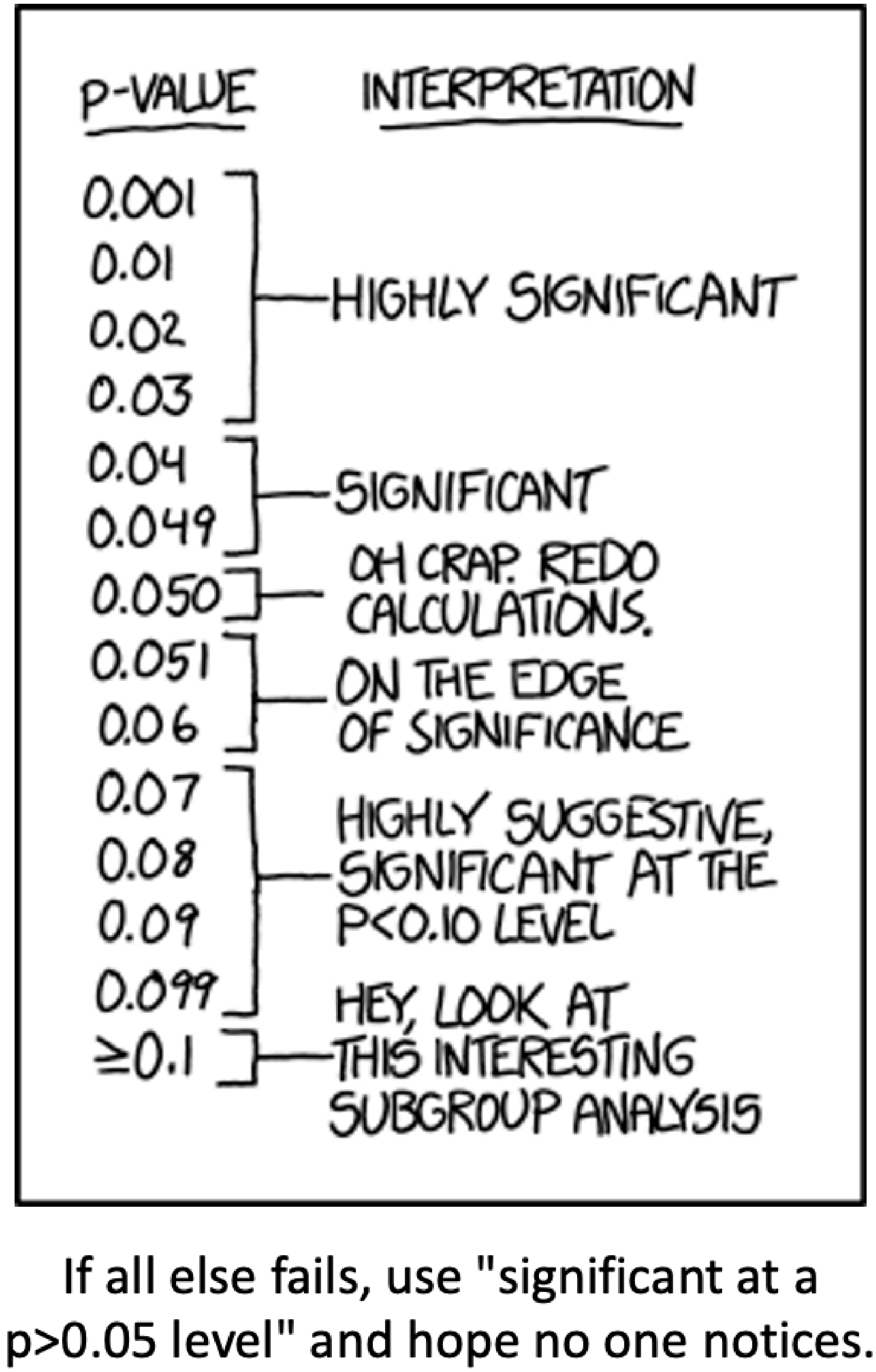
Lecture 12 - Factorial ANOVA of Limpet Egg Production

Bill Perry

Lecture 12: Review

ANOVA

- Analysis of variance: single and multi-factor designs
- Examples: diatoms, circadian rhythms
- Predictor variables: fixed vs. random
- ANOVA model
- Analysis and partitioning of variance
- Null hypothesis
- Assumptions and diagnostics
- Post F Tests - Tukey and others
- Reporting the results



