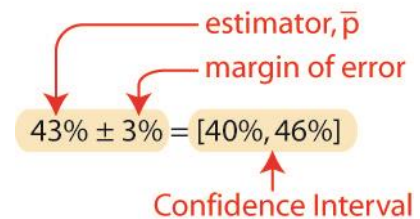# Chapter 6 – Confidence Intervals

This chapter introduces a technique for statistical inference, confidence intervals.

6.1 – Introduction

1. As stated in Chapter 1, statistical inference is when we take the information from a *sample* and use various statistical techniques to produce conclusions about a *population*. We will focus on two types of statistical inference:
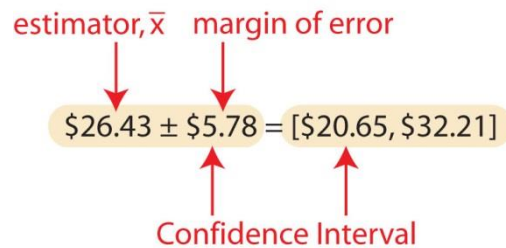
    1. Confidence Intervals (Chapter 6)
    2. Hypothesis Tests (Chapter 7)

2. Example – In a Gallup poll[1] taken June 8-10, 2015 of 1500 people, it was found that 43% of the sample approved of President Obama's job performance. A confidence interval allows us to make a conclusion about the population of voters:

    estimator, $\overline{p}$

    margin of error

    $43\% \pm 3\% = [40\%, 46\%]$

    Confidence Interval

    The confidence interval above is interpreted to mean that in the entire population, the true proportion of voters who approve of the president's job performance is between 40% and 46% at the 95% confidence level (we explain *confidence level* shortly). Notice that other (CNN, Fox, *etc.*) polling agencies[2] obtained results that are mostly in this range over the same time period and same question. The sample proportions in those polls ranged from 42% to 47%.

3. Example – From a student project. A 90% confidence interval on μ, the true, unknown average amount of money men spend at Walmart in Valdosta is:

    estimator, $\overline{x}$ margin of error

    $\$26.43 \pm \$5.78 = [\$20.65, \$32.21]$

    Confidence Interval

    How do we interpret this interval?

    *We have 90% confidence (again, we explain this shortly) that **on average** men spend between $20.65 and $32.21 at Walmart in Valdosta.*

4. A *confidence interval* (*CI*) is an *interval estimate* about a *population parameter*. In this chapter we study the population parameters:

    1. μ is the true, unknown mean of a population
    2. p is the true, unknown proportion of a population that have a characteristic of interest.

5. All confidence intervals have this form:

    $$Estimator \pm Margin\ of\ Error$$

    When we are drawing inference on μ, $\overline{x}$ is the estimator of μ. In other words, we estimate μ with the sample average. We will talk about the margin of error later.

---

[1] http://www.gallup.com/poll/113980/Gallup-Daily-Obama-Job-Approval.aspx
[2] http://www.realclearpolitics.com/polls/

6. Confidence intervals include two levels of uncertainty or error:

   a. Margin of error – sometimes called the *precision* of the estimate. A confidence interval is *more precise* when the margin of error is small and in that case we say that we have a *tighter* estimate of μ.

   b. Level of confidence – We you make a confidence interval, you must choose a *confidence level*. Typically one of these values is used: 80%, 90%, 95%, 99%. A natural question is why not 100%? We will address that later.

      Let's explain *confidence level* by way of an example. Suppose you have a population that you know has $\mu = 10$ and you take a sample of size 30 from that population. Next, you decide to us a confidence level of 90% and calculate a 90% confidence: [9,13]. We note that this CI *does* contain the true value of μ, *i.e.* $\mu = 10$ is in the range [9,13]. It is entirely possible that you take another sample of size 30 from this population and you calculate a CI that *does not* contain $= 10$, *e.g.* [6,9]. So, what does *confidence level* mean?

      If you took 100 samples, each of size 30 and calculated a 90% CI for each sample, you would expect 90 of the intervals to contain the true value of the μ and 10 to not. Of course in practice you never know whether a confidence interval you calculate is *correct* or not because in practice you do not know μ (if you did then you would be making a confidence interval!). Thus, we say that the technique we use to calculate a confidence interval works 90% of the time. This, however, is not correct:

      There is a 90% probability that the average amount men spend at Walmart in Valdosta is between $20.65 and $32.21.

7. A confidence interval about μ is a statement about the *true average (mean) of a population*. It does **not** concern individual items from the population. For example: consider the 90% confidence interval given earlier: [$20.65 and $32.21], this is **not** correct:

   90% of men shopping at Walmart in Valdosta spend between $20.65 and $32.21.

   Remember that the *empirical rule* can be used to make inferences (crude) about *individual* items provided that data is bell-shaped.

8. Example – 90% CI's for the mean amount men and women spend at Walmart are shown on the right. What conclusion can we draw?

| Men: | [$20.65, $32.21], $26.43 ± $5.78 |
| Women: | [$38.05, $48.09], $43.07 ± $5.02 |
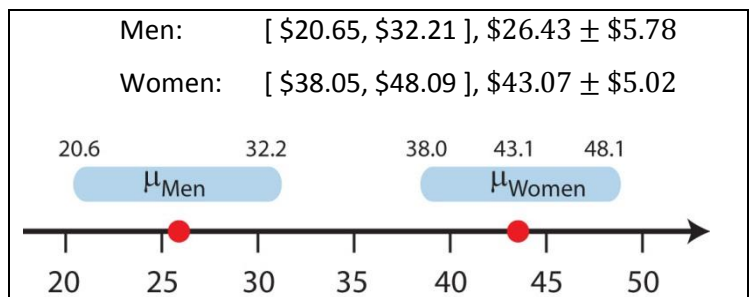
9. Example – 90% CI's for the mean amount people spend at Walmart in Valdosta and Savannah are shown on the right. What conclusion can we draw?

