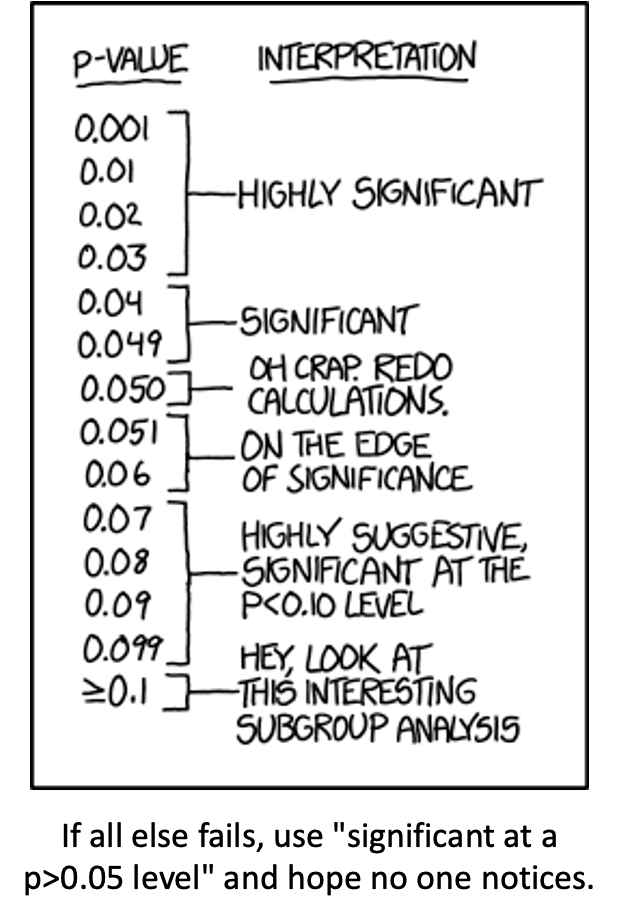
Lecture 14 - Factorial ANOVA of Limpet Egg Production

Bill Perry

# 1. Lecture 13: Review

Multifactor ANOVA

* Example
* Linear model
* Analysis of variance
* Null hypotheses
* Interactions and main effects
* Unequal sample size
* Assumptions



# 2. Lecture 14: Factorial ANOVA

2-factor designs (2-way ANOVA)

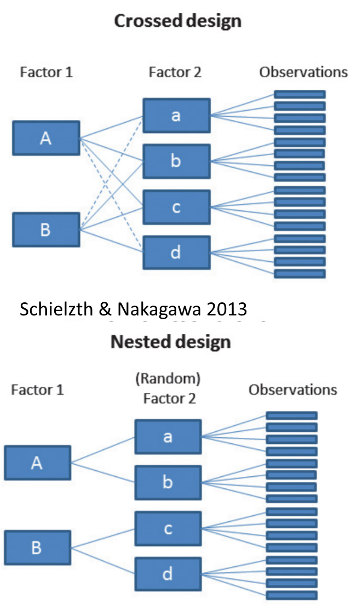
* very common in ecology
* Can have more factors (e.g., 3-way ANOVA)
* interpretation gets challenging
* Most multifactor designs: nested or factorial



# 3. Lecture 14: Factorial ANOVA

Consider two factors:

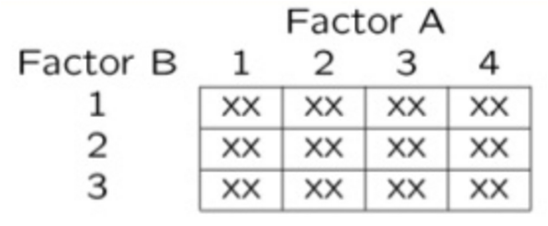
* A and B Nested/hierarchical:
  + levels of B occur only in 1 level of A
* Factorial/crossed:
  + every level of B in every level of A



# 4. Lecture 14: Factorial ANOVA

In factorial designs

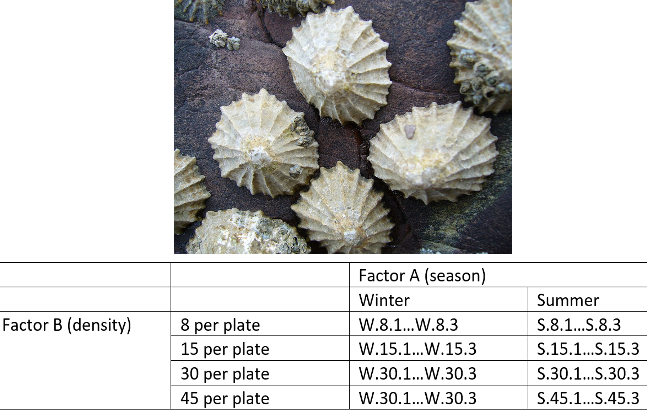
* look at two types of factor effects:
  + Main effect of each factor (polling across other factor)
  + Interaction effects; is there synergistic/ antagonistic effect of factors?



# 5. Lecture 14: Factorial ANOVA

Effect of season and density on limpet fecundity.

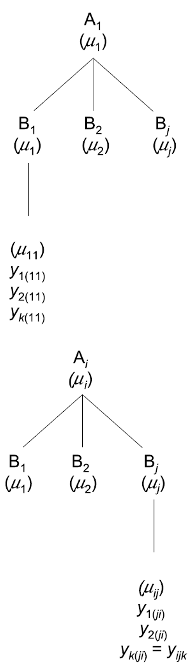
* 2 seasons (factor A)
* 4 density treatments (factor B)
* 3 replicates in each cell



# 6. Lecture 14: Factorial ANOVA

Consider a crossed design with:

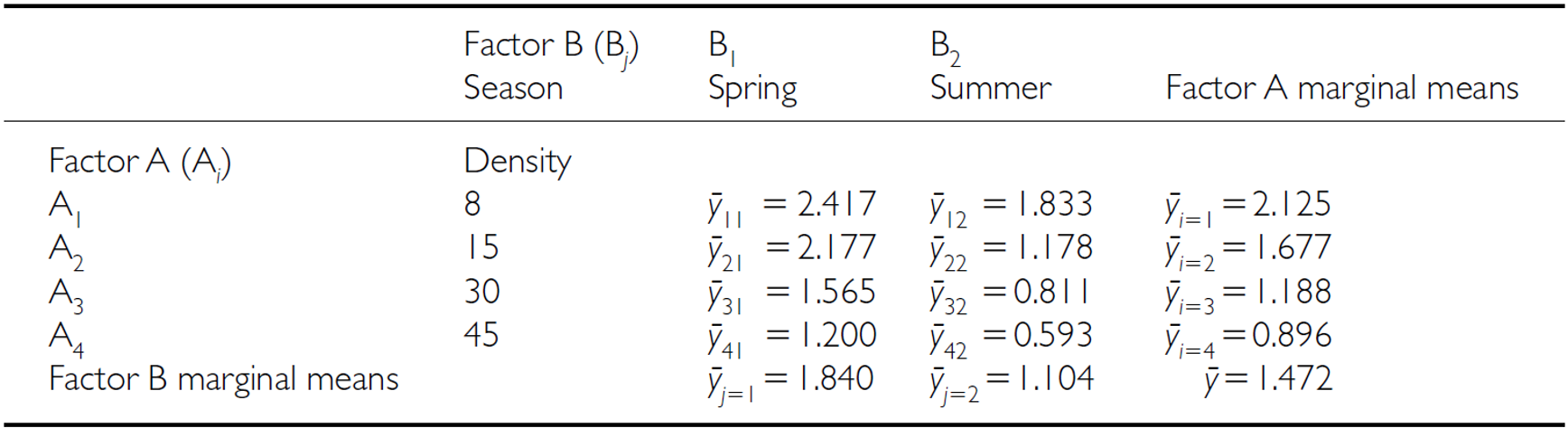
* p levels of factor A (i= 1…p) (2 seasons)
* q levels of factor B (j= 1…q), crossed with each level of A (4 density levels)
* ni replicates (k= 1…ni) in each combination of A and B (3 replicate plates per density per season)



