ISyE 2028, Fall 2015
Homework 4
100 points total.

This homework is due Thursday Oct. 22 in class.

- Please remember to staple if you turn in more than one page.
- Please make sure to **SHOW ALL WORK** in order to receive full credit.

1. **Mean Confidence Interval Estimation**

   (a) Temperature estimate.

   8-21. An article in the *Journal of Agricultural Science* ["The Use of Residual Maximum Likelihood to Model Grain Quality Characteristics of Wheat with Variety, Climatic and Nitrogen Fertilizer Effects" (1997, Vol. 128, pp. 135–142)] investigated means of wheat grain crude protein content (CP) and Hagberg falling number (HFN) surveyed in the UK. The analysis used a variety of nitrogen fertilizer applications (kg N/ha), temperature (°C), and total monthly rainfall (mm). The data shown below describe temperatures for wheat grown at Harper Adams Agricultural College between 1982 and 1993. The temperatures measured in June were obtained as follows:

15.2	14.2	14.0	12.2	14.4	12.5
14.3	14.2	13.5	11.8	15.2	

   Assume that the standard deviation is known to be $\sigma = 0.5$.

   (a) Construct a 99% two-sided confidence interval on the mean temperature.
   (b) Construct a 95% lower-confidence bound on the mean temperature.
   (c) Suppose that we wanted to be 95% confident that the error in estimating the mean temperature is less than 2 degrees Celsius. What sample size should be used?
   (d) Suppose that we wanted the total width of the two-sided confidence interval on mean temperature to be 1.5 degrees Celsius at 95% confidence. What sample size should be used?

   (b) The pH levels of a random sample of 16 chemical mixtures from a process were measured, and a sample mean $\bar{x} = 6.861$ and a sample standard deviation $s =$

0.440 were obtained. The scientists presented a confidence interval $(6.668, \infty)$ for the average pH level of chemical mixtures from the process. What is the confidence level of this confidence interval?

(c) A sample of 30 data observations has a sample mean $\bar{x} = 14.62$ and a sample standard deviation $s = 2.98$. Find the value of c for which $\mu \in (-\infty, c)$ is a one-sided 95% t-interval for the population mean μ. Is it plausible that $\mu \leq 16$?

2. **Proportion Parameter Estimation**

   (a) American's View on Illegal Drugs. A *CNN/ORC Poll* conducted in Jan. 2014, asked the following question: "Do you think the use of marijuana should be made legal, or not?". Go to:

   `http://www.pollingreport.com/drugs.htm`

   i. Based on the poll's results, calculate a 95% confidence interval for p, the proportion of all American adults who *oppose* the legalization, and interpret your interval in context.

   ii. The report provides the information that the margin of error equals ± 3 (note that it says in blue above the results" "Margin of error ± 3"). Write one or two sentences interpreting this value that could be understood by someone who does not know anything about statistics.

   (b) Recent Gallup Poll estimates that 88% Americans believe that cloning humans is morally unacceptable. Results are based on telephone interviews with a randomly selected national sample of $n = 1000$ adults, aged 18 and older, conducted May 2-4, 2004.

   i. Find 95% confidence interval for the true proportion? Does 0.9 fall in the interval?

   ii. Pretend that you want to replicate Gallup's inquiry in the Atlanta area. What sample size is needed so that the length of a 95% confidence interval for the unknown proportion of people in the area who believe that cloning humans is morally unacceptable does not exceed 0.02?

   iii. How would you change the sample size if you are sure that the true proportion exceeds 0.7?

3. **Computer Problem**

   Background. The deflection temperature under load for two different types of plastic pipes is being investigated. Two random samples for each pipe type (with sample sizes equal to 33 and 38 respectively) are tested, and the deflection temperature observed is recorded (in $^\circ F$).

   Instructions for reading the data. The data file is included with this assignment (file 'Deflection.txt'). In these data, the first column indexes whether the data is for the first pipe (pipe=1) or the second one (pipe=2). The 'Temperature' variable consists of the deflection temperature.

   To read the data in R, save the file in your working directory (make sure you have changed the directory if different from the R working directory) and read the data using the R function `read.table`.

```
data = read.table("Deflection.txt",header=FALSE)
```

We write the two variables (index for the first and second pipes) and the deflection temperature into two vectors:

```
index = data[,1]
deflection = data[,2]
```

In order to separate the deflection temperature for the two pipe types, we need to run the following commands:

```
index.1 = which(index==1)
index.2 = which(index==2)
pipe1 = deflection[index.1]
pipe2 = deflection[index.2]
```

The vector 'pipe1' consists of the deflected temperatures for pipes of type 1, whereas 'pipe2' is that for pipes of type 2.

(a) Is the deflected temperature with the 71 pipes normally distributed? You may check normality using the `hist` function R which constructs a histogram for your data:

```
hist(deflection)
```

Construct the two histograms of the deflection variable for pipes of type 1 and that of type 2. Use the function 'hist' and the defined vectors above to obtain these histograms. Can we assume normal distribution for the possible deflection for pipes of type 1? Comment on the shape of the two histograms.

(b) What are the point estimates for the following parameters:

- μ_1 = the deflection for pipes of type 1.
- μ_2 = the deflection for pipes of type 2?

What are standard deviations of the *estimates* of the two parameters?

To obtain these estimates you may use the function `mean` in R. For example, you want to compute the mean of the values in a vector called 'values', you will perform the following command:

```
mean(values)
```

To obtain the σ^2 variance of a vector of values (e.g. values), you can run the following R command

```
var(values)
```

The function in R for the square root of a value is `sqrt`. For example, to compute the square root of the variance of the vector 'values', the R command is

```
sqrt(var(values))
```

(c) Construct 95% confidence intervals for the mean deflection temperature for pipes of type 1, and that of type 2. Use the t-interval for that. Compare the two intervals. What are your concluding remarks for this case study?