Valdosta: [$27.31, $42.77], $35.04 \pm $7.73

Savannah: [$29.10, $39.80], $34.45 \pm $5.35
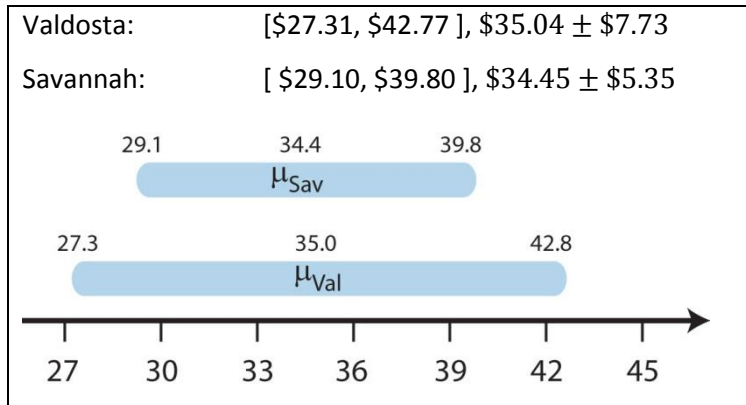
10. Definitions – Suppose a 95% confidence interval on the mean time (seconds) to open a combination lock is [5 *sec*, 11 *sec*].

   a. **Width** of the confidence interval is:

$$Width = Upper\ Limit - Lower\ Limit = 2E$$

$$= 11 - 5 = 6\ sec$$

   b. **Margin of error** is:

$$E = \frac{Width}{2} = \frac{6}{2} = 3\ sec$$

   c. **Point Estimate** of μ (given a confidence interval) is:

$$\bar{x} = Lower\ Limit + E = Upper\ Limit - E = \frac{Lower\ Limit + Upper\ Limit}{2}$$

$$= 5 + 3 = 11 - 3 = \frac{5+11}{2} = 8$$

6.2 – Calculator Directions, Large Sample Confidence Interval about the Mean

1. To calculate confidence intervals on the calculator for the situation where you have a large sample, $n \geq 30$, choose: Stats, Tests, ZInterval (#7). Your screen will look like one of the two depictions below, depending on whether *Stats* or *Data* is highlighted.

ZInterval
    Inpt: Data **Stats**
    σ:
    $\bar{x}$:
    n:
    C-Level:
    Calculate

If you have summary statistics (the standard deviation, average, and sample size), choose *Stats* as shown above. Supply the appropriate values, and press *Calculate*.

ZInterval
    Inpt: **Data** Stats
    σ:
    List:
   Freq: 1
    C-Level:
    Calculate

If you have raw data in a list, choose *Data* as shown above. At the *List* prompt, supply the list your data is in (*e.g.* L1, L2, *etc.*). Always use a *Freq* of 1. Notice, also that you must supply σ in which case you would normally use s_x from the numerical descriptive statistics (Stat, Calc, 1, List#, Enter).

2. Example – The life in hours of a 75-watt light bulb has unknown mean and distribution. A SRS of size 50 is collected and it is found that the sample mean is 1014 hours and the sample standard deviation is 25 hours. (a) Find 80%, 90%, 95%, 99% confidence intervals on the true mean lifetime of the bulbs (b) Find the width of each interval, (c) Find the margin-of error of each interval, (d) Make a table of the results and draw a conclusion, (e) Interpret the 95% interval.

Solution:

Calculator	Calculator	Calculator	Calculator
ZInterval	ZInterval	ZInterval	ZInterval
Inpt: Data *Stats*	Inpt: Data *Stats*	Inpt: Data *Stats*	Inpt: Data *Stats*
σ: 25	σ: 25	σ: 25	σ: 25
$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014
n : 50	n : 50	n : 50	n : 50
C-Level:0.8	C-Level:0.9	C-Level:0.95	C-Level:0.99
Calculate	Calculate	Calculate	Calculate
[1009.5, 1018.5]	[1008.2, 1019.8]	[1007.1, 1020.9]	[1004.9, 1023.1]

| Confidence Level	Confidence Interval	Width	Margin of Error
80%	[1009.5, 1018.5]	9.0	4.5
90%	[1008.2, 1019.8]	11.6	5.8
95%	[1007.1, 1020.9]	13.8	6.9
99%	[1004.9, 1023.1]	18.2	9.1
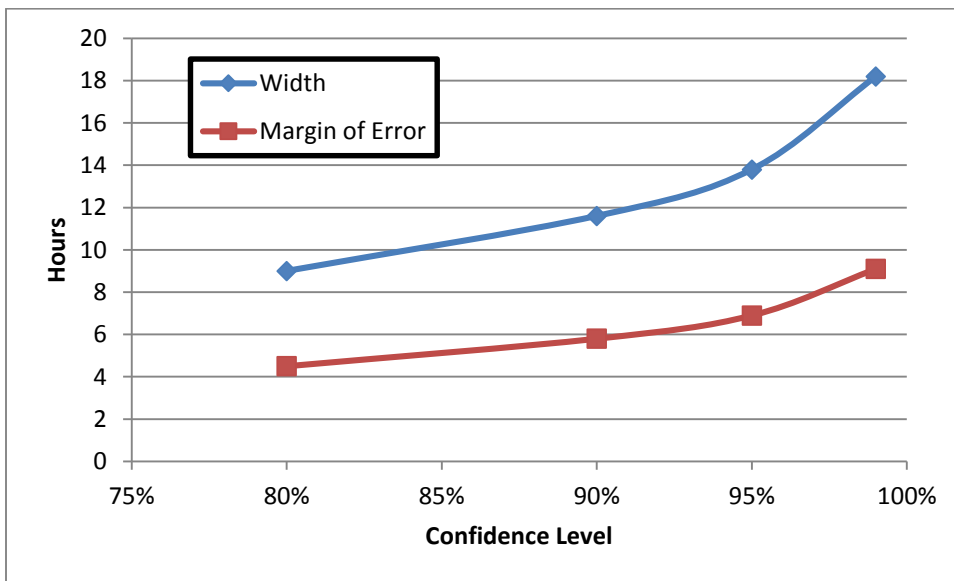
- The larger the confidence level the larger the margin of error (or width) is resulting in a *less precise* interval.

