Lecture 15 - ANCOVA

Bill Perry

# Lecture 14: Review

## Review

General Linearized Models (GLM)

* Gaussian GLMs (normal distribution)
* Poisson GLMs (count data)
* Logistic GLMs (binary outcomes)
* Model assumptions and diagnostics
* Connection to ANOVA and linear models

# Lecture 15: ANCOVA Overview

## Overview

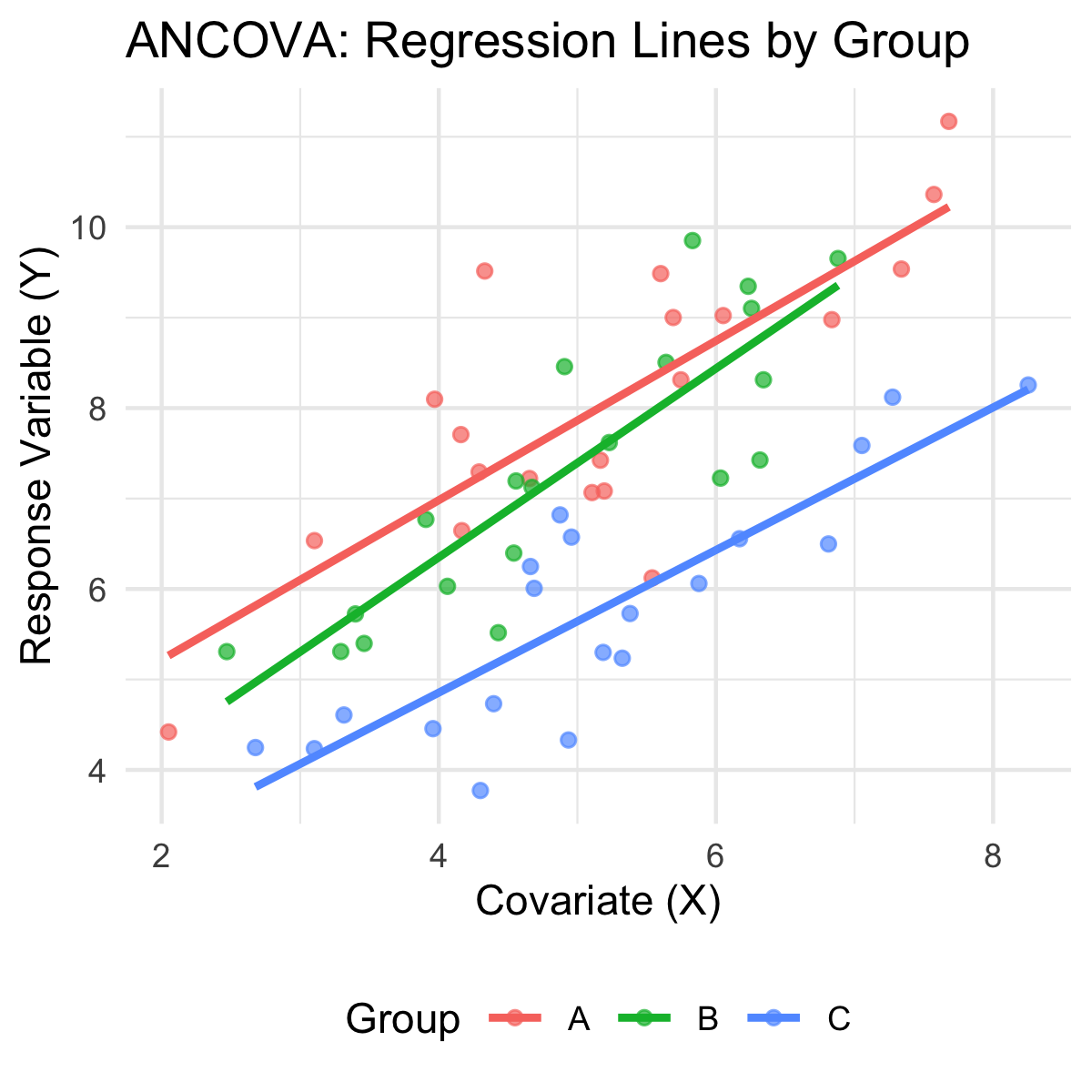
Analysis of covariance (ANCOVA):

* Introduction to ANCOVA
* When to use ANCOVA
* Linear model for ANCOVA
* Analysis of variance in ANCOVA
* Assumptions of ANCOVA
  + Homogeneous slopes
* Robust ANCOVA approaches
* Specific comparisons of means
* Examples and interpretation
* Scientific reporting of ANCOVA results

# Introduction to ANCOVA

## What is ANCOVA?

* ANCOVA = Analysis of COVAriance
* Combination of regression and ANOVA
* A continuous covariate is measured along with the response variable for each experimental unit
* Common use: compare means of factor levels (groups), adjusting for variance from continuous covariate
* Another use: determine whether two or more regression lines differ in slopes and intercepts



# When to Use ANCOVA

## Common Applications of ANCOVA

* **Increasing statistical power**
  + Removing variation associated with a covariate can reduce residual error
  + More powerful test of treatment effects
* **Adjusting for confounding variables**
  + When treatments might differ in some continuous variable
  + Need to separate treatment effects from covariate effects
* **Testing equality of regression lines**
  + Do treatments have the same relationship with a continuous variable?
  + Tests for both slopes and intercepts

# ANCOVA Example: Cricket Chirping

## Cricket Chirping Example

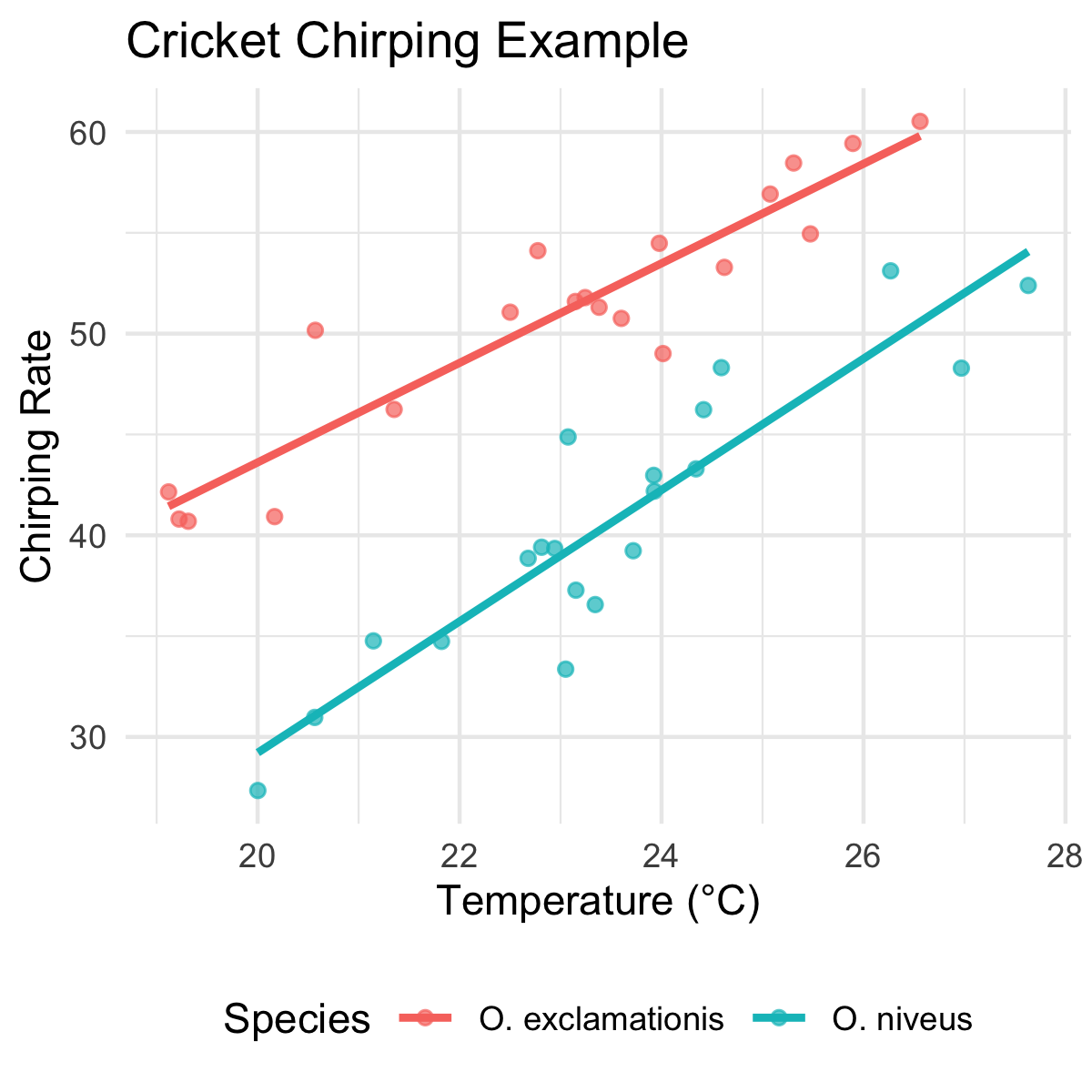
Want to compare chirping rate of two cricket species:

* *Oecanthus exclamationis*
* *Oecanthus niveus*

But:

* Measured rates at different temperatures
* Range of temperatures differed between species
* Apparent relationship between pulse rate and temperature

ANCOVA lets us adjust for temperature effect to get a more powerful test!



