Chad Huntebrinker’s Homework 4

Chad Huntebrinker

2024-09-24

Question 2.4b and c: With a 99% confidence interval for B1, test using the test statistic t\* to find a linear association. Find the P-value.

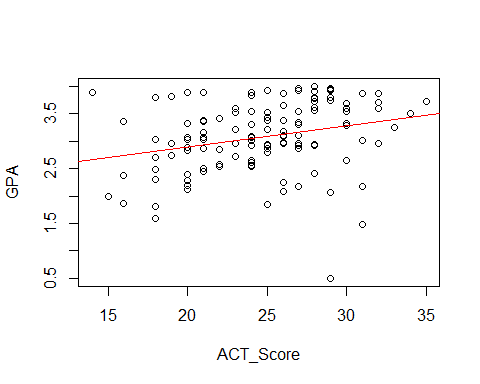
#Chad Huntebrinker  
#Problem 2.4b  
#Load the library and data in to R  
library(readxl)  
  
excel\_data <- read\_excel("Grade\_Point\_Average\_Data.xlsx")  
  
# GPA ~ ACT  
# Fit the model and get the summary of the model  
model\_1 <- lm(GPA~ACT\_Score,data=excel\_data)  
sum\_of\_model\_1 <- summary(model\_1)  
sum\_of\_model\_1

##   
## Call:  
## lm(formula = GPA ~ ACT\_Score, data = excel\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.74004 -0.33827 0.04062 0.44064 1.22737   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.11405 0.32089 6.588 1.3e-09 \*\*\*  
## ACT\_Score 0.03883 0.01277 3.040 0.00292 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6231 on 118 degrees of freedom  
## Multiple R-squared: 0.07262, Adjusted R-squared: 0.06476   
## F-statistic: 9.24 on 1 and 118 DF, p-value: 0.002917

sum\_of\_model\_1$coefficients[2,4]

## [1] 0.002916604

# Plot the graph just to see  
plot(GPA~ACT\_Score,data=excel\_data)  
abline(model\_1,col="red")



#Calculate the t score  
b1 <- sum\_of\_model\_1$coefficients[2, 1]  
se\_b1 <- sum\_of\_model\_1$coefficients[2, 2]  
  
#a = 0.01  
t\_1 <- b1 / se\_b1  
  
#If the t\_1 > than the t score, than we conclude that H1 is true (which is B1 != 0)  
t\_1 > qt(0.995, 118)

## [1] TRUE

#Problem 2.4c  
#Calculate the p\_value  
p\_1 <- 2 \* (1 - pt(t\_1, 118))  
  
#If the p score is less than alpha, than H1 is true (which is B1 != 0)  
p\_1 < 0.01

## [1] TRUE

Problem 2.6b: With a level of significance of .05, conduct a t test and find the P-value to see if there is a linear association

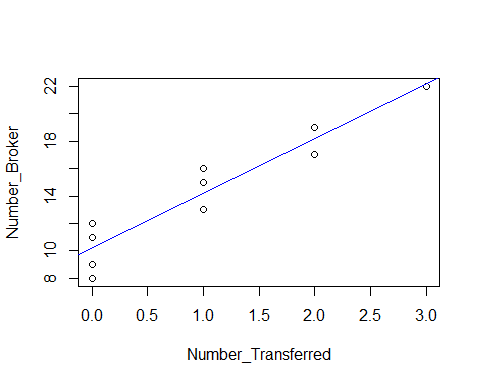
#Chad Huntebrinker  
#Problem 2.6b  
#Load the library and data in to R  
library(readxl)  
  
excel\_data <- read\_excel("Airfreight\_Breakage\_Data.xlsx")  
  
#Number\_Broker ~ Number\_Transferred  
#Fit the model  
model\_2 <- lm(Number\_Broker~Number\_Transferred,data=excel\_data)  
sum\_of\_model\_2 <- summary(model\_2)  
sum\_of\_model\_2

##   
## Call:  
## lm(formula = Number\_Broker ~ Number\_Transferred, data = excel\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.2 -1.2 0.3 0.8 1.8   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 10.2000 0.6633 15.377 3.18e-07 \*\*\*  
## Number\_Transferred 4.0000 0.4690 8.528 2.75e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.483 on 8 degrees of freedom  
## Multiple R-squared: 0.9009, Adjusted R-squared: 0.8885   
## F-statistic: 72.73 on 1 and 8 DF, p-value: 2.749e-05

sum\_of\_model\_2$coefficients[2,4]

## [1] 2.748669e-05

#Plot to see the graph  
plot(Number\_Broker~Number\_Transferred, data = excel\_data)  
abline(model\_2,col="blue")



#Calculate the t score  
b1 <- sum\_of\_model\_2$coefficients[2, 1]  
se\_b1 <- sum\_of\_model\_2$coefficients[2, 2]  
  
#a = 0.05  
t\_1 <- b1 / se\_b1  
  
#If the t\_1 > than the t score, than we conclude that H1 is true (which is B1 != 0)  
t\_1 > qt(0.975, 8)