(e) We are 95% certain that the true mean lifetime of the bulbs is between 1007 and 1020 hours.

3. Example – This is very similar to the last problem. The life in hours of a 75-watt light bulb has unknown mean and distribution. A SRS of size 50 is collected and it is found that the sample mean is 1014 hours and the sample standard deviation is 25 hours. (a) Find a 95% confidence interval on the true mean lifetime of the bulbs (b) Repeat using a sample sizes of 100, 200, and 500 (c) Find the width and margin-of error of each interval, (d) Make a table of the results and draw a conclusion

Calculator	Calculator	Calculator	Calculator
ZInterval	ZInterval	ZInterval	ZInterval
Inpt: Data **Stats**	Inpt: Data **Stats**	Inpt: Data **Stats**	Inpt: Data **Stats**
σ: 25	σ: 25	σ: 25	σ: 25
$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014
n : **50**	n : **100**	n : **200**	n : **500**
C-Level:0.95	C-Level: 0.95	C-Level: 0.95	C-Level: 0.95
Calculate	Calculate	Calculate	Calculate
[1007.1, 1020.9]	[1009.1, 1018.9]	[1010.5, 1017.5]	[1011.8, 1016.2]

Sample Size	Confidence Interval	Width	Margin of Error
50	[1007.1, 1020.9]	13.8	6.9
100	[1009.1, 1018.9]	9.8	4.9
200	[1010.5, 1017.5]	7.0	3.5
500	[1011.8, 1016.2]	4.4	2.2
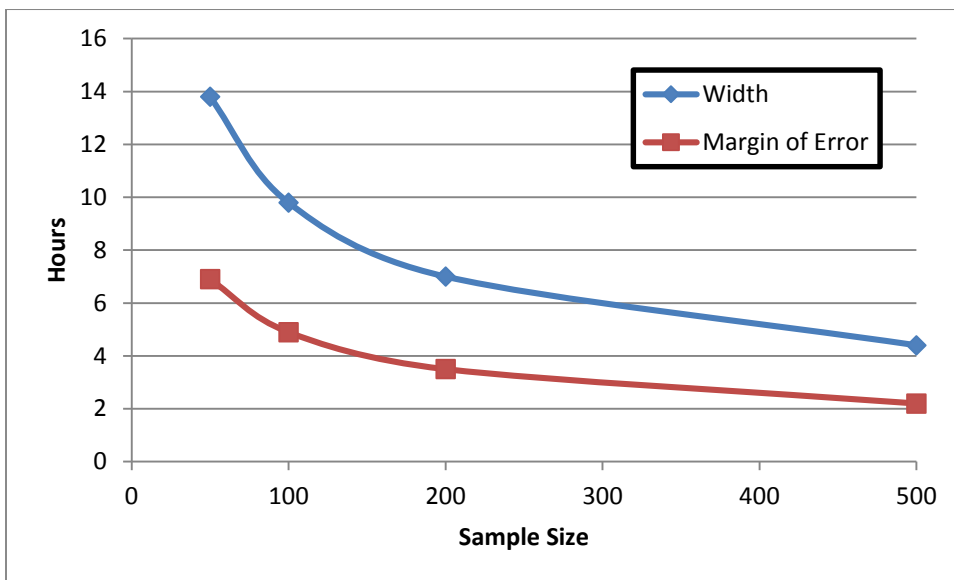
- The larger the sample size the more precise the interval is, *e.g.* the smaller the margin of error (width).

4. Example – Again, this is very similar to the last problem. The life in hours of a 75-watt light bulb has unknown mean and distribution. A SRS of size 50 is collected and it is found that the sample mean is 1014 hours and the sample standard deviation is 25 hours. (a) Find a 95% confidence interval on the true mean lifetime of the bulbs (b) Repeat using a sample standard deviation of 30, 35, and 40 (c) Find the width and margin-of error of each interval, (d) Make a table of the results and draw a conclusion

Calculator	Calculator	Calculator	Calculator
Inpt: Data **Stats**	Inpt: Data **Stats**	Inpt: Data **Stats**	Inpt: Data **Stats**
σ: **25**	σ: **30**	σ: **35**	σ: **40**
$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014	$\bar{x}$: 1014
n : 50	n : 50	n : 50	n : 50
C-Level:0.95	C-Level: 0.95	C-Level: 0.95	C-Level: 0.95
Calculate	Calculate	Calculate	Calculate
[1007.1, 1020.9]	[1005.7, 1022.3]	[1004.3, 1023.7]	[1002.9 1025.1]