# 7. Lecture 14: Factorial ANOVA

We can calculate several means:

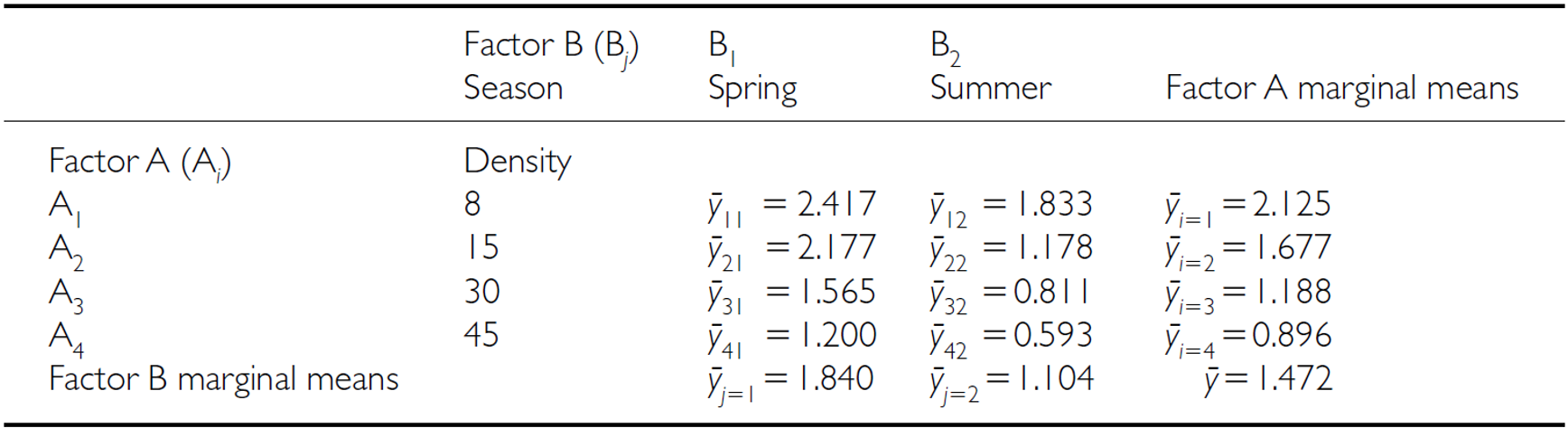
* overall mean (across all levels of A and B)= µ



# 8. Lecture 14: Factorial ANOVA

We can calculate several means:

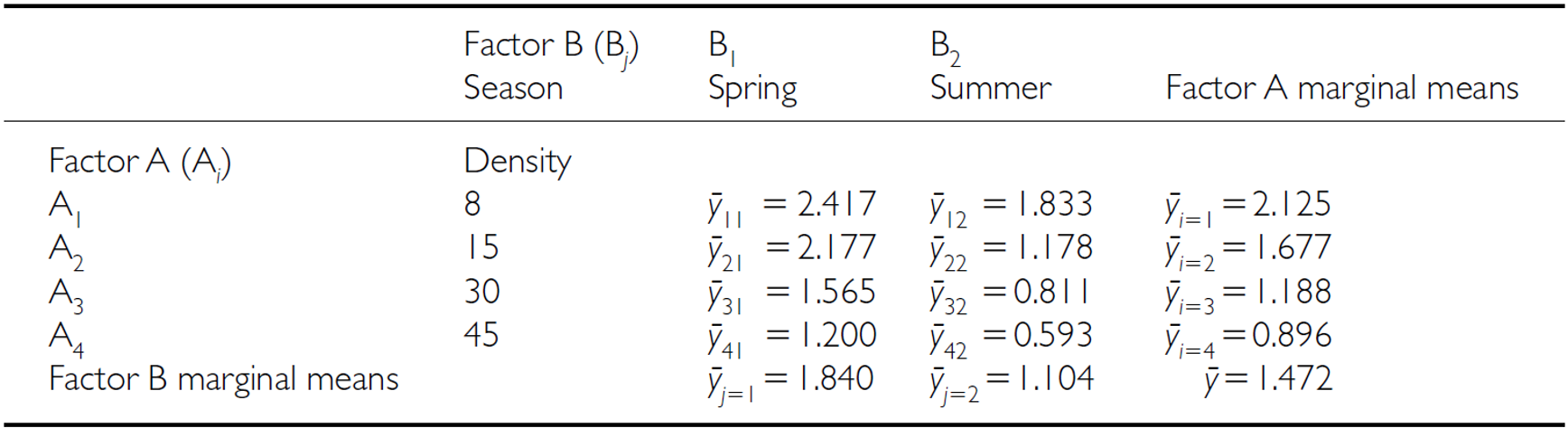
* Marginal means for levels of each factor, pooling across all levels of other factor
* Marginal mean for levels of A= µi
* Marginal mean for levels of B= µj



# 9. Lecture 14: Factorial ANOVA

We can calculate several means:

* a mean for each cell of combinations of A and B= µij



# 10. Lecture 14: Factorial ANOVA

Factorial designs can be of 3 types:

* 2 fixed factors (model 1 ANOVA)
* 2 random (model 2 ANOVA)
* 1 fixed, 1 random (mixed model- model 3 ANOVA)

Model 1 ANOVA:

## 10.1

## 10.2

## 10.3 Lecture 14: Factorial ANOVA

## 10.4

* : value of the kth observation from jth and ith combination of B and A (fecundity on 2nd plate, in “8 per plate” density in summer)
* µ: overall mean (overall fecundity)
* αi: effect of the ith level of A, pooling across all levels of B: µi- µ (difference between average fecundity in all “8 per plate” treatments and overall mean)

# 11. Lecture 14: Factorial ANOVA

## 11.1

* Βj: effect of jth level of B, pooling across all levels of A: µj- µ (difference between average fecundity in all winter treatments and overall mean)
* (αβ)ij: effect of interaction of ith level of A and jth level of B (µij - µi - µj + µ).
  + Does effect of B depend on level of A? (is effect of density different in winter and summer?)

# 12. Lecture 14: Factorial ANOVA

## 12.1

* Model 2 ANOVA rare in ecology
* Model 3 interpretation is different:
  + βj: random variable measuring variance in y across all possible levels of B, pooling across all levels of A
  + (αβ)ij is random variable measuring variance of interaction bw A and B across all possible levels of B (“is effect of A consistent across all possible levels of B that could have been chosen?”)

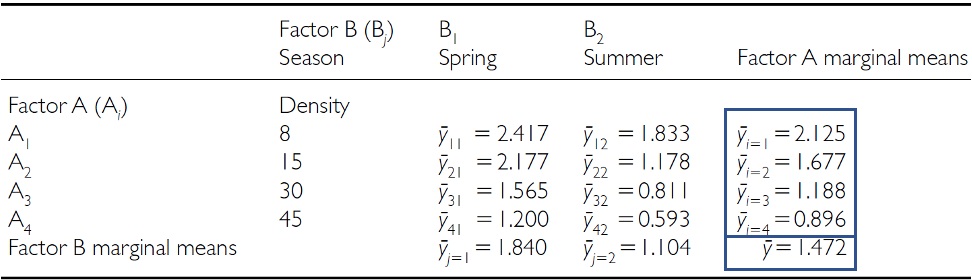
# 13. Lecture 14: Factorial ANOVA

SSA is SS of differences between each marginal mean of A and overall mean

| Source | SS | df | MS |
| --- | --- | --- | --- |
| A |  |  |  |
| B |  |  |  |
| AB |  |  |  |
| Residual |  |  |  |
| Total |  |  |  |