# ANCOVA Linear Model: Conceptual Framework

## The ANCOVA Model

Key concept in ANCOVA: the difference between “unadjusted” group means and “adjusted” means.

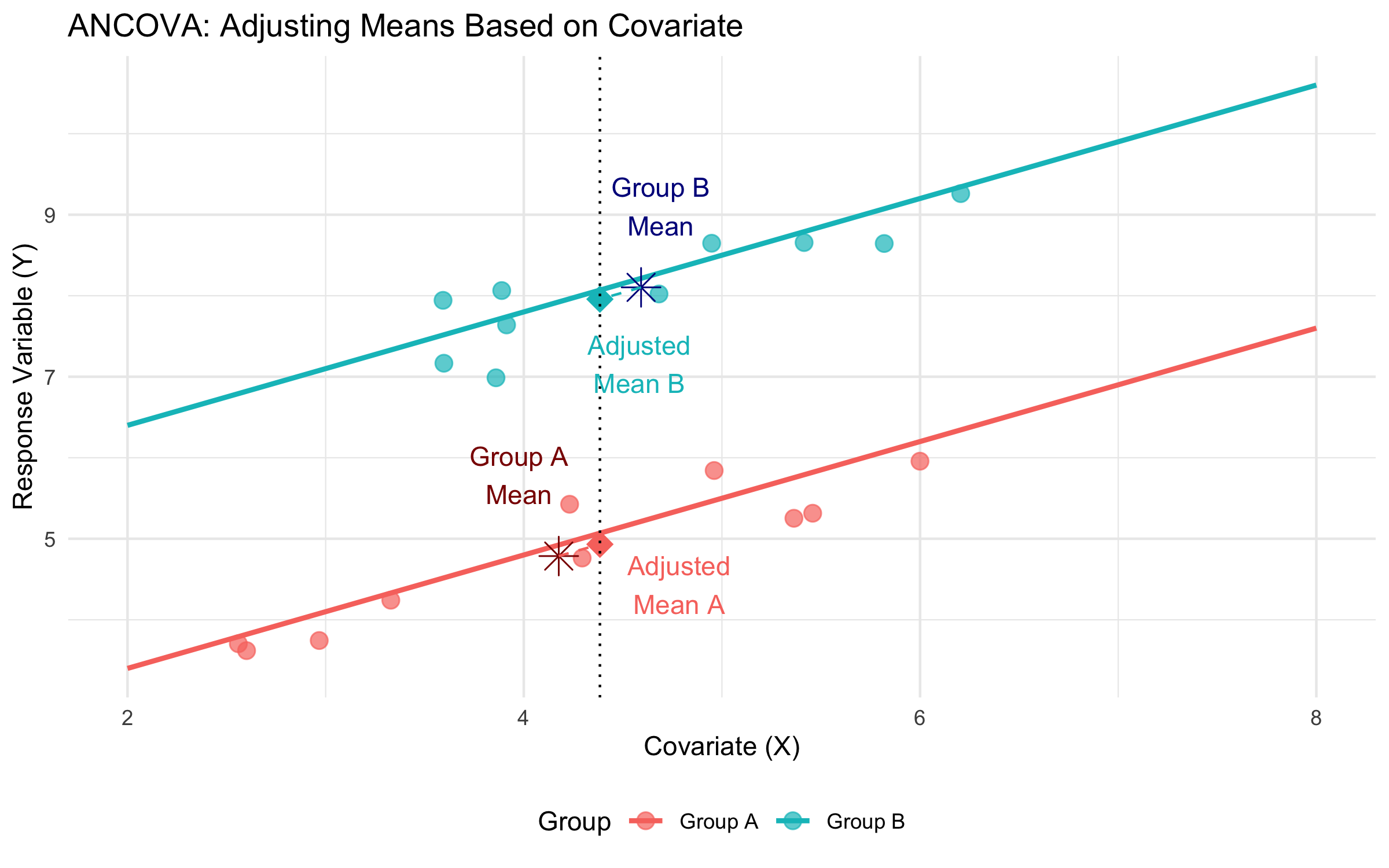
In this visualization:

* Group Means (shown as asterisks): raw/unadjusted means for each group - simply the average X value and average Y value for all points in that group. Notice that Group A and Group B have different mean X values (they’re positioned at different points along the X axis).
* Adjusted Means (shown as triangles): These are what ANCOVA actually compares. The adjusted means represent what each group’s mean would be if all groups had the same value of the covariate (in this case, the overall mean X).

The core purpose of ANCOVA is to make this adjustment. This is important because:

* When groups differ in their covariate values (as they often do in observational studies or even in experiments with random assignment), comparing raw means can be misleading
* The adjustment helps “level the playing field” by estimating what each group’s mean would be if they all had the same value of the covariate

# ANCOVA Model Visualization



# Mathematical Model for ANCOVA

For a single-factor ANCOVA with factor A (p levels, i = 1 to p), a continuous covariate (x), and response variable (y):

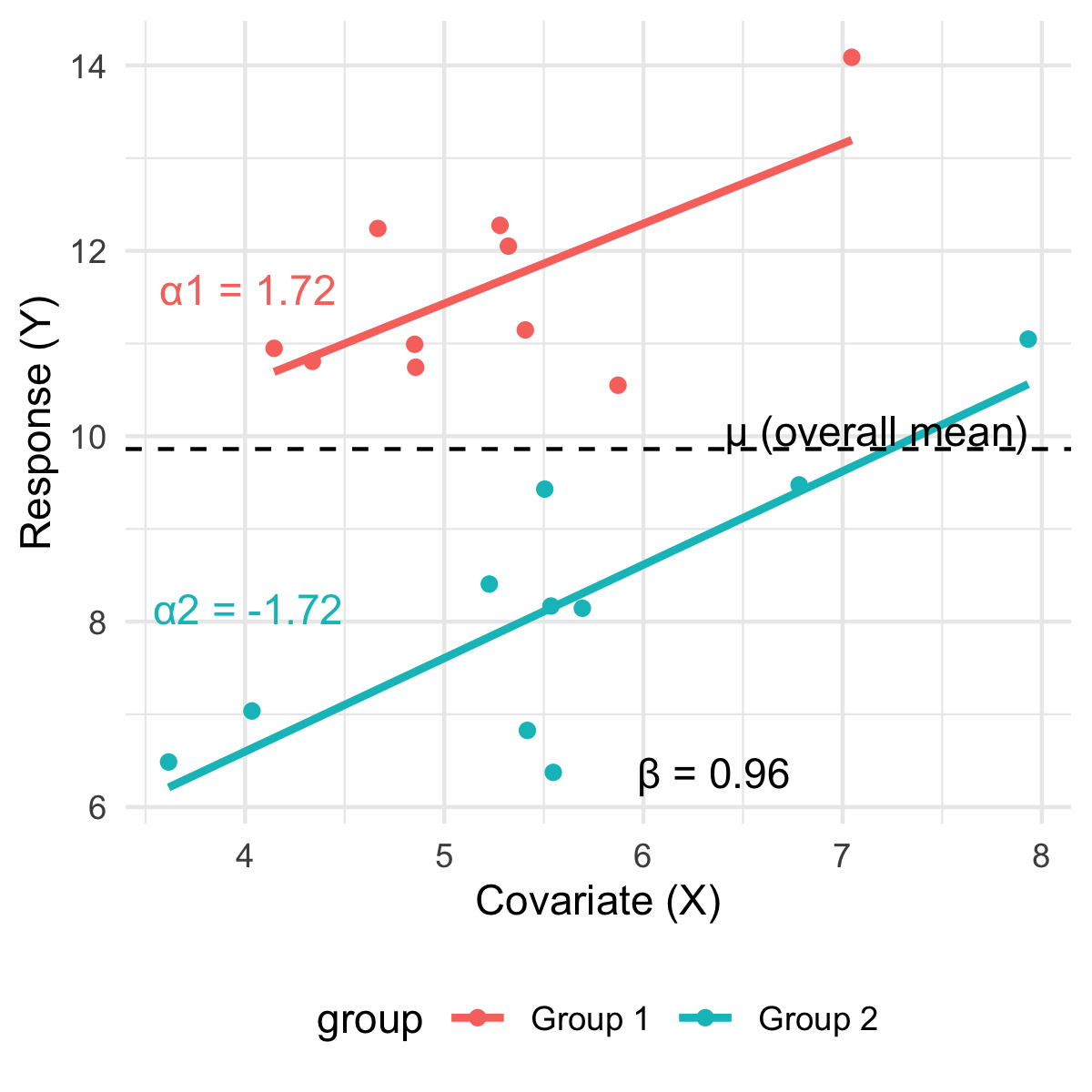
Where: - = response value for observation j in level i of factor A - = overall mean - = effect of level i of factor A - = common regression slope relating Y to X - = covariate value for observation j in level i of factor A - = mean value of covariate - = error term

# ANCOVA Parameters Interpretation

## Interpretation of Parameters

* = overall mean response
* = effect of level i (difference between group mean and overall mean)
* = pooled within-group regression coefficient
* = covariate value for observation j in group i
* = overall mean of covariate
* = unexplained error

This model assumes **homogeneous slopes** across all treatment groups (we’ll test this later).