Lecture 12: Factorial ANOVA

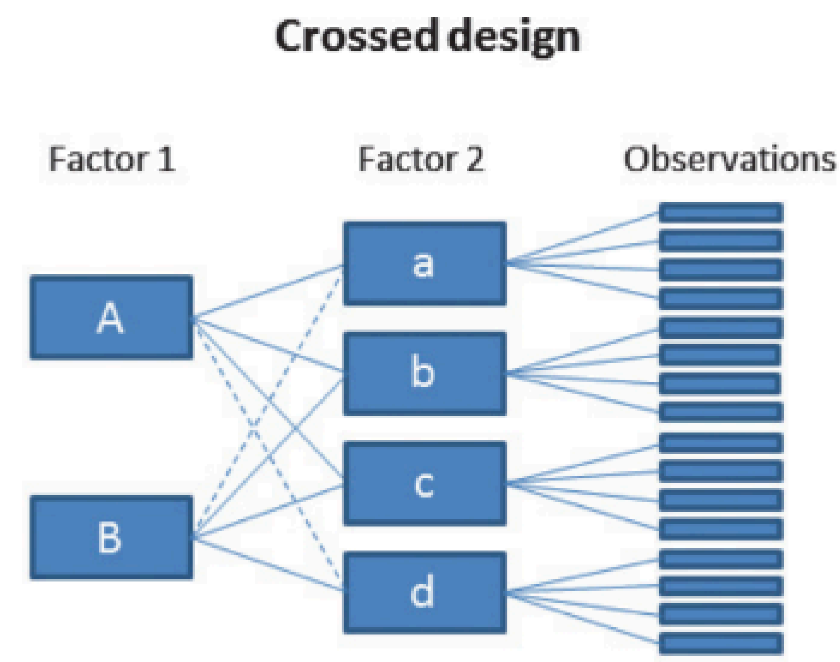
2-factor designs (2-way ANOVA)

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

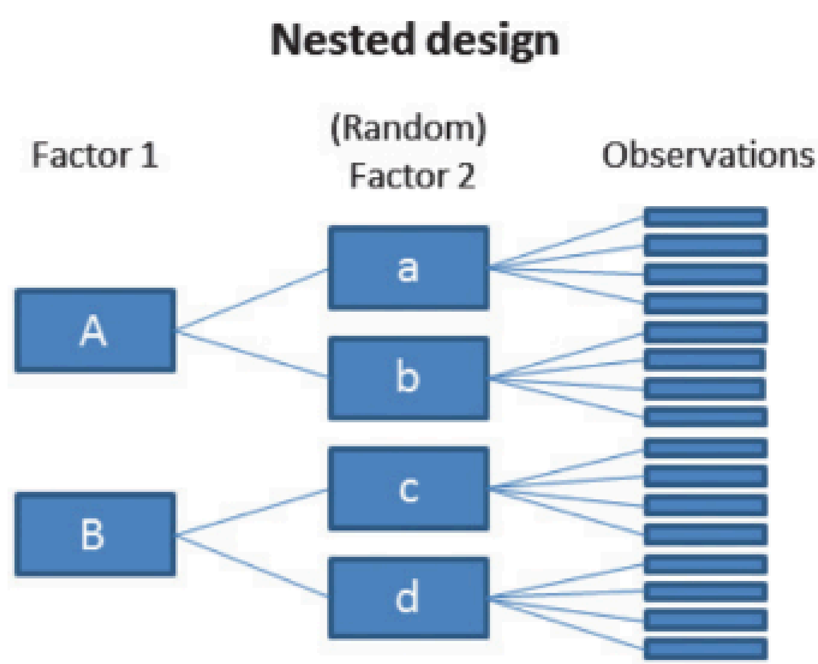
Factorial ANOVA: Design Structure

Consider two factors:

- Factorial/crossed:
  - every level of B in every level of A



Schielzth & Nakagawa 2013



Factorial ANOVA: Effect Types

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?