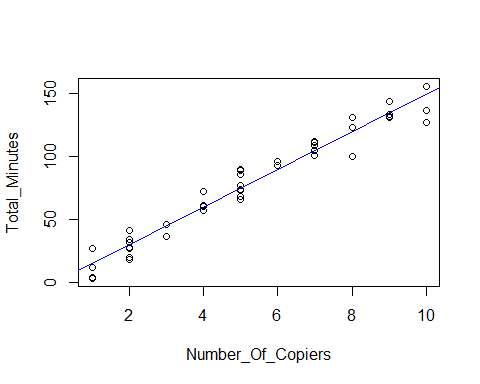
## [1] TRUE

#Calculate the p\_value  
p\_1 <- 2 \* (1 - pt(t\_1, 8))  
  
#If the p score is less than alpha, than H1 is true (which is B1 != 0)  
p\_1 < 0.01

## [1] TRUE

Problem 2.14a: Find the confidence interval with a 90% confidence interval for the mean service time on calls in which six copier are serviced

#Chad Huntebrinker  
#Problem 2.14a  
#Load the library and data in to R  
library(readxl)  
  
excel\_data <- read\_excel("Copier\_Maintenance\_Data.xlsx")  
  
#Total\_Minutes ~ Number\_Of\_Copiers  
#Fit the model  
model\_3 <- lm(Total\_Minutes~Number\_Of\_Copiers,data=excel\_data)  
sum\_of\_model\_3 <- summary(model\_3)  
  
#Plot to see the graph  
plot(Total\_Minutes~Number\_Of\_Copiers, data = excel\_data)  
abline(model\_3,col="blue")

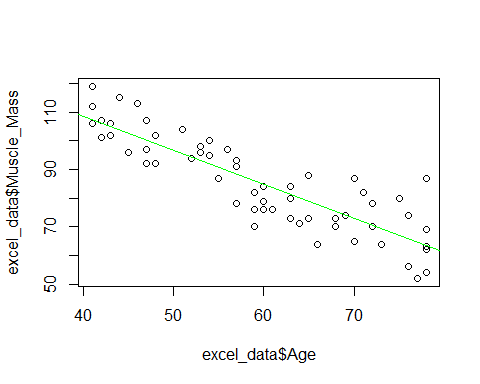


#Get b1 and se\_b1  
b1 <- sum\_of\_model\_3$coefficients[2, 1]  
se\_b1 <- sum\_of\_model\_3$coefficients[2, 2]  
  
#Get the mean service time for 6 copiers serviced  
Y\_hat <- predict(model\_3, data.frame(Number\_Of\_Copiers=6))  
  
#Now get the confidence interval  
predict(model\_3, data.frame(Number\_Of\_Copiers=6), interval = "confidence", level = 0.90)

## fit lwr upr  
## 1 89.63133 87.28387 91.9788

Problem 2.27a and c: With a control risk for type 1 error of .05, find if there is a negative linear association. Estimate with a 95% confidence interval the difference in expected mm for women whose ages differ by one year.

#Chad Huntebrinker  
#Problem 2.27a  
#Load the library and data in to R  
library(readxl)  
  
excel\_data <- read\_excel("Muscle\_Mass\_Data.xlsx")  
  
#Muscle\_Mass ~ Age  
#Fit the model  
model\_4 <- lm(Muscle\_Mass~Age,data=excel\_data)  
sum\_of\_model\_4 <- summary(model\_4)  
  
#Plot to see the graph  
plot(excel\_data$Muscle\_Mass~excel\_data$Age)  
abline(model\_4,col="green")



#Get b1 and se\_b1  
b1 <- summary(model\_4)$coefficients[2, 1]  
se\_b1 <- summary(model\_4)$coefficients[2, 2]  
  
#Find t score  
t\_1 <- b1 / se\_b1  
  
#If the t\_1 > than the t score, than we conclude that H1 is true (which is B1 != 0)  
t\_1 < qt(0.05, 58)

## [1] TRUE

#Find p score  
p\_1 <- pt(t\_1, 58)  
  
#Problem 2.27c  
qt(0.975, 58)

## [1] 2.001717

b1

## [1] -1.189996

se\_b1

## [1] 0.09019725

#Lower confidence  
b1 - (qt(0.975, 58) \* se\_b1)

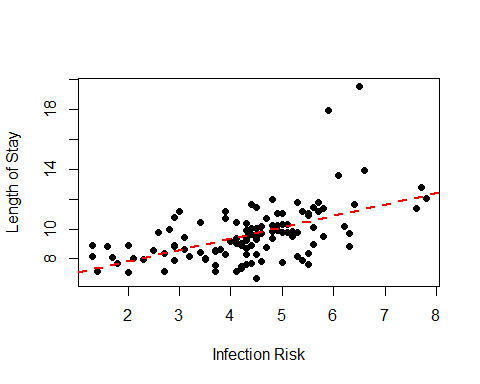
## [1] -1.370545

#Higher confidence  
b1 + (qt(0.975, 58) \* se\_b1)

