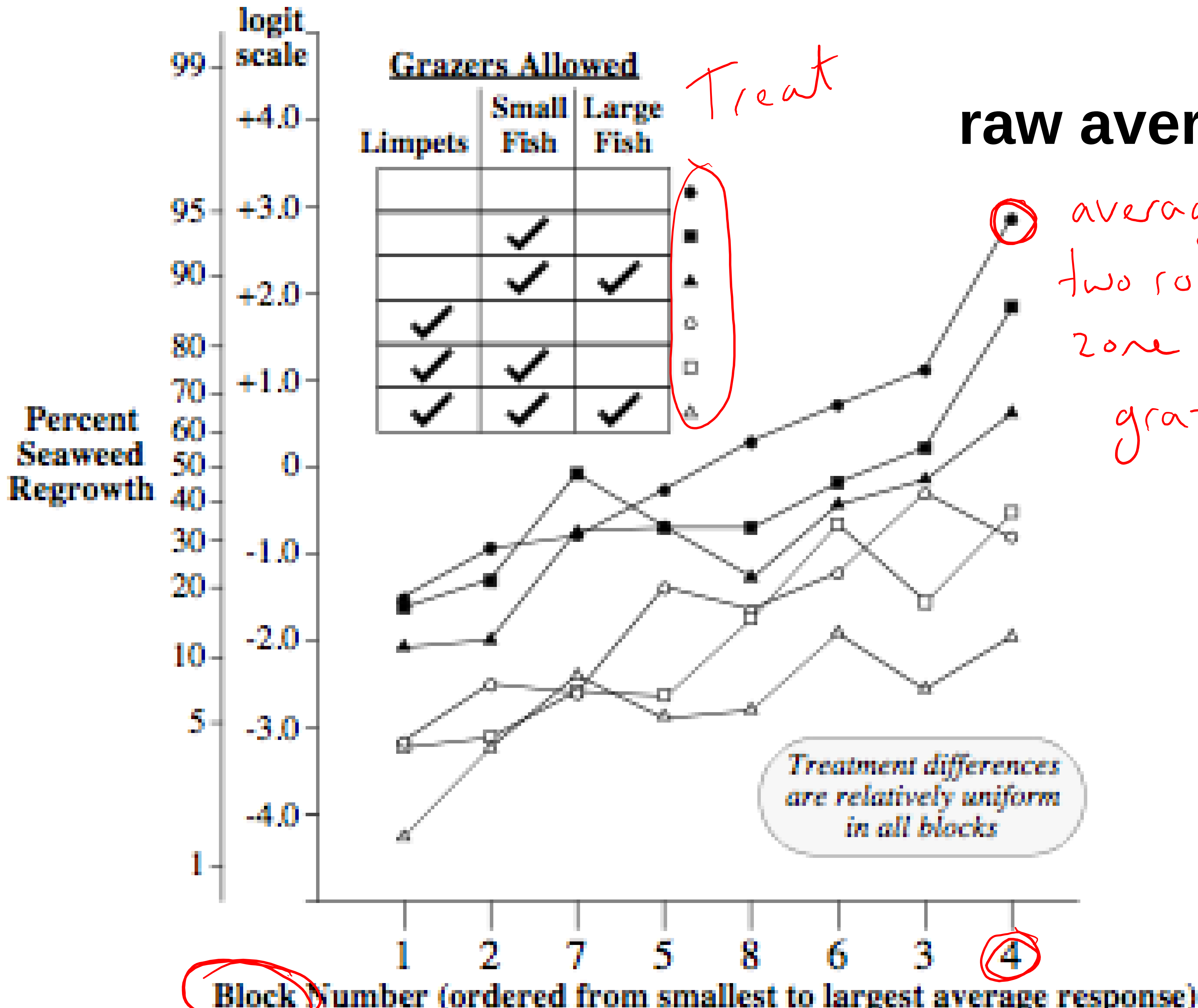
Stat 412/512

Two-way Anova Example cont.

