

# Blocking in Experimental Designs

## Lecture 23

---

Alexandra Chronopoulou



### **COLLEGE OF LIBERAL ARTS & SCIENCES**

Department of Statistics  
101 Illini Hall, MC-374  
725 S. Wright St.  
Champaign, IL 61820-5710

© Alexandra Chronopoulou. Do not distribute without permission of the author.

## Complete Randomized Block Design

---

## Complete Randomized Block Design

Experimental units are first sorted into *homogeneous groups (blocks)* and all treatments and combinations are then assigned at random to experimental units *within the blocks*.

## Criteria for Blocking

1. Characteristics associated with the **unit**.
2. Characteristics associated with the **experimental setting**.

- Consider an experiment to compare the *yield* of **8** different varieties of oats.
- The growing area was heterogeneous and so was grouped into **5** blocks.
- Each variety was sown once within each block and the *yield* in grams per 16ft row was recorded.
- This is the oats data set in the *faraway* library.

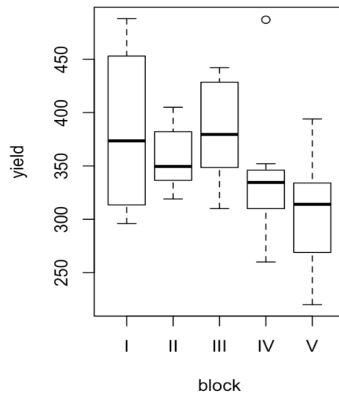
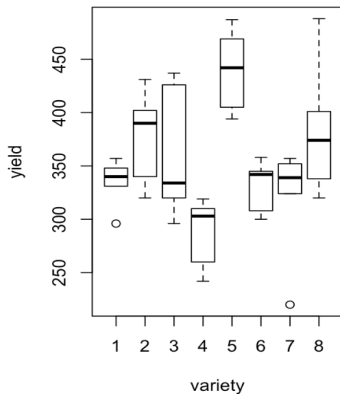
- Treatment Factor ‘variety’: 8 levels
- Block Factor ‘block’: 5 levels
- Observe that we have *only one* observation per treatment-block combination.

```
xtabs(yield~variety + block, oatvar)
```

```
##          block
## variety  I  II III  IV  V
##      1 296 357 340 331 348
##      2 402 390 431 340 320
##      3 437 334 426 320 296
##      4 303 319 310 260 242
##      5 469 405 442 487 394
##      6 345 342 358 300 308
##      7 324 339 357 352 220
##      8 488 374 401 338 320
```

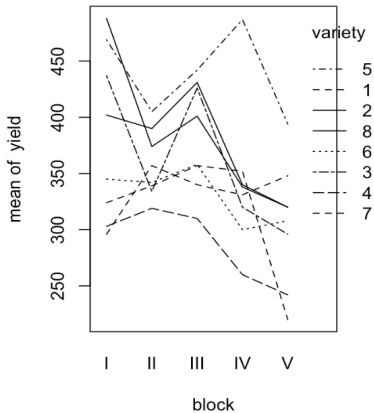
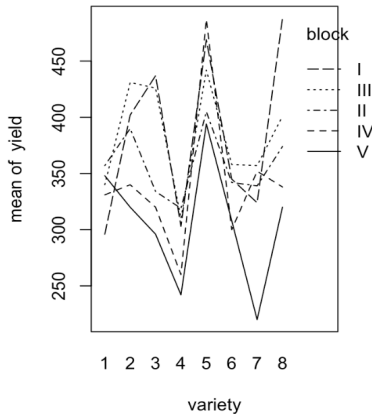
# Oats Experiment

Side-by-side box-plots for each factor:



# Oats Experiment

Interaction plots for each factor:



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

where

- $\mu_{..}$  is a constant
- $\tau_i$  are the treatment effects
- $\beta_j$  are the block effects
- $\varepsilon_{ij}$  are independent  $\mathcal{N}(0, \sigma^2)$
- $i = 1, \dots, n_b$  (total number of blocks),  $j = 1, \dots, r$  (total number of treatments)