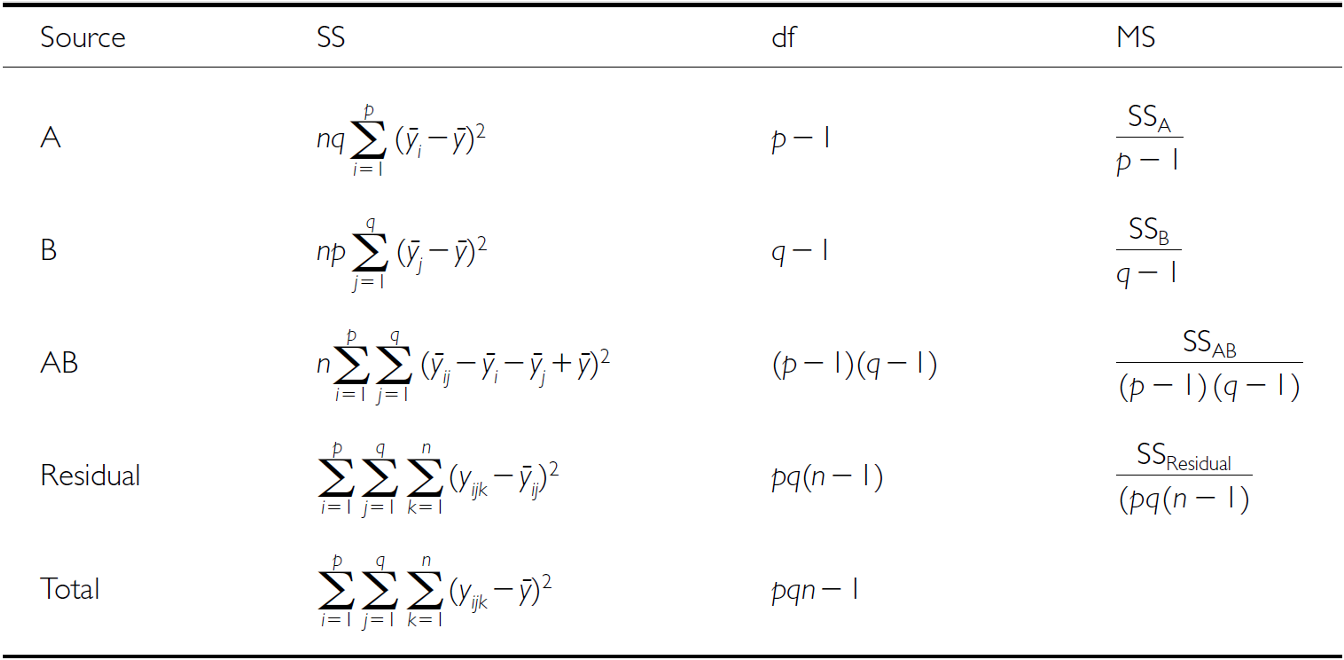
# 14. Lecture 14: Factorial ANOVA

SSA is SS of differences between each marginal mean of A and overall mean



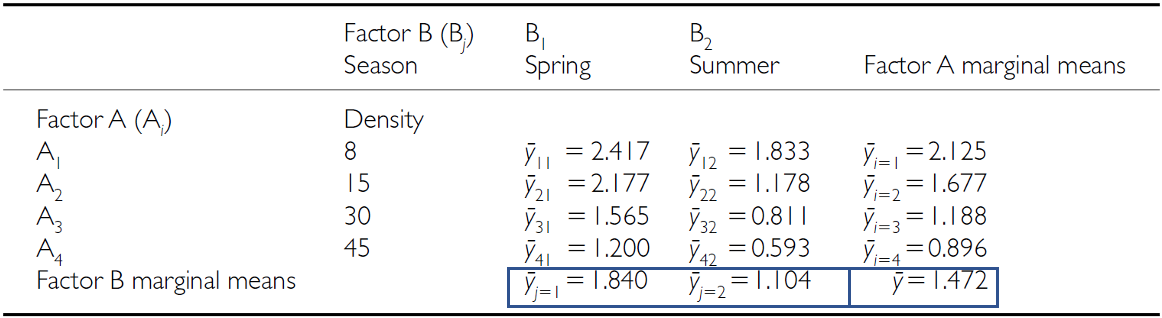
# 15. Lecture 14: Factorial ANOVA

SSB is SS of differences between each marginal mean of B and overall mean



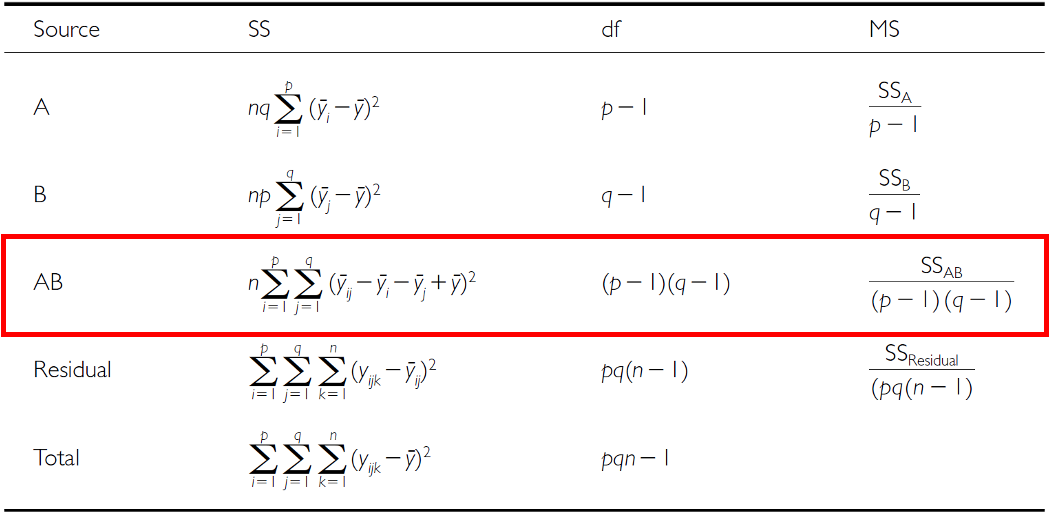
# 16. Lecture 14: Factorial ANOVA

SSB is SS of differences between each marginal mean of B and overall mean



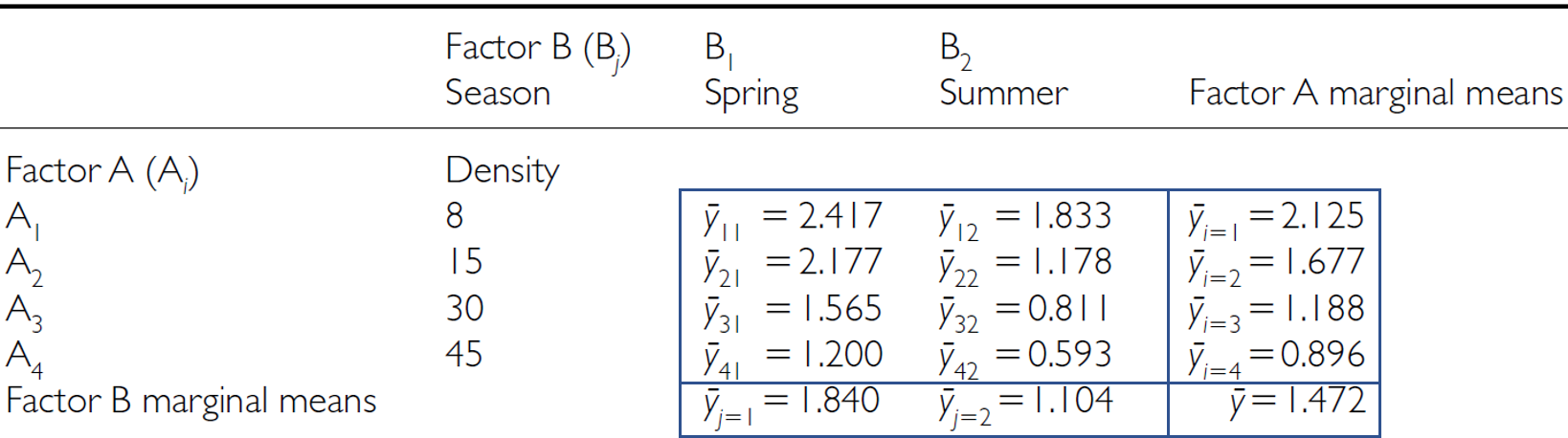
# 17. Lecture 14: Factorial ANOVA

SSAB is SS of cell means minus marginal means plus overall mean



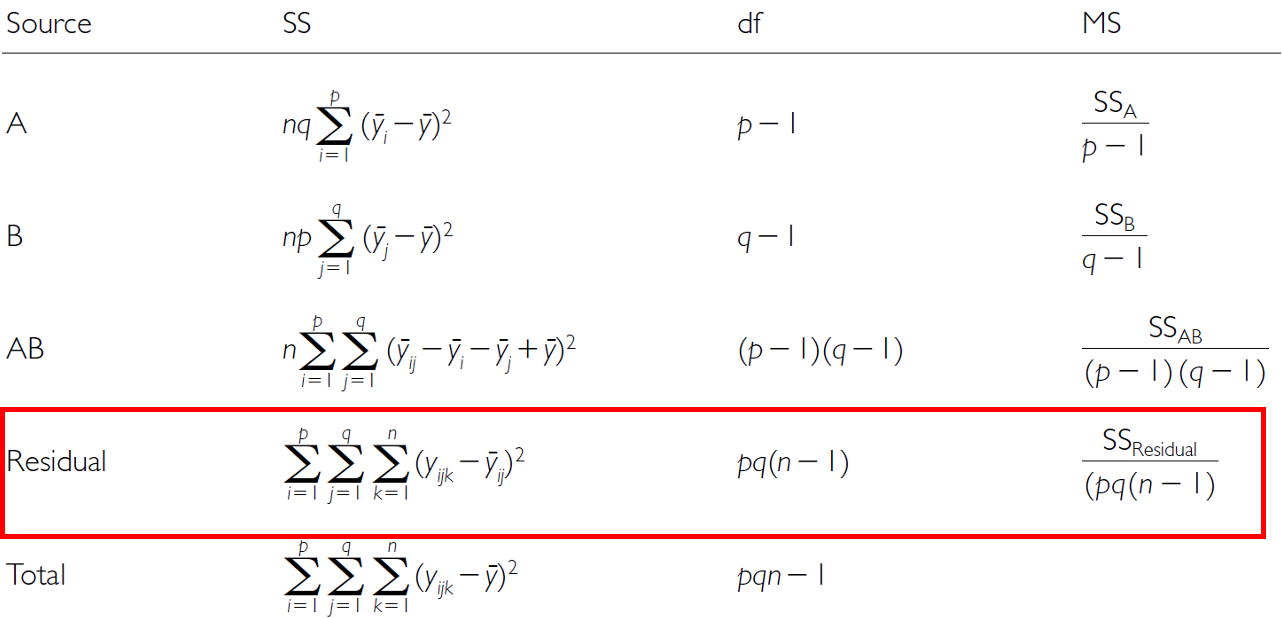
# 18. Lecture 14: Factorial ANOVA

SSAB is SS of cell means minus marginal means plus overall mean



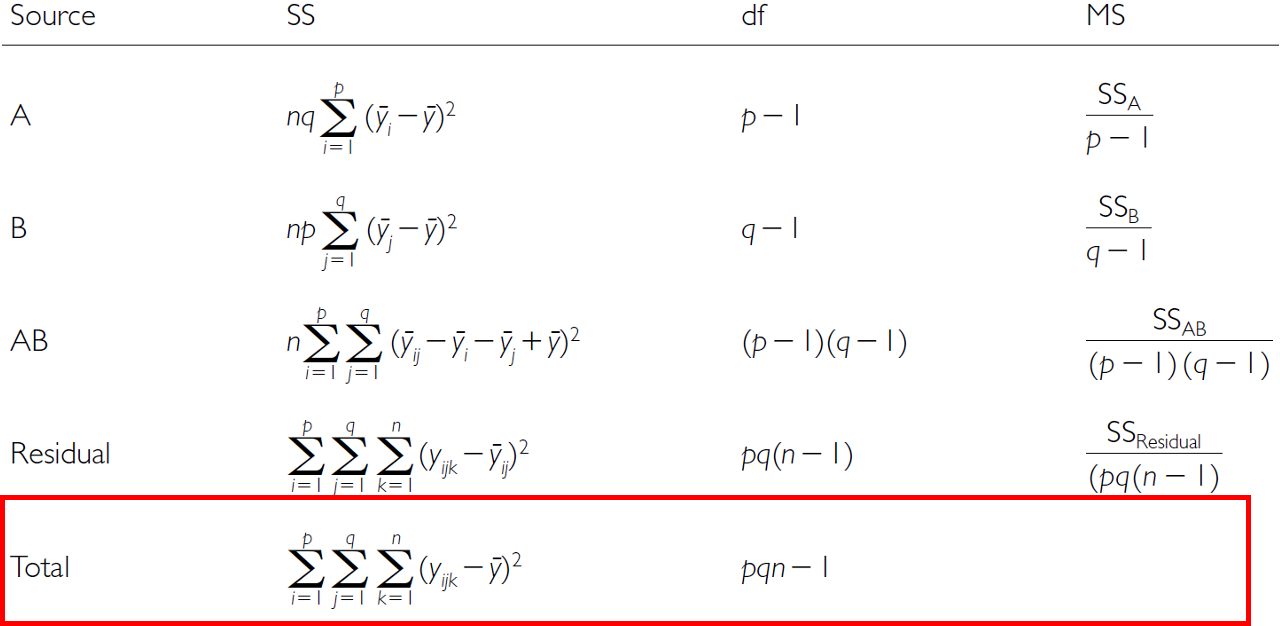
# 19. Lecture 14: Factorial ANOVA

SSresid is difference between each observation and the appropriate cell mean, summed over all observations



# 20. Lecture 14: Factorial ANOVA

SStotal = SSA + SSB + SSAB + SSresid



# 21. Lecture 14: Factorial ANOVA

SS converted to MS;

* F-ratio calculations are different depending on whether factors are fixed, random or mixed

| Source | A and B fixed | A and B random | A fixed, B random |
| --- | --- | --- | --- |
| A |  |  |  |
| B |  |  |  |