	Factor A			
Factor B	1	2	3	4
1	XX	XX	XX	XX
2	XX	XX	XX	XX
3	XX	XX	XX	XX

Example: Limpet Fecundity Study

Effect of season and density on limpet fecundity.

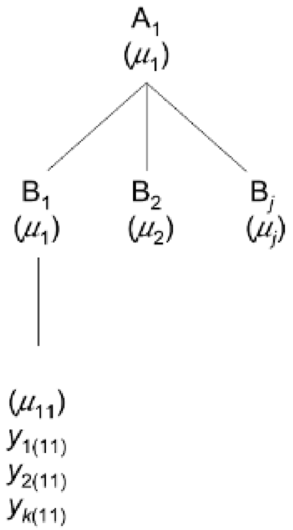
- 2 seasons (factor A)
- 4 density treatments (factor B)
- 3 replicates in each cell

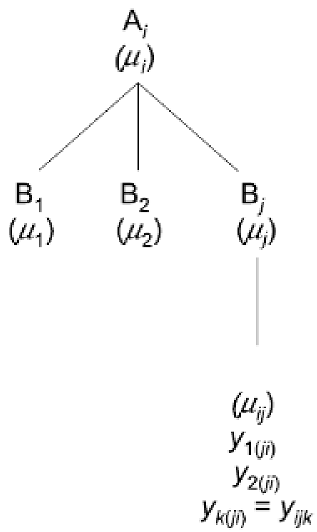
		Factor A (season)	
		Winter	Summer
Factor B (density)	8 per plate	W.8.1...W.8.3	S.8.1...S.8.3
	15 per plate	W.15.1...W.15.3	S.15.1...S.15.3
	30 per plate	W.30.1...W.30.3	S.30.1...S.30.3
	45 per plate	W.30.1...W.30.3	S.45.1...S.45.3

Factorial Design Structure

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)
- n_i replicates (k= 1...n_i) in each combination of A and B (3 replicate plates per density per season)





Calculating Different Means

We can calculate several means:

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

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal m
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

Marginal Means

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

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal m
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

Cell Means

We can calculate several means:

- a mean for each cell of combinations of A and B= μ_{ij}

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal m
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

Factorial ANOVA Models

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:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Model Components Explained

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- y_{ijk}: value of the k_{th} 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)

Interaction Effects

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- Bj: 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?)

Model Types and Interpretation

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- 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?")

ANOVA Table Structure

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

Source	SS	df	MS
A	$nq \sum_{i=1}^{p} \left(\bar{y}_{i.} - \bar{y}\right)^2$	$p-1$	$\frac{SS_A}{p-1}$
B	$np \sum_{j=1}^{q} \left(\bar{y}_{.j} - \bar{y}\right)^2$	$q-1$	$\frac{SS_B}{q-1}$
AB	$n \sum_{i=1}^{p} \sum_{j=1}^{q} \left(\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}\right)^2$	$(p-1)(q-1)$	$\frac{SS_{AB}}{(p-1)(q-1)}$
Residual	$\sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{n} \left(y_{ijk} - \bar{y}_{ij}\right)^2$	$pq(n-1)$	$\frac{SS_{\text{Residual}}}{pq(n-1)}$
Total	$\sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{n} \left(y_{ijk} - \bar{y}\right)^2$	$pqn-1$	

SSA: Factor A Effects

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

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal means
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

SSB: Factor B Effects

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

Source	SS	df	MS
A	$nq\sum_{i=1}^{p}(\bar{y}_i-\bar{y})^2$	$p-1$	$\frac{SS_A}{p-1}$
B	$np\sum_{j=1}^{q}(\bar{y}_j-\bar{y})^2$	$q-1$	$\frac{SS_B}{q-1}$
AB	$n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{y}_{ij}-\bar{y}_i-\bar{y}_j+\bar{y})^2$	$(p-1)(q-1)$	$\frac{SS_{AB}}{(p-1)(q-1)}$
Residual	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij})^2$	$pq(n-1)$	$\frac{SS_{Residual}}{(pq(n-1)}$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk}-\bar{y})^2$	$pqn-1$	

