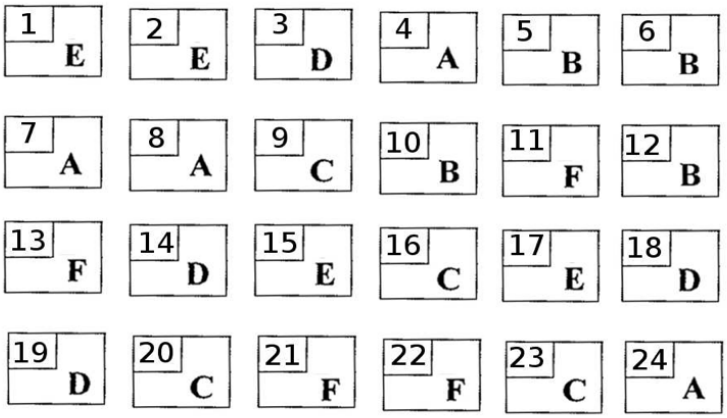
"Complete" means that each block includes all of the treatments.

Using an RCB design removes the block-to-block variability (related to the moisture gradient) from the error in the nitrogen data.

The figures on the next two slides show examples of these two designs for the wheat experiment, with letters indicating the six fertilization methods.

Completely randomized design

Completely randomized:



This figure shows the schematic layout of a completely randomized design. The six treatments are completely randomized to the 24 plots.

Note, for example, that treatment C does not appear in the first row, and appear two times in the last row.

Randomized complete block

Randomized complete block:

1	2	3	4	5	6
B	E	D	A	F	C
A	C	D	F	E	B
F	C	E	A	B	D
B	D	F	E	C	A

This figure show shows the schematic layout of an RCBD, or randomized complete block design.

The treatments are completely randomized within each row. Each row contains a complete set of replicates of all treatments.

Since experimental units within each block are more similar, comparisons of treatment means will be more precise using an RCB design.

Randomization procedure

Randomization is restricted, so that each treatment occurs once in each block.

Treatments are completely randomized to experimental units within each block.

To generate the randomization plan, for each block,

- label the plots as $1, \ldots, t$,
- obtain a random permutation of the numbers $1, \ldots, t$ (e.g., using `sample(1:t)` in R), and
- assign treatments to plots within blocks according to the permuted numbers.

In an RCBD experiment, randomization is restricted, so that each treatment occurs once in each block.

Treatments are completely randomized to experimental units within each block.

To generate the randomization plan, for each block,

label the plots as 1–t,

obtain a random permutation of the numbers 1–t, and

assign treatments to plots within blocks according to the permuted numbers.

Advantages of a randomized complete block design

1. If blocking is effective, we get more precise results than from a completely randomized design.

2. Variability in experimental units can be deliberately introduced to **increase generality** without sacrificing precision. (**Increasing generality is often overlooked in practice.**)

   - For example, if we find one pathogen is effective in an experiment carried out using one cultivar. We can repeat the experiment on other cultivars. If those additional experiments show same conclusions. We gain evidence that the pathogen is effective more generally.

There are two general advantages of a RCBD.

First, if blocking is effective, we get more precise results than from a completely randomized design.

Second, variability in experimental units can be deliberately introduced to **increase generality** without sacrificing precision. Increasing generality is often overlooked in practice.

For example, if we find one pathogen is effective in an experiment carried out using one cultivar. We can repeat the experiment on other cultivars. If those additional experiments show same conclusions. We gain evidence that the pathogen is effective more generally.

Disadvantages

1. We lose one degree of freedom for each block after the first, so the residual sum of squares has fewer degrees of freedom than in a completely randomized design.
2. If there is only one observation per treatment/block combination, one must assume there are no treatment/block interactions. (The treatment/block interaction is totally confounded with the error terms in the model.)
3. Missing observations require more complex calculations.

There are also some disadvantages associated with RCBD:

1. We lose one degree of freedom for each block after the first, so the residual sum of squares has fewer degrees of freedom than in a completely randomized design.