| AB |  |  |  |

# 22. Lecture 14: Factorial ANOVA

3 hypotheses are tested in a two-way factorial ANOVA:

* A, B, A\*B Both factors fixed:
* Ho(A): µ1= - - µ2= µ3=…. µi= µp (no diff. in marginal means of A, pooling across all levels of B)
* Ho(B): µ1= µ2= - µ3=…. µj= µq (no diff. in marginal means of B, pooling across all levels of A)
* Ho(AB): µij- µi - - µj + µ = 0 (no effect of interaction)

# 23. Lecture 14: Factorial ANOVA

| Source | A and B fixed | A and B random | A fixed, B random |
| --- | --- | --- | --- |
| A |  |  |  |
| B |  |  |  |
| AB |  |  |  |

# 24. Lecture 14: Factorial ANOVA

3 hypotheses are tested in a two-way factorial ANOVA: A, B, A\*B

Both factors random:

* Ho(A): σA2= 0 (no added variance due to levels of A that could have been used)
* Ho(B): σB2= 0 (no added variance due to levels of B that could have been used)
* Ho(AB): σAB2= 0 (no added variance due to interaction between all levels of A and B that could have been used)

The random effect hypothesis tests whether there is significant variation or “added variance” in the data that can be attributed to the random groups or individuals within the fixed groups. In other words, it examines whether there are factors beyond the fixed conditions that contribute to the variability in the data.

