STA-371G
Homework 5

---

This assignment contains 2 questions; there are 51 total points.

## Instructions

Submit your assignment as a PDF/Word document and an R script to Canvas by **10/20/2017 by 7:00PM**. It is your responsibility to follow the homework guidelines on Canvas and ensure that the file you upload is complete and uncorrupted.

## Grading

A small number of the problems on each assignment will be graded for correctness, and the remainder graded for completeness. A complete response answers the question posed and also shows your work. This means showing the steps of a mathematical calculation, or including the R code you used to arrive at your answer. For questions that are not just calculations (e.g., more than computing an expected value from a table) you should answer in complete, concise sentences. Detailed solutions will be available – you should always check your work against these solutions.

---

1. (24 total points)

   A company sets different prices for a particular stereo system in eight different regions of the country. The table below shows the numbers of units sold (in 1000s of units) and the corresponding prices (in hundreds of dollars).

   | Sales | 420 | 380 | 350 | 400 | 440 | 380 | 450 | 420 |
   |-------|-----|-----|-----|-----|-----|-----|-----|-----|
   | Price | 5.5 | 6.0 | 6.5 | 6.0 | 5.0 | 6.5 | 4.5 | 5.0 |

   (The R script accompanying this assignment contains this dataset).

   (a) (8 points) Let $X$ denote Price and $Y$ denote Sales. Using R, calculate the sample means of $X$ and $Y$, sample standard deviations of $X$ and $Y$, sample covariance between $X$ and $Y$, and sample correlation between $X$ and $Y$.

   (b) (4 points) In R, regress sales on price using the `lm` function and obtain the estimates of the intercept $b_0$, slope $b_1$ and coefficient of determination $R^2$. *In your Word document, give the answer (e.g.: $b_0$=3, $b_1$ = 10, $R^2$=0.4), followed by copying and pasting the relevant R output from the* **summary** *function.*

   (c) (2 points) Use the results obtained in (a) to calculate the intercept $b_0$, slope $b_1$ and the coefficient of determination $R^2$. Are the results the same as those obtained in (b)?

(d) (5 points) Present a plot with the data and the regression line. In R, you can add the regression fit to a scatterplot using the **abline** function:

```
# Example for plotting a scatterplot + regression line
# This is just an example, you should modify this to
# plot the sales/price data
dat = data.frame(x = c(1,2,3,4,5),y = c(-1, 2, 0, 5, 4))
example_fit = lm(y~x, data=dat)
plot(y~x, data=dat)
abline(example_fit)
```
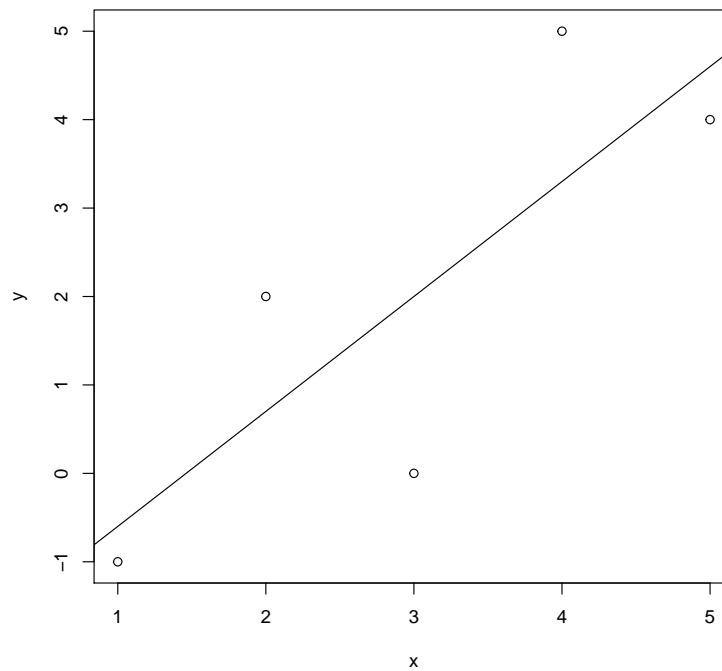


*To add your plot (which should be a plot of the sales and price data, not the above plot!) to your Word document, from RStudio use the Export button in the Plots pane and select "Copy to clipboard". Resize the image appropriately and then click the "Copy Plot" button. Open your Word file and paste the plot at the appropriate point in your document.*

(e) (5 points) Based on this analysis, briefly describe your understanding of the relationship between sales and prices.

2. (27 points)

Read the case "Waite First Securities" in the course packet.

The data for this case are available from the course website, and you can read them into R using the code below (also available in the R script posted with this assignment).

```
library(readr)
# This is a shortcut to download data directly from the course website.
# To work offline, download the file to a folder on your computer
# and edit the path variable to point to that folder. Or you
# can use RStudio's import tool like in the tutorials in HW 0
path = "https://jaredsmurray.github.io/sta371g_f17/data/"
waite = read_csv(paste0(path, 'waite.csv'))

## Parsed with column specification:
## cols(
##   No.   = col_integer(),
##   Yr.   = col_integer(),
##   Mo.   = col_integer(),
##   SP500 = col_double(),
##   HILTON = col_double(),
##   TI = col_double(),
##   GIANT = col_double()
## )
```

Consider the regression model

$$TI_t = \alpha + \beta SP500_t$$

where $TI_t$ represents the return on Texas Instruments in month $t$ and $SP500_t$ represents the return on the S&P 500 in month $t$.

(a) (9 points) If the returns on the SP500 are 0%, what are the predicted returns on TI? If the returns on the SP500 are 1%, what are the predicted returns on TI? If the returns on the SP500 are -1%, what are the predicted returns on TI? What is the interpretation of $\beta$ in terms of a measure of risk of the stock?

(b) (8 points) Plot $TI$ against $SP500$. What graphical evidence is there of a relationship between $TI$ and $SP500$? Does the relationship appear to be linear? Why or why not?

(c) (10 points) Estimate $\beta$. What is the interpretation of this estimate in terms of the risk of the stock? Why is Mr. Gagnon interested in this estimate?