2. If there is only one observation per treatment/block combination, one must assume there are no treatment/block interactions. The treatment/block interaction is totally confounded with the error terms in the model.

3. Missing observations require more complex calculations.

Examples of blocking factors

Age or weight of animals, batches of reagents, gender, physical location (e.g., neighboring plots), time.

Here we list some examples of commonly used blocking factors: age or weight in animal studies, batches of reagents is chemical or biological experiments, gender in animal or human studies, physical location in crop field experiments, time, and so on.

ANALYSIS

Model equation

The model equation for an RCBD experiment can be written as:

$$y_{ij} = \mu + \tau_i + \rho_j + e_{ij}$$

where

- y_{ij} = observation of ith treatment in jth block ($i = 1, \ldots, t; j = 1, \ldots, r$).
- μ = overall mean
- τ_i = effect of ith treatment (fixed)
- ρ_j = effect of jth block (fixed)
- e_{ij} = random error, assumed iid $N(0, \sigma^2)$.

Note there is no interaction term since interaction would be totally confounded with the error terms.

Since there is no interaction term in these model, whether we treat the block effect as random or fixed will not affect inferences on treatment contrasts.

The model equation for an RCBD experiment looks similar to the model equation for two-factor fixed-effects model with a completely randomized design.

But note that there is no interaction term since interaction would be totally confounded with the error terms.

Since there is no interaction term in these model, whether we treat the block effect as random or fixed will not affect inferences on treatment contrasts.

The key difference between an RCBD experiment and a completely randomized experiment are the randomization schemes. There are many subtle implications that are not fully reflected in the model equation.

Compare with a two-factor completely randomized experiment

Comparing the model equations for two-factor CRD and RCBD, besides the interaction, the only difference is that we replaced α, β there by τ, ρ here.

However, the key difference between the two types of experiments are the **randomization schemes**. There are many subtle implications that are not fully reflected in the model equation:

1. In a CRD experiment, it is reasonable to assume all error terms iid since all experimental units are "interchangeable" with regard to the randomization.

2. In an RCBD experiment, the iid assumption is no longer automatic: the randomization is within each block, there is no obvious reason why the error variance from different blocks should be similar.

3. Due to restricted randomization, it is also not reasonable to test for the block effects in an RCBD. (This is subtle. Intuitively, there is no randomization of blocks, and thus no reliable way to estimate between-block errors.)

The model equation for an RCBD experiment looks almost identical to the model equation for two-factor fixed-effects model with a completely randomized design. besides the interaction term, on the surface, the only difference is that we replaced alpha, beta there by tau, rho here.

However, the key difference between the two types of experiments are the randomization schemes. There are many subtle implications that are not fully reflected in the model equation:

1. In a completely randomized design, the treatments (all level combinations) are completely randomly assigned to experimental units.

It is reasonable to assume all error terms iid since all experimental units are "interchangeable" with regard to the randomization scheme.

2. In an RCBD experiment, there is one observation per treatment in each block, treatments are randomized within each block.

The iid assumption is no longer automatic: the randomization is within each block, there is no obvious reason why the error variance from different blocks should be similar.

3. Due to restricted randomization, it is also not reasonable to test for the block effects in an RCBD. This is subtle. But intuitively, there is no randomization of blocks, and thus no reliable way to estimate between-block errors.

Big picture: follow the randomization

Now that we have seen experiments with more than one factors, it is important to **follow the randomization**:

- The randomization determines the assumption on the error terms (think about the RCBD versus two-factor RCD differences).
- More generally, there can be more than one random terms in a model equation, each correspond to a randomization step or a random sampling step. Think about the random-effects models.

Now that we have seen experiments with more than one factors, it is important to **follow the randomization**.

The randomization determines the assumption on the error terms. For example, one can think about the differences between an RCBD experiment and a two-factor completely randomized experiment that we just talked about.

More generally, there can be more than one random terms in a model equation, each correspond to a randomization step or a random sampling step. For example, one can think about the random-effects models.

