

# Module 4F

# Supervised Machine

# Learning

Different flavors of REGRESSION and General Linear Models

# Blocking

Results in an additional variable, a block, that must be included in analysis  
Can no longer use simple one-factor ANOVA

Randomized block design

**Paired design** for > 2 treatments

Example:

Every treatment is replicated **once** within each block  
Minimize “noise”

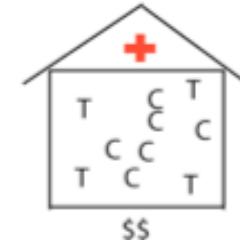
Accounting for any variation caused by blocking **can improve our treatment effect detection** (i.e., increase the power of our test)

Treatment effects are assessed by different treatments within each block so there is no interaction term

## Goals of experiments:

determine how explanatory variable (treatment) affects response variable

- Eliminate Bias
- Reduce Sampling Error
  - Blocking:



C = Control  
T = Treated



Variance among hospitals will not contribute to SE.

Only variance within hospitals will contribute to "noise"

# Main Principle of Blocking

Response = Constant + Treatment + Block

$H_0$ : Response = Constant + Block

$H_A$ : Response = Constant + Block + Treatment

- Determine significance via ANOVA table which includes a row for the **block**
- Calculates a F value for block - examines how much better fit is with the block versus without the block

# Example of Blocking:

Response = Constant + Treatment + Block

$H_0$ : Response = Constant + Block

$H_A$ : Response = Constant + Block + Treatment

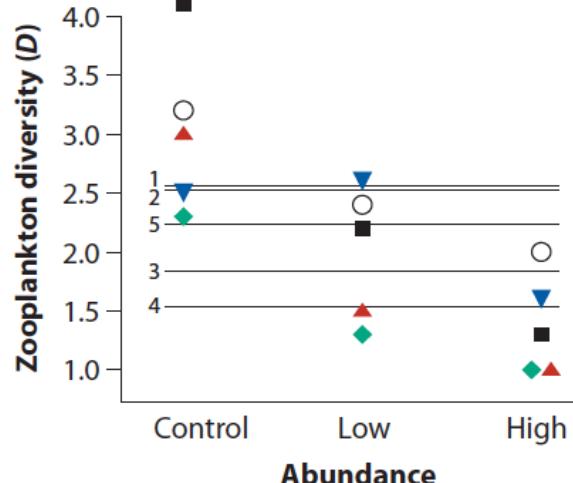
$$F = \frac{H_A - \text{Constant} + \text{Block} + \text{Treatment}}{\text{residual} + \text{location}}$$

$H_0$  Constant + Block

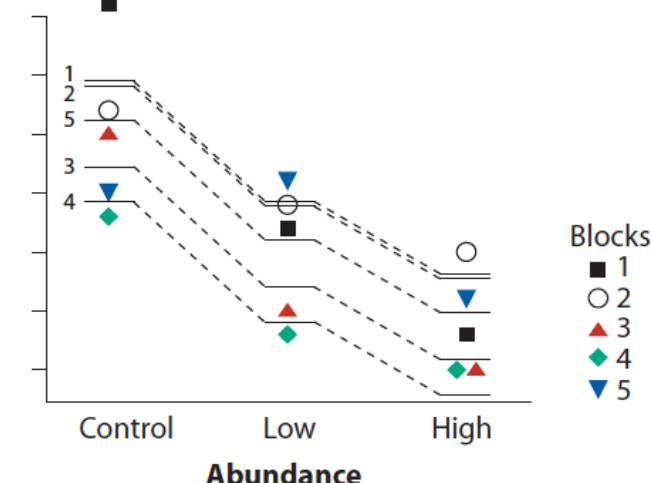
$$= \frac{\text{residual} + \text{location} + \text{fish Abundance}}{\text{residual} + \text{location}}$$

Source of variation	Sum of Squares	df	Mean Square	F	P
BLOCK	2.340	4	0.5850		
Treatment	6.8573	2	3.4287	16.37	0.001
Residual	1.6760	8	0.2095		
Total	10.8733	14			

DIVERSITY = CONSTANT + BLOCK



DIVERSITY = CONSTANT + BLOCK + ABUNDANCE



Blocks

- 1
- 2
- ▲ 3
- ◆ 4
- ▼ 5

# Example of Blocking:

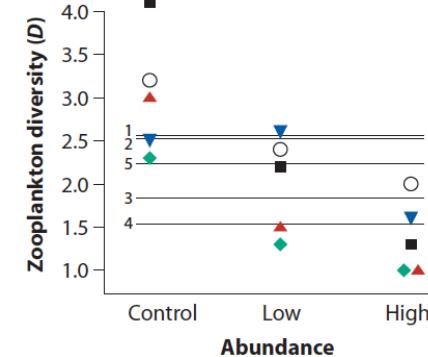
Response Full = Constant + Treatment + Block