Standard Deviation	Confidence Interval	Width	Margin of Error
25	[1007.1, 1020.9]	13.8	6.9
30	[1005.7, 1022.3]	16.6	8.3
35	[1004.3, 1023.7]	19.4	9.7
40	[1002.9 1025.1]	22.2	11.1
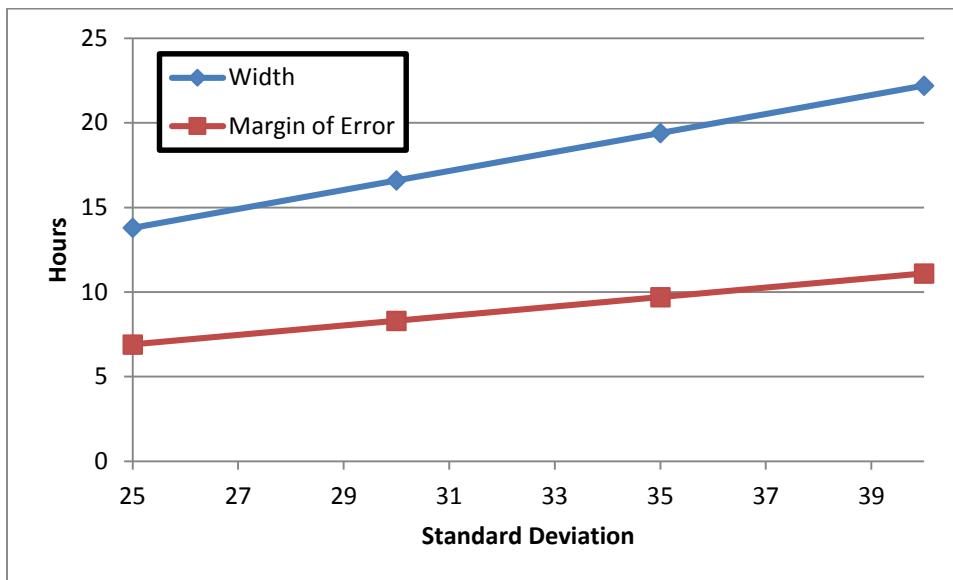
- The larger the standard deviation the less precise the interval is, *e.g.* the larger the margin of error (width).

Homework 6.1

1. In a study of sleeping habits of college students it was found that the sample average was 7.2 hours per night and the sample standard deviation was 0.44 hours per night. These descriptive statistics were computed from a random sample of size 81. (a) What is the standard error of the mean? (b) Find a 95% confidence interval on the true mean time that college students sleep.

2. A company wants to find an interval estimate of the true mean clock speed of their new microprocessor. A sample of 49 microprocessors revealed that the sample average was 2.95 *Ghz* with a standard deviation of 0.23 *Ghz.* (a) Find a 90% confidence interval on the true mean clock speed. (b) Find the margin of error. (c) Find a 95% confidence interval.

3. A sample of 36 Live Oak leaves is taken and the sample average is found to be 2 inches with a standard deviation of 0.5 inches. A sample of 100 Magnolia leaves is taken and the sample average is found to be 3.5 inches with a standard deviation of 0.7 inches. (a) Find a 95% confidence interval for the average width of Live Oak leaves. (b) Find a 95% confidence interval for the average width of Magnolia leaves. (c) The inferences on which population of leaves is more precise? (hint: which has the smallest margin of error?)

4. Answer each with either *increases* or *decreases*

   a. Hold the level of confidence and the standard deviation constant. As the sample size increases, the width of the CI _____.
   b. Hold the level of confidence and the standard deviation constant. As the sample size decreases, the width of the CI _____.

c. Hold the sample size and standard deviation constant. As the confidence level increases, the width of the CI _____.

d. Hold the sample size and standard deviation constant. As the confidence level decreases, the width of the CI _____.

e. Hold the level of confidence and the standard deviation constant. As the sample size increases, the margin of error _____.

f. Hold the level of confidence and the sample size constant. As the standard deviation increases, the width of the confidence interval _____.

g. Hold the level of confidence and the sample size constant. As the standard deviation increases, the precision of the confidence interval _____.

5. Fill in the blanks in the table below with the word *Increases* or *Decreases.*

		Margin of Error	**Width**	**Precision**
as *n*	Increases			
	Decreases			
As *s*	Increases			
	Decreases			
As Confidence Level	Increases			
	Decreases			

6.3 – Small Sample Confidence Interval for the Mean

1. To calculate confidence intervals on the calculator for the situation where you have a small sample, $n < 30$, choose: Stats, Tests, TInterval (#8). **However, you must assume that the data is normally distributed.** Your screen will look like one of the two depictions below, depending on whether *Stats* or *Data* is highlighted.

TInterval
    Inpt: Data **Stats**
    $\bar{x}$:
    s_x:
    n:
    C-Level:
    Calculate

If you have summary statistics (the standard deviation, average, and sample size), choose *Stats* as shown above. Supply the appropriate values, and press *Calculate*.

TInterval
    Inpt: **Data** Stats
    List:
    Freq: 1
    C-Level:
    Calculate

If you have raw data in a list, choose *Data* as shown above. At the *List* prompt, supply the list your data is in (*e.g.* L1, L2, *etc.*). Always use a *Freq* of 1.

2. Example – SAT scores are normally distributed. At the completion of GED program, 10 students are randomly selected to take the SAT. The results showed a mean of 740 and a standard deviation of 40. (a) Find a 95% CI on the true mean SAT score of a graduate of the GED. (b) Find the margin of error.

TInterval
    Inpt: Data **Stats**
    $\bar{x}$: 740
    s_x: 40
    n: 10
    C-Level: 95
    Calculate

a. $[711, 769]$

b. $\frac{769 - 711}{2} = 29$. Thus, the confidence interval is: 740 ± 29

3. Example – A company manufactures plastic containers and is interested in estimating the average melting temperature. A sample of 20 cups revealed that the average melting temperature was 190 degrees with a standard deviation of 10 degrees. Find a 99% confidence interval for the true mean melting temperature. Assume that melting temperatures are approximately normally distributed.

TInterval	$[183.6, 196.4] \approx [184, 196]$
Inpt: Data **Stats**	
$\bar{x}$: 190	
s_x: 10	
n: 20	
C-Level: 99	
Calculate	

4. The boiling point of a mixture is being investigated. Six batches of the mixture are produced and the boiling point of each is determined. The results are: 50.1, 51.3, 47.3, 52.5, 48.7, 49.4 degrees Celsius. Assume that boiling points of this mixture are normally distributed. Provide a 90% CI on the average boiling point of the mixture. Use your calculator.