Feb 11 2015

Charlotte Wickham

Averages of the log of the seaweed regeneration ratio versus block number, with code for treatment



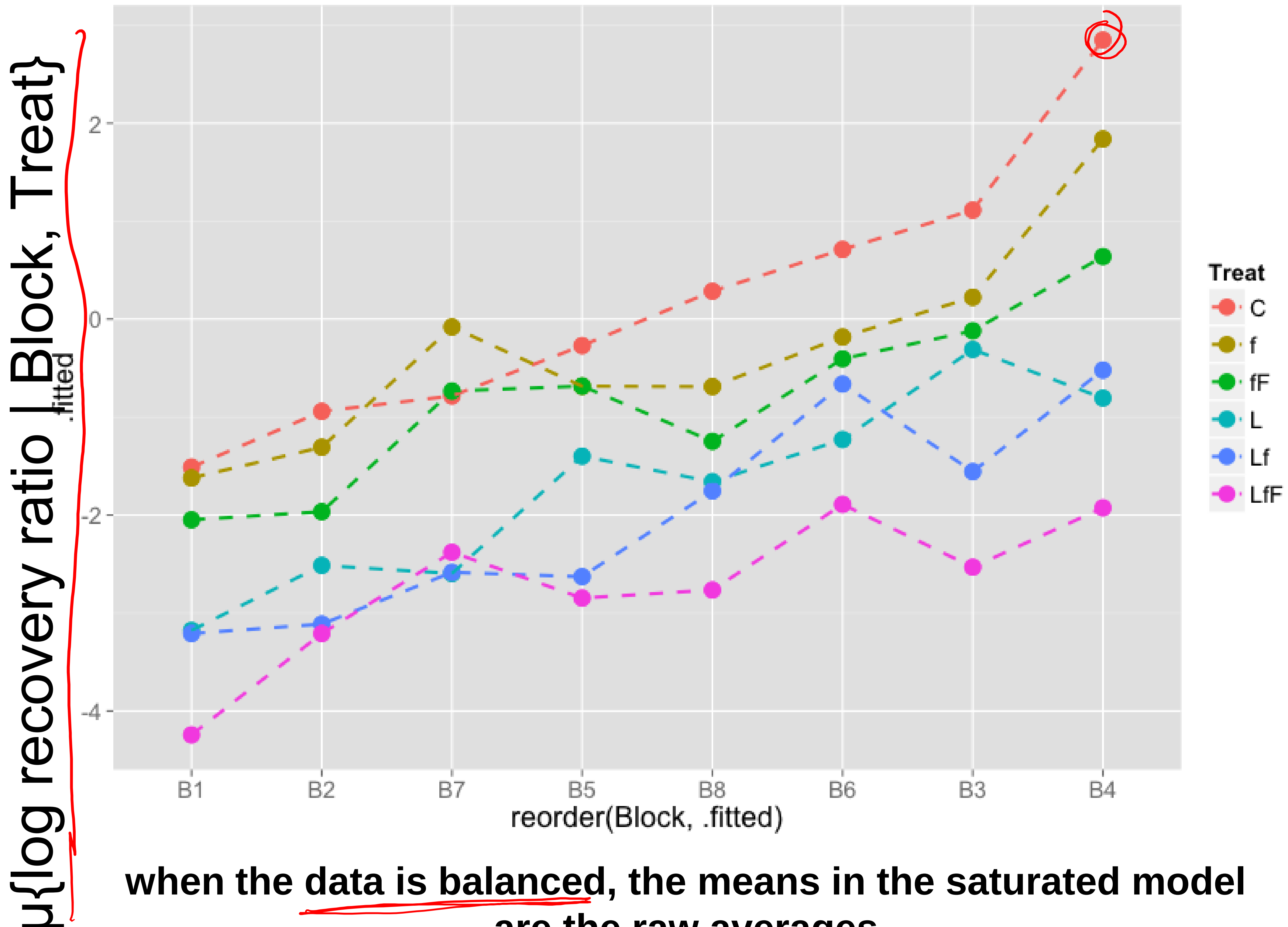
raw averages

average over two rocks in tidal zone #4, no grazers allowed

96 rocks

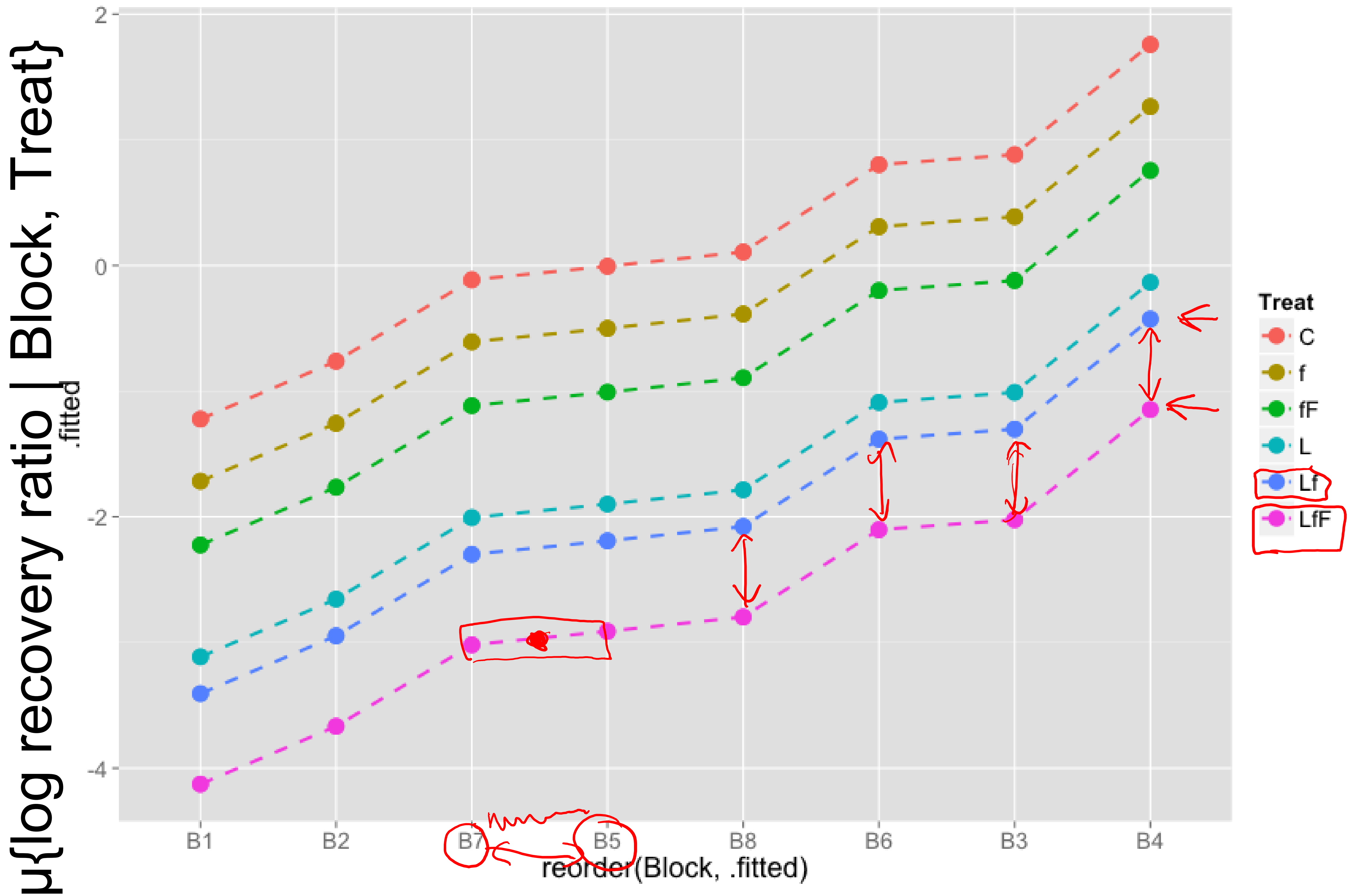
Saturated model

fitted means

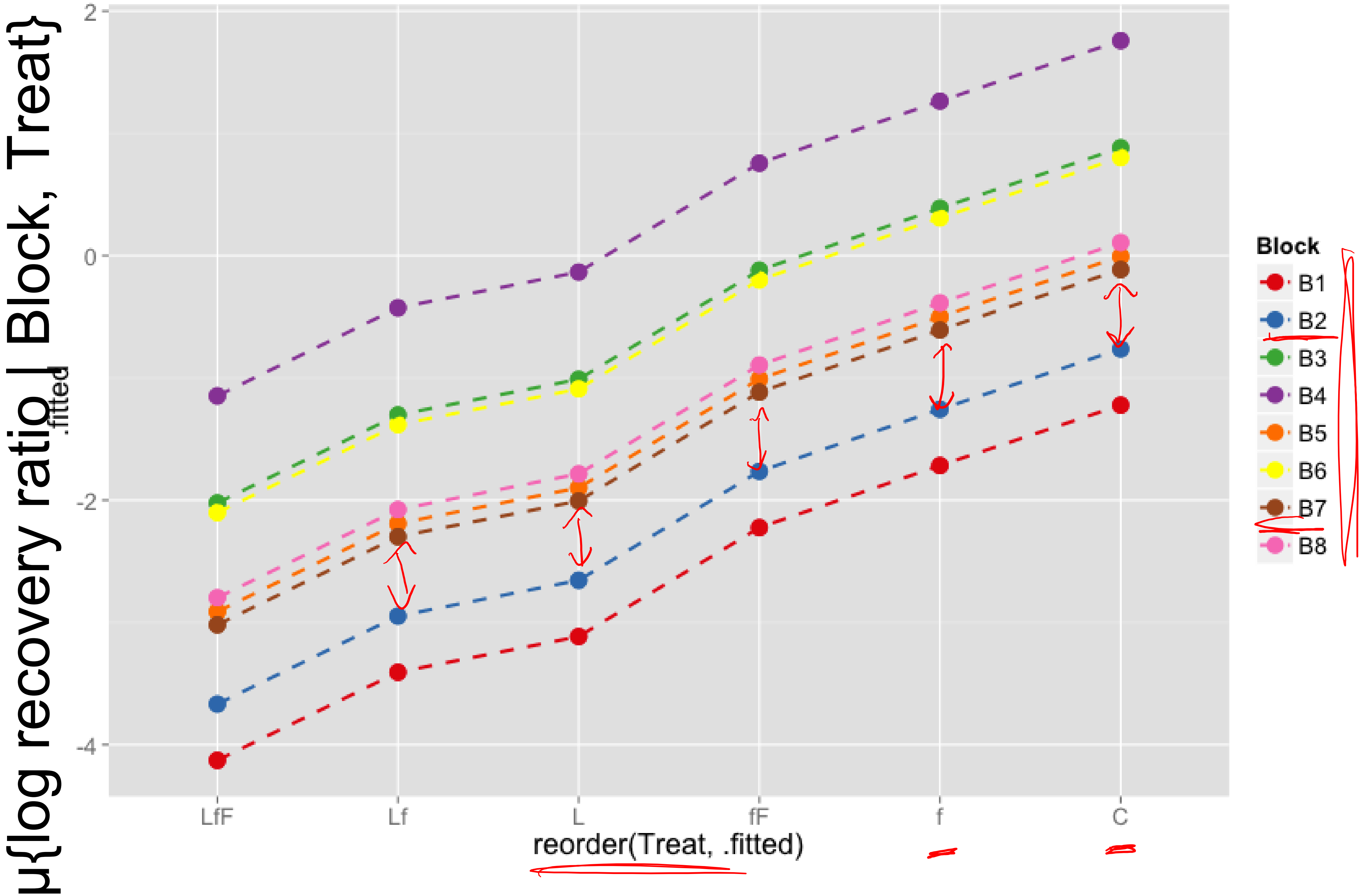


when the data is balanced, the means in the saturated model are the raw averages

Additive model
fitted means



Additive model
fitted means



Estimating effects

not of the treatments, but of the animals

Two approaches:

Using averages over cell, rows and columns.

HARD, and only relevant for balanced data

Using indicator variables and multiple regression.

A regression approach

Set up indicators:

sml = 1, small fish are present if f, Ff, Lf, LFf

big = 1, large fish are present if Ff, LFf

limp =1, limpets are present if L, Lf, LFf

Equivalent to the additive model (TREAT + BLOCK):

BLOCK + sml + big + limp + sml x limp + big x limp

sml x big : can't estimate, since big fish always present with little fish.

LfF ← → parameters in front of all terms

Lf ←

Ff ←

f sml fish allowed

L limpets allowed

$\boxed{\text{C no animals}}$ baseline

$+ \beta_9$ sml + big + limpet + sml × limpet + limpet × big

what is the baseline? sml = 0, big = 0, limpet = 0

why not? big × small & big × small × limpet by design assumed = 0

Animal Effects

Limpet effect: change in mean log recovery ratio in going from limp = 0, to limp = 1, holding other variables constant.

How much do limpets graze (holding access by other animals constant)?

