We use the following notation:

p = population proportion $\hat{p}$ = sample proportion
μ = population mean $\bar{x}$ = sample mean

Confidence Intervals
A confidence interval is an interval of plausible values for the population parameter of interest:

confidence interval for p is $\hat{p} \pm m$
confidence interval for μ is $\bar{x} \pm m$

where m = margin of error shows how accurate we believe our estimate is, based on the variability of the estimate.

The confidence level of a confidence interval is the probability that the interval covers the true unknown parameter *prior to collecting the data.* Once the data is collected, our interval either covers the true unknown parameter or it does not – something we do not know. That is why we say we are "95% confident" that the true parameter lies within a given interval. We don't know if it is in there or not but the method gives confidence intervals which do contain the parameter 95% of the time (or whatever the selected level is, such as 90%, 99%.)

- We can make a confidence interval smaller by choosing a larger sample size or a smaller confidence level.
- The data should be a simple random sample from the population of interest.
- The margin of error only indicates how much error can be expected because of chance variation – it does not cover errors from undercoverage, nonresponse, and other sampling difficulties.

Example 1
In a 2007 study at Michigan State University, researchers examined the relationship between use of Facebook and the formation and maintenance of social capital (defined to be the resources accumulated through the relationships among people). In a random sample of size 286, the researchers found that the average amount of time MSU undergraduate students spent on the internet was 2 hour and 56 minutes per day. Suppose the standard deviation is 1.54 hours. Construct a 95% CI for the population mean time spent per day on the internet. Interpret the confidence interval. Source: Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), article 1. http://jcmc.indiana.edu/vol12/issue4/ellison.html

We can interpret the CI in two ways:

a. We are 95% confident that the population mean time on the internet spent by MSU college students per day is between 2.8 hours and 3.1 hours (rounded to 1 decimal place).
b. If we constructed many of these CIs, 95% of them would contain the population mean time per day spent by MSU college students on the internet.

Hypothesis Tests (tests of significance)

A test of significance allows us to formally test an assumption about a population using observed data. There are two pieces: the Null Hypothesis = Ho and the Alternative Hypothesis = Ha. These are always statements about POPULATION parameters and should involve either p or μ.

The null hypothesis typically reflects the current state or the status quo. The alternative hypothesis is usually the effect we would like to support using our data.

There are one-sided and two-sided hypothesis tests:
One-sided:
Ho: $\mu = 14$ vs. Ha: $\mu > 14$
Ho: $\mu = 14$ vs. Ha: $\mu < 14$

Two-sided:
Ho: $\mu = 14$ vs. Ha: $\mu \neq 14$
The p-value reflects the degree of evidence against the null hypothesis. It is the probability of seeing the observed data or something more extreme, assuming the null hypothesis is true. A small p-value means this is unlikely so we REJECT the null hypothesis.

Typically p-values less than 0.05 are rejected but before collecting the data an investigator may decide the *level of significance* they would like for their test. This is the α-level of the test. An $\alpha = 0.05$ level test will reject the null hypothesis only if the p-value is less than 0.05. Otherwise we say there is insufficient evidence to reject the hull hypothesis and conclude Ha.

- The smaller the p-value, the stronger the evidence against the null hypothesis.

- When we reject a null hypothesis at a given level α, we usually say the test is "statistically significant" at that α level.

- A statistically significant result does not necessarily mean it is practically significant. See problem 6.82 as an example of this.

- If you reject Ho at $\alpha = 0.01$ then you will also reject it at $\alpha = 0.05$ because the p-value must be less than 0.01 (and hence 0.05). This generalizes to other α values.

- Two-sided hypothesis tests are related to confidence intervals. If you perform a two-sided hypothesis test and REJECT the null hypothesis at an $\alpha = 0.05$ level, then a 95% confidence interval for the parameter will NOT contain the null hypothesis value. If you reject at $\alpha = 0.01$ then a 99% CI will not contain the value.

  If you cannot reject Ho at a given level, then the corresponding CI will contain the value.

  For example, consider testing Ho: $\mu = 14$ vs. Ha: $\mu \neq 14$ at $\alpha = 0.01$ level. If the 99% CI for μ contains 14 then we will not reject Ho. That is, the data are consistent with the null hypothesis because the CI contains all of the plausible values for the population mean μ.

<u>Example 2</u>
Suppose that prior to collecting their data, the MSU researchers believed that MSU college students spent no more than 2 hours and 45 minutes per day on the internet, on average. Set up and test their hypothesis. Is this test statistically significant?

95% confidence interval results:
μ : population mean
Standard deviation = 1.54

Mean	n	Sample Mean	Std. Err.	L. Limit	U. Limit
μ	286	2.9333334	0.091062106	2.754855	3.1118119

Hypothesis test results:
μ : population mean
$H_0 : \mu = 2.75$
$H_A : \mu < 2.75$
Standard deviation = 1.54

Mean	n	Sample Mean	Std. Err.	Z-Stat	P-value
μ	286	2.9333334	0.091062106	2.013278	0.978

STAT 202
Lab #10 – Introduction to Confidence Intervals and Hypothesis Tests
Due: At the end of class today.

Turn in only one answer sheet per group please.

Software such as StatCrunch makes the computation of confidence intervals and hypothesis tests straight forward. For this activity you will be computing and interpreting the results of these statistical analyses.