Put data in List 1: choose: Stat, Edit, Edit, and type your data in L1

TInterval	$[48.357, 51.409] \approx [48.4, 51.4]$
Inpt: **Data** Stats	
List: L1	
Freq: 1	
C-Level: 90	
Calculate	

Homework 6.2

1. An engineers union in Columbus, Ohio claims that the average pay rate for women working in the auto industry is $16 per hour. A simple random sample of 9 women working in 1998 as engineers in the automotive industry in Columbus, Ohio revealed that the average pay rate was $15 per hour with a sample standard deviation of $2. Assume that the pay rate is normally distributed. (a) Find a 99% confidence interval on the true mean pay rate of women engineers. (b) What can you say about the union's claim? Why? Write a sentence or two.

2. A sample of 5 shrimp are weighed and the results showed weights of 12, 14, 10, 12, 13 grams. (a) Find an 80% confidence interval on the true weight of shrimp. (b) What is the margin of error? (c) What is the width of the confidence interval?

6.4 – Theory

1. Definitions

   Our goal is to use statistical inference to estimate μ, the *true mean* of some population.

   a. **Point Estimate (Estimator)** – $\bar{x}$ is a point estimate of μ is. How good is the estimate? In general, $\bar{x}$ is an *unbiased* estimator of μ. This means that "on average" it is correct, *i.e.* $E[\bar{x}] = \mu$.

   b. **Standard Error** – This is a measure of the variability in the estimator. The standard error of $\bar{x}$ is $\frac{\sigma}{\sqrt{n}}$ which we estimate with $\frac{s}{\sqrt{n}}$.

   c. **Interval Estimate** of μ is also called a **confidence interval** on μ.

   d. **Sampling Error** – The sampling error is defined as $|\bar{x} - \mu|$. This is the true error. Obviously we would like to know the sampling error, but it *unknown* (because μ is unknown). The margin of error estimates the maximum sampling error.

   e. **Confidence Level** – The confidence level is defined as $1 - \alpha$. For example, when the confidence level is 95%,
   $$0.95 = 1 - \alpha \implies \alpha = 0.05$$

   Similarly, when the confidence level is 99%,

   $$0.99 = 1 - \alpha \implies \alpha = 0.01$$

   But what is α? It is called the *significance level*. We will discuss this in the next chapter.

   As stated earlier, common values for confidence levels are: 99%, 95%, 90%, 80%. The person doing the confidence interval *picks* the confidence they want or require.

2. **Confidence Interval Form** – Earlier, we said all confidence intervals are of the form:

$$Estimator \pm Margin\ of\ Error$$

   Now, we expand that a bit by looking more carefully at the margin of error:

$$Estimator \pm (Table\ Statistic) * (Standard\ Error\ of\ Estimator)$$

3. **Confidence Interval Formulas**

	Large Sample (Z Interval)	Small Sample (T Interval)
Assumptions	$n \geq 30$, thus the sample average has an approximate normal distribution because of the Central Limit Theorem	$n < 30$ and data is normally distributed. In this case the sample average has a t-distribution
Point Estimate	$\bar{x}$	$\bar{x}$
Standard Error	$\dfrac{\sigma}{\sqrt{n}}$	$\dfrac{s}{\sqrt{n}}$
Sampling Distribution	Normal distribution	t-distribution
Confidence Interval	$\bar{x} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$	$\bar{x} \pm t_{\alpha/2,n-1} \dfrac{s}{\sqrt{n}}$
Margin of Error	$z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$	$t_{\alpha/2,n-1} \dfrac{s}{\sqrt{n}}$
Table Statistic	$z_{\alpha/2}$	$t_{\alpha/2,n-1}$

- $z_{\alpha/2}$ (or $t_{\alpha/2,n-1}$) is called a *table statistic* because in the past, you would find this value in a table in the back of your book. Now, we can use our calculator and it computes the table statistic for us (we will do this shortly).

- Note: we use s to estimate σ when it is unknown, which is usually the case.

- Note that we have not considered the case where the distribution of the data is unknown and the sample size is small. In this case the central limit theorem does not work, so there is nothing we can do except collect more data. None of the procedures we have learned will work for this case.

4. **Theory** – The sample average for a large sample has an approximate normal distribution because of the Central Limit Theorem. This is shown by the red curve in the figures below. Remember, even if the true average of the population is μ, when we take a sample, we will almost certainly NOT get exactly this value. The sample average will vary. How it varies is indicated by the red curve. The theory says that when the true mean is μ, the standard deviation is σ, and we take a SRS of size n, then the sample average $\bar{x}$ should fall in the yellow region below (left) with probability $1 - \alpha$. For example, the figure on the right below says that 95% of sample means should fall in the yellow region. Note that this follows the Empirical Rule we studied earlier: if you go out 2 standard deviations then that should be about 95% of the data. Instead of 2, the exact value is 1.96 as shown in the figure.
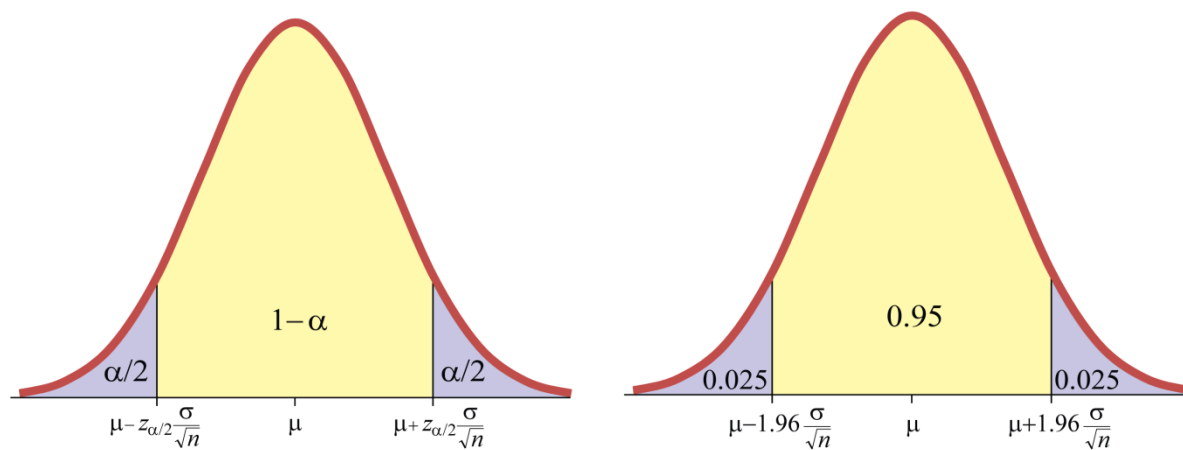
The lower and upper limits in the figure on the left,