SSB Visualization

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

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal means
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11}=2.417$	$\bar{y}_{12}=1.833$	$\bar{y}_{i=1}=2.125$
A_2	15	$\bar{y}_{21}=2.177$	$\bar{y}_{22}=1.178$	$\bar{y}_{i=2}=1.677$
A_3	30	$\bar{y}_{31}=1.565$	$\bar{y}_{32}=0.811$	$\bar{y}_{i=3}=1.188$
A_4	45	$\bar{y}_{41}=1.200$	$\bar{y}_{42}=0.593$	$\bar{y}_{i=4}=0.896$
Factor B marginal means		$\bar{y}_{j=1}=1.840$	$\bar{y}_{j=2}=1.104$	$\bar{y}=1.472$

SSAB: Interaction Effects

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

Source	SS	df	MS
A	$nq\sum_{i=1}^{p}(\bar{y}_i-\bar{y})^2$	$p-1$	$\frac{SS_A}{p-1}$
B	$np\sum_{j=1}^{q}(\bar{y}_j-\bar{y})^2$	$q-1$	$\frac{SS_B}{q-1}$
AB	$n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{y}_{ij}-\bar{y}_i-\bar{y}_j+\bar{y})^2$	$(p-1)(q-1)$	$\frac{SS_{AB}}{(p-1)(q-1)}$
Residual	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij})^2$	$pq(n-1)$	$\frac{SS_{Residual}}{(pq(n-1)}$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk}-\bar{y})^2$	$pqn-1$	

SSAB Visualization

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

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal means
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

SSresid: Residual Variance

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

Source	SS	df	MS
A	$nq\sum_{i=1}^{p}(\bar{y}_i - \bar{y})^2$	$p-1$	$\frac{SS_A}{p-1}$
B	$np\sum_{j=1}^{q}(\bar{y}_j - \bar{y})^2$	$q-1$	$\frac{SS_B}{q-1}$
AB	$n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2$	$(p-1)(q-1)$	$\frac{SS_{AB}}{(p-1)(q-1)}$
Residual	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij})^2$	$pq(n-1)$	$\frac{SS_{Residual}}{(pq(n-1)}$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y})^2$	$pqn-1$	

Total Sum of Squares

SStotal = SSA + SSB + SSAB + SSresid

Source	SS	df	MS
A	$nq\sum_{i=1}^{p}(\bar{y}_i - \bar{y})^2$	$p-1$	$\frac{SS_A}{p-1}$
B	$np\sum_{j=1}^{q}(\bar{y}_j - \bar{y})^2$	$q-1$	$\frac{SS_B}{q-1}$
AB	$n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2$	$(p-1)(q-1)$	$\frac{SS_{AB}}{(p-1)(q-1)}$
Residual	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij})^2$	$pq(n-1)$	$\frac{SS_{Residual}}{(pq(n-1)}$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y})^2$	$pqn-1$	

F-ratio Calculations

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	$\frac{MS_A}{MS_{Residual}}$	$\frac{MS_A}{MS_{AB}}$	$\frac{MS_A}{MS_{AB}}$

Source	A and B fixed	A and B random	A fixed, B random
B	$\frac{MS_B}{MS_{Residual}}$	$\frac{MS_B}{MS_{AB}}$	$\frac{MS_B}{MS_{AB}}$
AB	$\frac{MS_{AB}}{MS_{Residual}}$	$\frac{MS_{AB}}{MS_{Residual}}$	$\frac{MS_{AB}}{MS_{Residual}}$

Hypotheses: Fixed Factors

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)

Hypotheses: Random Factors

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.

Hypotheses: Mixed Model

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)

Example Study Details

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 cm^2 enclosure) on the production of egg masses by inter-tidal pulmonate limpets (*Siphonaria diemenensis*) as described in Quinn (1988).