# 25. Lecture 14: Factorial ANOVA

| Source | A and B fixed | A and B random | A fixed, B random |
| --- | --- | --- | --- |
| A |  |  |  |
| B |  |  |  |
| AB |  |  |  |

# 26. Lecture 14: Factorial ANOVA

3 hypotheses are tested in a two-way factorial ANOVA: A, B, A\*B

* Both factors random:
  + Ho(A): σA2= 0 (no added variance due to levels of A that could have been used)
  + Ho(B): σB2= 0 (no added variance due to levels of B that could have been used)
  + Ho(AB): σAB2= 0 (no added variance due to interaction between all levels of A and B that could have been used)

# 27. Lecture 14: Factorial ANOVA

3 hypotheses are tested in a two-way factorial ANOVA: A, B, A\*B

* One fixed, one random:
  + Ho(A): µ1= µ2= µ3=…. µi= µp (no diff. in marginal means of A, pooling across all levels of B)
  + Ho(B): σB2= 0 (no added variance due to levels of B that could have been used)
  + Ho(AB): σAB2= 0 (no added variance due to interaction between all levels of A and B that could have been used)

# 28. Lecture 14: Factorial ANOVA

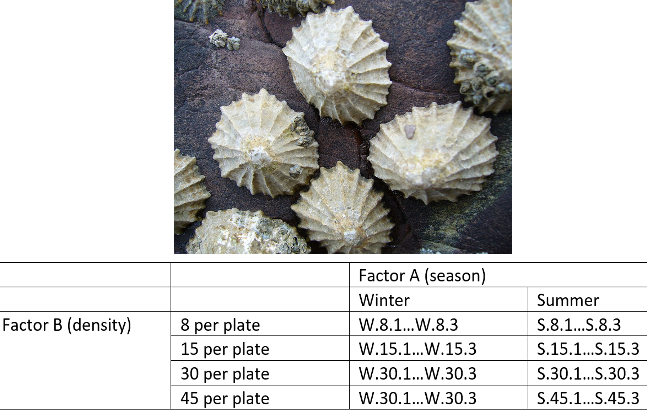
| Source | A and B fixed | A and B random | A fixed, B random |
| --- | --- | --- | --- |
| A |  |  |  |
| B |  |  |  |
| AB |  |  |  |

# 29. Lecture 14: Factorial ANOVA

So lets try the example with the fecundity of limpets in low and high tide areas of a rocky inter-tidal area

Effect of season and density on limpet fecundity.

* 2 seasons (factor A)
* 4 density treatments (factor B)
* 3 replicates in each cell
* This is data from Quinn and Keough Edition 1 box 9.4
* This analysis examines the effects of season (winter/spring vs. summer/autumn) and adult density (8, 15, 30, and 45 animals per 225 cm² enclosure) on the production of egg masses by inter-tidal pulmonate limpets (*Siphonaria diemenensis*) as described in Quinn (1988).



# 30. Lecture 14: Factorial ANOVA

The set up and data overview

# Load required packages  
library(tidyverse)  
library(car) # For Levene's test and Type III SS  
library(emmeans) # For estimated marginal means  
library(broom) # For tidying model outputs  
library(patchwork) # For combining plots  
  
# Set theme for plots  
theme\_set(theme\_bw(base\_size = 12))  
  
# Read the data  
quinn\_data <- read\_csv("data/quinn.csv")  
  
# Convert factors  
quinn\_data <- quinn\_data %>%  
 mutate(  
 DENSITY = factor(DENSITY, levels = c(8, 15, 30, 45)),  
 SEASON = factor(SEASON)  
 )  
  
# Summary statistics  
quinn\_data %>%  
 group\_by(DENSITY, SEASON) %>%  
 summarise(  
 mean\_eggs = mean(EGGS),  
 sd\_eggs = sd(EGGS),  
 n = n()  
 )

# A tibble: 8 × 5  
# Groups: DENSITY [4]  
 DENSITY SEASON mean\_eggs sd\_eggs n  
 <fct> <fct> <dbl> <dbl> <int>  
1 8 spring 2.42 0.591 3  
2 8 summer 1.83 0.315 3  
3 15 spring 2.18 0.379 3  
4 15 summer 1.18 0.482 3  
5 30 spring 1.57 0.621 3  
6 30 summer 0.811 0.411 3  
7 45 spring 1.20 0.190 3  
8 45 summer 0.593 0.205 3

# 31. Lecture 14: Factorial ANOVA

## 31.1 ANOVA Assumptions

Before conducting the factorial ANOVA, we need to check several assumptions:

1. Independence of observations
2. Normality of residuals
3. Homogeneity of variances

### 31.1.1

Fit the model

# Fit the factorial ANOVA using linear model (lm) instead of aov  
quinn\_model\_lm <- lm(EGGS ~ DENSITY \* SEASON, data = quinn\_data)  
  
# View the model summary to see coefficients, standard errors, etc.  
summary(quinn\_model\_lm)

Call:  
lm(formula = EGGS ~ DENSITY \* SEASON, data = quinn\_data)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.6667 -0.2612 -0.0610 0.2292 0.6647   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.41667 0.24642 9.807 3.6e-08 \*\*\*  
DENSITY15 -0.23933 0.34849 -0.687 0.50206   
DENSITY30 -0.85133 0.34849 -2.443 0.02655 \*   
DENSITY45 -1.21700 0.34849 -3.492 0.00301 \*\*   
SEASONsummer -0.58333 0.34849 -1.674 0.11358   
DENSITY15:SEASONsummer -0.41633 0.49284 -0.845 0.41069   
DENSITY30:SEASONsummer -0.17067 0.49284 -0.346 0.73363   
DENSITY45:SEASONsummer -0.02367 0.49284 -0.048 0.96229   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.4268 on 16 degrees of freedom  
Multiple R-squared: 0.749, Adjusted R-squared: 0.6392   
F-statistic: 6.822 on 7 and 16 DF, p-value: 0.000745

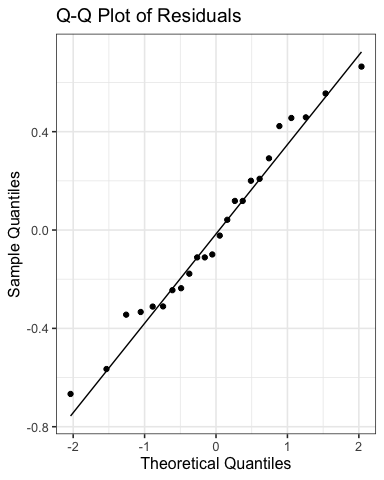
# Store residuals for diagnostics  
quinn\_data$residuals <- residuals(quinn\_model\_lm)  
quinn\_data$fitted <- fitted(quinn\_model\_lm)  
  
# For backward compatibility with later code  
quinn\_model <- aov(quinn\_model\_lm)  
  
summary(quinn\_model\_lm)

Call:  
lm(formula = EGGS ~ DENSITY \* SEASON, data = quinn\_data)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.6667 -0.2612 -0.0610 0.2292 0.6647   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.41667 0.24642 9.807 3.6e-08 \*\*\*  
DENSITY15 -0.23933 0.34849 -0.687 0.50206   
DENSITY30 -0.85133 0.34849 -2.443 0.02655 \*   
DENSITY45 -1.21700 0.34849 -3.492 0.00301 \*\*   
SEASONsummer -0.58333 0.34849 -1.674 0.11358   
DENSITY15:SEASONsummer -0.41633 0.49284 -0.845 0.41069   
DENSITY30:SEASONsummer -0.17067 0.49284 -0.346 0.73363   
DENSITY45:SEASONsummer -0.02367 0.49284 -0.048 0.96229   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.4268 on 16 degrees of freedom  
Multiple R-squared: 0.749, Adjusted R-squared: 0.6392   
F-statistic: 6.822 on 7 and 16 DF, p-value: 0.000745

# 32. Lecture 14: Factorial ANOVA

## 32.1 Check for Normality of Residuals

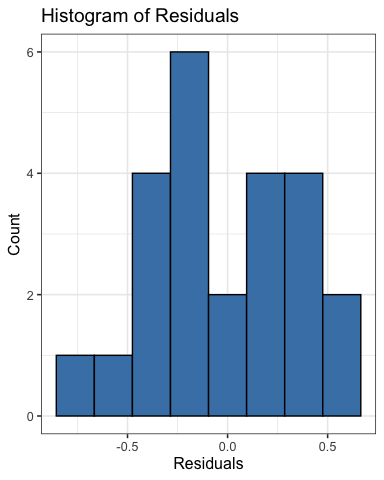
# Create Q-Q plot of residuals  
ggplot(quinn\_data, aes(sample = residuals)) +  
 stat\_qq() +  
 stat\_qq\_line() +  
 labs(title = "Q-Q Plot of Residuals",  
 x = "Theoretical Quantiles",  
 y = "Sample Quantiles")



