Chloe Lang

Professor Nelson

Models 2 Assignment

November 28, 2021

**Q1**. Conditional boxplot of penguin body mass, conditioned on sex and species.

Chart, box and whisker chart

Description automatically generated

**Q2**. Based on the boxplots, male penguins (of any species) are heavier than female penguins. Statistically, I would need to see more numerical values to determine if males are significantly heavier than females. However, when comparing each species specifically between male and female you are able to see that for each species, males appeared to have higher body masses, indicated by the boxplots being placed higher on the plot.

**Q3**. I think adding sex to a model that already includes species does improve the model fit. This is because by adding another variable to the plot, you are increasing homogeneity, thus making the models a better fit.

**Q4**. Code: fit\_both = lm(body\_mass\_g ~ species \* sex, data = penguins)

**Q5**. The base case for the two-way model that includes sex and species is female adelie penguins.

**Q6**. The two coefficients that you need to calculate the average mass of female Chinstrap penguins is the coefficient intercept and the coefficient “speciesChinstrap”.

**Q7**. The predicted average mass of female Chinstrap penguins according to this interactive model is 3527.21 grams.

**Q8**. The observed average mass of female Chinstrap penguins, calculated from the penguin data is 3527.206 grams.

Code Used:

##Walkthrough

require(palmerpenguins)

#1 sample t test

t.test(subset(penguins, species == "Gentoo")$flipper\_length\_mm)

t.test(

x = subset(penguins, species == "Gentoo")$flipper\_length\_mm,

mu = 218

)

t.test(

x = subset(penguins, species == "Gentoo")$flipper\_length\_mm,

mu = 218,

alternative = "less"

)

#2 sample t test

t.test(flipper\_length\_mm ~ species, data = subset(penguins, species != "Chinstrap"))

#1- Analysis of Variance (ANOVA)

#Model 1 Body Mass by Species

#1- Way Analysis of Variance Procedure

#Data exploration

#Graphical

par(mfrow = c(1, 2))

hist(penguins$body\_mass\_g, breaks = 80, main = "histogram of body mass", xlab = "body mass (g)")

plot(density(penguins$body\_mass\_g, na.rm = TRUE), main = "density plot of body mass")

boxplot(body\_mass\_g ~ species, data = penguins)

#Numerical

dat\_chinstrap = subset(penguins, species == "Chinstrap")

mean(dat\_chinstrap$body\_mass\_g, na.rm = TRUE)

shapiro.test(dat\_chinstrap$body\_mass\_g)

aggregate(body\_mass\_g ~ species, data = penguins, FUN = mean)

#Fit a linear model

fit\_species = lm(body\_mass\_g ~ species, data = penguins)

summary(fit\_species)

#Conduct the ANOVA

anova(fit\_species)

#1 way ANOVA Complete Walkthrough

fit\_species = lm(body\_mass\_g ~ species, data = penguins)

summary(fit\_species)

#Conduct the ANOVA

anova(fit\_species)

#2 way Additive ANOVA

boxplot(body\_mass\_g ~ species, data = penguins)

#Fit a 2 way additive model

fit\_additive = lm(body\_mass\_g ~ sex + species, data = penguins)

#2 way interactive (factorial) ANOVA

fit\_interactive = lm(body\_mass\_g ~ sex \* species, data = penguins)

#Simple linear regression: Penguin bills & body mass

lm(bill\_length\_mm ~ body\_mass\_g, data = penguins)

##Questions

#Interpreting Boxplots

#Q1 Conditional boxplot of penguin body mass conditioned on sex & species

boxplot(body\_mass\_g ~ species + sex, data = penguins,

main = "Body Mass by

Sex and Species",

ylab = "Body Mass (g)",

names = c("Female\nAdelie", "Female\nChinstrap", "Female\nGentoo", "Male\nAdelie", "Male\nChinstrap", "Male\nGentoo"),

col = "lightskyblue2")

#Model Fit 1

#Q4 factorial linear model of penguin mody mass predicted by sex & species

fit\_both = lm(body\_mass\_g ~ species \* sex, data = penguins)

summary(fit\_both)

#Model Coefficients

#Q6

summary(fit\_both)

#Q7 predicted avg mass of female chinstrap penguins in the interactive model

3368.84 + 158.37

#Q8 observed avg mass of female chinstrap penguins, from the data

aggregate(

body\_mass\_g ~ species + sex,

data = penguins,

FUN = function(x) mean(x, na.rm = TRUE)

)