Chapter 1 HW

Ahmad M. Osman

9/7/2017

knitr::opts\_chunk$set(echo = TRUE)

## Problem 1.19

#Supply the file name for problem 1.19  
gpa <- read.table("CH01PR19.txt",header=F,col.names=c("GPA","ACT"))  
# This command makes data frame column names available as variables  
attach(gpa)  
  
# Plot the data  
plot(ACT,GPA)  
  
# Fit the regression model  
fit <- lm (GPA ~ ACT)  
fit

##   
## Call:  
## lm(formula = GPA ~ ACT)  
##   
## Coefficients:  
## (Intercept) ACT   
## 2.11405 0.03883

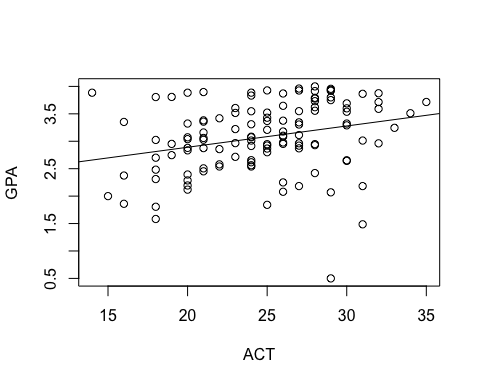
b1 <- fit$coeff[2]  
b0 <- fit$coeff[1]  
anova(fit)

## Analysis of Variance Table  
##   
## Response: GPA  
## Df Sum Sq Mean Sq F value Pr(>F)   
## ACT 1 3.588 3.5878 9.2402 0.002917 \*\*  
## Residuals 118 45.818 0.3883   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ACTat30 <- b0 + b1\*30  
ACTat30

## (Intercept)   
## 3.278863

abline (b0,b1)



ACTat31 <- b0 + b1\*31  
ACTat31

## (Intercept)   
## 3.31769

ACTat31 - ACTat30

## (Intercept)   
## 0.03882713

The intercept is 2.11 GPA units and the slope is 0.0388 GPA unit per ACT score increment.

Explanation:

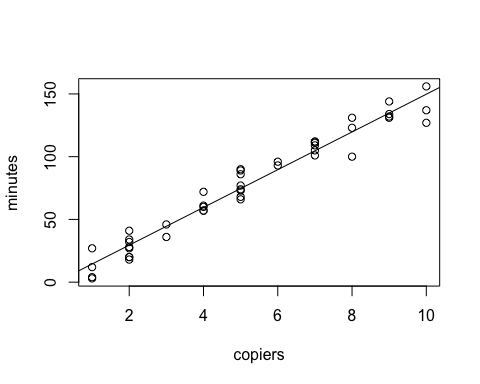
a- Β0 = 2.11405 Β1= 0.03883  
  
 Y = 2.11405 + 0.03883 \* X   
  
b- Yes. It looks good.   
  
c- GPA = 3.278863 for ACT = 30.  
  
d- The change = 0.03882713.

## Problem 1.20

#Supply the file name for problem 1.20  
machines <- read.table("CH01PR20.txt",header=F,col.names=c("minutes","copiers"))  
  
  
# This command makes data frame column names available as variables  
attach(machines)  
  
# Plot the data  
plot(copiers,minutes)  
  
# Fit the regression model  
fit <- lm (minutes ~ copiers)  
fit

##   
## Call:  
## lm(formula = minutes ~ copiers)  
##   
## Coefficients:  
## (Intercept) copiers   
## -0.5802 15.0352

b1 <- fit$coeff[2]  
b0 <- fit$coeff[1]  
abline (b0,b1)



anova(fit)

## Analysis of Variance Table  
##   
## Response: minutes  
## Df Sum Sq Mean Sq F value Pr(>F)   
## copiers 1 76960 76960 968.66 < 2.2e-16 \*\*\*  
## Residuals 43 3416 79   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Explanation:

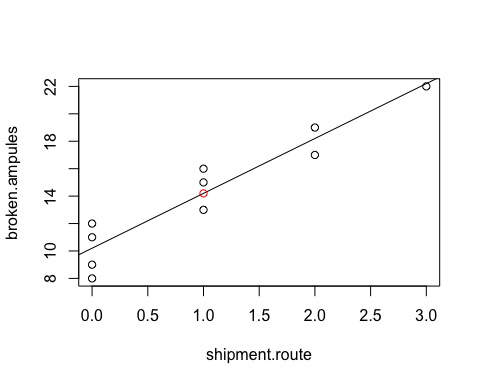
a- Y=-0.5802 + 15.0352 \*X  
  
b- The line fits well.  
  
c- For 0 copiers, it takes -0.58 minutes.  
  
d- The mean service time for 5 copiers = 74.59608 minutes.

## Problem 1.21

#Supply the file name for problem 1.21  
aircraft <- read.table("CH01PR21.txt",header=F,col.names=c("broken ampules","shipment route"))  
  
# This command makes data frame column names available as variables  
attach(aircraft)  
  
# Plot the data  
plot(shipment.route,broken.ampules)  
  
# Fit the regression model  
fit <- lm (broken.ampules ~ shipment.route)  
fit

##   
## Call:  
## lm(formula = broken.ampules ~ shipment.route)  
##   
## Coefficients:  
## (Intercept) shipment.route   
## 10.2 4.0

b1 <- fit$coeff[2]  
b0 <- fit$coeff[1]  
abline(b0,b1)  
points(mean(shipment.route),mean(broken.ampules), col='red')



Explanation:

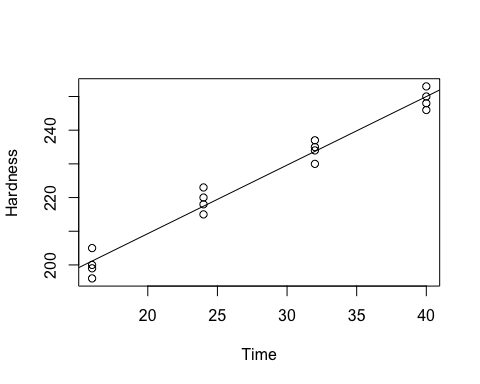
a- Y = 10.2 + 4.0 \* X. Fits well.   
  
b- 14.2  
  
c- As the slope is 4, once we plug it into the equation, every time x increases by 1, makes the answer 4.   
  
d- Explained through the graph.

## Problem 1.22

#Supply the file name for problem 1.22  
Plastic <- read.table("CH01PR22.txt",header=F,col.names=c("Hardness","Time"))  
  
  
# This command makes data frame column names available as variables  
attach(Plastic)  
  
# Plot the data  
plot(Time,Hardness)  
  
# Fit the regression model  
fit <- lm (Hardness ~ Time)  
fit

##   
## Call:  
## lm(formula = Hardness ~ Time)  
##   
## Coefficients:  
## (Intercept) Time   
## 168.600 2.034

b1 <- fit$coeff[2]  
b0 <- fit$coeff[1]  
abline(b0,b1)



Explanation:

a- Y = 2.034x + 168.600   
  
b- 249.975  
  
c- The point of estimate when hardness increases by one is equal the slope (2.034)

## Problem 1.23

a- Using the formula "sum(fit$residuals)", the residual sum = -2.942091e-15. The residual is close to zero. The residual squared is 45.81761.  
b- The anova table stated that the MSE = 0.3883. Standard deviation is the results of the square root of the MSE = 0.623GPA.

## Problem 1.24

a- Using the formula "sum(fit$residuals)", the residual sum = -1.176836e-14. The residual is close to zero. The residual squared is 3416.377.  
b- The anova table stated that the MSE = 79. Standard deviation is the results of the square root of the MSE, SQRT(79) = 8.888 minutes.

## Problem 1.25

a- I residual for the first case is 1.8.  
b- MSE= 2.2. The variance is estimated by the MSE.