# ANCOVA in R: Basic Implementation

## Running ANCOVA in R

Basic ANCOVA model: - Response: continuous variable (y) - Predictor: categorical factor (A)  
- Covariate: continuous variable (x)

The simplest ANCOVA model is:

model <- lm(y ~ A + x, data = mydata)  
anova(model)

Alternative: use aov() function

model <- aov(y ~ A + x, data = mydata)   
summary(model)

Both approaches use Type I SS (sequential). For unbalanced designs, you may want Type III SS using car package.

# Load the partridge dataset  
partridge <- read.csv("data/partridge.csv")  
  
# Look at the first few rows  
head(partridge)

PARTNERS TYPE TREATMEN LONGEV LLONGEV THORAX RESID1 PREDICT1 RESID2  
1 8 0 1 35 1.544068 0.64 -5.868456 40.86846 -0.04743024  
2 8 0 1 37 1.568202 0.68 -9.301196 46.30120 -0.07105067  
3 8 0 1 49 1.690196 0.68 2.698804 46.30120 0.05094369  
4 8 0 1 46 1.662758 0.72 -5.733936 51.73394 -0.02424867  
5 8 0 1 63 1.799341 0.72 11.266064 51.73394 0.11233405  
6 8 0 1 39 1.591065 0.76 -18.166676 57.16668 -0.14369601  
 PREDICT2  
1 1.591498  
2 1.639252  
3 1.639252  
4 1.687007  
5 1.687007  
6 1.734761

# Create factors for treatment  
partridge$TREATMEN <- as.factor(partridge$TREATMEN)  
  
# Basic ANCOVA model  
model1 <- lm(LONGEV ~ THORAX + TREATMEN,   
 data = partridge)  
  
# View ANOVA table with Type I SS  
anova(model1)

Analysis of Variance Table  
  
Response: LONGEV  
 Df Sum Sq Mean Sq F value Pr(>F)   
THORAX 1 15496.6 15496.6 140.293 < 2.2e-16 \*\*\*  
TREATMEN 4 9611.5 2402.9 21.753 1.719e-13 \*\*\*  
Residuals 119 13144.7 110.5   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# ANCOVA in R: Type III Sum of Squares

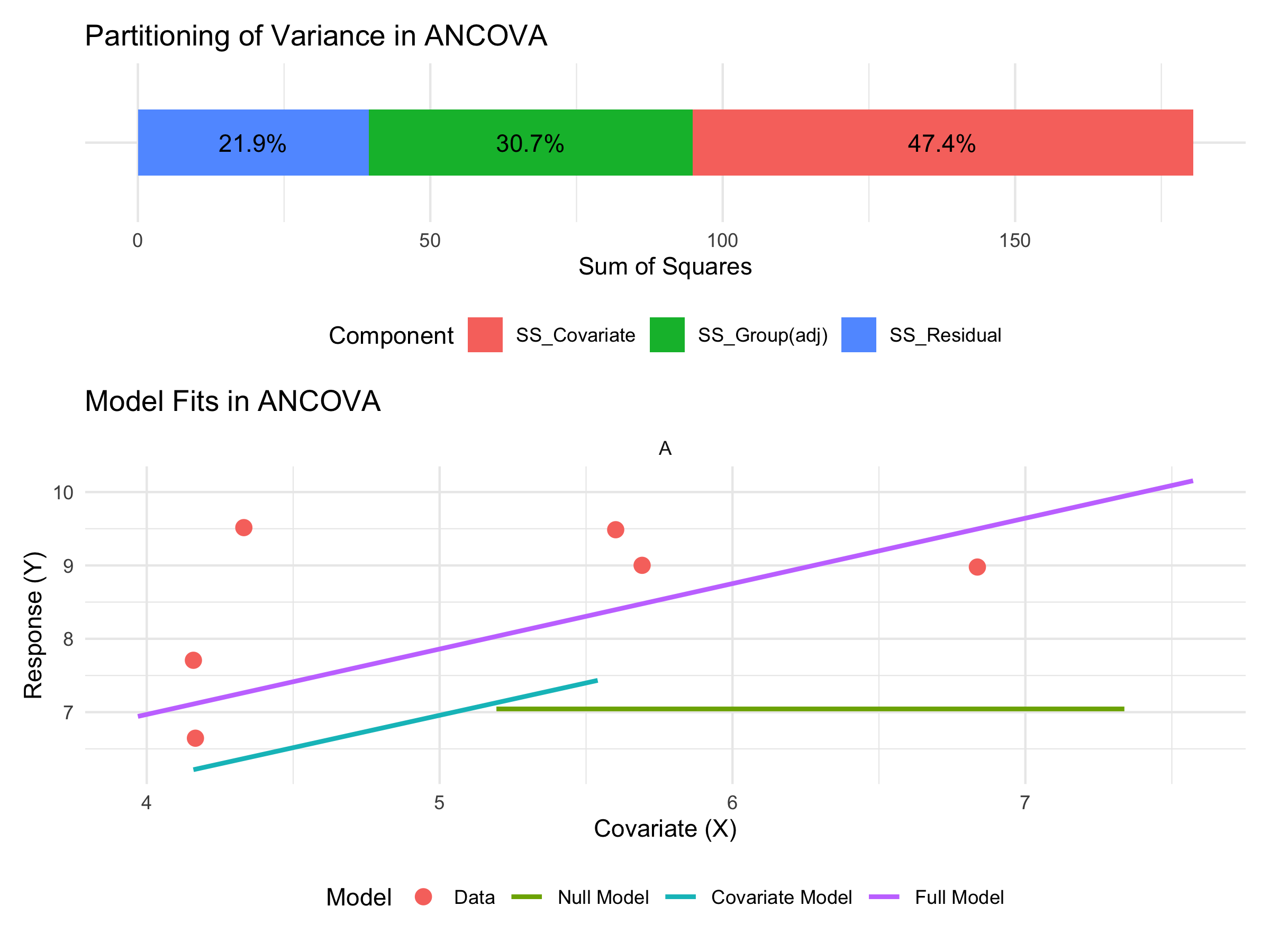
Using Type III SS from car package:

# Using Type III SS from car package  
model2 <- lm(LONGEV ~ TREATMEN + THORAX,   
 data = partridge)  
Anova(model2, type = "III")

Anova Table (Type III tests)  
  
Response: LONGEV  
 Sum Sq Df F value Pr(>F)   
(Intercept) 2234.9 1 20.233 1.605e-05 \*\*\*  
TREATMEN 9611.5 4 21.753 1.719e-13 \*\*\*  
THORAX 13168.9 1 119.219 < 2.2e-16 \*\*\*  
Residuals 13144.7 119   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The Type III approach is often preferred for: - Unbalanced designs - When you want to test each effect adjusted for all others - More conservative approach when groups differ in covariate values

# Analysis of Variance for ANCOVA: Partitioning



# ANOVA Table for ANCOVA

The ANOVA table for a single-factor ANCOVA has these components:

| **Source** | **df** | **Sum of Squares** | **Mean Square** | **F-ratio** | **Expected MS** |
| --- | --- | --- | --- | --- | --- |
| Factor A (adjusted) | (p-1) | SS\_A(adj) | MS\_A(adj) = SS\_A(adj)/(p-1) | MS\_A(adj)/MS\_Residual | σ² + n∑α²/(p-1) |
| Covariate | 1 | SS\_Covariate | MS\_Covariate = SS\_Covariate/1 | MS\_Covariate/MS\_Residual | σ² + β²∑(X-X̄)² |
| Residual | n-p-1 | SS\_Residual | MS\_Residual = SS\_Residual/(n-p-1) |  | σ² |
| Total | n-1 | SS\_Total |  |  |  |

# Null Hypotheses in ANCOVA

1. **Treatment Effect (adjusted for covariate)**
   * Are the adjusted group means equal?
   * Test with F = MS\_A(adj)/MS\_Residual
2. **Covariate Effect**
   * Is there a relationship between the covariate and the response?
   * Test with F = MS\_Covariate/MS\_Residual
3. **Homogeneity of Slopes** (test this first!)
   * Are the regression slopes the same for all groups?
   * Test by adding group\*covariate interaction term

# Testing Homogeneity of Slopes

## Testing for Homogeneous Slopes

ANCOVA assumes the regression slopes are the same for all groups (parallel regression lines)

To test this assumption:

1. Fit model with interaction term:

model\_int <- lm(y ~ A \* x, data = mydata)

1. Test significance of interaction:

anova(model\_int)

1. If interaction is significant (p < 0.05):
   * Slopes are not homogeneous
   * Standard ANCOVA inappropriate
   * Use alternative approaches

# Test homogeneity of slopes in partridge data  
# by adding interaction term  
model\_int <- lm(LONGEV ~ TREATMEN \* THORAX,   
 data = partridge)  
  
