LGN 5822 - Biometrical Genetics

L08 – Randomized Complete Block Design

Michele Jorge Silva Siqueira

2023

Introduction

"We shall need to judge of the magnitude of the differences introduced by testing our treatments upon the **different plots** by the discrepancies between the performances of the same treatment in **different blocks**"

R.A Fisher, 1935

The Design of Experiments, Section 26

Introduction

- If the researcher finds that any factor disturbs the homogeneity of the experimental units or environmental conditions

Introduction

Controlled environmental conditions

Introduction

- How can we assure, then, that an observed agronomic difference is the result of a specific treatment, rather than the result of the experimental units to which it was allocated?

- In other words, how do we prevent our treatment results from being confounded with our experimental units?

Introduction

The **heterogeneity** of experimental units presents a problem:

- Difficulty recognizing differences between experimental units!

- This could lead us to conclude that the differences in our variables are the result of the treatments applied, when in fact they were caused by the pre existing condition of the experimental units

Introduction

Randomized Complete Block Design (RCBD)

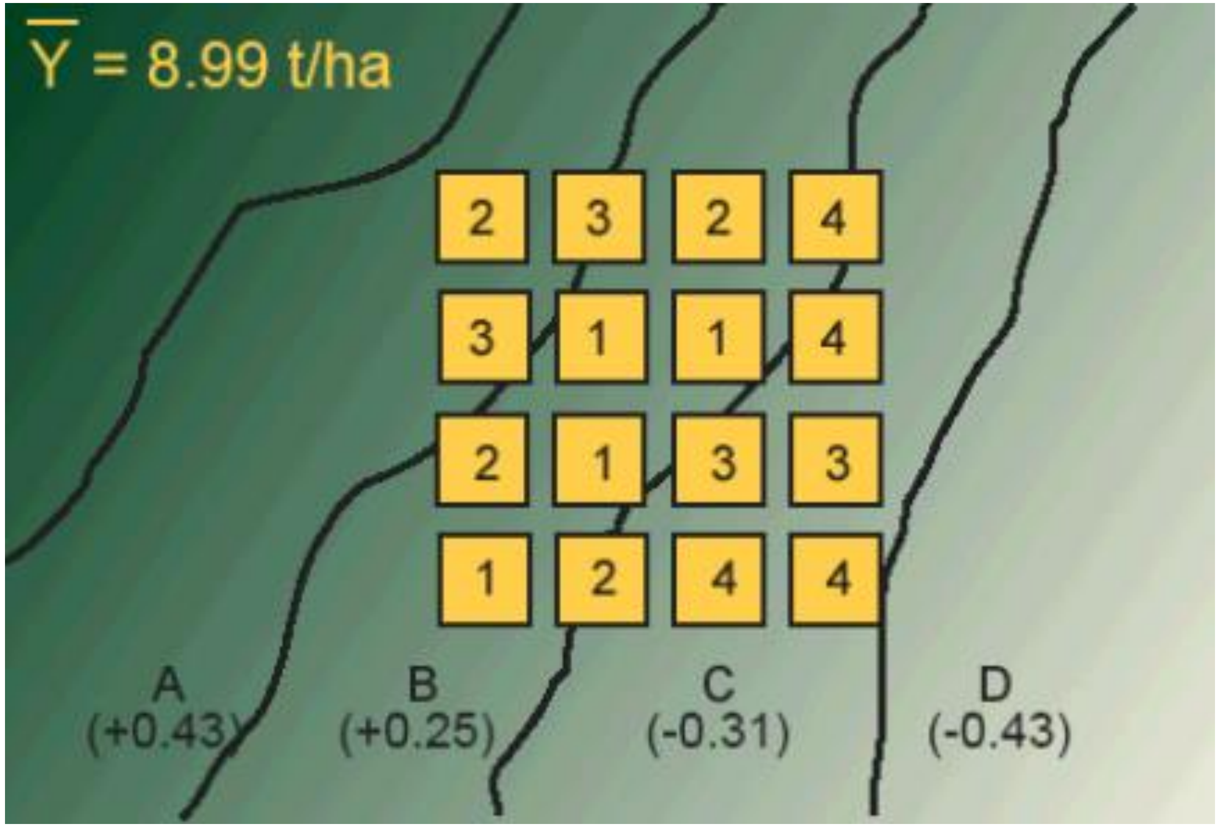
- How to control undesirable variation between experimental units?

