

R Take-Home Assignment 5 (Instructor Key): Statistics with Penguins

Instructor

2025-10-01

Overview

In this assignment you will practice simple statistical tests in R using the **palmerpenguins** dataset. We will ask biological questions about penguin body mass, learn to run standard tests in R, and interpret the results. Each section starts with a short story to explain why we are doing this step.

Source

palmerpenguins package

Skills

t-tests, ANOVA, post-hoc tests, effect sizes, diagnostics

1. Load the dataset

Before we can analyze penguins, we need to load the dataset. Real-world data often contain missing values, which can cause errors in analysis, so we will remove them first.

```
data(penguins)
penguins <- penguins %>% drop_na()
glimpse(penguins)

## Rows: 333
## Columns: 8
## $ species      <fct> Adelie, Adelie, Adelie, Adelie, Adelie, Adelie, Adel-
## $ island       <fct> Torgersen, Torgersen, Torgersen, Torgersen, Torgerse-
## $ bill_length_mm <dbl> 39.1, 39.5, 40.3, 36.7, 39.3, 38.9, 39.2, 41.1, 38.6~
## $ bill_depth_mm <dbl> 18.7, 17.4, 18.0, 19.3, 20.6, 17.8, 19.6, 17.6, 21.2~
## $ flipper_length_mm <int> 181, 186, 195, 193, 190, 181, 195, 182, 191, 198, 18~
## $ body_mass_g   <int> 3750, 3800, 3250, 3450, 3650, 3625, 4675, 3200, 3800~
## $ sex          <fct> male, female, female, female, male, female, male, fe-
## $ year         <int> 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007, 2007~
```

2. Two-sample t-test

Imagine you are a biologist asking: *Do male and female penguins of the same species differ in body mass?* The **t-test** compares the means of two groups. It answers whether the difference is large enough that it is unlikely to be due to chance.

```
adelie <- penguins %>% filter(species == "Adelie")
t.test(body_mass_g ~ sex, data = adelie)
```

```
##
## Welch Two Sample t-test
##
## data: body_mass_g by sex
## t = -13.126, df = 135.69, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group female and group male is not equal to 0
## 95 percent confidence interval:
## -776.3012 -573.0139
## sample estimates:
## mean in group female mean in group male
## 3368.836 4043.493
cohens_d(body_mass_g ~ sex, data = adellie)

## Cohen's d | 95% CI
## -----
## -2.17 | [-2.58, -1.76]
##
## - Estimated using pooled SD.
```

3. ANOVA

Now suppose we want to compare *all three species at once*. Do Gentoo, Adelie, and Chinstrap penguins differ in body mass?

An **ANOVA (Analysis of Variance)** extends the t-test to more than two groups.

```
anova_model <- aov(body_mass_g ~ species, data = penguins)
summary(anova_model)
```

```
##           Df    Sum Sq Mean Sq F value Pr(>F)
## species      2 145190219 72595110  341.9 <2e-16 ***
## Residuals   330  70069447  212332
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4. Post-hoc tests

ANOVA tells us if there is *some* difference between groups, but not *which groups differ*. For that, we run a **Tukey HSD post-hoc test**.

```
TukeyHSD(anova_model)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = body_mass_g ~ species, data = penguins)
##
## $species
##           diff      lwr      upr    p adj
## Chinstrap-Adelie 26.92385 -132.3528 186.2005 0.916431
## Gentoo-Adelie    1386.27259 1252.2897 1520.2554 0.000000
## Gentoo-Chinstrap 1359.34874 1194.4304 1524.2671 0.000000
```

5. Effect sizes

Statistical significance is not enough: we also want to know *how large the difference is*.

- For t-tests, we use **Cohen's d** (small 0.2, medium 0.5, large 0.8).

- For ANOVA, we use ² (**eta squared**) (small .01, medium .06, large .14).