# Test significance of interaction  
anova(model\_int)

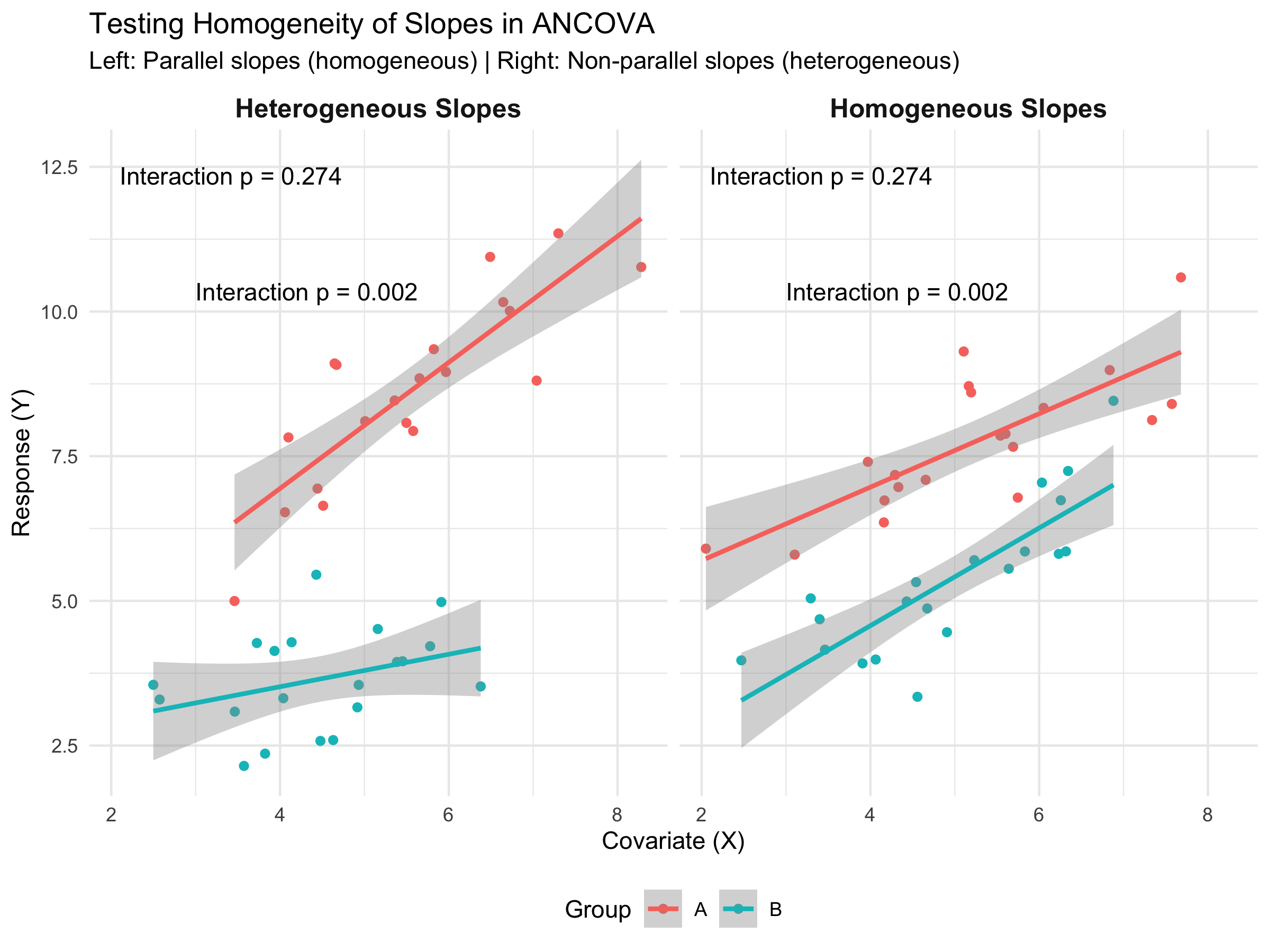
Analysis of Variance Table  
  
Response: LONGEV  
 Df Sum Sq Mean Sq F value Pr(>F)   
TREATMEN 4 11939.3 2984.8 26.1983 1.896e-15 \*\*\*  
THORAX 1 13168.9 13168.9 115.5855 < 2.2e-16 \*\*\*  
TREATMEN:THORAX 4 42.5 10.6 0.0933 0.9844   
Residuals 115 13102.1 113.9   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Extract p-value for interaction  
interaction\_pvalue <- anova(model\_int)[3, "Pr(>F)"]

The p-value for the interaction is 0.984. Since p > 0.05, we can proceed with standard ANCOVA (assuming homogeneous slopes).

# Visualization of Homogeneity of Slopes

## Parallel vs. Non-Parallel Slopes



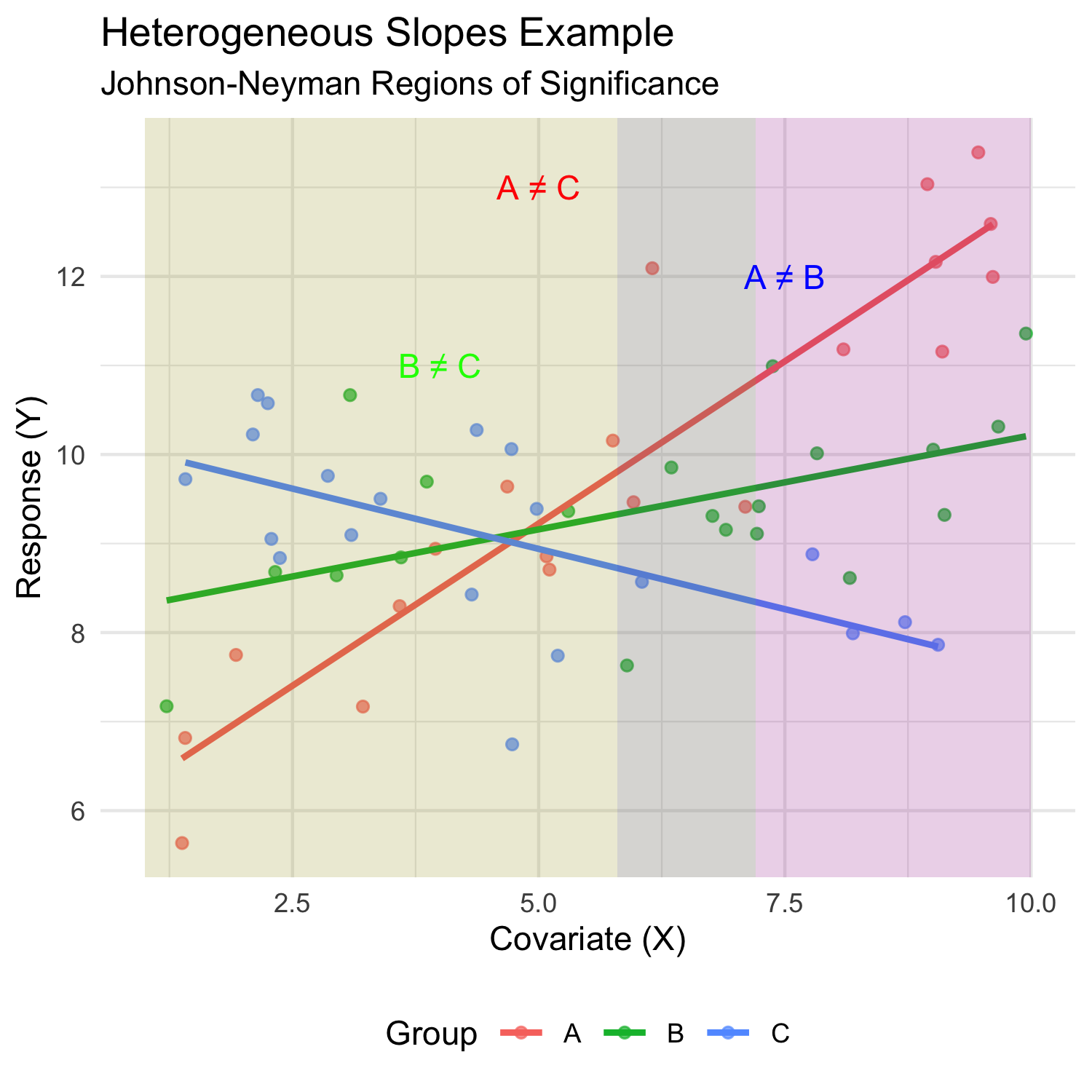
# Handling Heterogeneous Slopes

## When Slopes Are Not Homogeneous

If the interaction term is significant (p < 0.05), the slope-group relationship is not the same across groups.

Options:

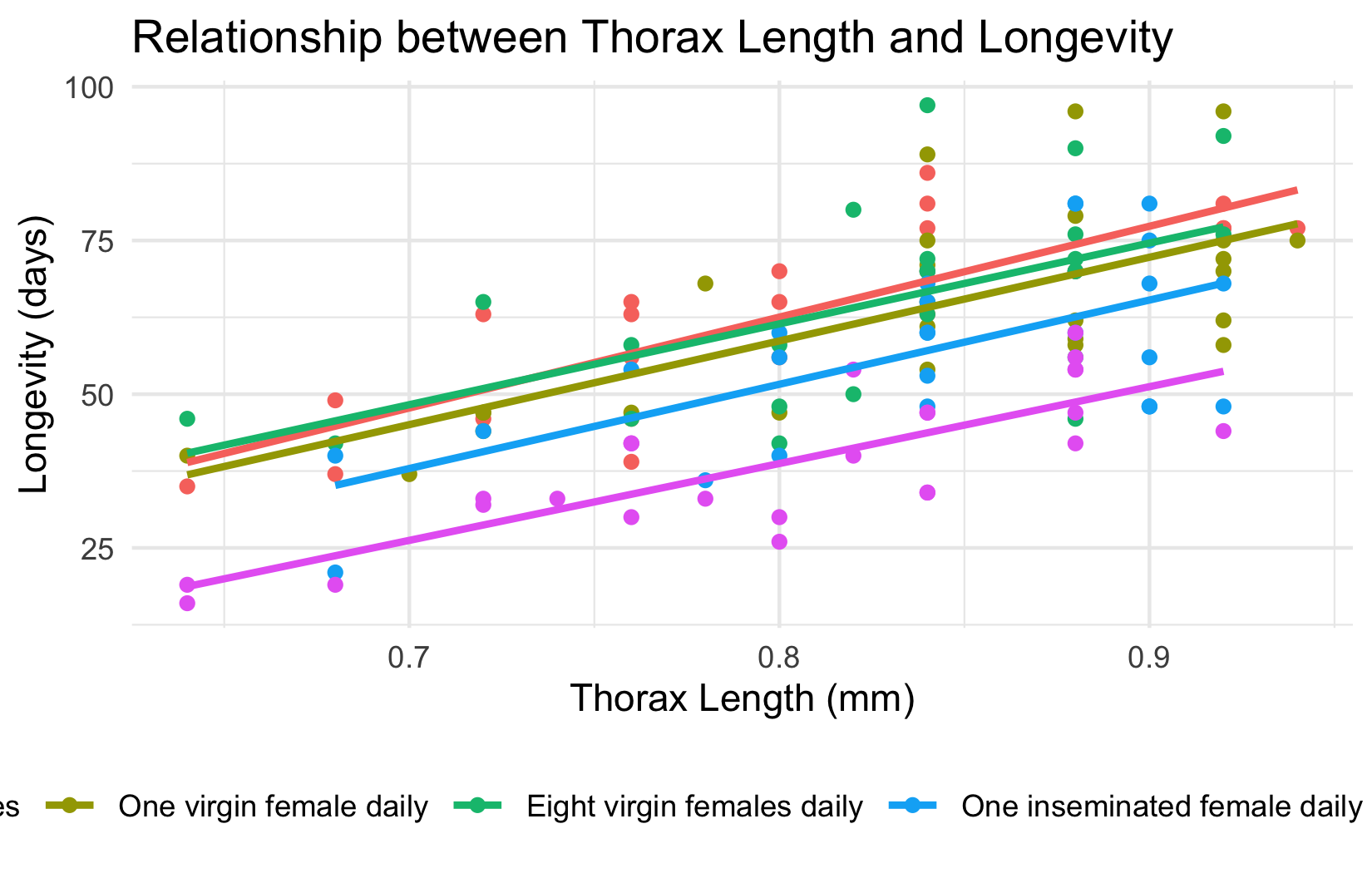
1. **Report the interaction** - this is a biologically interesting result!
2. **Separate regressions** - analyze each group separately
3. **Johnson-Neyman procedure** - identifies regions of the covariate where groups differ significantly
4. **Alternative models** - consider transformation, polynomial terms, or more complex models



# Partridge Example: Data Overview

## ANCOVA on Longevity of Male Fruitflies

# 1. Examine the partridge data structure  
# str(partridge)  
  
# Create better names for treatments  
partridge$treatment <- factor(partridge$TREATMEN,  
 levels = 1:5,  
 labels = c("No females",   
 "One virgin female daily",  
 "Eight virgin females daily",  
 "One inseminated female daily",  
 "Eight inseminated females daily"))  
  
# 2. Create a plot of the data showing relationship  
ggplot(partridge, aes(x = THORAX, y = LONGEV, color = treatment)) +   
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE) +  
 labs(title = "Relationship between Thorax Length and Longevity",  
 x = "Thorax Length (mm)",  
 y = "Longevity (days)",  
 color = "Treatment") +  
 theme\_minimal() +  
 theme(legend.position = "bottom")



# Partridge Example: Testing Homogeneity

## Testing Homogeneity of Slopes

# Test for homogeneity of slopes  
homo\_slopes\_model <- lm(LONGEV ~ THORAX \* treatment, data = partridge)  
anova(homo\_slopes\_model)

Analysis of Variance Table  
  
Response: LONGEV  
 Df Sum Sq Mean Sq F value Pr(>F)   
THORAX 1 15496.6 15496.6 136.0170 < 2.2e-16 \*\*\*  
treatment 4 9611.5 2402.9 21.0905 4.617e-13 \*\*\*  
THORAX:treatment 4 42.5 10.6 0.0933 0.9844   
Residuals 115 13102.1 113.9   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Extract the p-value for the interaction  
p\_interaction <- anova(homo\_slopes\_model)[3, "Pr(>F)"]

The p-value for the interaction term (treatment × THORAX) is 0.984. Since this value is > 0.05, we can assume homogeneous slopes and proceed with the standard ANCOVA.

# Partridge Example: Full ANCOVA Analysis

# Fit the ANCOVA model (without interaction)  
ancova\_model <- lm(LONGEV ~ THORAX + treatment, data = partridge)  
  
# View ANOVA table  
anova(ancova\_model)

Analysis of Variance Table  
  
Response: LONGEV  
 Df Sum Sq Mean Sq F value Pr(>F)   
THORAX 1 15496.6 15496.6 140.293 < 2.2e-16 \*\*\*  
treatment 4 9611.5 2402.9 21.753 1.719e-13 \*\*\*  
Residuals 119 13144.7 110.5   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Get more detailed summary  
summary(ancova\_model)

Call:  
lm(formula = LONGEV ~ THORAX + treatment, data = partridge)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-26.189 -6.599 -0.989 6.408 30.244   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)  
(Intercept) -46.055 10.239 -4.498 1.61e-05  
THORAX 135.819 12.439 10.919 < 2e-16  
treatmentOne virgin female daily -3.929 2.997 -1.311 0.192347  
treatmentEight virgin females daily -1.276 2.983 -0.428 0.669517  
treatmentOne inseminated female daily -10.946 2.999 -3.650 0.000391  
treatmentEight inseminated females daily -23.879 2.973 -8.031 7.83e-13  
   
(Intercept) \*\*\*  
THORAX \*\*\*  
treatmentOne virgin female daily   
treatmentEight virgin females daily   
treatmentOne inseminated female daily \*\*\*  
treatmentEight inseminated females daily \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 10.51 on 119 degrees of freedom  
Multiple R-squared: 0.6564, Adjusted R-squared: 0.6419   
F-statistic: 45.46 on 5 and 119 DF, p-value: < 2.2e-16

# Get adjusted means using emmeans  
adjusted\_means <- emmeans(ancova\_model, "treatment")  
adjusted\_means

treatment emmean SE df lower.CL upper.CL  
 No females 65.4 2.11 119 61.3 69.6  
 One virgin female daily 61.5 2.11 119 57.3 65.7  
 Eight virgin females daily 64.2 2.10 119 60.0 68.3  
 One inseminated female daily 54.5 2.11 119 50.3 58.7  
 Eight inseminated females daily 41.6 2.12 119 37.4 45.8  
  
Confidence level used: 0.95

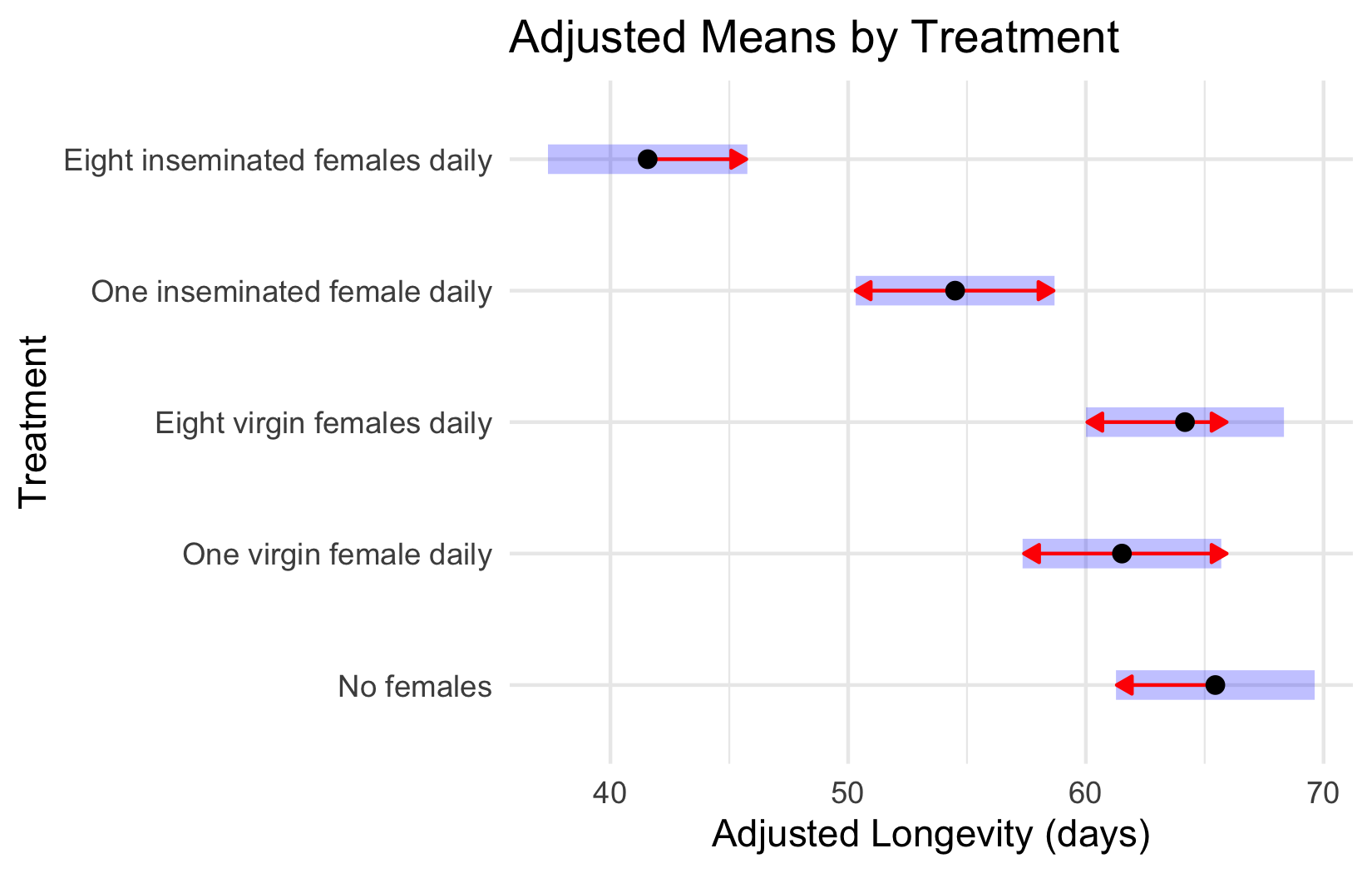
# Partridge Example: Pairwise Comparisons

