Biometry: Lab 3   Fall 2017
Due: midnight, Sep 13 (Thursday labs); midnight, Sep 14 (Friday labs)

Lab 3: The normal probability distribution

Objectives
- Plot normal probability distributions and cumulative distributions.
- Calculate normal probabilities densities, and quantiles.
- Standardize a normal random variable.
- Use the following R functions: `dnorm, pnorm, qnorm, rnorm, lines, legend, range`.

Exercises
1. In many animals, island populations diverge in size from mainland populations of the same species. You are interested in finding out if this is the case for monitor lizards, which are widespread on the Australian continent and surrounding islands. Based on earlier studies, you believe that the mean and variance for the lengths of adult mainland lizards are as follows: $\mu = 100$ cm, $\sigma^2 = 400$. You also make the reasonable assumption that length is a normal random variable. Before collecting any data on the Kangaroo Island lizards, you will use this prior knowledge to say something about their expected probability distribution. Write an R script to do the following:

   - Plot the probability density function of lizard lengths. The plot must include informative axis titles and an overall title.

   - Plot the cumulative distribution function of lizard lengths, again including informative axis titles and an overall title.

2. Write another script to answer the following questions about what you expect to find when you measure Kangaroo island lizards, assuming that they follow the same probability distribution as mainland lizards. Inspect the plots you made in 1 to check if your answers make sense.
   - What is the probability density for a length of 75 cm?
   - What is the probability that a lizard will be less than or equal to 75 cm?
   - Greater than 120 cm?
   - Between 95 and 115cm?
   - At least 40 cm different from the mean?
   - Closer than 1.3 σ to the mean?
   - Further than 1.5 σ from the mean?
   - Further than 0.7 σ *below* the mean?
   - What are the quartiles of the distribution?
   - 2/3 of observations are expected to lie below what value?
   - 80% of observations are expected to lie above what value?

3. Write a script to explore the effect of changing the variance of length. Your script should make two plots, one above the other:
   - The first should show three PDFs on the same graph, each with a mean of 100, but with different variances (100, 400, and 625). Plot each line in a different color. If necessary, adjust the axis lengths to make the best use of space.
   - The second plot should show the three corresponding CDF's.
   - Add a legend to each plot to indicate the identity of each curve.

4. Britain's domestic intelligence service MI5 places an upper limit on the height of its spies, on the assumption that people who are too tall do not blend in well with the crowd. To be a spy, men must be no taller than 180 cm (~5 feet 11 inches) and women no taller than 173 cm (~5 feet 8 inches).
   - If the mean height of British men is 177 cm, with a standard deviation of 7.1 cm, what proportion of British men are excluded from being spies by this height restriction? Assume that height follows a normal distribution.
   - The mean height of British women is 163.3 cm, with a standard deviation of 6.4 cm. Assuming a normal distribution of female height, what fraction of women meet MI5's height standard?
   - Imagine that MI5 wants to change its maximum height for female spies. Its goal is to exclude the same proportion of women as men. What should the new maximum height for women be? (Round your answer to the nearest centimeter.)
   - Sean Connery, the original James Bond, is 183 cm tall. By how many standard deviations does he exceed the height limit for spies?

5. It is often helpful to think of a probability distribution in terms of standard deviations, rather than the actual units of measure of a random variable. That way you can answer questions that apply to any normal variable, regardless of its mean or variance. Write an R script to answer the following questions:
   - What is the probability that a normal random variable will have a value within 1 standard deviation of the mean?
   - What is the probability that it will be within 5 standard deviations of the mean?
   - Fill in the blank: A normal random variable has a 50% probability of lying within____ standard deviations of the mean.
   - Fill in the blank: A normal random variable has a 95% probability of lying within____ standard deviations of the mean.
   - Fill in the blank: A normal random variable has a 99% probability of lying within____ standard deviations of the mean.

Assignment: Turn in a single Microsoft Word document with a heading that includes your name, the lab number (Lab 2 for this lab), and your section (day and time). Create a subheading for each exercise. For each exercise, include (1) your annotated R script, (2) your annotated output, and (3) your answer to the final

"Answer" item for each exercise (where applicable). In addition to the word document, you must also submit a separate .R file with all of your R scripts.