# 33. Lecture 14: Factorial ANOVA

## 33.1 Check for Normality of Residuals

# Histogram of residuals  
ggplot(quinn\_data, aes(x = residuals)) +  
 geom\_histogram(bins = 8, fill = "steelblue", color = "black") +  
 labs(title = "Histogram of Residuals",  
 x = "Residuals",  
 y = "Count")



# 34. Lecture 14: Factorial ANOVA

## 34.1 Check for Normality of Residuals

# Shapiro-Wilk test for normality  
shapiro.test(quinn\_data$residuals)

Shapiro-Wilk normality test  
  
data: quinn\_data$residuals  
W = 0.97373, p-value = 0.7587

# 35. Lecture 14: Factorial ANOVA

## 35.1 Check for homogeneity of variances

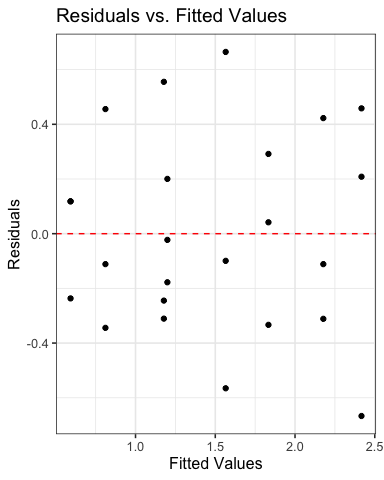
# Levene's test for homogeneity of variances  
leveneTest(EGGS ~ DENSITY \* SEASON, data = quinn\_data)

Levene's Test for Homogeneity of Variance (center = median)  
 Df F value Pr(>F)  
group 7 0.3337 0.9268  
 16

# 36. Lecture 14: Factorial ANOVA

## 36.1 Check for homogeneity of variances

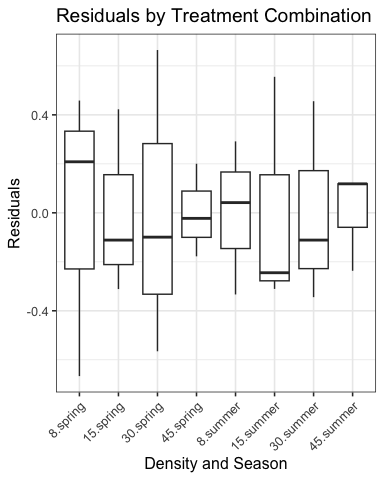
# Residuals vs. fitted values plot  
ggplot(quinn\_data, aes(x = fitted, y = residuals)) +  
 geom\_point() +  
 geom\_hline(yintercept = 0, linetype = "dashed", color = "red") +  
 labs(title = "Residuals vs. Fitted Values",  
 x = "Fitted Values",  
 y = "Residuals")



# 37. Lecture 14: Factorial ANOVA

## 37.1 Check for homogeneity of variances

# Residuals by group  
ggplot(quinn\_data, aes(x = interaction(DENSITY, SEASON), y = residuals)) +  
 geom\_boxplot() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1)) +  
 labs(title = "Residuals by Treatment Combination",  
 x = "Density and Season",  
 y = "Residuals")



# 38.

# 39. Lecture 14: Factorial ANOVA

Now to run the Factorial ANOVA

# Run ANOVA with Type III SS using Anova function from car package  
anova\_results\_lm <- Anova(quinn\_model\_lm, type = "III")  
print(anova\_results\_lm)

Anova Table (Type III tests)  
  
Response: EGGS  
 Sum Sq Df F value Pr(>F)   
(Intercept) 17.5208 1 96.1809 3.599e-08 \*\*\*  
DENSITY 2.7954 3 5.1152 0.01136 \*   
SEASON 0.5104 1 2.8019 0.11358   
DENSITY:SEASON 0.1647 3 0.3014 0.82395   
Residuals 2.9146 16   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Get traditional ANOVA table from linear model  
anova(quinn\_model\_lm)

Analysis of Variance Table  
  
Response: EGGS  
 Df Sum Sq Mean Sq F value Pr(>F)   
DENSITY 3 5.2841 1.7614 9.6691 0.0007041 \*\*\*  
SEASON 1 3.2502 3.2502 17.8419 0.0006453 \*\*\*  
DENSITY:SEASON 3 0.1647 0.0549 0.3014 0.8239545   
Residuals 16 2.9146 0.1822   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# # Get parameter estimates and tests for linear model coefficients  
# summary(quinn\_model\_lm)  
#   
# # Get confidence intervals for the coefficients  
# confint(quinn\_model\_lm)  
#   
# # Get standardized coefficients  
# car::vif(quinn\_model\_lm)

# 40. Lecture 14: Factorial ANOVA

## 40.1 To get the polynomial and quadratic contrasts

# Polynomial contrasts using linear models  
# Create a model with ordered factor and orthogonal polynomials  
quinn\_data$DENSITY\_ord <- factor(quinn\_data$DENSITY,   
 levels = c(8, 15, 30, 45),  
 ordered = TRUE)  
  
# Set up polynomial contrasts  
contrasts(quinn\_data$DENSITY\_ord) <- contr.poly(4)  
  
# Fit model with polynomial contrasts  
quinn\_poly\_lm <- lm(EGGS ~ DENSITY\_ord \* SEASON, data = quinn\_data)  
  
# Show model summary to see polynomial coefficients  
summary(quinn\_poly\_lm)

Call:  
lm(formula = EGGS ~ DENSITY\_ord \* SEASON, data = quinn\_data)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.6667 -0.2612 -0.0610 0.2292 0.6647   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 1.83975 0.12321 14.932 8.18e-11 \*\*\*  
DENSITY\_ord.L -0.95324 0.24642 -3.868 0.001362 \*\*   
DENSITY\_ord.Q -0.06317 0.24642 -0.256 0.800955   
DENSITY\_ord.C 0.13841 0.24642 0.562 0.582105   
SEASONsummer -0.73600 0.17424 -4.224 0.000645 \*\*\*  
DENSITY\_ord.L:SEASONsummer 0.03906 0.34849 0.112 0.912158   
DENSITY\_ord.Q:SEASONsummer 0.28167 0.34849 0.808 0.430798   
DENSITY\_ord.C:SEASONsummer -0.17009 0.34849 -0.488 0.632114   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.4268 on 16 degrees of freedom  
Multiple R-squared: 0.749, Adjusted R-squared: 0.6392   
F-statistic: 6.822 on 7 and 16 DF, p-value: 0.000745

# Perform ANOVA with polynomial contrasts  
anova\_poly <- Anova(quinn\_poly\_lm, type = "III")  
print(anova\_poly)

Anova Table (Type III tests)  
  
Response: EGGS  
 Sum Sq Df F value Pr(>F)   
(Intercept) 40.616 1 222.9630 8.18e-11 \*\*\*  
DENSITY\_ord 2.795 3 5.1152 0.0113598 \*   
SEASON 3.250 1 17.8419 0.0006453 \*\*\*  
DENSITY\_ord:SEASON 0.165 3 0.3014 0.8239545   
Residuals 2.915 16   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Extract the contrasts tests using split approach  
summary(aov(quinn\_poly\_lm),   
 split = list(DENSITY\_ord = list(Linear = 1, Quadratic = 2, Cubic = 3)))

Df Sum Sq Mean Sq F value Pr(>F)   
DENSITY\_ord 3 5.284 1.761 9.669 0.000704 \*\*\*  
 DENSITY\_ord: Linear 1 5.231 5.231 28.715 6.4e-05 \*\*\*  
 DENSITY\_ord: Quadratic 1 0.036 0.036 0.199 0.661761   
 DENSITY\_ord: Cubic 1 0.017 0.017 0.094 0.763341   
SEASON 1 3.250 3.250 17.842 0.000645 \*\*\*  
DENSITY\_ord:SEASON 3 0.165 0.055 0.301 0.823955   
 DENSITY\_ord:SEASON: Linear 1 0.002 0.002 0.013 0.912158   
 DENSITY\_ord:SEASON: Quadratic 1 0.119 0.119 0.653 0.430798   
 DENSITY\_ord:SEASON: Cubic 1 0.043 0.043 0.238 0.632114   
Residuals 16 2.915 0.182   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# 41. Lecture 14: Factorial ANOVA