## Pairwise Comparisons of Adjusted Means

# Pairwise comparisons of adjusted means  
pairs(adjusted\_means, adjust = "tukey")

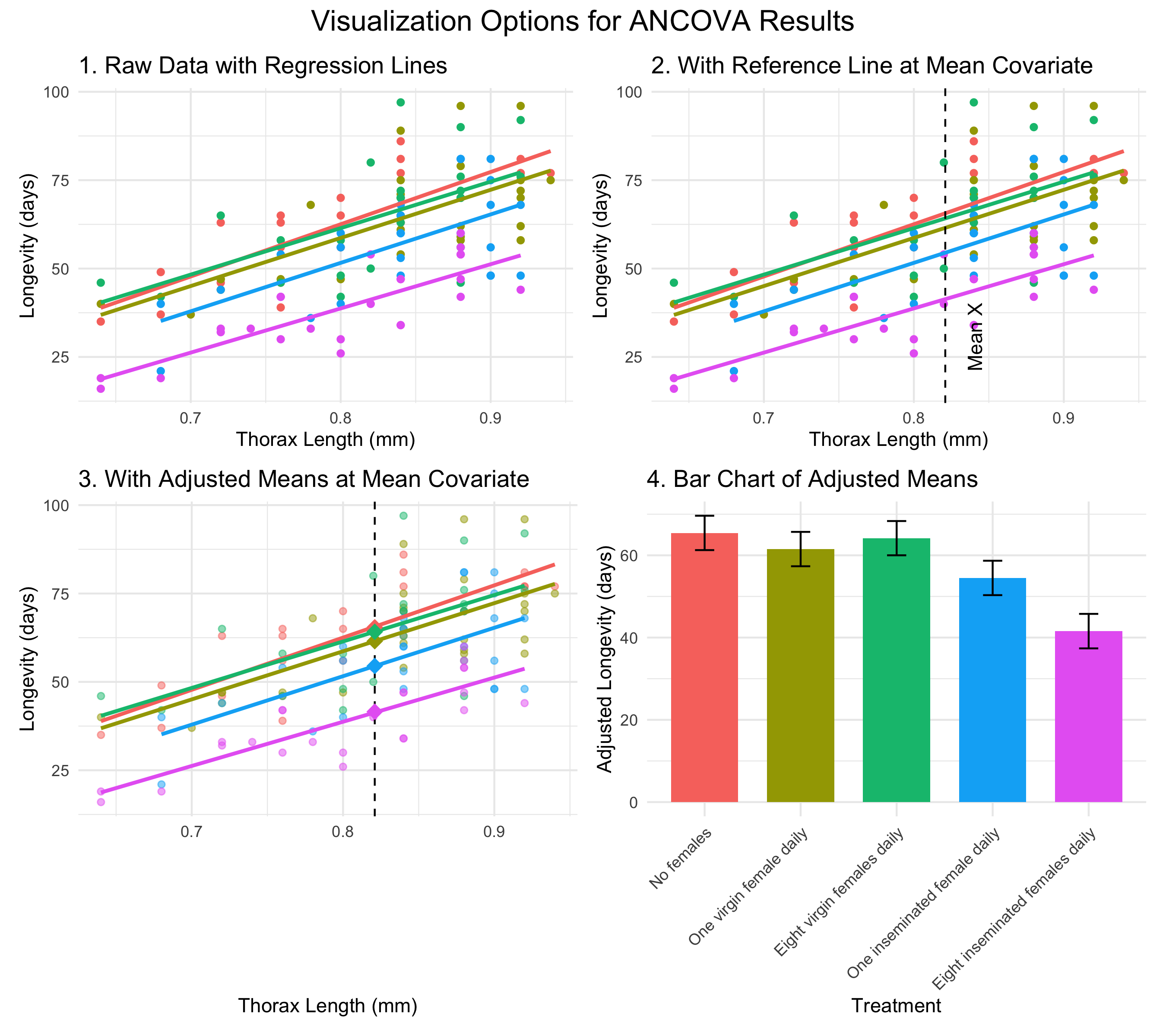
contrast estimate SE  
 No females - One virgin female daily 3.93 3.00  
 No females - Eight virgin females daily 1.28 2.98  
 No females - One inseminated female daily 10.95 3.00  
 No females - Eight inseminated females daily 23.88 2.97  
 One virgin female daily - Eight virgin females daily -2.65 2.98  
 One virgin female daily - One inseminated female daily 7.02 2.97  
 One virgin female daily - Eight inseminated females daily 19.95 3.01  
 Eight virgin females daily - One inseminated female daily 9.67 2.98  
 Eight virgin females daily - Eight inseminated females daily 22.60 2.99  
 One inseminated female daily - Eight inseminated females daily 12.93 3.01  
 df t.ratio p.value  
 119 1.311 0.6849  
 119 0.428 0.9929  
 119 3.650 0.0035  
 119 8.031 <.0001  
 119 -0.891 0.8996  
 119 2.361 0.1336  
 119 6.636 <.0001  
 119 3.249 0.0129  
 119 7.560 <.0001  
 119 4.298 0.0003  
  
P value adjustment: tukey method for comparing a family of 5 estimates

# Plot adjusted means with confidence intervals  
plot(adjusted\_means, comparisons = TRUE) +   
 labs(title = "Adjusted Means by Treatment",  
 x = "Adjusted Longevity (days)",  
 y = "Treatment") +  
 theme\_minimal()



# Visualizing ANCOVA Results

## Visualization Options for ANCOVA

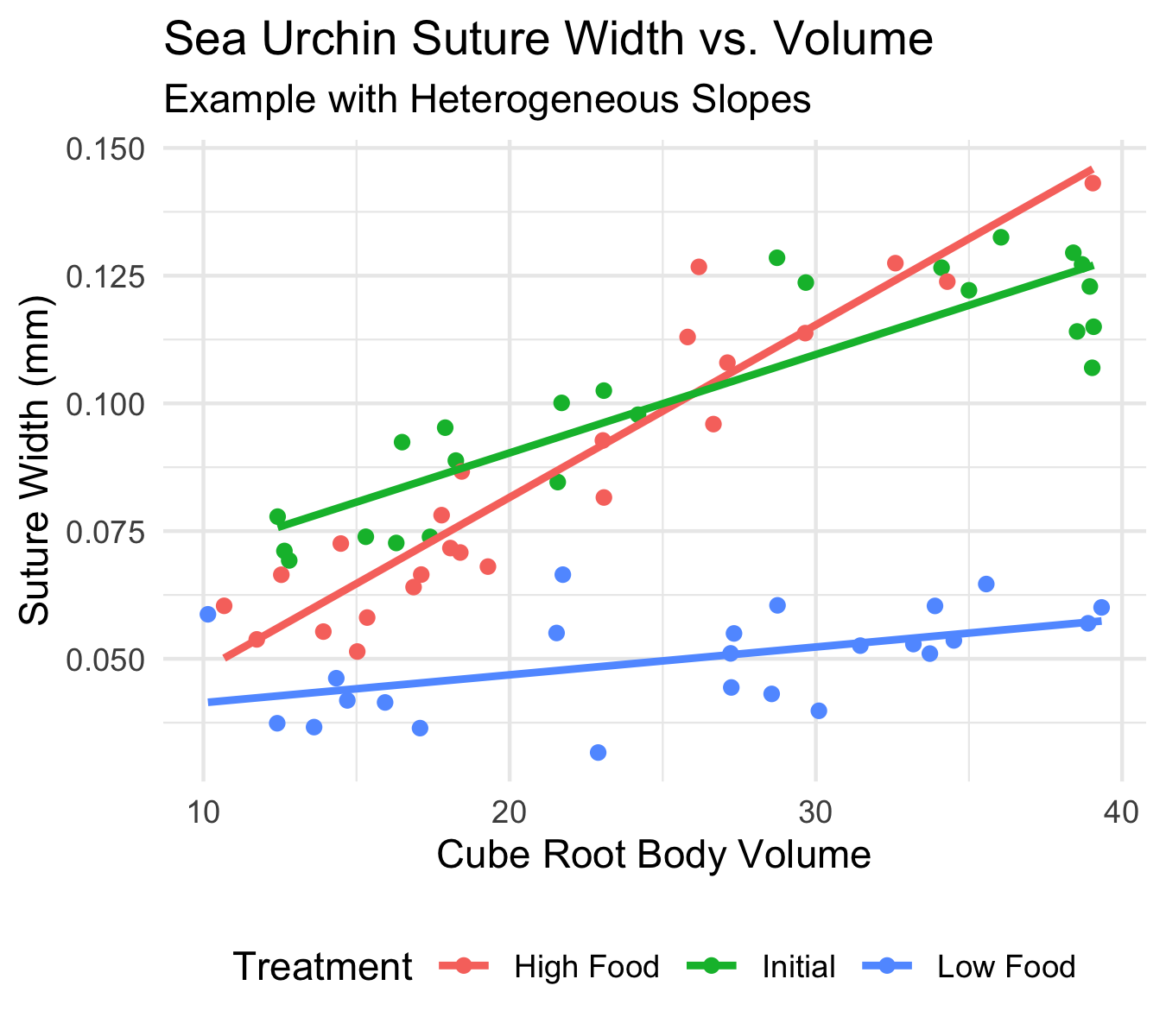


# Sea Urchin Example: Heterogeneous Slopes

## Heterogeneous Slopes Example

Constable (1993) studied shrinking in sea urchin test: - Compared suture widths between treatments - Three groups: high food, low food, initial sample - Covariate: body volume (cube root transformed)