## Remarks

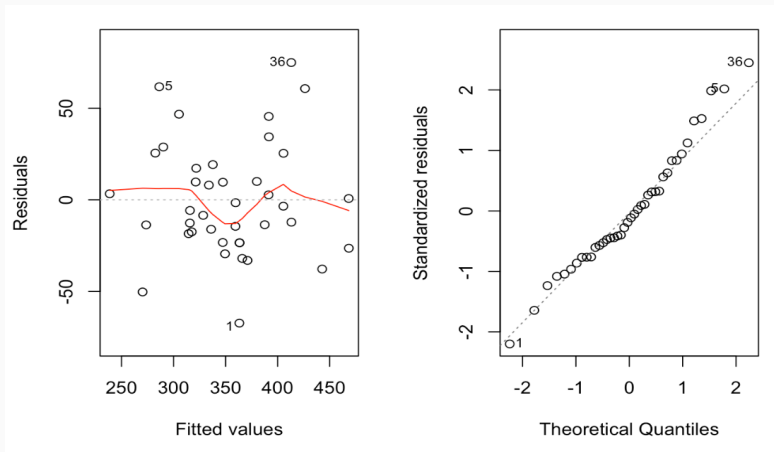
- $y_{ij}$  is the response for the  $j$ th treatment in the  $i$ th block.
- There is a single observation per block. This implies that we have a limited ability to detect an interaction between treatment and block. So, we are working with the **additive** model.
- We can check for treatment and block main effects, but blocking is a feature of the design which means that if insignificant, we cannot gain the degrees of freedom.

```
anova(lm(yield~ block + variety, oatvar))
```

```
## Analysis of Variance Table
##
## Response: yield
##      Df Sum Sq Mean Sq F value    Pr(>F)
## block    4  33396   8348.9    6.2449 0.001008 **
## variety  7  77524  11074.8    8.2839 1.804e-05 ***
## Residuals 28  37433   1336.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Based on the sequential ANOVA table, we observe that *both* effects are statistically significant.
- In this example, we have a *balanced* (or *orthogonal*) design, so changing the order of the terms in ANOVA will not affect our results.

# Oats Experiment



The diagnostics of the fitted model look good.

- We can use *Tukey's additivity test* to check whether interactions are significant.

```
varcoefs <- c(0, coef(lmodcrbd)[6:12])
blockcoefs <- c(0,coef(lmodcrbd)[2:5])
ab <- rep(varcoefs, each=5)*rep(blockcoefs, 8)
h <- update(lmodcrbd, .-.+ab)
anova(h)
```

```
## Analysis of Variance Table
##
## Response: yield
##           Df Sum Sq Mean Sq F value    Pr(>F)
## block      4  33396   8348.9    6.0563 0.001299 **
## variety    7  77524  11074.8    8.0337 2.787e-05 ***
## ab         1    213    212.9    0.1544 0.697428
## Residuals 27  37220   1378.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The  $p$ -value is 0.6974, so we fail to reject the null hypothesis of no interaction.

- If we want to assess the relative advantage of performing a RCBD over a CRD, we compare the  $\hat{\sigma}^2$  for both designs. Specifically, define the **relative efficiency** as

$$\hat{\sigma}_{CRD}^2 / \hat{\sigma}_{RCBD}^2$$

- In **R**, we compute it by fitting the model with and without the blocking effect:

```
lmcrd <- lm(yield-variety, oatvar)
(summary(lmcrd)$sig / summary(lmodcrbd)$sig)^2
```

```
## [1] 1.655617
```

This implies that a CRD would require 66% more observations to obtain the same level of precision as an RCBD.