Partitioning of the sum of squares

We can write

$$y_{ij} - \overline{y}_{..} = (\overline{y}_{i.} - \overline{y}_{..}) + (\overline{y}_{.j} - \overline{y}_{..}) + (y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})$$

This leads to the partitioning of the total sum of squares and degrees of freedom:

$$\text{SS Total} = \text{SS Treatment} + \text{SS Block} + \text{SSE}$$
$$rt - 1 = (t - 1) + (r - 1) + (t - 1)(r - 1).$$

The total sum of squares from an RCBD experiment can be partitioned as the sum of

SS Treatment, SS Block and SSE

Note that we do not have an interaction term in the RCBD model, and there is no sum of squares for interaction either.

ANOVA table

Table 8.2 Analysis of variance for an experiment in a randomized complete block design

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*Expected Mean Square*
Total	$rt-1$	$\sum_i \sum_j (y_{ij} - \overline{y}_{..})^2$		
Blocks	$r-1$	$t \sum_j (\overline{y}_{.j} - \overline{y}_{..})^2$	$MS\boldsymbol{B}$	
Treatments	$t-1$	$r \sum_i (\overline{y}_{i.} - \overline{y}_{..})^2$	$MS\boldsymbol{T}$	$\sigma^2 + r\theta_t^2$
Error	$(r-1)(t-1)$	SS Error	MSE	σ^2

Here is the ANOVA table for an RCBD experiment. In particular, note that the expected mean square is not computed for MSB, Mean Square Block.

Also, here the treatments are assumed fixed-effects.

Test the equality of treatment means

To test the equality of treatment means

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_t = 0,$$

calculate $F_0 = MST/MSE$.

If H_0 is true, the F-statistic has the $F_{t-1,(t-1)(r-1)}$ distribution.

The p-value is

$$Pr(F_{t-1,(t-1)(r-1)} > F_0) = 1 - \mathtt{pf}(F_0, t-1, (t-1)(r-1)).$$

(`pf` refers to the R function.)

To test for no treatment effect, the null hypothesis is that all treatment parameters are equal. The F-test statistic is MST over MSE.

This part is the same as in the two-factor fixed-effects model.

Note 1: The form of the error sum of squares

In the RCBD experiment, the fitted value is

$$\hat{y}_{ij} = (\text{overall mean}) + (\text{treatment effect}) + (\text{block effect})$$
$$= \overline{y}_{..} + (\overline{y}_{i.} - \overline{y}_{..}) + (\overline{y}_{.j} - \overline{y}_{..}) = \overline{y}_{i.} + \overline{y}_{.j} - \overline{y}_{..},$$

and

$$\text{SSE} = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \hat{y}_{ij})^2 = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2.$$

This SSE has the same form as the interaction sum of squares (SS Interaction) in a two-way analysis of variance (if the interaction term were fitted).

In the RCBD experiment, SSE has the same form as the interaction sum of squares in a two-way analysis of variance.

Note 1 continued

SSBlock + SSE is equal to the error sum of squares we would get if we fit the single-factor model, erroneously ignoring the blocking factor:

$$\text{SSBlock} + \text{SSE} = \sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{.j} - \overline{y}_{..})^2 + \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.})^2$$

In a sense, when we block, the error sum of squares from a single-factor completely randomized design is replaced by a sum of squares due to blocking and a new, smaller error sum of squares for the RCB design.

In an RCBD experiment, (SS Block + SSE) is equal to the error sum of squares we would get if we fit the single-factor model, erroneously ignoring the blocking factor.

In a sense, when we block, the error sum of squares from a single-factor completely randomized design is replaced by a sum of squares due to blocking and a new, smaller error sum of squares for the RCB design.

Note 2: treatment block interaction

In the usual RCB design, we have no replication within treatment/block combinations, so we cannot test for a treatment × block interaction. **There is not way to tell apart pure error from the treatment by block interaction.** We have to assume there is no interaction, i.e., that the effects of treatments and blocks are additive.

Another way to look at this is that we use the treatment×block interaction as the error terms for assessing the significance of the treatment effects. Recall that the treatment by block interaction simply means the variation of treatment effects among blocks.