$$\mu \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

This *looks* like the confidence interval formula we presented earlier. Of course, we do not know μ. So, what we do in practice is to substitute $\bar{x}$ for μ which gives us the confidence interval formula:

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Making this substitution we leave the realm of probability. Thus, the interval, $\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ either *contains* the true value of μ or *it does not*. If we were to take another sample, we would have another value of $\bar{x}$ and thus another interval, which may or may not contain μ. The theory says that if we repeat this process, $100(1-\alpha)\%$ of intervals will contain the true value of μ. In the case where $\alpha = 0.05$, 95% of intervals will contain μ.

As stated previously, thus, when we compute a 95% confidence interval, we say that we are using a procedure that works (contains μ) 95% of the time. This is the true interpretation of a confidence interval. It is incorrect to say that there is a 95% probability that the interval contains μ. To bridge the gap between the previous two statements, we say things like: We are 95% *sure* that the interval contains μ. Notice also that this approach does not say that μ is more likely to be close to $\bar{x}$ and less likely to towards the end of the interval. It says that μ can be anywhere in the interval.

6.5 – Sample Size Calculations

1. An important practical problem is, "how much data do we need to collect?" In many situations, each data value costs a significant amount of money to collect. Thus, we want to collect just enough data to meet our needs.

   For instance, suppose that we want to estimate the average length of local phone calls to within 1 minute. How much data would be required? More generally, suppose that we want to estimate the mean of some population to within E units with a certain amount of confidence when the standard deviation is σ, how much data is required?

2. We remember that:

$$Margin\ of\ Error = E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Solving for the sample size, n:

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$$

Thus, given:

    a. Margin of error, E
    b. Confidence level, $1 - \alpha$, (from which we can calculate $z_{\alpha/2}$)
    c. Standard deviation: σ

We can calculate the required sample size.

3. How do we calculate: $z_{\alpha/2}$? – The z-value in the formula above (and for the confidence interval formula presented earlier) is found using your calculator. Follow these steps:

    1. $Confidence\ Level = 1 - \alpha$.
    2. Solve for α.
    3. Compute $\alpha/2$.
    4. Calculate: $z_{\alpha/2} = -invNorm(\alpha/2)$,

       Remember: *invNorm* is found on the calculator by pressing: 2nd+Distr, 3.

    a. Example – Suppose that we want to calculate a 99% Confidence Interval. What is the appropriate z value?

       $1 - \alpha = 0.99 \iff \alpha = 0.01 \iff \alpha/2 = 0.005$

       $z_{\alpha/2} = z_{0.005} = -invNorm(0.005) = 2.576$

    b. Example – Suppose that we want to calculate a 95% Confidence Interval. What is the appropriate z value?

       $1 - \alpha = 0.95 \iff \alpha = 0.05 \iff \alpha/2 = 0.025$

       $z_{\alpha/2} = z_{0.025} = -invNorm(0.025) = 1.96$

4. The formula for calculating the sample size requires that he have σ, but where do we get σ? After all, we haven't collected any data yet.

    a. Previous or similar study – Many times a production manager will know the standard deviation of the products they produce because similar studies have been done before, maybe even on a daily basis.

    b. Pilot Study – Sometimes we will do a small experiment, sometimes collecting as few as three data values and using those to calculate the standard deviation. Then we would use this value in the formula to calculate the required sample size.

c. Guesstimate – We could use the empirical rule to estimate the standard deviation. For instance, suppose that someone tells you that, "most of the times my parts are between 20 *mm* and 28 *mm* long." If you assume the data is bell-shaped and that *"most"* means 95%, then you can use the empirical rule and work backwards to find that an estimate of the standard deviation.

Remember that with 95%, the empirical rule says that about 95% of data should fall within: $\bar{x} \pm 2s$. Thus, with this problem we see that $\bar{x} \approx \frac{20+28}{2} = 24$. Thus, we see that $24 \pm 4 = [20, 28]$, which means that $2s = 4$, and finally that: $s = 2$. So, we could use $\sigma = 2$ in the formula above. In general, we can say, using the 95% assumptions, that $s = \frac{range}{4}$.

5. Example – How much data is necessary in order to estimate the mean annual income of U.S. factory workers to within $500 using 95% confidence? Assume a standard deviation of $5000.

$$1 - \alpha = 0.95 \iff \alpha = 0.05 \iff \alpha/2 = 0.025$$

$$z_{\alpha/2} = z_{0.025} = -invNorm(0.025) = 1.96$$

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left(\frac{1.96(5000)}{500}\right)^2 = 384.16 \to 385$$

Thus, with a sample of 385, we are 95% certain that the largest the margin of error will be is 500.

