

This lab is a bit different from previous ones. Your answer will be a 2-3 page write-up of your solutions to the problems below. For this reason, **you are welcome to form groups of 2-3 individuals if you want**. You have been given more time to complete this assignment, and I have also increased the **total points to 20** (worth four “normal” labs).

**Commands needed for R:**

By and large, this lab will utilize confidence intervals. In real-world statistics, the Z-test is seldom used, since to use it you must know  $\sigma$ , which requires knowledge of  $\mu$ , which you don't know.

(1) **Confidence interval for  $\mu$  when  $\sigma$  unknown (“t-interval”):**

- The command `t.test()` should be used. This will provide you with more information than you need right now (we'll use it more in later chapters), but this seems to be the easiest way for now.
- If you have a data set `y`, then a 99% confidence interval can be found using

```
t.test(y, conf.level=.99)
```

After running this command, you should see a large amount of verbiage (most of which you don't need), and about halfway down you should see

```
99 percent confidence interval:
```

and the next line has your two bounds

(2) **Confidence interval for  $p$ :**

- There is a `prop.test()` command in R, but it constructs its confidence intervals using another method from what we use in class. To get the same intervals we use, you will need to define a new function. Copy and paste the following into your RStudio console:

```
p.interval=function(x,n,conf.level=.95)
{c((x/n)-qnorm(1-(1-conf.level)/2)*sqrt(((x*n)-x^2)/n^3),
(x/n)+qnorm(1-(1-conf.level)/2)*sqrt(((x*n)-x^2)/n^3))}
```

This command will (temporarily) create a new command (“`p.interval`”) which you can use to calculate confidence intervals for proportions. You will need to input `x` (the number of individuals in the category of interest), `n` (the sample size, as always) and the confidence level (the default is 95%, so if you do not enter a third argument, it will assume 95%). For example,

```
> p.interval(1272,2230)
```

has output

```
[1] 0.5498581 0.5909491.
```

The output here is the 95% confidence interval. To change this, use something like

```
> p.interval(1272,2230,conf.level=.99)
```

(3) **Confidence interval for  $\sigma$ :**

- Once again, RSudio does not have a single command for these intervals. Instead, use the following command to create the temporary function `var.interval()`:

```
var.interval=function(data, conf.level=.95)
{c((length(data)-1)*var(data)/qchisq((1-conf.level)/2,length(data)-1),
  (length(data)-1)*var(data)/qchisq((1-conf.level)/2,
  length(data)-1,lower.tail=FALSE))}
```

You can now feed `var.interval()` a data set and a confidence level, and it should return the appropriate  $\chi^2$  interval for you.

**Problems:**

- (1) Your company has been hired by the government to analyze the rivets used in the construction of the F-22 Raptor. The construction requires that the mean diameter of the rivets be .375cm, with a standard deviation of at most .005cm. Also, any rivets under .372cm or over .378cm will not be able to be used, and the government wants to know how much of the shipment will be unusable, if it accepts the shipment. The government has given you measurements of the diameters of 200 such rivets (provided in the spreadsheet), and is asking your company whether the F-22 project should accept delivery of 300,000 of these rivets from the manufacturer.

Your answer to this problem should be a Word document which provides the following information:

- (a) What should be your company's recommendation to the government?
- (b) A technical write up of how you arrived at your answer. This should be no more than one page in length. Clarity of expression is valued much more than the length of the response.
- (2) Computer processors of different speeds are manufactured on a single assembly line. Variations and errors in the process of manufacturing alter the speed at which these processors can operate. The company tests each processor as it leaves the assembly line to determine its speed and sells it at an appropriate price. Here is the revenue index:

Tested Speed $x$ (GHz)	Advertised Speed	Selling price
$x < 3.2$	Discarded	N/A
$3.2 \leq x < 3.4$	3.2GHz	\$125
$3.4 \leq x < 3.6$	3.4GHz	\$155
$3.6 \leq x < 3.8$	3.6GHz	\$195
$3.8 \leq x < 4.0$	3.8GHz	\$255
$4.0 \leq x$	4GHz	\$325

A company has produced a sample of 100 processors, whose speeds have been provided to you in the spreadsheet. They want to produce 450,000 processors on the same assembly line, and would like you to estimate their potential revenue if they go ahead with their plan.

Your answer to this problem should be a part of the Word document you submit for number (1). However, you should start a new page for this problem. Your answer should include the following:

- (a) A range of potential revenues that the company might make, and how confident you are that the actual revenue will be in this range.
  
- (b) A 1-2 page explanation of how you arrived at your estimates. Describe the statistical methods that brought you to your final answer. There are a number of methods you might try to employ, all of which will involve a number of different steps. There will be a variety of correct answers, depending on the choices you make. That means you need to tell me what you did clearly so I can evaluate your method. Keep your answer brief and precise, but also provide the necessary details.