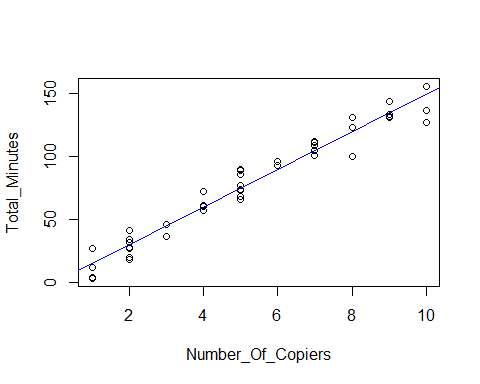
Chad Huntebrinker’s Homework 6

Chad Huntebrinker

2024-10-08

Question 2.14

#Chad Huntebrinker  
  
#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")



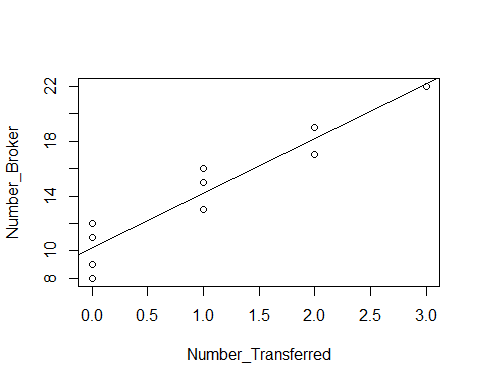
#Get the mean service time for 6 copiers serviced  
Y\_hat <- predict(model\_3, data.frame(Number\_Of\_Copiers=6))  
  
#Now get the confidence interval  
CI <- predict(model\_3, se.fit = TRUE, data.frame(Number\_Of\_Copiers=6), interval = "confidence", level = 0.90)  
  
#Problem 2.14b  
#We're getting a 90% prediction interval when 6 copiers are serviced  
predict(model\_3, data.frame(Number\_Of\_Copiers=6),interval="prediction",level=0.90)

## fit lwr upr  
## 1 89.63133 74.46433 104.7983

#The interval 74.46433 <= Yh <= 104.7983 is wider than   
#the confidence interval in part a and we would expect it to be.  
  
#Problem 2.14d  
#Getting the boundary values of the 90 percent confidence band of the regression line at Xh = 6  
W <- sqrt(2\*qf(1-0.1,2,45-2))  
lower\_bound <- Y\_hat - W \* CI$se.fit  
upper\_bound <- Y\_hat + W \* CI$se.fit  
  
#This confidence band (86.6, 92.7) is wider than the confidence interval in part a and we would expect it to be.

Question 2.15

#Chad Huntebrinker  
  
#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\_4 <- lm(Number\_Broker~Number\_Transferred,data=excel\_data)  
sum\_of\_model\_4 <- summary(model\_4)  
  
#Plot to see the graph  
plot(Number\_Broker~Number\_Transferred, data = excel\_data)  
abline(model\_4)



#Problem 2.15a  
#Getting the mean breakage at X = 2 and 4. Also getting a 99% confidence interval at those values too.  
predict(model\_4, data.frame(Number\_Transferred=2),interval="confidence",level=0.99)

## fit lwr upr  
## 1 18.2 15.97429 20.42571

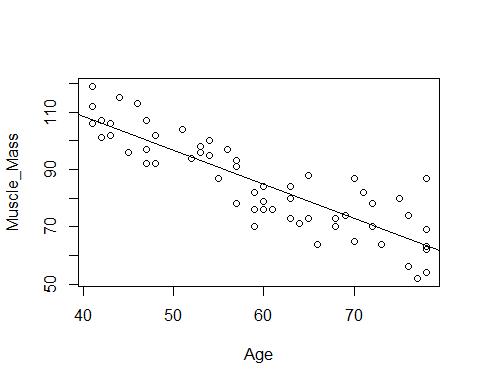
predict(model\_4, data.frame(Number\_Transferred=4),interval="confidence",level=0.99)

## fit lwr upr  
## 1 26.2 21.22316 31.17684

#Problem 2.15d  
#Getting the boundary values of the 99 percent confidence band for the regression line at X = 2 and 4.  
Y\_hat2 <- predict(model\_4, data.frame(Number\_Transferred=2))  
Y\_hat4 <- predict(model\_4, data.frame(Number\_Transferred=4))  
  
W <- sqrt(2\*qf(1-0.01,2,10-2))  
  
CI2 <- predict(model\_4, se.fit = TRUE, data.frame(Number\_Transferred=2), interval = "confidence", level = 0.99)  
lower\_bound2 <- Y\_hat2 - W \* CI2$se.fit  
upper\_bound2 <- Y\_hat2 + W \* CI2$se.fit  
  
CI4 <- predict(model\_4, se.fit = TRUE, data.frame(Number\_Transferred=4), interval = "confidence", level = 0.99)  
lower\_bound4 <- Y\_hat4 - W \* CI4$se.fit  
upper\_bound4 <- Y\_hat4 + W \* CI4$se.fit  
  
#We find that both of these confidence bands are wider than the confidence  
#interval in part a and we would expect that.