		Factor A (season)	
		Winter	Summer
Factor B (density)	8 per plate	W.8.1...W.8.3	S.8.1...S.8.3
	15 per plate	W.15.1...W.15.3	S.15.1...S.15.3
	30 per plate	W.30.1...W.30.3	S.30.1...S.30.3
	45 per plate	W.30.1...W.30.3	S.45.1...S.45.3

Data Setup and Overview

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
```

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

Fit the model

```r
# 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
```

```r
# Store residuals for diagnostics
quinn_data$residuals <- residuals(quinn_model_lm)
quinn_data$fitted <- fitted(quinn_model_lm)
```

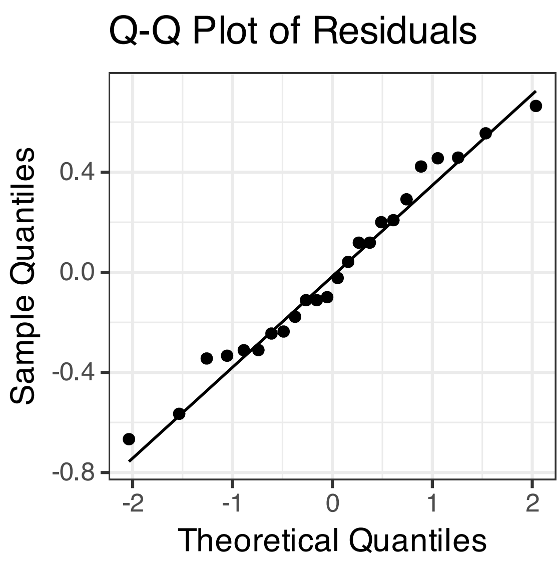
```r
# For backward compatibility with later code
quinn_model <- aov(quinn_model_lm)

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
```

Check for Normality of Residuals: Q-Q Plot

Q-Q Plot of Residuals

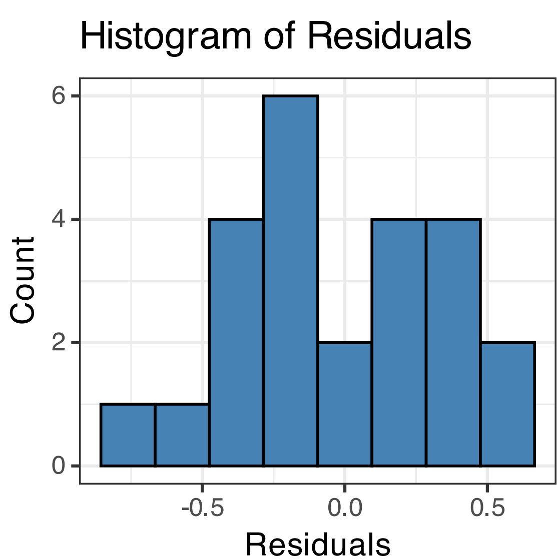
```r
# Create Q-Q plot of residuals
ggplot(quinn_data, aes(sample = residuals)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "Q-Q Plot of Residuals",
       x = "Theoretical Quantiles",
       y = "Sample Quantiles")
```



Check for Normality of Residuals: Histogram

Histogram 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")
```



Check for Normality of Residuals: Shapiro-Wilk Test

Shapiro-Wilk Test for Normality

```
# 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
```

Check for Homogeneity of Variances: Levene's Test

Levene's Test for Homogeneity of Variances

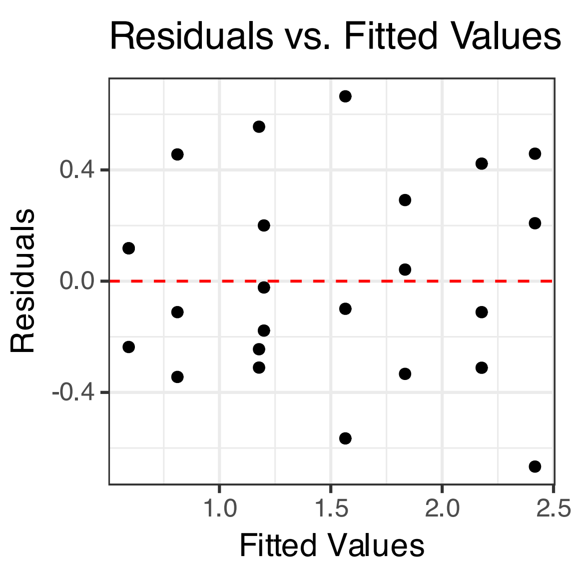
```
# Levene's test for homogeneity of variances
leveneTest(EGGS ~ DENSITY * SEASON, data = quinn_data)
```

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

Check for Homogeneity of Variances: Residuals Plot

Residuals vs. Fitted Values

```
# 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")
```