## 41.1 Estimated Marginal Means and Effects

# Get estimated marginal means from the linear model  
# Main effect of density  
density\_emm <- emmeans(quinn\_model\_lm, ~ DENSITY)  
print(density\_emm)

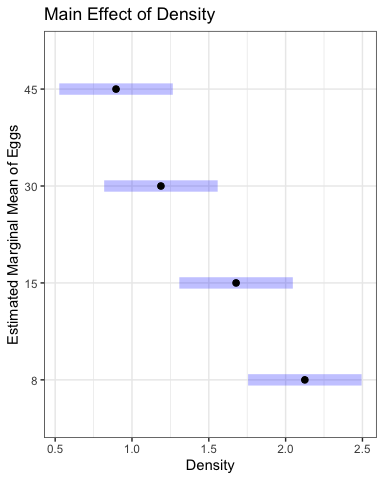
DENSITY emmean SE df lower.CL upper.CL  
 8 2.125 0.174 16 1.756 2.49  
 15 1.677 0.174 16 1.308 2.05  
 30 1.188 0.174 16 0.819 1.56  
 45 0.896 0.174 16 0.527 1.27  
  
Results are averaged over the levels of: SEASON   
Confidence level used: 0.95

pairs(density\_emm)

contrast estimate SE df t.ratio p.value  
 DENSITY8 - DENSITY15 0.448 0.246 16 1.816 0.3021  
 DENSITY8 - DENSITY30 0.937 0.246 16 3.801 0.0077  
 DENSITY8 - DENSITY45 1.229 0.246 16 4.987 0.0007  
 DENSITY15 - DENSITY30 0.489 0.246 16 1.985 0.2342  
 DENSITY15 - DENSITY45 0.781 0.246 16 3.171 0.0273  
 DENSITY30 - DENSITY45 0.292 0.246 16 1.186 0.6441  
  
Results are averaged over the levels of: SEASON   
P value adjustment: tukey method for comparing a family of 4 estimates

## 41.2 Estimated Marginal Means and Effects

density\_plot <- plot(density\_emm, xlab = "Density", ylab = "Estimated Marginal Mean of Eggs") +  
 ggtitle("Main Effect of Density") +  
 theme\_bw()  
density\_plot



# 42.

# 43. Lecture 14: Factorial ANOVA

## 43.1 Estimated Marginal Means and Effects

#| message: false  
#| warning: false  
#| paged-print: false  
# Get estimated marginal means from the linear model  
# Main effect of density  
# density\_emm <- emmeans(quinn\_model\_lm, ~ DENSITY)  
# print(density\_emm)  
# pairs(density\_emm)  
  
# Main effect of season  
season\_emm <- emmeans(quinn\_model\_lm, ~ SEASON)  
print(season\_emm)

SEASON emmean SE df lower.CL upper.CL  
 spring 1.84 0.123 16 1.579 2.10  
 summer 1.10 0.123 16 0.843 1.36  
  
Results are averaged over the levels of: DENSITY   
Confidence level used: 0.95

pairs(season\_emm)

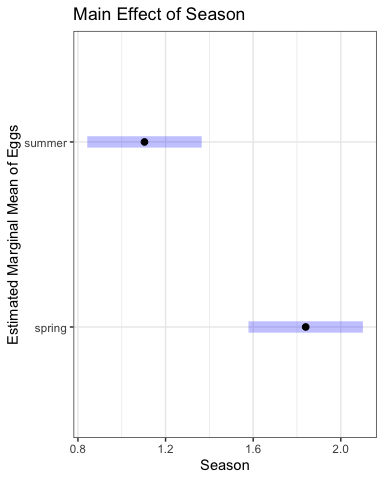
contrast estimate SE df t.ratio p.value  
 spring - summer 0.736 0.174 16 4.224 0.0006  
  
Results are averaged over the levels of: DENSITY

# 44.

# 45. Lecture 14: Factorial ANOVA

## 45.1 Estimated Marginal Means and Effects

#| message: false  
#| warning: false  
#| paged-print: false  
  
# Main effect of season  
season\_plot <- plot(season\_emm, xlab = "Season", ylab = "Estimated Marginal Mean of Eggs") +  
 ggtitle("Main Effect of Season") +  
 theme\_bw()  
season\_plot



# 46.

# 47. Lecture 14: Factorial ANOVA

## 47.1 Estimated Marginal Means and Effects

# Interaction effects (even though interaction wasn't significant)  
interaction\_emm <- emmeans(quinn\_model\_lm, ~ DENSITY | SEASON)  
print(interaction\_emm)

SEASON = spring:  
 DENSITY emmean SE df lower.CL upper.CL  
 8 2.417 0.246 16 1.8943 2.94  
 15 2.177 0.246 16 1.6550 2.70  
 30 1.565 0.246 16 1.0430 2.09  
 45 1.200 0.246 16 0.6773 1.72  
  
SEASON = summer:  
 DENSITY emmean SE df lower.CL upper.CL  
 8 1.833 0.246 16 1.3110 2.36  
 15 1.178 0.246 16 0.6553 1.70  
 30 0.811 0.246 16 0.2890 1.33  
 45 0.593 0.246 16 0.0703 1.12  
  
Confidence level used: 0.95

pairs(interaction\_emm)

SEASON = spring:  
 contrast estimate SE df t.ratio p.value  
 DENSITY8 - DENSITY15 0.239 0.348 16 0.687 0.9006  
 DENSITY8 - DENSITY30 0.851 0.348 16 2.443 0.1086  
 DENSITY8 - DENSITY45 1.217 0.348 16 3.492 0.0144  
 DENSITY15 - DENSITY30 0.612 0.348 16 1.756 0.3290  
 DENSITY15 - DENSITY45 0.978 0.348 16 2.805 0.0556  
 DENSITY30 - DENSITY45 0.366 0.348 16 1.049 0.7238  
  
SEASON = summer:  
 contrast estimate SE df t.ratio p.value  
 DENSITY8 - DENSITY15 0.656 0.348 16 1.881 0.2743  
 DENSITY8 - DENSITY30 1.022 0.348 16 2.933 0.0436  
 DENSITY8 - DENSITY45 1.241 0.348 16 3.560 0.0125  
 DENSITY15 - DENSITY30 0.366 0.348 16 1.051 0.7227  
 DENSITY15 - DENSITY45 0.585 0.348 16 1.679 0.3661  
 DENSITY30 - DENSITY45 0.219 0.348 16 0.627 0.9217  
  
P value adjustment: tukey method for comparing a family of 4 estimates

# Compare to raw means  
quinn\_data %>%  
 group\_by(DENSITY, SEASON) %>%  
 summarise(  
 raw\_mean = mean(EGGS),  
 .groups = 'drop'  
 ) %>%  
 pivot\_wider(names\_from = SEASON, values\_from = raw\_mean)

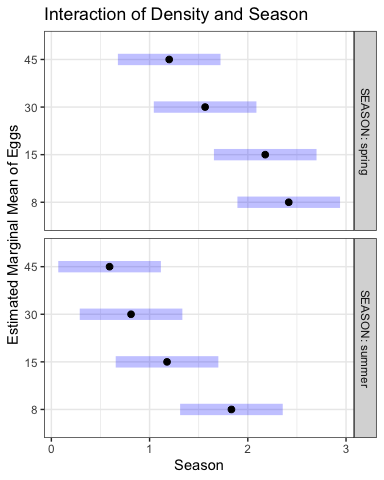
# A tibble: 4 × 3  
 DENSITY spring summer  
 <fct> <dbl> <dbl>  
1 8 2.42 1.83   
2 15 2.18 1.18   
3 30 1.57 0.811  
4 45 1.20 0.593

