Lab 3

David Wiley / Duy Truong

January 30, 2019

Reading in Data

```
# Reading in the data from the file
dat <- read.csv("C:\\Users\\Nick\\Documents\\0_Spring 2019\\Applied Regression\\Labs_HW\\Data_Sets\\Chap

x = dat[,2]
y = dat[,1]
n = length(x)
```

Calculating the Variables

```
# Calculating the values to be used later

Sxx = sum(x^2) - sum(x)^2/n
Sxy = sum(x*y) - sum(x) * sum(y) / n

B1H = Sxy / Sxx

B0H = mean(y)  - mean(x) * B1H

SSres = sum((y-mean(y))^2)-B1H*Sxy
MSres = SSres/(n-2)

seB1H = sqrt(MSres/Sxx)
seB0H = sqrt(MSres * (1 / n + mean(x)^2 / Sxx))
```

Problem 1

For the soft drink delivery time data in Problem 2.9, calculate the 90% confidence interval for the intercept, slope, and variance.

```
# Setting Alpha
alp = 0.1

# %90 for B0
B0L = B0H - qt(1 - alp/2, n - 2) * seB0H
B0U = B0H + qt(1 - alp/2, n - 2) * seB0H

# %90 for B1
B1L = B1H - qt(1 - alp/2, n-2) * seB1H
B1U = B1H + qt(1 - alp/2, n-2) * seB1H

# %90 for sigma^2
sig2L = (n-2) * MSres / qchisq(1 - alp/2, df = n - 2)
sig2U = (n-2) * MSres / qchisq(alp/2, df = n - 2)
```

%90 Confidence Interval

Alpha = `0.1`

The confidence interval for $\hat{\beta}_0$ is `0.9709358 to 5.670624`

The confidence interval for $\hat{\beta}_1$ is `1.963596 to 2.3887374`

The confidence interval for σ^2 is `11.4332004 to 30.7194811`

This means that there is a %90 chance that the CI contains the true population mean.

Problem 2

Plot the data points, the fit, and the 95% C.I. for the mean response in one graph.

The two curved lines are the upper and lower %95 confidence limits for the soft drink delivery time data.

```r
data = data.frame(x = seq(from = 0, to = 30, length.out = 30))
fit = lm(y~x)
lim = as.data.frame(predict(fit, data, level = 0.95, interval = "conf"))
plot(x, y, xlim = c(0, 30), ylim = c(0, max(y)))
abline(fit, col = "blue")

# Plotting the two intervals
lines(cbind(data,lim$lwr, col = "red", lty = "dashed"))
lines(cbind(data,lim$upr, col = "red", lty = "dashed"))
```
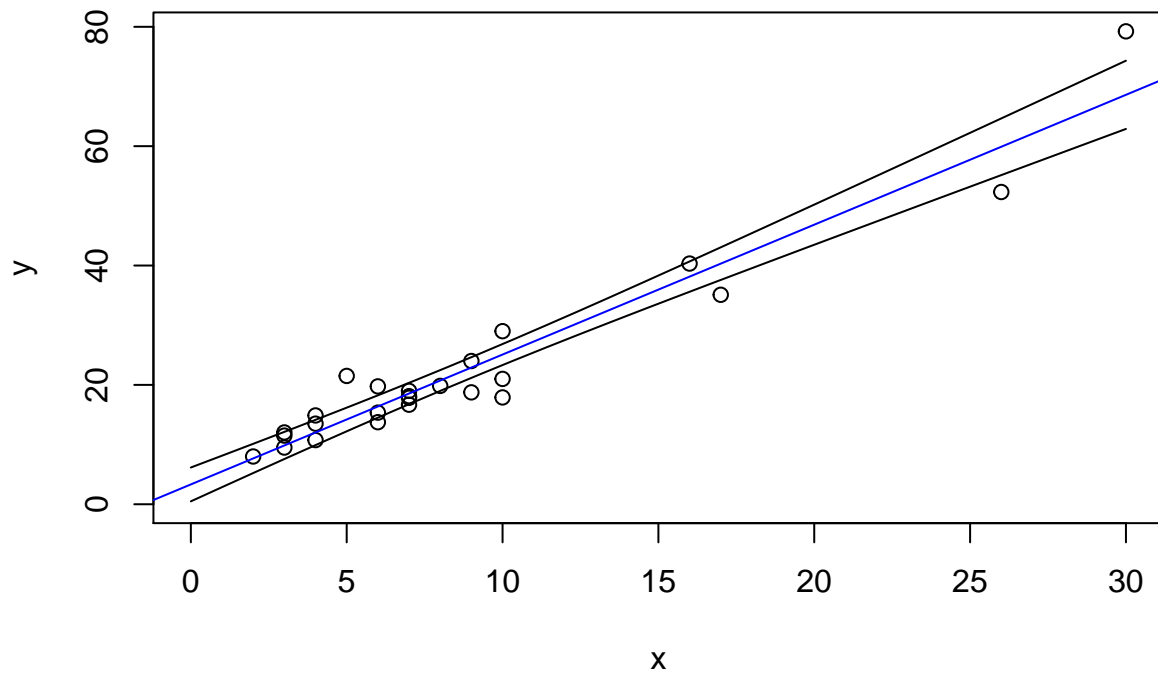
The two curved lines are the upper and lower %95 confidence limits for the soft drink delivery time data. This means that is multiple samples were taken and another CI was calculated, %95 of the intvervals would contain the population mean.