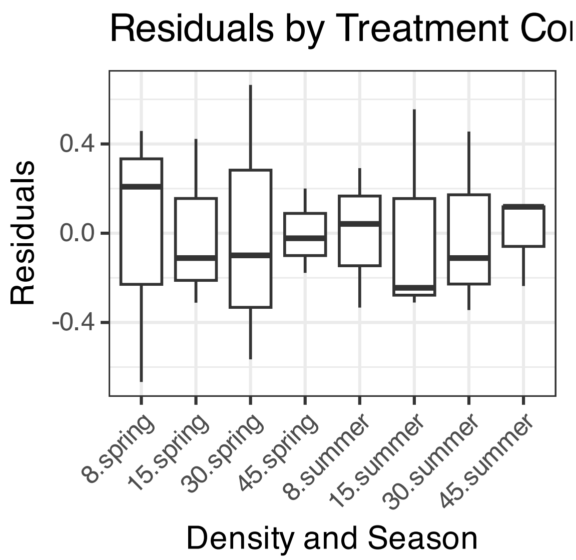


Check for Homogeneity of Variances: Boxplot

Residuals by Treatment Combination

```
# 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")
```



Factorial ANOVA Results

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
```

Polynomial and Quadratic Contrasts

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
```

Estimated Marginal Means: Density Effect

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)
```

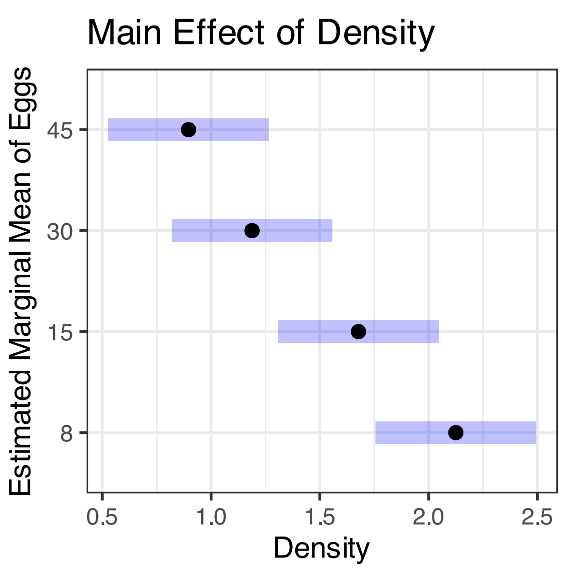
```
 contrast              estimate    SE df t.ratio p.value
 DENSITY8 - DENSITY15     0.448 0.246 16   1.816  0.3021
```

```
 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
```

Density Effect Visualization

Main Effect of Density Plot

```
density_plot <- plot(density_emm, xlab = "Density", ylab = "Estimated Marginal Mean of Eggs")
+
  ggtitle("Main Effect of Density") +
  theme_bw()
density_plot
```



Estimated Marginal Means: Season Effect

Main Effect of Season

```
# 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
```

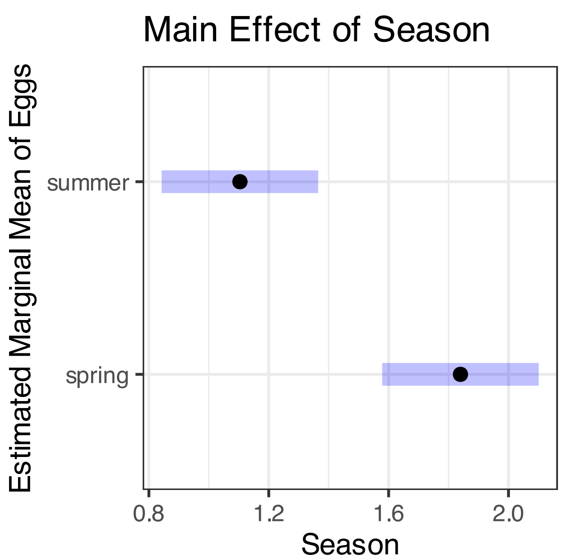
```
pairs(season_emm)
```

```
 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
```

Season Effect Visualization

Main Effect of Season Plot

```
# 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
```



Interaction Effects Analysis

Interaction Effects and Raw Means Comparison