Minimize Field Differences

Introduction

Randomized Complete Block Design (RCBD)

- The first step in using the RCBD is to recognize the source(s) of potential heterogeneity among plots (experimental units)



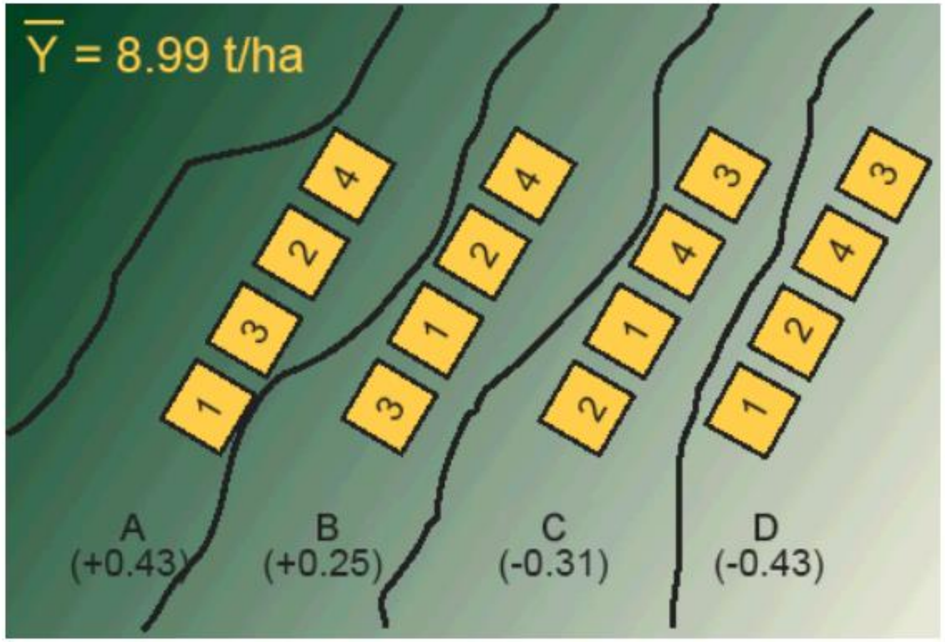
"Production gradient"

Introduction

How to block

- In blocking, we generally place an equal-sized block on every map unit. Each block, in this case, contains four experimental units (plots)

- Each treatment is applied to one experimental unit within the block



Blocking of randomized treatments to account for a known yield potential gradient

Design Characterization

Randomized Complete Block Design (RCBD)

Design Characterization

Randomized Complete Block Design (RCBD)

- RCBD is the standard design for agricultural experiments where similar experimental units are grouped into blocks or replicates

- A **RCBD** is the most basic blocking design

RCDB uses the basic principles of repetition, randomization and local control

Design Characterization

Randomized Complete Block Design (RCBD)

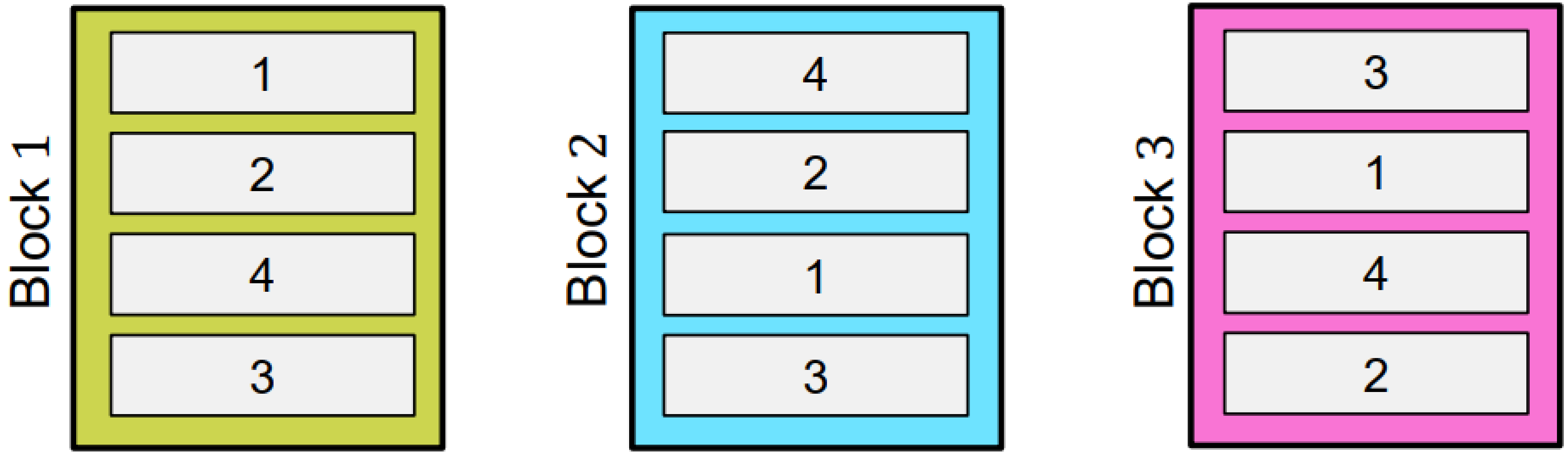
- It is used to control for variation in an experiment not predicted by the researcher (random error)