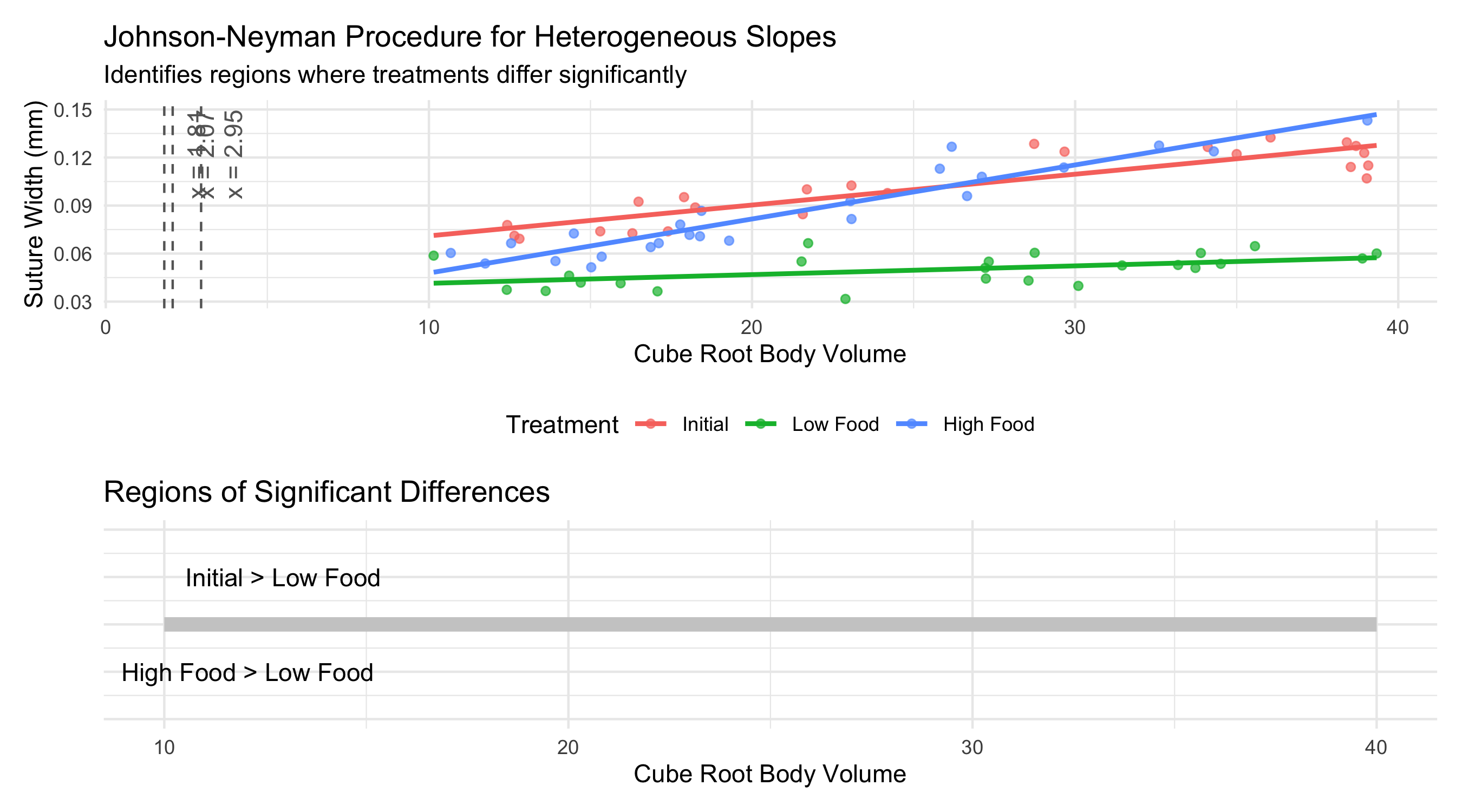
The analysis showed: - Significant interaction between treatment and covariate - **Heterogeneous slopes** across treatments - Can’t use standard ANCOVA



# Johnson-Neyman Procedure

## Johnson-Neyman Procedure for Heterogeneous Slopes

Analysis of Variance Table  
  
Response: suture\_width  
 Df Sum Sq Mean Sq F value Pr(>F)   
volume 1 0.015724 0.0157238 176.91 < 2.2e-16 \*\*\*  
treatment 2 0.036482 0.0182411 205.23 < 2.2e-16 \*\*\*  
volume:treatment 2 0.006213 0.0031064 34.95 4.453e-11 \*\*\*  
Residuals 66 0.005866 0.0000889   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



# Interpretation of Heterogeneous Slopes

When you have heterogeneous slopes, the Johnson-Neyman procedure identifies regions of the covariate where groups differ:

1. **Initial > Low Food** when cube root body volume > 2.95
   * For large urchins, the initial sample has wider sutures than low food urchins
2. **High Food > Initial** when cube root body volume > 1.81
   * For most urchins, high food treatment results in wider sutures than initial samples
3. **High Food > Low Food** when cube root body volume > 2.07
   * For most medium to large urchins, high food results in wider sutures than low food

The biological interpretation is that **food regime affects suture width differently depending on urchin size**. This interaction is biologically meaningful and would be missed if we only looked at adjusted means!

# Assumptions of ANCOVA

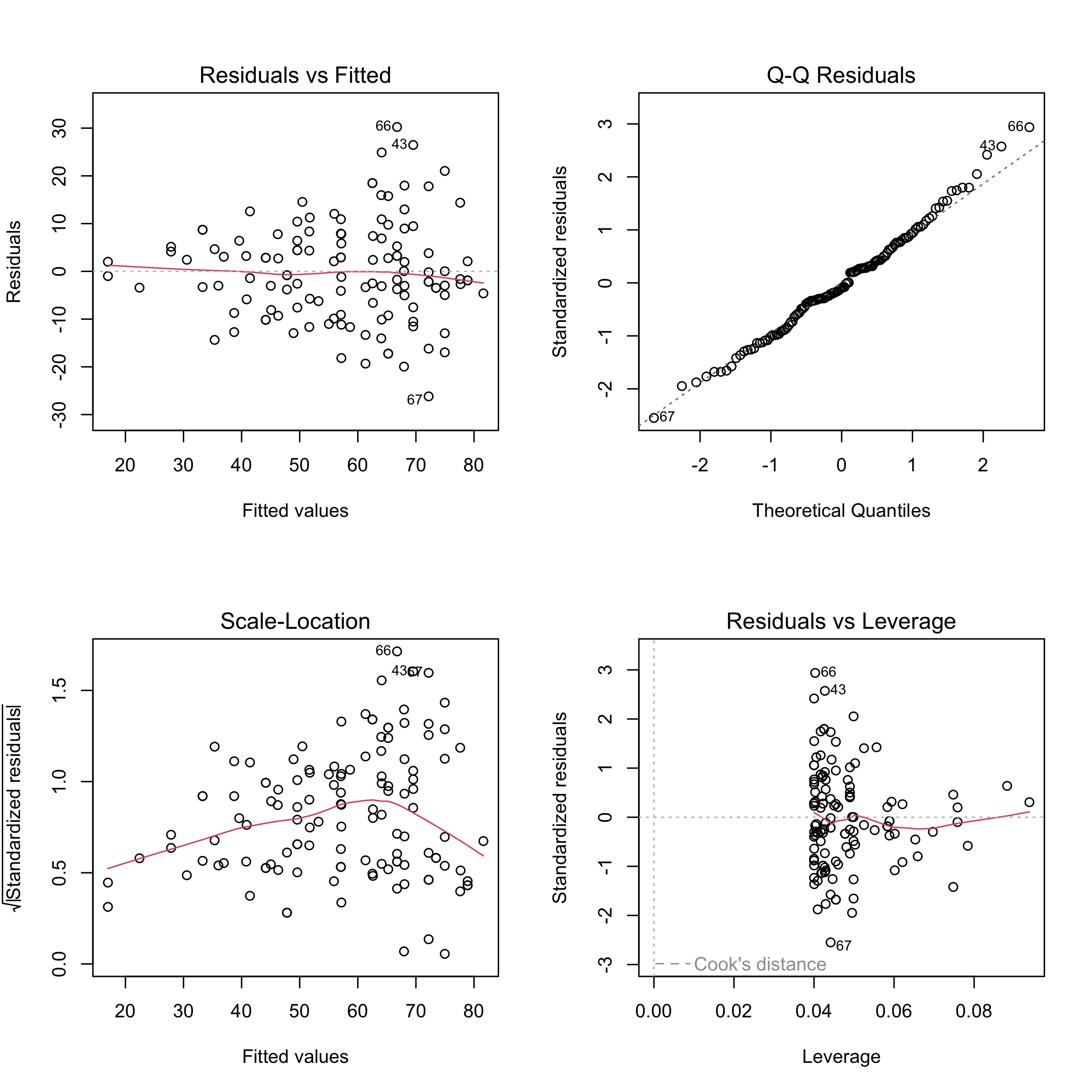
## Key Assumptions

1. **Independence of observations**
   * Samples are random and independent
   * No clustering or repeated measures
2. **Normal distribution of residuals**
   * Residuals follow a normal distribution
   * Check with QQ plots or formal tests
3. **Homogeneity of variances**
   * Equal variances across groups
   * Check with residual plots vs. fitted values
4. **Linearity**
   * Linear relationship between Y and X within each group
   * Check with scatterplots
5. **Homogeneity of regression slopes** (critical!)
   * Regression slopes equal across all groups
   * Test with interaction term