```
# 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
```

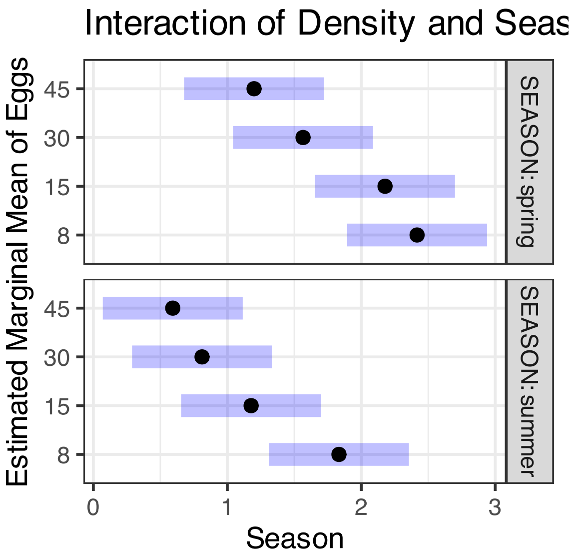
```
# Compare to raw means
quinn_data %>%
  group_by(DENSITY, SEASON) %>%
  summarise(
    raw_mean = mean(EGGS),
    .groups = 'drop'
  ) %>%
  pivot_wider(names_from = SEASON, values_from = raw_mean)
```

```
# 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
```

Interaction Plot: Standard

Standard Interaction Plot

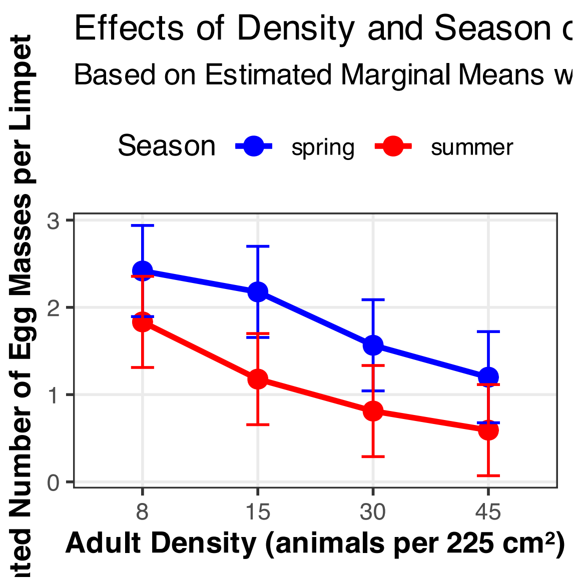
```
# Interaction effects (even though interaction wasn't significant)
interaction_plot <- plot(interaction_emm, xlab = "Season", ylab = "Estimated Marginal Mean of
Eggs") +
  ggtitle("Interaction of Density and Season") +
  theme_bw()

interaction_plot
```



Interaction Plot: Custom

Custom Interaction Plot with Error Bars

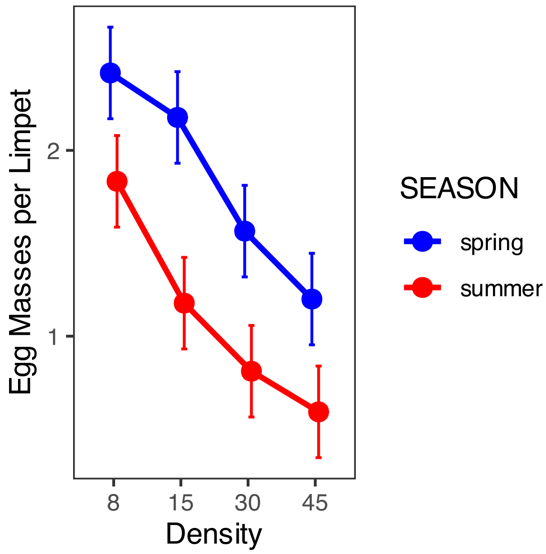
```
# Alternative approach using ggplot2 for more customization
# Convert emmeans object to data frame
interaction_df <- as.data.frame(interaction_emm)

# 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
```



Publication-Quality Plot

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
```



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.

Results Interpretation: Effect Sizes

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:

4. **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

Results Interpretation: Model Performance

5. **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.

6. **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.

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.