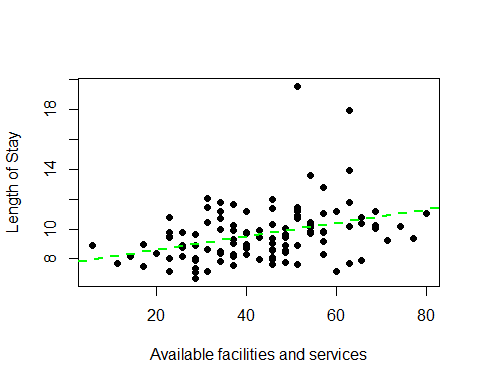
## [1] -1.009446

Problem 2.65:Obtain a separate interval estimate of B1 for each egion. Use a 95% confidence coefficient in each case.

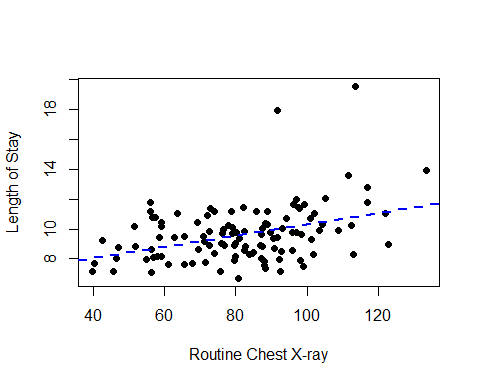
#Chad Huntebrinker  
#Problem 2.65  
#Load the library and data in to R  
library(readxl)  
  
excel\_data <- read\_excel("SENIC\_Data.xlsx")  
  
#First  
#For Infection\_Risk, plot the Data  
plot(Length\_of\_Stay ~ Infection\_Risk, data = excel\_data, pch = 16, xlab = "Infection Risk", ylab = "Length of Stay")  
  
#Fit a Linear Regression model  
model\_7 <- lm(Length\_of\_Stay ~ Infection\_Risk, data = excel\_data)  
  
#Find b coefficients  
b7\_coefficients <- model\_7$coefficients  
  
#Add the fitted line  
abline(b7\_coefficients, lwd =2, lty =2, col="red")



#Second  
#For Available\_facilities\_and\_services, plot the Data  
plot(Length\_of\_Stay ~ Available\_facilities\_and\_services, data = excel\_data, pch = 16, xlab = "Available facilities and services", ylab = "Length of Stay")  
  
#Fit a Linear Regression model  
model\_8 <- lm(Length\_of\_Stay ~ Available\_facilities\_and\_services, data = excel\_data)  
  
#Find b coefficients  
b8\_coefficients <- model\_8$coefficients  
  
#Add the fitted line  
abline(b8\_coefficients, lwd =2, lty =2, col="green")



#Third  
#For Routine\_Chest\_X-ray, plot the Data  
plot(Length\_of\_Stay ~ `Routine\_Chest\_X-ray`, data = excel\_data, pch = 16, xlab = "Routine Chest X-ray", ylab = "Length of Stay")  
  
#Fit a Linear Regression model  
model\_9 <- lm(Length\_of\_Stay ~ `Routine\_Chest\_X-ray`, data = excel\_data)  
  
#Find b coefficients  
b9\_coefficients <- model\_9$coefficients  
  
#Add the fitted line  
abline(b9\_coefficients, lwd =2, lty =2, col="blue")



#Get the summary of each model  
sum7 <- summary(model\_7)  
sum8 <- summary(model\_8)  
sum9 <- summary(model\_9)  
sum7$coefficients[2,4]

## [1] 1.176961e-09

sum8$coefficients[2,4]

## [1] 0.000111339

sum9$coefficients[2,4]

## [1] 2.905559e-05

#Calculate the t score  
b7 <- summary(model\_7)$coefficients[2, 1]  
se\_b7 <- summary(model\_7)$coefficients[2, 2]  
  
b8 <- summary(model\_8)$coefficients[2, 1]  
se\_b8 <- summary(model\_8)$coefficients[2, 2]  
  
b9 <- summary(model\_9)$coefficients[2, 1]  
se\_b9 <- summary(model\_9)$coefficients[2, 2]  
  
#a = 0.05  
t\_7 <- b7 / se\_b7  
t\_8 <- b8 / se\_b8  
t\_9 <- b9 / se\_b9  
  
#Calculate the lower and upper confidence bounds for each variable  
b7 - (qt(0.975, 112) \* se\_b7)

## [1] 0.5336665

b7 + (qt(0.975, 112) \* se\_b7)

## [1] 0.9871753

b8 - (qt(0.975, 112) \* se\_b8)

## [1] 0.02260447

b8 + (qt(0.975, 112) \* se\_b8)

## [1] 0.06681087

b9 - (qt(0.975, 112) \* se\_b9)

## [1] 0.02060314

b9 + (qt(0.975, 112) \* se\_b9)

## [1] 0.05490852