Assignment 4: Due to Oct 21, 2019

*** Please print your assignment and submit it to the lecturer on the due date, thanks for your cooperation.**

1. Following are 16 samples from a normal population with mean μ and unknown standard deviation:

   2.59 2.67 2.16 1.95 2.61 1.11 2.62 2.06 2.06 1.66 2.16 3.35 2.46 2.55 3.12 1.92

(a) Compute an estimate for σ.
(b) Compute an estimate for μ.
(c) Find the 95% confidence interval for μ derived from the 16 samples.
(d) Find the 80% confidence interval for μ derived from the 16 samples.

2. Suppose birth weights in a population are normally distributed with a mean of 109 oz and a standard deviation of 13 oz,

   (a) Simulate (randomly take) 5 subjects (n=5) from the distribution of the birth weights (generate a normal distribution with mean of 109 and standard devialtion of 13), compute the mean of the sample, and then repeat the process 500 times in order to approximate the population mean and calculate the standard error. And then perform same simulations with n=20, 100 and 500. Display the boxplots for the distribution of means of samples with different sample sizes.

   (b) Estimate the chance of obtaining a birth weight of 141 oz or heavier by taking 1000 random samples from the distribution of the birth weight above, each with 100 subjects. If the standard deviation increases to 26 oz, do estimation again by simulation.

   * Hint: R package: ggplot; rnorm(n, mean, sd); boxplot ()

3. A student collected a large amount of demographic data from school children in a depressed area. Since this population was possibly malnourished, she was concerned that the children would have a hemoglobin level below the healthy average. The healthy average is 13 g/dL.

   She had collected a sample of size 120 children. Sample hemoglobin levels: Mean = 11.7 g/dL, Standard deviation = 3.2 g/dL. Please firstly generate a normal distribution (μ=13, σ=3.2), and then perform random sampling with a size of 120. Repeat it 1000 times and estimate the probability of the **sample mean** equal to or less than 11.7 g/dL.

4. Recall: the example 1 in Lecture 4 was shown as below:

Assume the true difference is zero. Please answer the question using computer simulation: is

	Cases (n=263)	Controls (n=1241)
Any antidepressant drug ever	120 (46%)	448 (36%)
	46%	**36%**

Difference=10%

10% bigger or smaller than the expected sampling variability? What's the result if the study samples 100 cases and 100 controls?

* Hint: rbinom(n, size, prob)