## Latin Squares

---

## Latin Squares Example

- $A$ ,  $B$ ,  $C$ : 3 treatments
- Day of week (Monday, Wednesday, Friday): Blocking Variable
- Operator ID: 1, 2, 3: Blocking Variable

Day	Operator		
	1	2	3
Monday	$B$	$A$	$C$
Wednesday	$A$	$C$	$B$
Friday	$C$	$B$	$A$

\* Each operator runs each treatment, and all treatments are run on each day

- There are  $r$  treatments.
- There are 2 blocking variables, each containing  $r$  classes.
- Each row and each column in the design square contains all treatments.
- Each treatment is assigned to each block only once.



- Advantages

- Reduces more experimental error than with 1 blocking factor.
- Small scale studies can isolate important treatment features.
- Repeated measures designs can remove order effects.effects.

- Disadvantages

- Each blocking factor must have  $r$  levels.
- No interactions among factors.
- With small  $r$ , we have very few error degrees of freedom.
- Complex Randomization.

- Determine  $r$ , the number of treatments, row blocks, and column blocks.
- Select a *Standard Latin Square* (from tables or with software).
- Use Capital Letters to represent treatments (A,B,C,...) and randomly assign treatments to labels.
- Randomly assign Row Block levels to Square Rows.
- Randomly assign Column Block levels to Square Columns.

- In **R** we use the package `magic` to generate such squares.

### A 4×4 Latin Square

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + e_{ijk}$$

where

- $\mu$  is a constant
- $\tau_i$  treatment effect (latin letter)
- $\beta_j$  (column) blocking effect
- $\gamma_k$  (row) blocking effect
- $e_{ijk}$  are independent  $\mathcal{N}(0, \sigma^2)$
- $i, k, j = 1, \dots, r$

## Abrasion Example

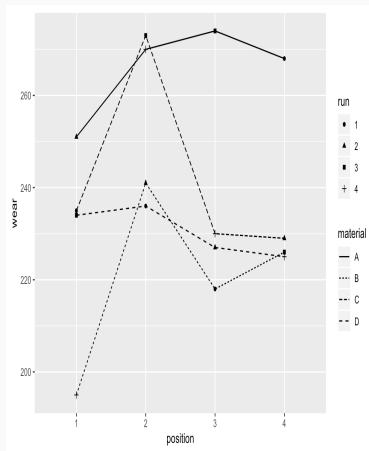
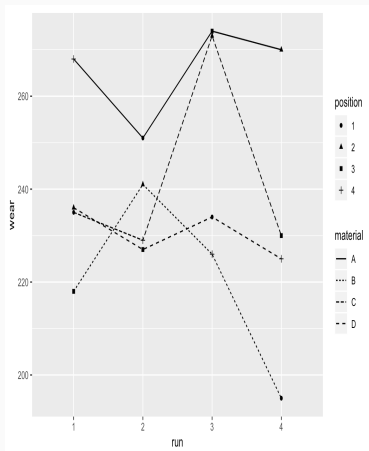
- Four materials were fed into a wear testing machine and the amount of wear recorded.
- The composite data frame has 16 rows and 4 columns and is the abrasion data set in the *faraway* library.
- The following Latin Square design was used:

```
matrix(abrasion$material, 4, 4, byrow=TRUE)
```

```
##      [,1] [,2] [,3] [,4]  
## [1,] "C"  "D"  "B"  "A"  
## [2,] "A"  "B"  "D"  "C"  
## [3,] "D"  "C"  "A"  "B"  
## [4,] "B"  "A"  "C"  "D"
```

# Abrasion Example

The interaction plots for each factor are:



We fit a Latin Square model and test each variable *relative to the full model*:

```
lmodls <- lm(wear ~ material+run+position, abrasion)
drop1(lmodls, test="F")

## Single term deletions
##
## Model:
## wear ~ material + run + position
##           Df Sum of Sq    RSS      AIC F value    Pr(>F)
## <none>                 367.5   70.146
## material  3    4621.5 4989.0 105.878 25.1510 0.0008498 ***
## run       3     986.5 1354.0  85.012  5.3687 0.0390130 *
## position  3    1468.5 1836.0  89.884  7.9918 0.0161685 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The variables are statistically significant.

- We want to look further in order to understand differences among treatment levels.

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  254.750      6.187  41.174 1.37e-08 ***
## materialB    -45.750      5.534  -8.267 0.000169 ***
## materialC    -24.000      5.534  -4.337 0.004892 **
## materialD    -35.250      5.534  -6.370 0.000703 ***
## run2         -2.250      5.534  -0.407 0.698423
## run3         12.500      5.534   2.259 0.064657 .
## run4         -9.250      5.534  -1.671 0.145658
## position2     26.250      5.534   4.743 0.003180 **
## position3      8.500      5.534   1.536 0.175454
## position4      8.250      5.534   1.491 0.186608
##
```

- The model coefficients show that material B looks best (in terms of wear) followed by material C, but we cannot say whether the difference is significant or not.



- We perform *Tukey* pairwise comparisons. The width of the Tukey interval computes as 19.143 and a table of the material differences is:

```
scoefs <- c(0, coef(lmodls)[2:4])  
outer(scoefs, scoefs, "-")
```

```
##           materialB materialC materialD  
##           0.00      45.75      24.00      35.25  
## materialB -45.75       0.00     -21.75     -10.50  
## materialC -24.00      21.75       0.00      11.25  
## materialD -35.25      10.50     -11.25       0.00
```

- This indicates that (B, D) and (D, C) differences are not significant at the 5% level, while all the other differences are.

- We compare the relative efficiency of the Latin Square design over the CRD:

```
lmodr <- lm(wear-material, abrasion)
(summary(lmodr)$sig/summary(lmodls)$sig)^2
```

```
## [1] 3.840136
```

- We conclude that the Latin square is 3.84 times more efficient than the CRD.
- We can do similar comparisons with other designs that contain blocks.

## Balanced Incomplete Block Design

---

## Example: Rabbit Weight Gain

- A nutritionist studied the effects of six diets, on weight gain of domestic rabbits.
- From past experience with sizes of litters, it was felt that only 3 uniform rabbits could be selected from each available litter.
- There were ten litters available forming blocks of size three.

- What design would be appropriate if 6 uniform rabbits could be selected?
- In this experiment, only 3 uniform rabbits were available.
- With this restriction, the three rabbits represent a *single incomplete block*.

The data we obtained are as follows:

```
xtabs(gain~treat+block, rabbit)
```

```
##      block
## treat  b1 b10 b2  b3  b4  b5  b6  b7  b8  b9
##   a  0.0 37.3 40.1 0.0 44.9 0.0 0.0 45.2 44.0 0.0
##   b 32.6 0.0 38.1 0.0 0.0 0.0 37.3 40.6 0.0 30.6
##   c 35.2 0.0 40.9 34.6 43.9 40.9 0.0 0.0 0.0 0.0
##   d 0.0 42.3 0.0 37.5 0.0 37.3 0.0 37.9 0.0 27.5
##   e 0.0 0.0 0.0 0.0 40.8 32.0 40.5 0.0 38.5 20.6
##   f 42.2 41.7 0.0 34.3 0.0 0.0 42.8 0.0 51.9 0.0
```

The zero values correspond to **no observation**.

- [→] Why Balanced?  
Each pair of treatments occur together  $\lambda$  times.
- [→] Why Incomplete?  
Cannot fit all treatments in each block.

## Notation

- $t$  treatments
- $b$  blocks
- $k$  treatments per block (block size)
- $r$  times each treatment occurs
- $N = t \cdot r = b \cdot k$  observations in total

- Treatment  $i$  occurs in  $r$  blocks.
- To have balance, each other treatment is equally likely to be treatment  $i$  in a block.
- Since there are  $k - 1$  other units in a block and  $t - 1$  other treatments, the number of times each pair of treatments appears in the same block is

$$\lambda = \frac{r(k-1)}{t-1},$$

where  $\lambda$  is an integer.