My model:

$\mu\{\text{log recovery ratio} \mid \text{Block, L, f, F}\} =$

$\beta_0 + \beta_1 B2 + \ldots + \beta_8 \underline{\text{limp}} + \beta_9\,\text{sml} + \beta_{10}\,\text{big}$

Limpet effect = β_8

If there were animal interactions (e.g. limp x sml) then the effect of limpets would depend on whether small (or big) fish also had access.

```
Analysis of Variance Table

Model 1: log(Cover/(100 - Cover)) ~ Block + L + f + F
Model 2: log(Cover/(100 - Cover)) ~ Block + L + f + F + L:F + L:f
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     85 29.996
2     83 29.767  2   0.22928 0.3197 0.7273

Call:
lm(formula = log(Cover/(100 - Cover)) ~ Block + L + f + F, data = case1301)

Residuals:
     Min       1Q   Median       3Q      Max
-1.47682 -0.40585  0.03001  0.33617  1.30143

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.2545     0.2011  -6.238 1.66e-08 ***
BlockB2       0.4600     0.2425   1.897  0.06127 .
BlockB3       2.1046     0.2425   8.678 2.42e-13 ***
BlockB4       2.9807     0.2425  12.291  < 2e-16 ***
BlockB5       1.2160     0.2425   5.014 2.87e-06 ***
BlockB6       2.0251     0.2425   8.350 1.11e-12 ***
BlockB7       1.1085     0.2425   4.571 1.64e-05 ***
BlockB8       1.3300     0.2425   5.484 4.19e-07 ***
L            -1.8288     0.1213 -15.082  < 2e-16 ***
f            -0.3933     0.1485  -2.648  0.00965 **
F            -0.6140     0.1485  -4.135 8.31e-05 ***
```

no evidence for animal interactions

estimates of effects

There is no evidence that the grazing effects differ depending on microhabitat (extra SS F-test on interaction between grazers and blocks, p-value = 0.12).

There is no evidence that the different grazers impact each other (extra SS F-test on interactions between limpets and fish, p-value = 0.72).

rocks

Allowing limpets access to plots caused significant changes in the regeneration of seaweed (two sided p-value < 0.00001 from a t-test on the effect of limpets). It is estimated that the median regeneration ratio when limpets were present is 0.161 times as large as the median regeneration time when they are excluded (95% CI: 0.126 to 0.205).

confint

exp

exp(-1.82) = 0.161

... two more, one for small fish, one for big fish

Estimating effects

not of the treatments, but of the animals

Two approaches:

Using averages over cell, rows and columns.

HARD, and only relevant for balanced data

Using indicator variables and multiple regression.

Balanced data only

Basic Idea: Averages over cells, rows and columns estimate means of interest

Display 13.12 p. 388

Table of averages of log percent seaweed regeneration ratio with different grazer combinations in eight blocks

average log recovery ratio for all treatments in first block

average log recovery ratio for control treatment in first block

Block	Control	L	f	Lf	fF	LfF	Block Average	Block Effect
	Treatment: Grazers with Access							
1	-1.51	-3.18	-1.62	-3.21	-2.05	-4.24	-2.64	-1.40
2	-0.94	-2.51	-1.31	-3.11	-1.97	-3.21	-2.18	-0.94
3	1.11	-0.31	0.22	-1.56	-0.12	-2.53	-0.53	0.70
4	2.85	-0.81	1.84	-0.52	0.64	-1.93	0.34	1.58
5	-0.27	-1.40	-0.69	-2.63	-0.68	-2.83	-1.42	-0.19
6	0.71	-1.23	-0.18	-0.66	-0.41	-1.89	-0.61	0.62
7	-0.79	-2.60	-0.08	-2.59	-0.74	-2.38	-1.53	-0.29
8	0.28	-1.66	-0.64	-1.75	-1.25	-2.77	-1.31	-0.07
Treatment Average	0.18	-1.71	-0.31	-2.00	-0.82	-2.72	-1.23	
Treatment Effect	1.41	-0.48	0.92	-0.77	0.41	-1.49		

average log recovery ratio for control treatments over all blocks

average log recovery ratio over all treatments and all blocks

Balanced data only

Basic Idea: Our best guess for the mean recovery ratio for the control treatment, is the average recovery ratio for the control treatment, over all the blocks

Display 13.12 p. 388

Table of averages of log percent seaweed regeneration ratio with different grazer combinations in eight blocks