Treatment:

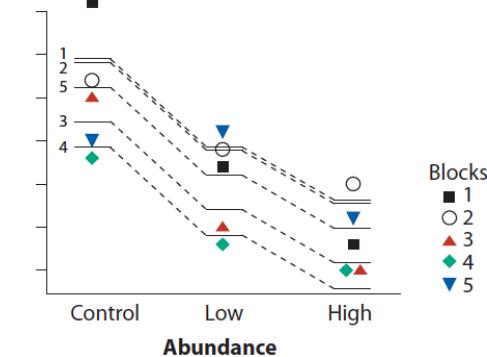
$H_0$ : Response = Constant + Block

$H_A$ : Response = Constant + Block + Treatment

DIVERSITY = CONSTANT + BLOCK



DIVERSITY = CONSTANT + BLOCK + ABUNDANCE



$$F = H_A = \text{Constant+Block+Treatment}$$

$H_0$  Constant + Block

$$= \frac{\text{residual+location+fish Abundance}}{\text{residual+location}} = \frac{MS_{\text{treatment}}}{MS_{\text{block}}} = \frac{3.43}{0.59} = 16.37$$

$F_{0.05(1),2,14} = 3.74$  so we reject the  $H_0$

Source of variation	Sum of Squares	df	Mean Square	F	P
BLOCK	2.340	4	0.5850		
Treatment	6.8573	2	3.4287	16.37	0.001
Residual	1.6760	8	0.2095		
Total	10.8733	14			

# Example of Blocking:

Treatment:

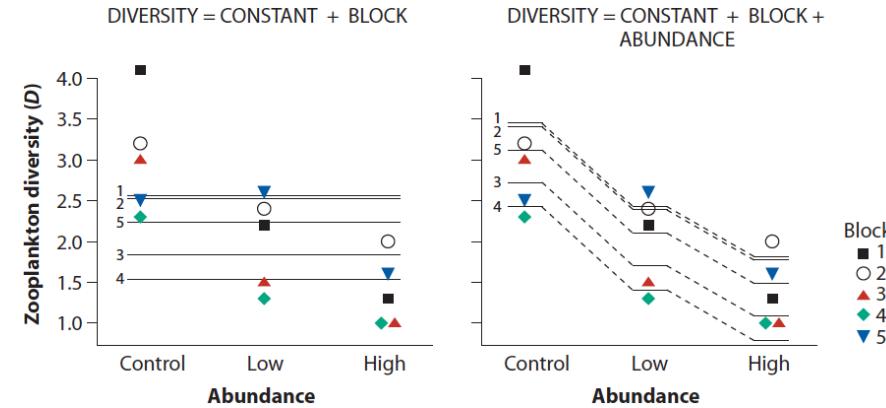
$H_0$ : Response = Constant + Block

$H_A$ : Response = Constant + Block + Treatment

$$F = \frac{H_A - \text{Constant} + \text{Block} + \text{Treatment}}{\text{Constant} + \text{Block}}$$

$$H_0 \quad \text{Constant} + \text{Block}$$

$$= \frac{\text{residual} + \text{location} + \text{fish Abundance}}{\text{residual} + \text{location}} = \frac{MS_{\text{treatment}}}{MS_{\text{block}}} = \frac{3.43}{0.59} = 16.37$$



$F_{0.05(1),2,14} = 3.74$  so we reject the  $H_0$

Block:

$$F_{\text{Block}} = H_A = \frac{\text{Residual} + \text{treatment} + \text{Block}}{\text{Residual} + \text{treatment}} = \frac{MS_{\text{block}}}{MS_{\text{residual}}} = \frac{0.5850}{0.2095} = 2.79$$

$$H_0 \quad \text{Residual} + \text{treatment} \quad MS_{\text{residual}} \quad 0.2095$$

$F_{0.05(1),4,8} = 3.84$  so we fail to reject the  $H_0$

Source of variation	Sum of Squares	df	Mean Square	F	P
BLOCK	2.340	4	0.5850		
Treatment	6.8573	2	3.4287	16.37	0.001
Residual	1.6760	8	0.2095		
Total	10.8733	14			