Homework 9, Your Name.

~~Problem #1 (Central Limit Theorem)~~

~~In similar fashion to what I did for sample proportion in the lecture code, proceed to use R and generate the following values:~~

$$\cancel{X_1, X_2, ..., X_{10,000} \sim Uniform(-2, 2)}$$

~~subsequently demonstrating (once again, using R) the aspects of sampling distribution for $\bar{X}$, such as its~~

- ~~shape (is it bell-shaped, or not quite?)~~
- ~~mean (is it unbiased?)~~
- ~~standard deviation (does it correspond to theoretical st. dev.?)~~

~~For more details on Uniform distribution (e.g. its mean, variance, etc) check https://en.wikipedia.org/wiki/Uniform_distribution_(continuous).~~

~~Do it for:~~

a. ~~$n = 2$;~~

b. ~~$n = 50$.~~

Problem #2 (Coverage of Confidence Interval)

1. ~~Write your own function that takes as arguments:~~

- ~~total # of observations,~~
- ~~number of successes~~
- ~~confidence level (default at 0.95)~~

~~and calculates the confidence interval for population proportion, outputting a 2-element vector (left and right end of the interval, respectively).~~

2. ~~Proceed to generate 10,000 values from $Bin(n = 1000, p = 0.6)$ distribution, and use your function from part 1 in order to calculate & record the 95% and 90% intervals for each of those 10,000 values. Calculate the % of times your confidence intervals actually contain the true population proportion $p = 0.6$. Is it what you expect for 95% and 90% intervals, accordingly?~~

~~Some code to start you off:~~

```r
set.seed(1)

n.sim <- 10000
prob <- 0.6
size <- 100
# Placeholders for left (first column) and right (second column) ends
# of our CIs.
ci.95 <- matrix(0, nrow=n.sim, ncol=2)
ci.90 <- matrix(0, nrow=n.sim, ncol=2)

# Here you will need to
# 1) generate the 10,000 values from Bin(1000,0.6),
# 2) loop through those and feed them as input to your confidence level function from part 1
#     (for cases of 95% and 90%)
#...

# That's an example of how you calculate the % of times your confidence interval
# contains the true parameter
mean(ci.95[,1] < prob & ci.95[,2] > prob)
#...
```

Problem #3

We will now consider the *Boston* housing data set, from the *MASS* library.

(a) ~~Based on this data set, provide an estimate for the population mean of *medv*. Call this estimate $\hat{\mu}$.~~

~~(b) Provide an estimate of the standard error for $\hat{\mu}$ via~~

- ~~central limit theorem (use theoretical formula);~~
- ~~bootstrap (over 10,000 replicates).~~

~~How do these compare to each other?~~

(c) ~~Based on this data set, provide an estimate, $\hat{\mu}_{med}$, for the median value of *medv* in the population.~~

~~(d) We now would like to estimate the standard error of $\hat{\mu}_{med}$. Unfortunately, there is no simple formula for computing the standard error of the median. Instead, estimate the standard error of the median using the bootstrap.~~

Problem #4 (Make sure to use R when appropriate for carrying out the calculations, showing your work)

~~**7.7**~~

~~**7.14**~~

~~**7.15 (not necessary to provide the graph, but make sure to play with the app at https://istats.shinyapps.io/sampdist_cont/)**~~

7.20

~~**8.6**~~

~~**8.13**~~

~~**8.16**~~

~~**8.25-8.26**~~

~~**8.29-8.30**~~

~~**8.37**~~