Roughly speaking, from an RCBD design, a treatment effect is only significant if it is consistent across blocks and its magnitude is large relative to its variation among blocks. (i.e., the treatment by block interaction is not significant.)

In the usual RCB design, we have no replication within treatment/block combinations, so we cannot test for a treatment-block interaction. There is not way to tell apart pure error from the treatment by block interaction. We have to assume there is no interaction, i.e., that the effects of treatments and blocks are additive.

Another way to look at this is that we use the treatment by block interaction as the error terms for assessing the significance of the treatment effects. Recall that the treatment by block interaction simply means the variation of treatment effects among blocks.

Roughly speaking, from an RCBD design, a treatment effect is only significant if it is consistent across blocks and its magnitude is large relative to its variation among blocks (i.e., the treatment by block interaction is not significant.)

Notes 3: testing block effects?

What about testing $H_0: \rho_1 = \cdots = \rho_r = 0$ (no block effects)?

An approximate F-test is done by comparing $F_0 = \text{MS Blocks/MSE}$ to the $F_{r-1,(t-1)(r-1)}$ distribution.

This is only approximate, because blocking represents a "restriction on randomization" that may affect the distribution of the F-ratio.

A large value of F_0 implies that blocking was helpful in improving the precision of the comparison of treatment means.

We mentioned earlier that due to restricted randomization, there is no valid test for the block effect form an RCBD experiment.

But nothing stops you from doing an approximate F-test by comparing the ratio of MS Block over MSE to a corresponding F distribution. This is only approximate, because blocking represents a "restriction on randomization" that may affect the distribution of the F-ratio.

A large value of F statistic value implies that blocking was helpful in improving the precision of the comparison of treatment means.

Example.

Example. Wheat data (nitrogen content as a function of 6 fertilization methods tested in 4 blocks; see p. 269 and data set `wheat` on Blackboard). Fit as a 2-way ANOVA without interaction: `lm(nitro ~ fert.f + block.f)` .

	Df	SS	Mean Sq	F	Pr(F)
fert method	5	201.3	40.26	5.59	0.0042
block	3	197.0	65.67	9.12	0.0011
error	15	108.0	7.20		

We have convincing evidence that mean nitrogen content varies among fertilization methods, and it appears that there is a strong block effect as well.

Now let's look at an example.

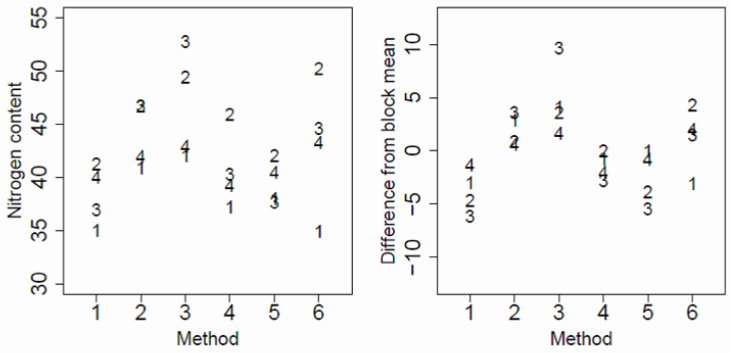
A researcher studied the effects of 6 fertilization methods on nitrogen content of wheat using an RCBD experiment. There are 4 blocks based on the irrigation gradient.

For the F-test in the ANOVA table, we see that there is convincing evidence that the mean nitrogen content varies among fertilization methods, and it appears that there is strong block effect as well.

Note, as we discussed earlier, the F-test for the block effect is only approximate since the assumptions for the F-test may not hold.

How blocking can help improve precision

The following figure shows, informally, how removing the block-to-block variation in the response gives us a more powerful assessment of the treatment effects. (The plotted numbers are the blocks.)



The figure shows, informally, how removing the block-to-block variation in the response gives us a more powerful assessment of the treatment effects.

The numbers in the plots indicate blocks.

On the left, we see the response values for the six treatments in four blocks.

On the right, we see the response values subtracting the block means.