—  $a = 3, b = 3, k = 2 \rightarrow r = 2, \lambda = 1$

Block

1	2	3
A	B	A
B	C	C

—  $a = 4, k = 2, b = 6 \rightarrow r = 3, \lambda = 1$

Block

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

- A BIBD enables us to run an experiment when the size of the available blocks of experimental units is smaller than the number of treatments.
- Estimates of treatment effects have equal precision and expressions for the variances of the estimated cell means and of contrasts of treatment means or effects are relatively simple.
- The presence of balance permits the use of Scheffé and Tukey procedures for the analysis of treatment effects.

- BIBD exist only for certain combinations of numbers of treatments, block sizes, and numbers of blocks.
- The assumption that there are no interactions between the blocking variable and the treatments is restrictive.
- The analysis of a BIBD is more complex than that of a RCBD.

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

- $\mu$  constant
- $\tau_i$  treatment effects
- $\beta_j$  the block effects
- $e_{ij}$  independent  $N(0, \sigma^2)$

## Remarks

- Not all  $y_{ij}$  exist, because of incompleteness.
- Non-orthogonality of treatments and blocks.

- We fit the model and test for the significance of the results.
- We use the `drop1` function instead of `anova`, because of the lack of orthogonality due to the incompleteness of the design.

```
lmodb1d <- lm(gain~block+treat, rabbit)
drop1(lmodb1d, test="F")
```

```
## Single term deletions
##
## Model:
## gain ~ block + treat
##           Df Sum of Sq    RSS   AIC F value    Pr(>F)
## <none>                 150.77  78.437
## block    9     595.74  746.51 108.426  6.5854 0.0007602 ***
## treat    5     158.73  309.50  90.013  3.1583 0.0381655 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## We look at the model coefficients

```
summary(lmodbibd)
```

```
##
## Call:
## lm(formula = gain ~ block + treat, data = rabbit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9583 -1.6146 -0.6083  1.9396  4.3028
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  36.01389    2.58863   13.912 5.59e-10 ***
## blockb10     3.29722    2.79604    1.179  0.25667
## blockb2      4.13333    2.69433    1.534  0.14583
## blockb3     -1.80278    2.69433   -0.669  0.51360
## blockb4      8.79444    2.79604    3.145  0.00667 **
## blockb5      2.30556    2.79604    0.825  0.42253
## blockb6      5.40833    2.69433    2.007  0.06309 .
## blockb7      5.77778    2.79604    2.066  0.05651 .
## blockb8      9.42778    2.79604    3.372  0.00419 **
## blockb9     -7.48056    2.79604   -2.675  0.01729 *
## treatb     -1.74167    2.24182   -0.777  0.44930
## treatc      0.40000    2.24182    0.178  0.86078
## treatd      0.06667    2.24182    0.030  0.97667
## treate     -5.22500    2.24182   -2.331  0.03413 *
## treatf      3.30000    2.24182    1.472  0.16169
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.17 on 15 degrees of freedom
## Multiple R-squared:  0.855, Adjusted R-squared:  0.7197
## F-statistic: 6.318 on 14 and 15 DF,  p-value: 0.000518
```

and we extract the necessary information for Tukey multiple comparisons of the treatment differences.

```
tcoefs<- (c(0, coef(lmodbibd)[11:15]))  
abs(outer(tcoefs, tcoefs, "-"))>7.27
```

```
##           treatb treatc treatd treat e treatf  
##      FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  
## treatb FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  
## treatc FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  
## treatd FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  
## treat e FALSE  FALSE  FALSE  FALSE  FALSE  TRUE  
## treatf FALSE  FALSE  FALSE  FALSE  TRUE  FALSE
```

- We observe that only (Treatment e – Treatment f) is statistically significant.

- We compute the relative efficiency of the BIBD over the CRD

```
lmot <- lm(gain ~ treat, rabbit)
(summary(lmot)$sig/summary(lmodbibd)$sig)^2
```

```
## [1] 3.094508
```

and we conclude that blocking was worthwhile here.