6. Example – A study of the average credit card bill in a restaurant revealed a margin of error of $10 and a standard deviation of $22. The manager would like to do another study in order to obtain a tighter estimate of the average credit card bill. He would like to cut the margin of error in half. How much data should he collect. Use 90% confidence.

$$1 - \alpha = 0.9 \iff \alpha = 0.1 \iff \alpha/2 = 0.05$$

$$z_{\alpha/2} = z_{0.05} = -invNorm(0.05) = 1.645$$

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left(\frac{1.645(22)}{5}\right)^2 = 52.39 \to 53$$

Homework 6.3

1. The Statewide Insurance Company used a simple random sample of 36 policy-holders to estimate the mean age of the population of policy holders. The resulting margin of error was reported to be 2.35 years using a 95% confidence level. The population standard deviation was assumed to be 7.2 years. (a) How large of a sample would be necessary to reduce the margin of error to 2 years? (b) How large of a sample would be required in order to reduce the margin of error to 1 year? (c) How much larger, using a percentage, is the sample size from (b) compared to the sample size from (a)? (d) How large of a sample would be required to estimate the true mean with a margin of error of no more than 2 years if we require 99% confidence? (e) How much larger, using a percentage, is the sample required from (d) compared to the sample required from (a)? Hint: refer to the Ch. 2 Appendix.

6.4 – Proportions

1. **Proportion, *p*** – The *true* (generally unknown) proportion (fraction or percentage) of a population that have a characteristic of interest is denoted by *p*.

$$p = \frac{Count \text{ of items in population that have a characteristic}}{Total \text{ number of items in population}}$$

2. **Sample Proportion, $\bar{p}$** – The proportion of a sample that have a characteristic of interest.

$$\bar{p} = \frac{\# \text{ with characteristic}}{total \text{ number in sample}} = \frac{Count}{n}$$

We use $\bar{p}$ as an estimate of p.

3. Example – Survey 200 people during the morning commute on a subway and find that 136 will eat lunch out at a restaurant or shop, 52 brought their lunch, and 16 do not plan to eat lunch. Thus,

$$\bar{p}_{eat\ out} = \frac{136}{200} = 0.68 = 68\% \qquad\qquad \bar{p}_{skip\ lunch} = \frac{12}{200} = 0.06 = 6\%$$

$$\bar{p}_{brought\ lunch} = \frac{52}{200} = 0.26 = 26\%$$

4. Thus, we will use $\bar{p}$ to estimate p in a similar way that we used $\bar{x}$ to estimate μ.

5. Casual inferences on proportions are ubiquitous. Generally, we should ask these questions:

   - How was the data collected? Was it a simple random sample?
   - What population does the sample represent?

6.7 Confidence Intervals About Proportions

1. Our goal is to use statistical inference to estimate p, the *true proportion* of some population that some characteristic of interest.

2. **Confidence Interval Formula – Theory**

	1 Proportion Z Interval
Assumptions	At least 10 items have the characteristic of interest and at least 10 items do not.
Point Estimate	$\bar{p}$
Standard Error	$\sigma_{\bar{p}} = \sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$
Sampling Distribution	Normal distribution

Confidence Interval	$\bar{p} \pm z_{\alpha/2}\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$
Margin of Error	$z_{\alpha/2}\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$
Table Statistic	$z_{\alpha/2}$

- Notice that $z_{\alpha/2}$ is the same value as from the large-sample confidence interval on μ. In fact, both confidence intervals are based on the Normal distribution.

- This procedure is valid when at least 10 elements of the sample *have* the characteristic of interest *and* at least 10 elements *do not have* the characteristic of interest. For instance, suppose that you survey 100 Valdosta residents and find that 5 people own a vacation home in the North Georgia mountains. You *cannot* draw inference on the true percentage of residents of Valdosta that own homes in the North Georgia mountains.

3. Calculator Directions – To calculate a confidence interval for a proportion on the calculator, choose: Stat, Tests, 1-PropZInt (Alpha+A). Your screen will look like this:

  1-PropZInt
      x:
      n:
      C-Level:
      Calculate

- *X* is the count of the number of items in the sample that have the characteristic of interest and *n* is the sample size.

4. Example – A large polling agency polls 1600 adult Americans at random and the results are shown in the table on the right. (a) Find a 95% CI on the true proportion of Americans that make $60,000 to $90,000 a year. (b) Find a 95% CI on the true proportion of Americans that make more than $60,000 a year.

1997 Yearly Salary	Number
less than $30,000	480
between $30,000 and $60,000	800
between $60,000 and $90,000	224
between $90,000 and $120,000	64
more than $120,000	32

a.

Calculator	Result
x:224	[0.123, 0.157] =
n : 1600	[12.3%, 15.7%] =
C-Level:0.95	[12%, 16%]
Calculate	

b.

Calculator	Result
x:320	[0.18, 0.22] =
n : 1600	[18%, 22%]
C-Level:0.95	
Calculate	

5. Example – A sample of lengths (minutes) of local phone calls made by people who work for a local company is shown below. Find a 90% CI on the true proportion of local calls that are 15 or more minutes in length.