We see that after we have subtracted the block means from the response values, the difference among methods become more obvious.

Mean contrasts

Compare the mean nitrogen content in method 3 vs. method 6, with and without adjusting for block. Given: $\overline{y}_{3\cdot} = 46.77$, $\overline{y}_{6\cdot} = 43.225$, $r = 4$.

The point estimation is straightforward:

$$\overline{y}_{3\cdot} - \overline{y}_{6\cdot} = 46.77 - 43.225 = 3.545.$$

How to correctly estimation the standard error?

Compare the mean nitrogen content in method 3 vs. method 6.

The point estimation is straightforward: we estimate the population means with the corresponding sample means.

The key question is how to correctly estimation the standard error.

The right way

1. *The right way.* From the model NITRO = FERT + BLOCK, we get MSE = 7.201, with $(r-1)(t-1) = 3 \cdot 5 = 15$ residual degrees of freedom.

$$\text{SE}(\bar{y}_{3\bullet} - \bar{y}_{6\bullet}) = \sqrt{\text{MSE}\,(1/4 + 1/4)} = \sqrt{7.201(0.5)} = 1.897\ .$$

A 95% CI for $\mu_3 - \mu_6$ is then

$$\begin{aligned} \bar{y}_{3\bullet} - \bar{y}_{6\bullet} &\pm t_{0.025,\,15} \cdot \text{SE}(\bar{y}_{3\bullet} - \bar{y}_{6\bullet}) \\ &= (46.77 - 43.225) \pm 2.13 \cdot 1.897 \\ &= 3.545 \pm 4.041 \quad = \quad (-0.50, 7.59)\ . \end{aligned}$$

If we fit the correct RCBD model to the data, and get the MSE = 7.201 with 15 d.f. from the ANOVA table.

Then we can use the formula for standard error of a contrast to get

(SE equals 1.897)

and a 95 percent CI for the (mu 3 - mu 6) is (-0.50, 7.59).

The wrong way

2. *The wrong way.* From the model NITRO = FERT, we get MSE = 16.945, with $rt - t = 24 - 6 = 18$ residual degrees of freedom.

$$\text{SE}(\bar{y}_{3\bullet} - \bar{y}_{6\bullet}) = \sqrt{\text{MSE}\,(1/4 + 1/4)} = \sqrt{16.954(0.5)} = 2.91\ .$$

95% CI for $\mu_3 - \mu_6$:

$$\begin{aligned} \bar{y}_{3\bullet} - \bar{y}_{6\bullet} \ &\pm\ t_{0.025,\,18} \cdot \text{SE}(\bar{y}_{3\bullet} - \bar{y}_{6\bullet}) \\ &= (46.77 - 43.225) \pm 2.10 \cdot 2.91 \\ &= 3.545 \pm 6.11 \quad = \quad (-2.57, 9.66)\ , \end{aligned}$$

which is much wider than the interval from (1).

If, however, we forgot to include the block effect in the fitted model, we will get an MSE of 16.945, with 18 d.f.

The resulting CI will be much wider than the correct one.

This is because we essentially use (SSE + SS Block) as our error sum of squares.

Data analysis in R

For analyzing an RCBD experiment using R, study the R code and notes in `exmaple8.1.html`

For analyzing an RCBD experiment using R, please study the R code and notes in exmaple8.1.html on

Summary

- RCBD:
  - Why?
  - How to randomize treatment assignment?
  - Pros and Cons?
- Follow the randomization
- ANOVA table
- Data analysis:
  - How to test the treatment effect?
  - How to make inference (e.g, CI) on mean and contrasts?
- Data analysis using R

In this lecture, we discussed the design and analysis of an RCBD experiment.

We answered why do we need an RCBD, how to generate an randomization plan for the RCBD experiment, what are the advantages and disadvantages of the RCBD design.

We discussed the general idea that the randomization scheme will determine the random terms in a model equation.

As usual, the ANOVA table summarizes many of the key information.

We should be able perform hypothesis test for the treatment effects and make inference on the mean or contrast parameters.

Finally, you should study the R code and notes in example 8.1.html.