1. Recently, 73% of first-year college students responding to a national survey identified "being very well-off financially" as an important personal goal. A state university finds that 132 of an SRS of 200 of its first-year students say that this goal is important. Test the hypothesis that the proportion of all first-year students at the state university who identify being well-off as an important goal is different from the national survey proportion. (Take 73% to be the TRUE population proportion of all first-year college students who identify "being very well-off financially" as an important personal goal." Technically this national survey is also based on a sample of students and we'll learn in Chapter 8 how to use two samples to compare proportions.)

    a. State the hypotheses you will test.

    b. Perform a test of these hypotheses in StatCrunch using the following steps.

        1. Got to the STAT menu.
        2. Select PROPORTIONS.
        3. Select ONE SAMPLE.
        4. Select WITH SUMMARY.
        5. Enter the NUMBER OF SUCCESSES = 132
        6. Enter the NUMBER OF OBSERVATIONS (this is the sample size) = 200
        7. Click on NEXT.
        8. Click on HYPOTHESIS TEST.
        9. Enter the appropriate NULL PROPORTION, 0.73 for this example.
        10. Choose the correct ALTERNATIVE hypothesis.
        11. Click on CALCULATE.

    The resulting table contains the estimate, the test statistic (called Z-stat), and the p-value. Identify these values from the table of results.

    $\hat{p}$ =

    z* =

    p-value =

    c. Perform an $\alpha = 0.05$ significance level test. State your conclusions.

d. Give a 95% confidence interval for the proportion of all first-year students at the state university who would identify being well-off as an important personal goal. To do this in StatCrunch, repeat the steps above but select CONFIDENCE INTERVAL in step (8) and enter 0.95 for the CONFIDENCE LEVEL.

e. Interpret your confidence interval.

f. What is the relationship between the results of your hypothesis test in and the confidence interval you found in (d)?

2. The distribution of the amount of medication in a particular pill is known to vary normally with unknown mean μ and standard deviation 10 mg. A pharmaceutical company that manufactures this pill advertises that they contain 500 mg of a particular medication. The plant manager where the medication is produced believes the machinery is incorrectly calibrated and the amount of medication in the pills is actually more than 500 mg. She takes a simple random sample of 15 pills and has them analyzed for the amount of medication they contain. The amount in the 15 pills, in mg, is:

507	503	493	500	488	514	502	520	489	503	511	495	496	498	504

Perform the appropriate hypothesis test by answering the following questions:

a. State the null and alternative hypotheses.

b. Enter the data into the StatCrunch data table.

c. Go to STAT/Z STATISTICS/ONE SAMPLE/WITH DATA. Select the variable. Enter 10 for the standard deviation and select NEXT. Enter the correct NULL: MEAN and ALTERNATIVE hypothesis. Click on CALCULATE.

d. From the StatCrunch output:

   i. What is the sample mean?

   ii. What is the test statistic?

   iii. What is the p-value?

   iv. What conclusions can you draw about the machinery?

3. In the previous problem, suppose instead that the plant manager had taken an SRS of size n = 50 pills. Repeat the previous problem using the same sample mean (501.53) and same population standard deviation (10) but assume that the sample mean was found from an SRS of size n = 50.

   Go to STAT/Z STATISTICS/ONE SAMPLE/WITH SUMMARY.
   Enter 501.53 for the mean, 10 for the standard deviation, and 50 for the sample size.
   Select NEXT. Enter the correct NULL: MEAN and ALTERNATIVE hypothesis. Click on CALCULATE.

   What conclusions can you draw about the machinery?

4. Repeat the previous problem with n = 200. What conclusions do you now draw about the machinery? How does changing the sample size change the hypothesis test results?

Rejection Regions: Lab

1. The level of calcium in the blood of healthy young adults follows a normal distribution with mean μ = 10 milligrams per deciliter and standard deviation σ = 1.0. A clinic measures the blood calcium of 100 healthy pregnant young women at their first visit for prenatal care. The mean of these 100 measurements is $\overline{x}$ = 9.8. Is this evidence that the mean calcium level in the population from which these women come is not equal to10?

   a. State the hypotheses you will test.

   b. If the null hypothesis is true then draw the density curve of the sample means (You need to find the standard deviation of the $\overline{x}$ first)

   c. If we are using a .05 significance level, use the normal calculator on statcrunch to find the rejection levels for the sample means. (Mean=10 and Stan dev.= standard deviation of $\overline{x}$.)

   Also with a two tailed test, you will have two rejection regions of area .025

   Lower Rejection Level__________ Upper Rejection Level__________

   d. Shade in the rejection regions on the density curve in part b. Explain what you have found.

   e. Calculate the test statistic without using statcrunch (How many standard deviations away from the proposed mean is the sample mean?)

   f. Would you reject with a .05 significance. State your conclusions.

   g. Verify that you have found the correct test-statistic and find the p-value by running the hypothesis test in statcrunch

2. The national mean SAT Math Score is 514. The mayor claims that DCPS students are *below* the national average. The standard deviation of the population is 100.

The national mean SAT Math Score is 514. The mayor claims that DCPS students are below the national average

a. What would the null hypothesis be? The alternative hypothesis?

b. Assume the null hypothesis is true and you sample 400 random DCPS students. Draw the density curve for the sample means.

c. If we are using a .05 significance level, use the normal calculator on statcrunch to find a rejection level for the sample means. Shade in the rejection region on the density curve in part b. Explain what you have found. (This is just a one tailed test so there will be only one rejection region)

e. Now assume that you have calculated a sample mean of 506. Should you reject the null hypothesis. Why?

g. Find the p-value and interpret

f. What conclusions can you make? What would you tell the mayor?