- e.g. variation in fertility or drainage differences in a field

Design Characterization

Randomized Complete Block Design (RCBD)

- Assume we have r blocks containing g units each



Here, $r = 3$ blocks with $g = 4$ units

In every of the r blocks we randomly assign the g treatments to the g units, **independently** of the other blocks

Design Characterization

Description of the Design: RCBD

- Probably the most used and useful of the experimental designs
- Takes advantage of grouping similar experimental units into blocks or replicates
- The blocks of experimental units should be as uniform as possible
- The purpose of grouping experimental units is to have the units in a block as **uniform as possible** so that the observed differences between treatments will be largely due to **“true” differences between treatments**
- Each block gets its “own” randomization

Design Characterization

Description of the Design: RCBD

- We call a blocking design **complete** if every treatment is used in every block
- Each treatment must appear at least once per replicate
- In general we observe every treatment (only) **once** in every block, hence we have a total of $\boldsymbol{r}$ **(the number of blocks)** observations per treatment

Design Characterization

Example

- Researchers wanted to evaluate the effect of several different fertilization (nitrogen) timing schedules on leaf tissue of maize

- Treatment: Six different nitrogen application timing and rate schedules (including a control treatment of no nitrogen)

- Response: Leaf tissue nitrogen amount

Design Characterization

Example

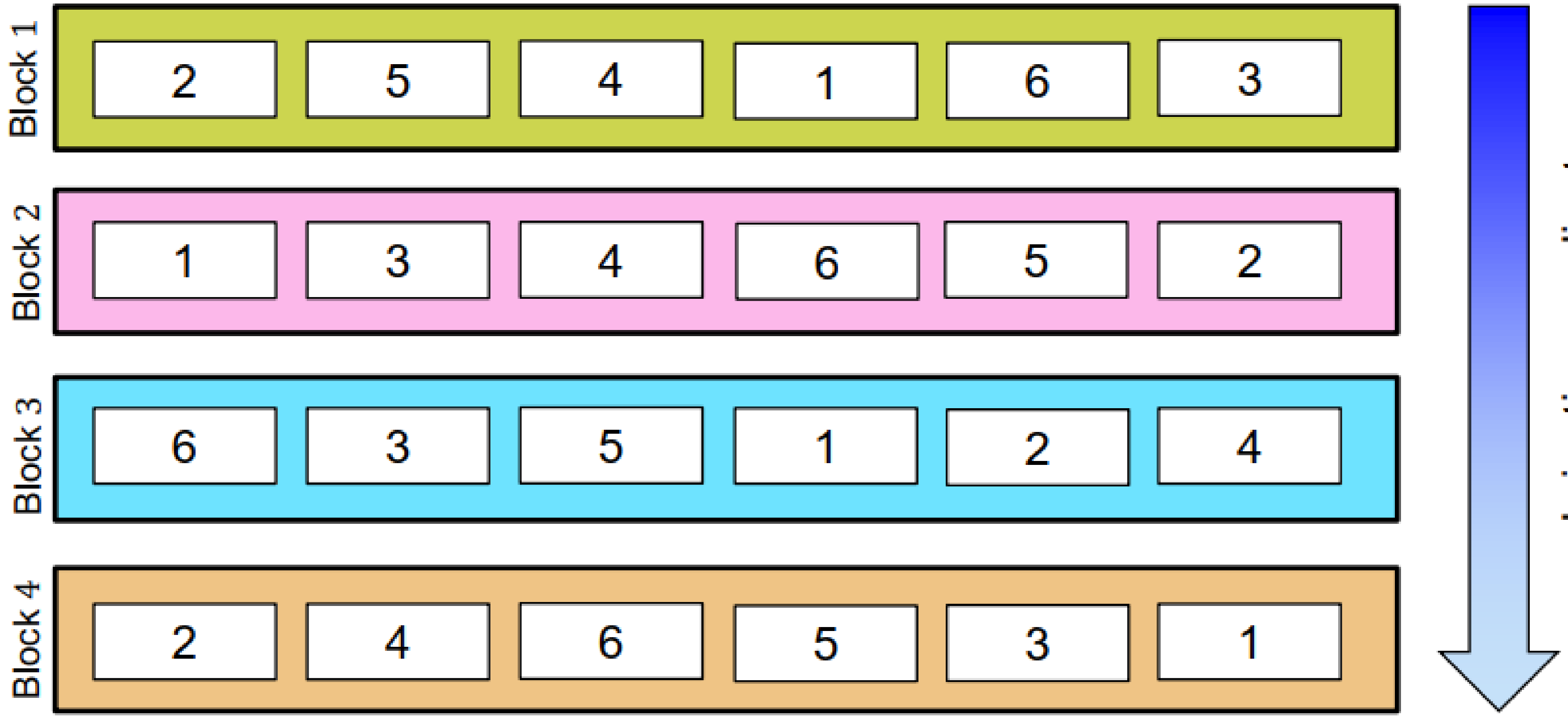
- Experiment design: irrigated field with a **water gradient** along one direction, see next slide

