Homework Assignment 2
(Due Wednesday, February 22, 2017)

The homework is due at 10.30am in the dropbox on the Course Plus page (you can find the dropbox under the 'Resources' tab in the upper right). For exercises involving R code, please knit a document from your R markdown (Rmd) file. Generate a single pdf file for your entire submission and give it a name that makes it identifiable (calling it `140.615.HW.Number.Lastname.Firstname` or similar).

1. Suppose $X_1, X_2, X_3, \ldots, X_{10}$ are independent and identically distributed (iid), with mean=3 and SD=3. Calculate the following:
   (a) E$(X_1 + X_2 + \cdots + X_{10})$
   (b) SD$(X_1 + X_2 + \cdots + X_{10})$
   (c) E$[(X_1 + X_2 + \cdots + X_{10})/10]$
   (d) SD$[(X_1 + X_2 + \cdots + X_{10})/10]$

2. Make your results reproducible using the R command `set.seed(10)` at the beginning of your code.
   (a) Generate 10 random numbers from a Uniform distribution on [0,10]. Use R to find the maximum and minimum values.
   (b) Generate 10 random normal numbers with mean 5 and standard deviation 5. Use R to find out how many of these are less than 0.
   (c) Generate 100 random normal numbers with mean 100 and standard deviation 10. How many are 2 standard deviations from the mean?
   (d) Toss a fair coin 50 times using the R function `sample()`. How many heads do you have?
   (e) Roll a fair six-sided die 100 times. How many 6's did you see?
   (f) For $X \sim$ Normal(0,1), find the number x so that $P(X \leq x) = 0.05$.
   (g) For $X \sim$ Normal(0,1), find the number x that solves $P(-x \leq X \leq x) = 0.10$.
   (h) How much area (probability) is to the right of 1.5 for a Normal(0,1)?

3. Suppose a new drug (B) improves upon an older one (A) by increasing the probability of self-reported relief from gastric distress from 40% to 65%. Suppose drug A is given to 10 randomly-sampled patients and drug B is given to a second 10. Let Y_A and Y_B be the number out of 10 patients that will experience relief in each group.
   (a) What is the set of possible outcomes for Y_A?
   (b) Plot the probability distributions for Y_A and Y_B.

4. Suppose that fasting plasma glucose concentrations (FPG) in some population are normally distributed with a mean of 90 mg/dl and standard deviation of 5 mg/dl.

   (a) Among 1,000 people randomly selected from this population, how many would you expect to have FPG less than 80 mg/dl?

   (b) If you draw 25 people from this population, what is the probability that the sample average is larger than 92 mg/dl?

   (c) If you select 5 people from this population, what is the probability that 4 or more of them have a FPG larger than 100 mg/dl?

5. We are interested in estimating the concentration of substance X in the Baltimore water supply on the basis of measurements of a number of samples. Suppose measurements of such samples will be approximately normally distributed with unknown mean (the true concentration) and *known* SD $=$ 1.5 ppb. How many samples should we measure if we wish our 95% confidence interval for the true concentration to have width $<$ 1 ppb?

6. Suppose I measure some treatment response on a set of 10 mice, and receive the following data:

   ```
   84 96 89 103 100 97 125 105 111 108
   ```

   Calculate a 95% confidence interval for the population mean treatment response. Derive the answers analytically (by calculating the sample mean, the sample standard deviation, etc), as well as a call to the R function `t.test()`.