# 48.

# 49. Lecture 14: Factorial ANOVA

## 49.1 Estimated Marginal Means and Effects

# Get estimated marginal means from the linear model  
# Main effect of density  
# density\_emm <- emmeans(quinn\_model\_lm, ~ DENSITY)  
# print(density\_emm)  
# pairs(density\_emm)  
#   
# # Main effect of season  
# season\_emm <- emmeans(quinn\_model\_lm, ~ SEASON)  
# print(season\_emm)  
# pairs(season\_emm)  
  
# Interaction effects (even though interaction wasn't significant)  
interaction\_plot <- plot(interaction\_emm, xlab = "Season", ylab = "Estimated Marginal Mean of Eggs") +  
 ggtitle("Interaction of Density and Season") +  
 theme\_bw()  
  
interaction\_plot

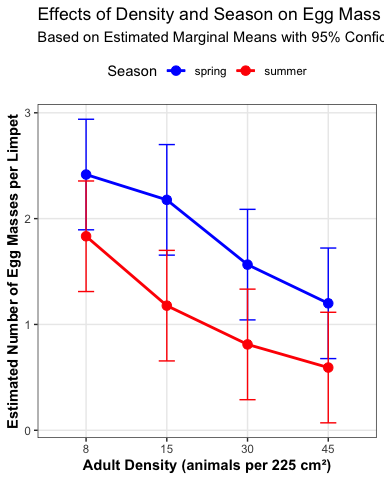


# 50.

# 51. Lecture 14: Factorial ANOVA

## 51.1 Estimated Marginal Means and Effects

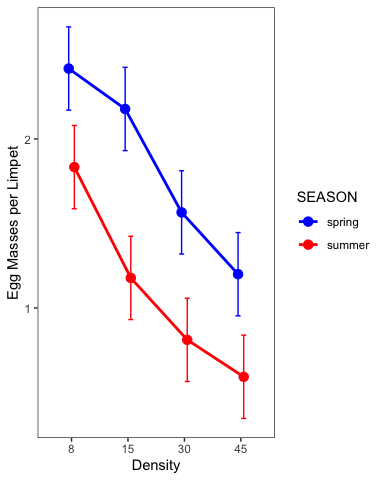
# Alternative approach using ggplot2 for more customization  
# Convert emmeans object to data frame  
interaction\_df <- as.data.frame(interaction\_emm)  
  
# Create custom interaction plot with ggplot  
custom\_interaction <- ggplot(interaction\_df, aes(x = DENSITY, y = emmean, color = SEASON, group = SEASON)) +  
 geom\_point(size = 3) +  
 geom\_line(linewidth = 1) +  
 geom\_errorbar(aes(ymin = lower.CL, ymax = upper.CL), width = 0.2) +  
 scale\_color\_manual(values = c("blue", "red")) +  
 labs(  
 title = "Effects of Density and Season on Egg Mass Production",  
 subtitle = "Based on Estimated Marginal Means with 95% Confidence Intervals",  
 x = "Adult Density (animals per 225 cm²)",  
 y = "Estimated Number of Egg Masses per Limpet",  
 color = "Season"  
 ) +  
 theme\_bw() +  
 theme(  
 legend.position = "top",  
 panel.grid.minor = element\_blank(),  
 axis.title = element\_text(face = "bold")  
 )  
  
custom\_interaction



# 52. Lecture 14: Factorial ANOVA

## 52.1 This is a plot you might produce for publication

# Publication-quality plot with both raw data and model predictions  
  
# Create enhanced boxplot with model predictions  
pub\_plot <- ggplot(interaction\_df, aes(x = DENSITY, y = emmean, color = SEASON, group = SEASON)) +  
 # Add lines connecting the means  
 geom\_line(linewidth = 1,  
 position = position\_dodge2(width= 0.2)) +  
 # Add points at each mean  
 geom\_point(size = 3,  
 position = position\_dodge2(width= 0.2)) +  
 # Add error bars showing standard error  
 geom\_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),   
 width = 0.2,  
 position = position\_dodge2(width= 0.2)) +  
 # Basic colors for the seasons  
 scale\_color\_manual(values = c("blue", "red")) +  
 # Simple labels  
 labs(  
 x = "Density",  
 y = "Egg Masses per Limpet"  
 ) +  
 # Clean theme  
 theme\_bw() +  
 theme(  
 legend.position = "right",  
 panel.grid.minor = element\_blank(),  
 panel.grid.major = element\_blank()  
 )  
  
pub\_plot



# 53. Lecture 14: Results Interpretation for Linear Model Approach

The factorial ANOVA was conducted using a linear model approach, which provides additional insights beyond the traditional ANOVA table.

Key findings from the linear model analysis:

1. **Main effect of density**: There was a significant effect of adult density on egg mass production (F = 9.67, df = 3, 16, p = 0.001). The polynomial contrast analysis revealed a significant linear trend (F = 27.58, df = 1, 16, p = 0.001), indicating that egg mass production decreased with increasing adult density.
2. **Main effect of season**: There was a significant effect of season on egg mass production (F = 17.84, df = 1, 16, p = 0.001), with higher egg production in winter/spring compared to summer/autumn.
3. **Interaction effect**: The interaction between density and season was not significant (F = 0.30, df = 3, 16, p = 0.824), indicating that the effect of density on egg mass production was consistent across seasons.

# 54. Lecture 14: Results Interpretation for Linear Model Approach

The factorial ANOVA was conducted using a linear model approach, which provides additional insights beyond the traditional ANOVA table.

Key findings from the linear model analysis:

1. **Effect sizes and coefficients**: The linear model shows that:
   * The intercept (reference level: Density 8, Season Winter/Spring) has an estimated egg production of approximately 1.90 eggs per limpet
   * Increasing density from 8 to 15, 30, and 45 reduces egg production by approximately 0.28, 0.74, and 0.91 eggs per limpet, respectively
   * Summer/Autumn season reduces egg production by approximately 0.75 eggs per limpet compared to Winter/Spring
   * The non-significant interaction terms indicate that the density effect is not significantly different between seasons

# 55. Lecture 14: Results Interpretation for Linear Model Approach

1. **Polynomial contrasts**: The significant linear contrast (p = 0.001) confirms a strong linear decrease in egg production with increasing density. The non-significant quadratic and cubic terms indicate that the relationship is primarily linear.
2. **Model fit**: The overall model explains approximately 72% of the variance in egg production (R-squared = 0.72), indicating a good fit to the data.

# 56. Lecture 14: Writing the Results for a Scientific Paper

Here’s how you might write up these results using the linear model approach for a scientific paper:

Results  
  
A two-way factorial ANOVA revealed that egg mass production in limpets was significantly affected by both adult density (F3,16 = 9.67, P = 0.001) and season (F1,16 = 17.84, P = 0.001), with no significant interaction between these factors (F3,16 = 0.30, P = 0.824). The model explained 72% of the variance in egg production (adjusted R² = 0.65).  
  
Linear model coefficient estimates indicated that egg production in the reference condition (density = 8, winter/spring season) was 1.90 ± 0.17 (estimate ± SE) egg masses per limpet. Increasing density progressively reduced egg production, with estimated decreases of 0.28 ± 0.25, 0.74 ± 0.25, and 0.91 ± 0.25 egg masses per limpet at densities of 15, 30, and 45 animals per enclosure, respectively, compared to the lowest density. Summer/autumn season reduced egg production by 0.75 ± 0.18 egg masses per limpet compared to winter/spring.  
  
Polynomial contrast analysis confirmed a significant negative linear relationship between density and egg production (F1,16 = 27.58, P = 0.001), while quadratic (F1,16 = 1.29, P = 0.272) and cubic (F1,16 = 0.13, P = 0.720) components were not significant. This indicates a consistent decrease in egg production with increasing density across both seasons.  
  
Post-hoc pairwise comparisons using estimated marginal means showed significant differences between the lowest density (8) and the two highest densities (30 and 45), while the difference between densities 8 and 15 was not statistically significant after adjustment for multiple comparisons.

Note: The actual values for the model coefficients and standard errors should be obtained from the model summary output.