# Checking Assumptions in R

## Checking Assumptions in R

# Fit ANCOVA model for partridge data  
ancova\_model <- lm(LONGEV ~ THORAX + treatment, data = partridge)  
  
# Create a 2x2 panel of diagnostic plots  
par(mfrow = c(2, 2))  
plot(ancova\_model)



# Robust ANCOVA Approaches

## Non-Parametric Alternatives

When ANCOVA assumptions are violated, consider:

1. **Rank Transformation**
   * Rank transform both Y and X variables
   * Run standard ANCOVA on ranked data
   * Simple but may not handle interactions well
2. **ANCOVA on Bootstrapped Data**
   * Use bootstrapping to estimate parameters
   * Doesn’t require normality assumption
3. **Quantile Regression**
   * Models relationships at different quantiles
   * Robust to outliers and heteroscedasticity
4. **Permutation Tests**
   * Randomization tests of treatment effects
   * No distributional assumptions

# Example of rank-transformed ANCOVA  
partridge$rank\_LONGEV <- rank(partridge$LONGEV)  
partridge$rank\_THORAX <- rank(partridge$THORAX)  
  
# Fit ANCOVA on ranked data  
rank\_ancova <- lm(rank\_LONGEV ~   
 rank\_THORAX + treatment,   
 data = partridge)  
anova(rank\_ancova)

Analysis of Variance Table  
  
Response: rank\_LONGEV  
 Df Sum Sq Mean Sq F value Pr(>F)   
rank\_THORAX 1 63622 63622 129.511 < 2.2e-16 \*\*\*  
treatment 4 40447 10112 20.584 6.536e-13 \*\*\*  
Residuals 119 58458 491   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Compare p-values with parametric ANCOVA  
cat("P-value for treatment effect (parametric):",   
 round(anova(ancova\_model)[2, "Pr(>F)"], 4), "\n")

P-value for treatment effect (parametric): 0

cat("P-value for treatment effect (rank-based):",   
 round(anova(rank\_ancova)[2, "Pr(>F)"], 4))

P-value for treatment effect (rank-based): 0

# Permutation test example using lmPerm package  
# library(lmPerm)  
# perm\_ancova <- lmp(LONGEV ~   
# THORAX + treatment,  
# data = partridge,   
# perm = "Prob")  
# summary(perm\_ancova)

Note: The permutation test is commented out as it requires the lmPerm package, which may not be installed. The rank-based approach is shown as a simple alternative.

# Writing Up ANCOVA Results

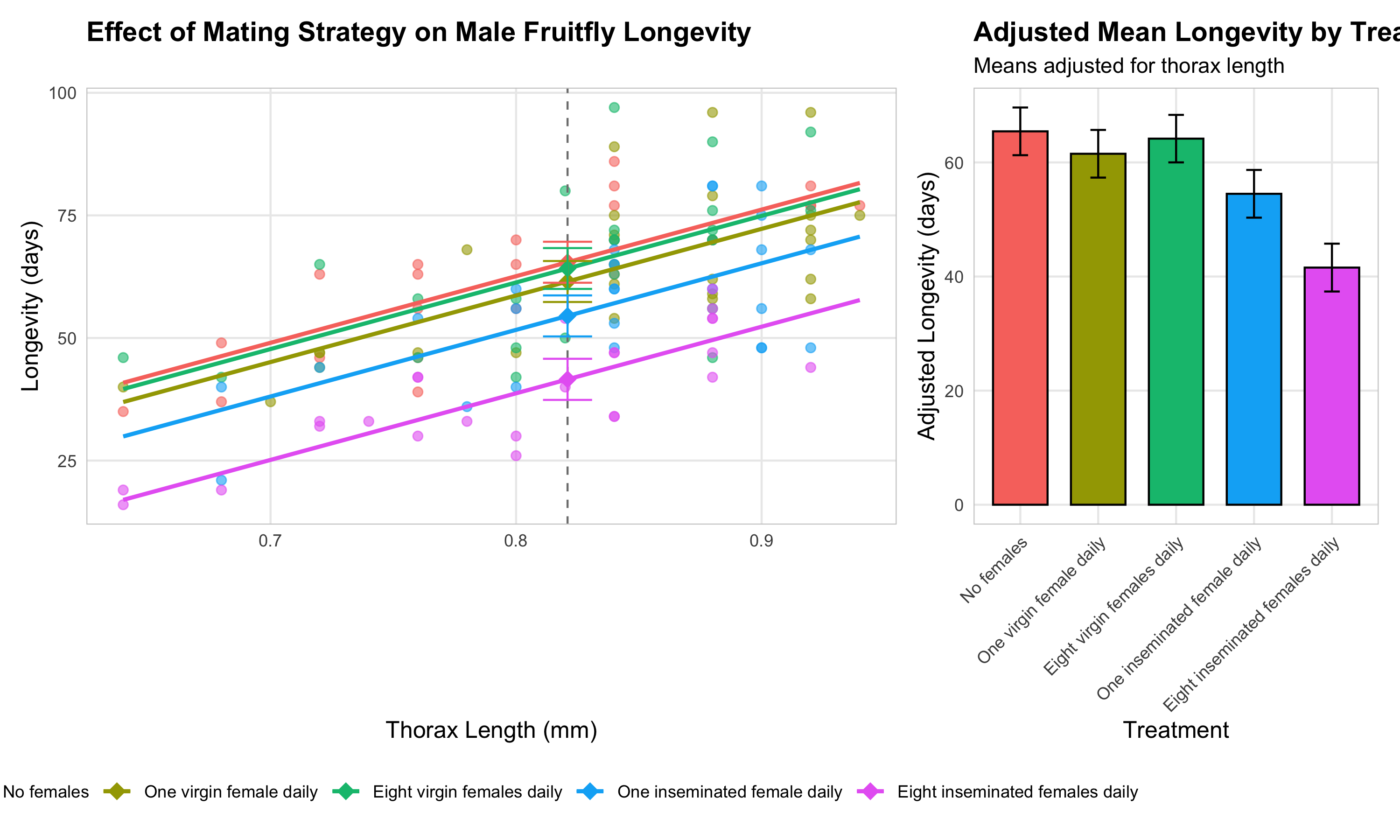
## Scientific Writing Example

Here’s how you might write up ANCOVA results for publication:

“We analyzed the effects of mating strategy on male fruitfly longevity using analysis of covariance (ANCOVA), with thorax length as a covariate. Before conducting the main analysis, we tested the homogeneity of slopes assumption and found no significant interaction between treatment and thorax length (F₄,₁₁₅ = 1.56, P = 0.19), indicating that the effect of body size on longevity was consistent across treatments.

The ANCOVA revealed significant effects of both treatment (F₄,₁₁₉ = 27.97, P < 0.001) and thorax length (F₁,₁₁₉ = 145.44, P < 0.001) on longevity. Thorax length was positively associated with longevity (b = 1.19), with larger males living longer. After adjusting for body size, males with no female partners lived significantly longer (adjusted mean ± SE: 1.81 ± 0.02 log₁₀ days) than males in any other treatment group. Males provided with a single virgin female daily (1.77 ± 0.02) or a single inseminated female daily (1.79 ± 0.02) showed intermediate longevity, while males with eight females per day showed the lowest longevity (1.72 ± 0.02 for inseminated females; 1.59 ± 0.02 for virgin females). Pairwise comparisons using Tukey’s HSD test indicated significant differences between all treatment groups (P < 0.05) except between the two treatments with a single female per day (P = 0.42).”

# Publication Quality Figure



# Summary

## Key Principles

1. **Purpose**:
   * ANCOVA combines regression and ANOVA approaches
   * Increases power by accounting for continuous covariates
   * Allows comparison of adjusted means
2. **The Analysis**
   * Always test for homogeneity of slopes first!
   * If slopes are homogeneous, proceed with standard ANCOVA
   * If slopes are heterogeneous, use alternatives (Johnson-Neyman procedure)
3. **Interpretation**
   * Focus on adjusted means (at mean covariate value)
   * Consider both statistical and biological significance
   * Visualize results clearly with appropriate graphs

## Assumptions

1. Independence of observations
2. Normal distribution of residuals
3. Homogeneity of variances
4. Linearity of relationships within groups
5. Homogeneity of regression slopes