HW4

Howard

2025-03-22

Problem 5

Create the data

```
# Data for Problem 5 - Painted Turtles

# Female turtles data
female <- data.frame(
  Length = c(98, 103, 103, 105, 109, 123, 123, 133, 133, 133, 134, 136, 138, 138, 141, 147, 149, 153, 155, 155, 158,
159, 162, 177),
  Width  = c(81, 84, 86, 86, 88, 92, 95, 99, 102, 102, 100, 102, 98, 99, 105, 108, 107, 107, 115, 117, 115, 118, 12
4, 132),
  Height = c(38, 38, 42, 42, 44, 50, 46, 51, 51, 51, 48, 49, 51, 51, 53, 57, 55, 56, 63, 60, 62, 63, 61, 67)
)

# Male turtles data
male <- data.frame(
  Length = c(93, 94, 96, 101, 102, 103, 104, 106, 107, 112, 113, 114, 116, 117, 117, 119, 120, 120, 121, 125, 127, 1
28, 131, 135),
  Width  = c(74, 78, 80, 84, 85, 81, 83, 83, 82, 89, 88, 86, 90, 90, 91, 93, 89, 93, 95, 93, 96, 95, 95, 106),
  Height = c(37, 35, 35, 39, 38, 37, 39, 39, 38, 40, 40, 40, 43, 41, 41, 41, 40, 44, 42, 45, 45, 45, 46, 47)
)

# Combine the female and male data into one data frame with a grouping variable
turtles <- rbind(
  data.frame(Group = "Female", female),
  data.frame(Group = "Male", male)
)

# View the combined dataset
head(turtles)
```

```
##      Group Length Width Height
## 1 Female     98    81     38
## 2 Female    103    84     38
## 3 Female    103    86     42
## 4 Female    105    86     42
## 5 Female    109    88     44
## 6 Female    123    92     50
```

Log-transform the measurements

```
turtles$logLength <- log(turtles$Length)
turtles$logWidth  <- log(turtles$Width)
turtles$logHeight <- log(turtles$Height)
```

Look at group means on the log scale

```
group_means <- aggregate(cbind(logLength, logWidth, logHeight) ~ Group, data = turtles, FUN = mean)
print(group_means)
```

```
##    Group logLength logWidth logHeight
## 1 Female  4.900659 4.622909  3.940286
## 2   Male  4.725444 4.477574  3.703186
```

$$H_0 : \mu_F = \mu_M \; versus \; H_1 : \mu_F \neq \mu_M$$

Fit a MANOVA model using the log-transformed variables

```
fit <- manova(cbind(logLength, logWidth, logHeight) ~ Group, data = turtles)

# Summary using Wilks' lambda test
summary_fit <- summary(fit, test = "Wilks")
print(summary_fit)
```

```
##             Df  Wilks approx F num Df den Df     Pr(>F)
## Group        1 0.35101   27.118      3     44 4.355e-10 ***
## Residuals   46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclution

Based on the extremely small p-value (on the order of 10^{-10}), we would **reject** the null hypothesis H_0 that female and male turtles have the same mean (log-transformed) carapace measurements. In other words, **there is a statistically significant difference** between the mean vectors of female and male turtles at the 5% significance level.

Problem 6

Data setup

```
# 1. Create vectors for species, nutrient, and the two reflectance variables
species <- factor(c("SS", "JL", "LP", "SS", "JL", "LP"))
nutrient <- factor(c("+", "+", "+", "-", "-", "-"))
x560 <- c(10.35, 13.41, 7.78, 10.40, 17.78, 10.40)
x720 <- c(25.93, 38.63, 25.15, 24.25, 41.45, 29.20)

# 2. Combine into a single data frame
seedlings <- data.frame(x560, x720,species, nutrient)
print(seedlings)
```

```
##     x560  x720 species nutrient
## 1 10.35 25.93      SS        +
## 2 13.41 38.63      JL        +
## 3  7.78 25.15      LP        +
## 4 10.40 24.25      SS        -
## 5 17.78 41.45      JL        -
## 6 10.40 29.20      LP        -
```

(a)Separate MANOVA TEST

MANOVA for the Effect of Species

- Test the effect of species (SS, JL, LP), ignoring nutrient.

```
# MANOVA for species only
fit_species <- manova(cbind(x560, x720) ~ species, data = seedlings)
summary_species <- summary(fit_species, test = "Wilks")
summary_species
```

```
##            Df   Wilks approx F num Df den Df  Pr(>F)
## species     2 0.015927   6.9238      4      4 0.04376 *
## Residuals   3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- **Interpretation**: The p-value is < 0.05, we reject the null hypothesis that all **three species** have the same mean vector (μ_{560}, μ_{720})

MANOVA for the Effect of Nutrient

- Test The effect of nutrient (+, −), ignoring species.

```
# MANOVA for nutrient only
fit_nutrient <- manova(cbind(x560, x720) ~ nutrient, data = seedlings)
summary_nutrient <- summary(fit_nutrient, test = "Wilks")
summary_nutrient
```

```
##            Df   Wilks approx F num Df den Df Pr(>F)
## nutrient    1 0.61136  0.95353      2      3  0.478
## Residuals   4
```

- **Interpretation**: The p-value is > 0.05, we do reject the null hypothesis that **the two nutrient levels** have the same mean vector (μ_{560}, μ_{720})

(b) Two-Way ANOVA for 560CM and 720CM Separately

ANOVA for 560CM

```
# Two-way ANOVA for 560CM (no interaction due to no replication)
fit_aov_560 <- aov(x560 ~ species + nutrient, data = seedlings)
summary(fit_aov_560)
```

```
##            Df Sum Sq Mean Sq F value Pr(>F)
## species     2  47.48  23.738  10.055 0.0905 .
## nutrient    1   8.26   8.260   3.499 0.2023
## Residuals   2   4.72   2.361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation:

- *Species (p = 0.0905)*: At the conventional 5% significance level, we do not reject the null hypothesis that all three species (SS, JL, LP) have the same 560 nm mean reflectance.

- *Nutrient (p = 0.2023)*: There is no evidence at the 5% level to conclude that the two nutrient levels (+, −) differ in their 560 nm mean reflectance.

ANOVA for 720CM

```
# Two-way ANOVA for 720CM (no interaction due to no replication)
fit_aov_720 <- aov(x720 ~ species + nutrient, data = seedlings)
summary(fit_aov_720)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## species      2 262.24  131.12  28.820 0.0335 *
## nutrient     1   4.49    4.49   0.987 0.4252
## Residuals    2   9.10    4.55
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation:

- *Species (p = 0.0335)*: At the 5% significance level, we reject the null hypothesis that the three species (SS, JL, LP) have the same mean reflectance at 720 nm. In other words, there is evidence of a significant difference among species in their 720 nm reflectance means.

- *Nutrient (p = 0.4252)*: There is no evidence at the 5% level to conclude that the two nutrient levels (+, −) differ in their 720 nm mean reflectance.

Comparing MANOVA vs. Separate ANOVAs

- MANOVA results tell you if the factor(s) has/have a joint effect on the vector x_{560}, x_{720}

- Separate ANOVAs on `x560` and `x720` test each response individually.