

10_homework_ANCOVA

Bill Perry

```
# Load required packages
library(skimr)

library(janitor)      # For cleaning variable names
library(broom)        # For tidy statistical output
library(car)          # For regression diagnostics and ANCOVA
library(emmeans)      # For estimated marginal means
library(patchwork)    # For combining plots
library(tidyverse)    # For data manipulation and visualization

# Set ggplot theme
theme_set(theme_bw())
```

Assignment Overview

This homework assignment focuses on Analysis of Covariance (ANCOVA) to examine the relationship between temperature and cricket chirp rate across two different cricket species. ANCOVA allows us to test whether the relationship between temperature and chirp rate differs between species while controlling for temperature as a covariate.

Learning Objectives

By completing this assignment, you will be able to:

1. Understand ANCOVA concepts and applications
2. Perform and interpret ANCOVA analysis
3. Test statistical assumptions for ANCOVA
4. Compare slopes and intercepts between groups
5. Create publication-quality figures for ANCOVA results
6. Write scientific methods and results sections

Data Description

The dataset `cricket_chirp_rate.csv` contains chirp rate measurements from two cricket species at different temperatures:

- species: Cricket species (ex = *Oecanthus exclamationis*, niv = *Oecanthus niveus*)
- temp: Temperature (°C)
- pulse: Chirp rate (chirps per second)

Important Background

Cricket chirp rate is known to be temperature-dependent, following a near-linear relationship. Different cricket species may have different thermal sensitivities, leading to different slopes or intercepts in the temperature-chirp rate relationship.

Part 1: Data Loading and Initial Exploration

1.1 Load and Clean the Data

```
# Load the cricket data

# # Convert species to factor with meaningful labels
# cr_df <- cr_df %>%
#   mutate(
#     species = factor(species,
#                       levels = c("0. exclamationis", "0. niveus"))
#   )
#
# # Check the cleaned data
# cr_df
```

Part 2: Statistical Model Setup

2.1 ANCOVA Model Statement

! Model Specification

Type of Analysis: Analysis of Covariance (ANCOVA)

Model Equation: $\text{pulse_rate} = \beta_0 + \beta_1(\text{temperature}) + \beta_2(\text{species}) + \beta_3(\text{temperature} \times \text{species}) + \varepsilon$

Where: - pulse_rate = chirp rate (response variable) - β_0 = intercept for reference species - β_1 = slope for temperature effect in reference species - β_2 = difference in intercept between species - β_3 = difference in slope between species (interaction term) - ε = random error term

Hypotheses Being Tested:

1. **Main effect of temperature:** $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$
2. **Main effect of species:** $H_0: \beta_2 = 0$ vs $H_1: \beta_2 \neq 0$
3. **Interaction effect:** $H_0: \beta_3 = 0$ vs $H_1: \beta_3 \neq 0$

ANCOVA Assumptions: 1. Independence of observations 2. Normality of residuals 3. Homogeneity of variance (homoscedasticity) 4. Linear relationship between covariate and response 5. Homogeneity of regression slopes (if no interaction)

Part 3: Descriptive Statistics and Exploratory Analysis

3.1 Summary Statistics by Group

```
# # Calculate summary statistics by species
# cr_df %>%
#   group_by(species) %>%
#   skim()
```

3.2 Exploratory Data Visualization

```
# Create histograms for temperature by species
t
```

```
function (x)
  UseMethod("t")
<bytecode: 0x135bf75b0>
<environment: namespace:base>
```

```
# Create histograms for pulse rate by species
```

Part 4: ANCOVA Analysis

4.1 Fit the ANCOVA Model

```
# Fit the full ANCOVA model with interaction
```

4.2 Model Interpretation

```
# Create ANOVA table
```

Model Interpretation:

- **Intercept (β_0):** Estimated pulse rate for *O. exclamationis* at 0°C
 - **temp (β_1):** For every 1°C increase in temperature, *O. exclamationis* chirp rate increases by β_1 chirps/sec
 - **species *O. niveus* (β_2):** Difference in intercept between *O. niveus* and *O. exclamationis*
 - **temp:species *O. niveus* (β_3):** Difference in temperature slope between the two species
-

Part 5: Assumption Testing

5.1 Diagnostic Plots

```
# # Create diagnostic plots
# par(mfrow = c(2, 2))
# plot(model)
# par(mfrow = c(1, 1))
```

5.2 Formal Assumption Tests

```
# Shapiro-Wilk test for normality of residuals
```

```
# Levene's test for homogeneity of variances
```

Part 7: Publication-Quality Figure

7.1 Create Final Figure

```
# Create publication-quality plot
```

Submission Guidelines

What to turn in -

1. a quarto markdown file and dataframe if you modified the original. All of the code should be able to run with what you turn in. **(2 points)**
2. a self-contained html file showing the code and output **(2 points)**
3. annotations in the quarto file that shows or tells what is being done in the r code chunks describing what you are trying to do - credit will be given even if it does not work as long as you detail what you are doing. As we start to move into more statistics you will be expected to interpret the results. **(2 points)**

Points

- summary stats - 2 point
- assumptions and hypotheses - 3 points
- exploratory graphs - 2 point
- interpretation - 4 points
- Final figure - 1 point
- Results - 2 points