Question 2.28

#Chad Huntebrinker  
  
#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\_5 <- lm(Muscle\_Mass~Age,data=excel\_data)  
sum\_of\_model\_5 <- summary(model\_5)  
  
#Plot to see the graph  
plot(Muscle\_Mass~Age, data = excel\_data)  
abline(model\_5)



#Problem 2.28a  
#Get a 95% confidence interval for the mean of the muscle mass (Y) when X = 60  
predict(model\_5, data.frame(Age=60),interval="confidence",level=0.95)

## fit lwr upr  
## 1 84.94683 82.83471 87.05895

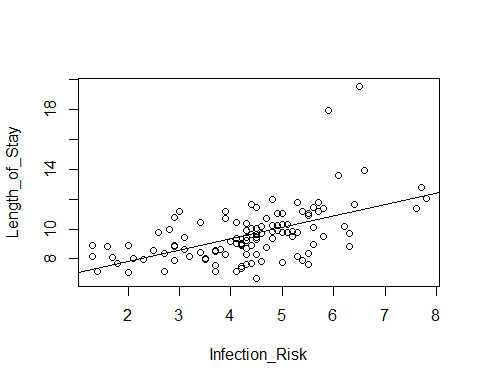
#We find the mean is 84.94683 with an upper bound of 87.05895 and lower bound of 82.83471  
#This means that with 95% confidence we can say that the mean muscle mass is between 82.83 and 87.05  
  
#Problem 2.28b  
#Get a 95% prediction interval for the muscle mass when X = 60  
predict(model\_5, data.frame(Age=60),interval="prediction",level=0.95)

## fit lwr upr  
## 1 84.94683 68.45067 101.443

#With the mean being 84.94683, we find the lower bound is 68.45067 and the upper bound is 101.443  
#This seems to indicate the prediction interval is relatively precise  
  
#Problem 2.28c  
#Get the boundary values of the 95% confidence band for the regression line at Xh = 60  
CI <- predict(model\_5, se.fit = TRUE, data.frame(Age=60), interval = "confidence", level = 0.95)  
  
W <- sqrt(2\*qf(1-0.05, 2, 60-2))  
lower\_bound <- CI$fit[1] - W \* CI$se.fit  
upper\_bound <- CI$fit[1] + W \* CI$se.fit  
  
#We find the confidence band is wider than the confidence interval in part a.  
#Yes, we would expect that.

Question 4.27

#Chad Huntebrinker  
  
#Load the library and data in to R  
library(readxl)  
excel\_data <- read\_excel("SENIC\_Data.xlsx")  
  
#Length\_of\_Stay ~ Infection\_Risk  
#Fit the model  
model\_6 <- lm(Length\_of\_Stay~Infection\_Risk,data=excel\_data)  
sum\_of\_model\_6 <- summary(model\_6)  
  
#Plot to see the graph  
plot(Length\_of\_Stay~Infection\_Risk, data = excel\_data)  
abline(model\_6)



#Problem 4.27a  
#1 - alpha/(2\*m) [where m = 2 since we need one for B0 and one for B1]  
1-0.1/4

## [1] 0.975

confint(model\_6, level = 0.975)

## 1.25 % 98.75 %  
## (Intercept) 5.1523372 7.521236  
## Infection\_Risk 0.5003816 1.020460

#We have the joint confidence interval for as 0.5003816 <= B1 <= 1.020460  
#and 5.1523372 <= B0 <= 7.521236  
  
#Problem 4.27b  
#The researcher is saying that B0 should be about 7 and B1 should be about 1.  
#According to Problem 2.27a, this is supported. Both B0=7 and B1=1 are in the  
#confidence interval. Thus, we can say that this is true when the family confidence  
#coefficient is at least 0.9 that the procedure leads to correct pairs of interval estimates.  
  
#Problem 4.27c and d  
W <- sqrt(2\*qf(1-0.05,2,113-2))  
B <- qt(1-0.05/8, 113-2)  
B > W

## [1] TRUE

#The more efficient method should be the Working-Hotelling. The reason being that we have  
#a larger value of m (m = 4) which means the Working-Hotelling will be more precise. We  
#also see that the W value is less than the B value, leading us to believe that the  
#Working-Hotelling method will lead to a tighter confidence limits.  
#But we'll calculate using both just to see.  
  
X.h <- c(2, 3, 4, 5)  
  
family\_con <- 1-0.05/8  
  
# Estimation  
CI.bf <- predict(model\_6,data.frame(Infection\_Risk = X.h), interval="confidence",level=family\_con) # Adjust level  
  
# Working-Hotelling multiplier for 95% simultaneous confidence band  
# Extract standard errors of Y-hat  
CI.wh <- predict(model\_6,data.frame(Infection\_Risk = X.h), se.fit = TRUE, interval = "confidence",   
 level = 0.95)  
wh.LB <- CI.wh$fit[,1] - W\*CI.wh$se.fit # WH lower bound  
wh.UB <- CI.wh$fit[,1] + W\*CI.wh$se.fit # WH upper bound  
  
#Bonferroni:  
#2: 6.994082 <-> 8.721175: dif = 1.7270924   
#3: 8.011273 <-> 9.224825: dif = 1.2135524   
#4: 8.937812 <-> 9.819129: dif = 0.8813172   
#5: 9.665901 <-> 10.611881: dif = 0.9459808   
  
#Working-Hotelling:  
#2: 7.088991 <-> 8.626266: dif = 1.537275  
#3: 8.077961 <-> 9.158137: dif = 1.080176  
#4: 8.986242 <-> 9.770698: dif = 0.784456  
#5: 9.717885 <-> 10.559897: dif = 0.842012