- We already know:
  Available moisture in the soil (humidity) will have an influence on the response

Design Characterization

Example

- Layout of Experimental Design



Design Characterization

Example

- Layout of Experimental Design

  - The differences in plant responses caused by the **water gradient** will be associated with **blocks**

  - We also say: we **control for the water gradient**

Design Characterization

Advantages of the RCBD

- Generally more precise than the CRD
- No restriction on the number of treatments or replicates
- Some treatments may be replicated more times than others
- Missing plots are easily estimated

Design Characterization

Disadvantages of the RCBD

- In experiments with a large number of treatments, a RCDB may become inefficient because each block must contain all treatments. This can significantly increase the size of the experiment, require more resources, and make the analysis more complex

- If the variability between blocks is too large, this can negatively affect the ability of the RCBD to detect differences between treatments

Design Characterization

Data table

- Consider an experiment installed at the RCDB with i treatments and j replicates (blocks)

Blocks	Treatments				Total
	1	2	...	I	
1	Y_{11}	Y_{21}	...	Y_{I1}	B_1
2	Y_{12}	Y_{22}	...	Y_{I2}	B_2
...	...	...	...	...	...
J	Y_{1J}	Y_{2J}	...	Y_{IJ}	B_J
Total	T_1	T_2	...	T_I	G

Design Characterization

Data table

- Number of experimental units: $N = I \times J$
- Total for treatment i: $T_i = \sum_{j=1}^{j} Y_{ij} = Y_{i.}$
- Total for block j: $B_i = \sum_{i=1}^{i} Y_{ij} = Y_{.j}$
- Mean for treatment i: $\widehat{m_i} = \frac{T_i}{J}$
- Mean for block j: $\widehat{m_j} = \frac{B_j}{I}$
- General mean of the experiment: $\widehat{m} = \frac{G}{IJ}$

Design Characterization

Statistical model

Model

- Data from the RCBD can be described with the following model:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where μ is the intercept (an overall mean), α_i is the effect of treatment i, β_j is the effect of block j and ε_{ij} is the associated random error term

Design Characterization

Anova table

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-stat
Treatment	*SST*	$t-1$	*MST*	$F^*=MST/MSE$
Block	*SSB*	$b-1$	*MSB*	
Error	*SSE*	$(t-1)(b-1)$	*MSE*	
Total	*TSS*	$tb-1$		

- Remember about the sum of squares!

- Typically, we are **not** making inference about blocks (we already know that blocks are different!)

- It is interesting to evaluate whether there is a difference between the treatments, which can be verified using the F test for treatments

Design Characterization

Hypotheses

- Treatments

$$H_0:\ t_i = 0, i = 1,2, \dots I$$

$$H_1: at\ least\ one\ value\ of\ t_k \neq 0$$

- Blocks

$$H_0:\ b_j = 0, j = 1,2, \dots J$$

$$H_1: at\ least\ one\ value\ of\ b_k \neq 0$$

In cases where the variation between blocks is doubtful, the researcher can perform the F test for blocks, to serve as guidance for setting up future experiments!

Let´s Practice 01!

#Load the "agridat" package

#Choose the dataset "federer.tobacco" from the agridat package

- Data: RCB of tobacco, height plants exposed to radiation: 56 observations
- Evaluation of plant growth in different doses of radiation

#Perform an analysis of variance (ANOVA) for the RCDB

#Tukey test for multiple comparisons of treatments

Let´s Practice!


```
> head(federer.tobacco)
  row block dose height
1   1     1 2500 1299.2
2   1     2  250 1369.2
3   1     3    0 1169.5
4   1     4 2500 1219.1
5   1     5 2500 1120.0
6   1     6 5000 1031.5
> 
```

Let´s Practice!


```
> summary(model_anova)

Error: block
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  1 278088  278088

Error: Within
          Df  Sum Sq Mean Sq F value Pr(>F)
dose       1  185289  185289   6.688 0.0125 *
Residuals 53 1468400   27706
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

References

- Chapter 9 - Analysis of Variance II: The Randomized Complete Block Design[1] (for a more classical view)

- Chapter 3 – Complete Block Designs[2]

1. Steel, R. G. & Torrie, J. H. Principles and Procedures of Statistics: A Biometrical Approach. 2nd Edition. (1980).

2. Casella, G. Statistical Design. (2008).