Block	Treatment: Grazers with Access: Control	L	f	Lf	fF	LfF	Block Average	Block Effect
1	-1.51	-3.18	-1.62	-3.21	-2.05	-4.24	-2.64	-1.40
2	-0.94	-2.51	-1.31	-3.11	-1.97	-3.21	-2.18	-0.94
3	1.11	-0.31	0.22	-1.56	-0.12	-2.53	-0.53	0.70
4	2.85	-0.81	1.84	-0.52	0.64	-1.93	0.34	1.58
5	-0.27	-1.40	-0.69	-2.63	-0.68	-2.83	-1.42	-0.19
6	0.71	-1.23	-0.18	-0.66	-0.41	-1.89	-0.61	0.62
7	-0.79	-2.60	-0.08	-2.59	-0.74	-2.38	-1.53	-0.29
8	0.28	-1.66	-0.64	-1.75	-1.25	-2.77	-1.31	-0.07
Treatment Average	0.18	-1.71	-0.31	-2.00	-0.82	-2.72	-1.23	
Treatment Effect	1.41	-0.48	0.92	-0.77	0.41	-1.49		

$\hat{\mu}_C$ $\hat{\mu}_L$ $\hat{\mu}_f$ $\hat{\mu}_{Lf}$ $\hat{\mu}_{fF}$ $\hat{\mu}_{LfF}$

Balanced data only

Your turn

Which means can we compare to tell us about the big fish (F) effect?

$\hat{\mu}_C$ $\hat{\mu}_L$ $\hat{\mu}_f$ $\hat{\mu}_{Lf}$ $\hat{\mu}_{fF}$ $\hat{\mu}_{LfF}$

$\hat{\mu}_F - \hat{\mu}_c$

$\mu_{LfF} - \mu_f - \mu_L$?

Large fish effect = change in mean associated with letting large fish access rock, (holding access by other animals constant).

Balanced data only

Large fish: $\frac{1}{2}(\hat{\mu}_{Ff} - \hat{\mu}_{f}) + \frac{1}{2}(\hat{\mu}_{FfL} - \hat{\mu}_{fL})$

Small fish: $\frac{1}{2}(\hat{\mu}_{Lf} - \hat{\mu}_{L}) + \frac{1}{2}(\hat{\mu}_{f} - \hat{\mu}_{C})$

Limpets: $\frac{1}{3}(\hat{\mu}_{LfF} - \hat{\mu}_{fF}) + \frac{1}{3}(\hat{\mu}_{Lf} - \hat{\mu}_{f}) + \frac{1}{3}(\hat{\mu}_{L} - \hat{\mu}_{C})$

Only makes sense if each term is estimating the same thing, i.e. the effect of one species doesn't depend on the presence of another (no interactions)

Limpets x small fish:

$\frac{1}{2}\left((\hat{\mu}_{LfF} - \hat{\mu}_{fF}) + \frac{1}{2}(\hat{\mu}_{Lf} - \hat{\mu}_{f})\right) - (\hat{\mu}_{L} - \hat{\mu}_{C})$

Limpets x large fish:

$(\hat{\mu}_{LfF} - \hat{\mu}_{fF}) - (\hat{\mu}_{Lf} - \hat{\mu}_{f}) \approx 0$ no interaction

Balanced data only

Display 13.13 p. 389

Separate effects of grazers using linear combinations of treatment means

Treatment:	LfF	fF	Lf	f	L	C	Estimate	Standard Error	t-Stat
Sample size:	16	16	16	16	16	16			
Average:	-2.7247	-0.8214	-2.0044	-0.3137	-1.7120	+0.1805			
Large Fish:	$+\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0	0	-0.6140	0.1497	4.10
Small Fish:	0	0	$+\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	-0.3933	0.1497	2.63
Limpets:	$+\frac{1}{3}$	$-\frac{1}{3}$	$+\frac{1}{3}$	$-\frac{1}{3}$	$+\frac{1}{3}$	$-\frac{1}{3}$	-1.8288	0.1222	14.97
Limpets x Small:	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	- 1	+ 1	+0.0955	0.2593	0.37
Limpets x Large:	+ 1	- 1	- 1	+ 1	0	0	-0.2126	0.2994	0.71

(Contrast Summary: Estimate, Standard Error, t-Stat)

not significant

Like in the one-way case (from ST411/511)

$\gamma = C_1\mu_1 + C_2\mu_2 + C_3\mu_3 + \ldots + C_I\mu_I$

$g = C_1\bar{Y}_1 + C_2\bar{Y}_2 + C_3\bar{Y}_3 + \ldots + C_I\bar{Y}_I$

$$\mathrm{SE}_g = s_p\sqrt{\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \ldots + \frac{C_I^2}{n_I}}$$

A two-way ANOVA

Sometimes only one factor is of interest, sometimes both are, sometimes the interaction is the primary interest.

The general approach is the same:

Start with the non-additive/saturated model

Use F-tools to simplify

Then answer specific questions about means