```
eta_squared(anova_model)
```

```
## # Effect Size for ANOVA
##
## Parameter | Eta2 |      95% CI
## -----
## species   | 0.67 | [0.63, 1.00]
##
## - One-sided CIs: upper bound fixed at [1.00].
```

6. Diagnostics (checking assumptions)

Every statistical test makes assumptions. For ANOVA, the two most important are:

1. **Equal spread of errors (homoscedasticity):**

The variation in the data should be roughly the same across all groups.

2. **Normally distributed errors:**

The differences between the observed data and the model's predictions (residuals) should follow a bell-shaped curve.

We can check these assumptions using two standard diagnostic plots:

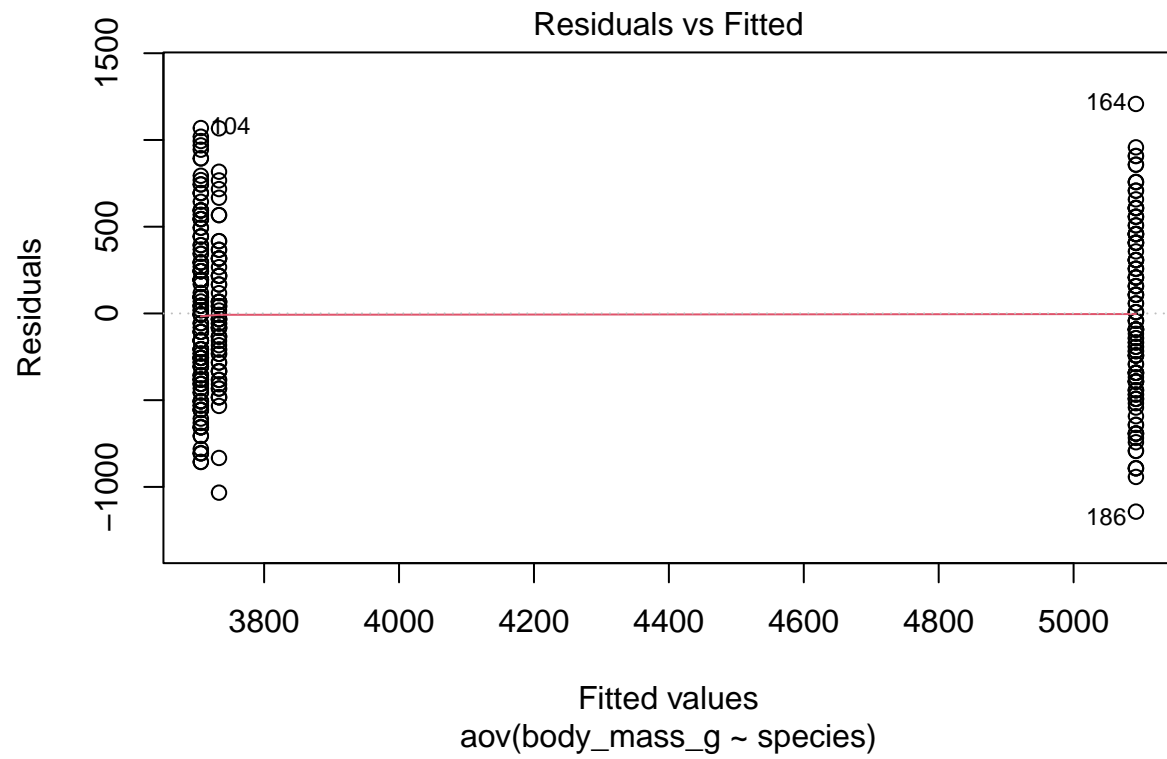
- **Residuals vs Fitted plot**

- Each dot is one observation.
- The x-axis shows what the model predicts, the y-axis shows the “error” (residual).
- What to look for: the dots should look like random noise.
 - * If you see a curve → the relationship might not be captured well.
 - * If you see a funnel shape → group variances may not be equal.

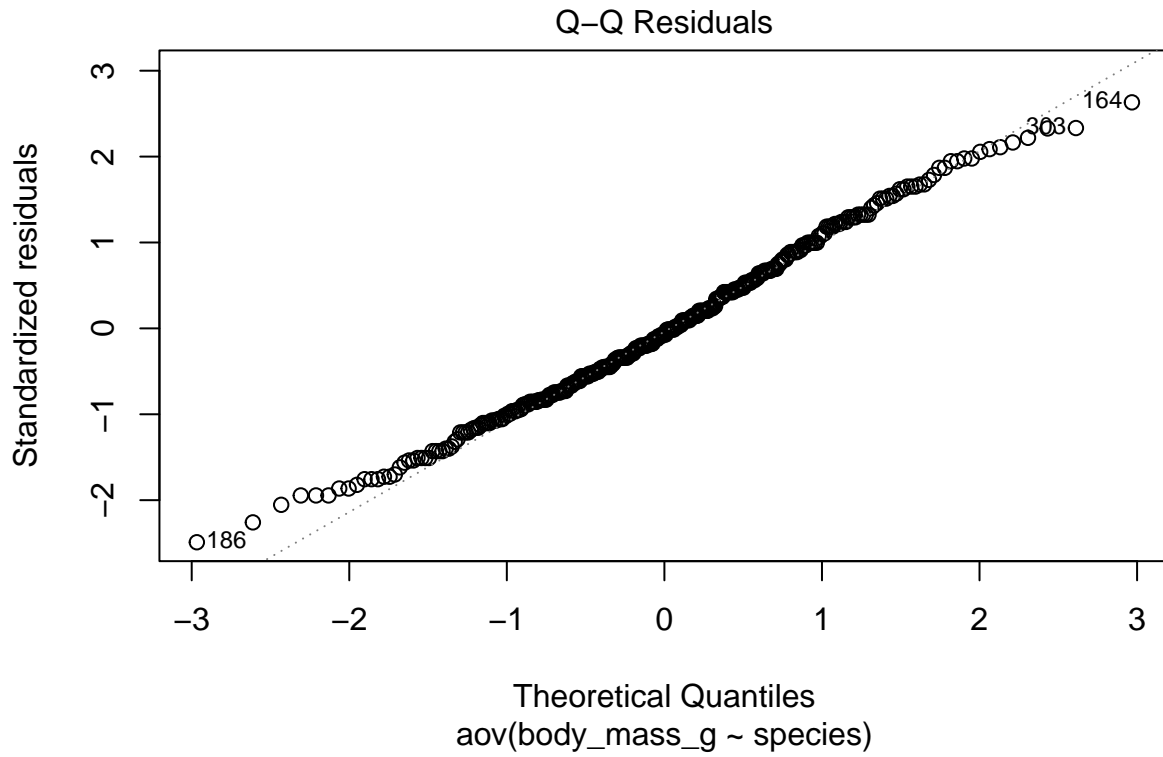
- **Normal QQ plot**

- This compares your residuals to what would be expected if they were perfectly normal.
- What to look for: the dots should fall roughly along the diagonal line.
 - * If they bend away strongly → residuals may not be normally distributed.

```
plot(anova_model, which = 1)
```



```
plot(anova_model, which = 2)
```



7. Reflection

Statistics are not just numbers: they help answer real questions. Think like a biologist preparing a short report.

Across species, body mass differs significantly, with Gentoo penguins being much heavier. Within a species (e.g. Adelie), males are significantly heavier than females. Effect sizes (Cohen's d and r^2) show these are large effects. Diagnostics are usually acceptable; residuals look roughly normal and variance is fairly equal.