8.7	9.6	10.0	17.5	14.0	6.2	6.0	7.0	14.8	11.0
4.5	17.9	6.1	8.0	6.2	19.8	17.7	12.6	16.3	15.6
12.1	17.6	9.1	12.8	7.7	19.9	13.4	6.2	6.9	14.6
8.5	13.5	11.8	18.7	16.1	25.3	7.1	19.0	8.6	7.1
9.5	8.6	15.0	9.8	5.0	6.4	6.8	11.2	9.1	12.6

| Calculator	Result
x:13	[0.16, 0.36] =
n : 50	[16%, 36%]
C-Level:0.90	
Calculate	

6. **Relationship between Sample Size, $\bar{p}$ and Margin of Error**

Notice that the margin of error for a confidence interval on a proportion depends on the sample size and the sample proportion.

$$E = z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

This is a bit different from inference on μ. There, the margin of error does not depend on the sample average $\bar{x}$. Thus, with a proportion, the value of the estimator itself influences the precision. Consider a 95% confidence on a proportion (thus, $z_{\alpha/2} = 1.96$). The blue curve shows the margin of error when the sample size is 2000 and $\bar{p}$ varies from 1% to 99%. The other curves show the margin of error for different sample sizes. We observe two things:

a.   The margin of error is largest, for a fixed sample size when $\bar{p}$ is 50% and falls off on either side.
b.   As the sample size decreases, the margin of error increases.
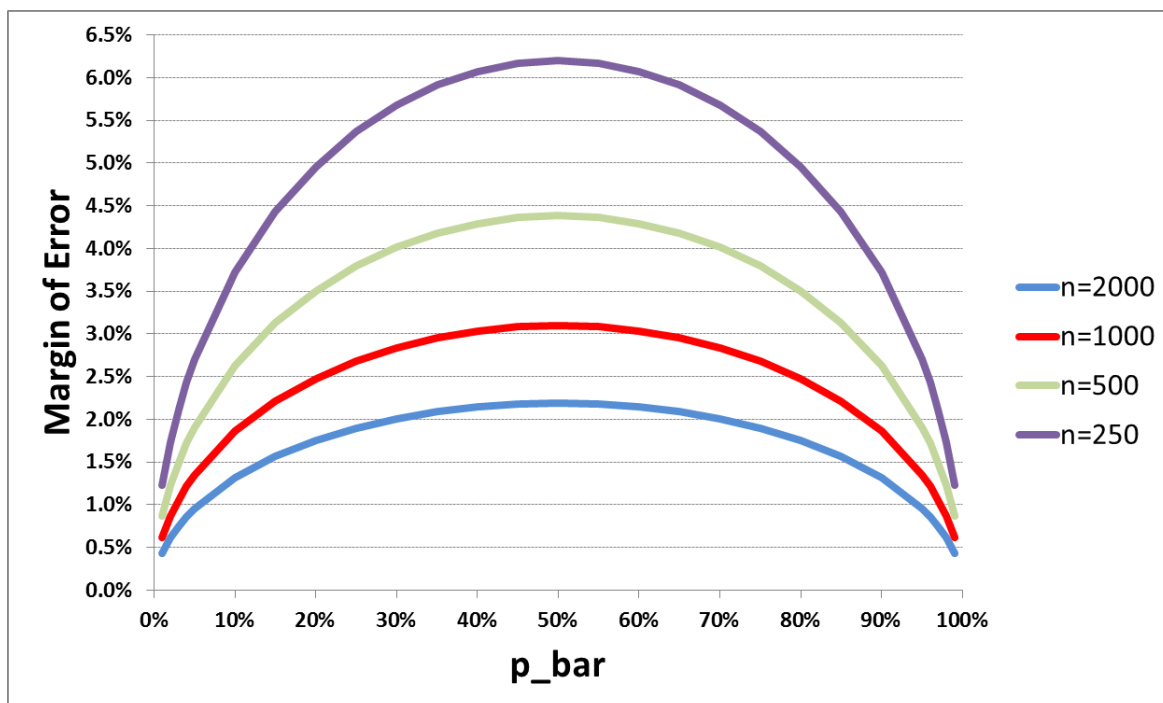
Homework 6.4

1. 200 people were surveyed to see if the watched the Superbowl on TV (or in person). It was found that 124 people did watch the Superbowl. (a) find a point estimate of the true proportion of people who watched the Superbowl (b) Find a 90% confidence interval on the true proportion of people who watched the Superbowl. (c) find the margin of error. (d) Find a 95% confidence interval on the true proportion of people who watched the Superbowl. (e) find the margin of error (f) Compare the two margins of error.

2. A recent survey found the following data about how far people travel to get to work each day in NYC (Manhattan-Financial District).

Distance to work	Number
0 to 1 mile	23
1-5 miles	68
5-15 miles	145
15-30 miles	84
More than 30 miles	448

   a. Find the point estimate, margin of error and a 90% confidence interval on the true proportion of people that live within 5-15 miles of work.
   b. Find the point estimate, margin of error and a 90% confidence interval on the true proportion of people that live within 5 miles of work.
   c. Compare the results of *a* and *b*.
   d. Any distance 15 miles or beyond is considered outside the city limits. Find the point estimate, margin of error and a 90% confidence interval on the true proportion of people that live outside the city limits.

3. A survey was conducted to see what 19 year-olds were doing. Survey participants had to choose from: College (C), Working (W), Unemployed (U), none of the above (N). The results are shown in the table below:

W	C	U	N	C	C	C	U	W	N	C	W
W	C	C	C	W	C	C	W	W	W	U	C
C	C	C	W	C	W	U	W	C	C	C	C
W	C	C	W	C	W	C	W	C	C	C	W
N	W	W	C	C	W	N	C	C	N	C	C

   Find the point estimate, margin of error and a 90% confidence interval on the true proportion of 19 year-olds that go to college. Comment on the margin of error.

4. A survey of 500 voters revealed that 300 would vote for Ron Borders. Find a 95% confidence interval on the true proportion of voters that will vote for Mr. Borders.

   (a) [0.52, 0.62] (b) [0.28, 0.32] (c) [0.54, 0.66] (d) [0.56, 0.64] (e) [0.26, 0.34]

6.8 Sample Size Calculations for Estimating a Proportion

1. We ask the same question was asked earlier about μ, "how much data do we need to collect to estimate a proportion to within a prescribed margin of error?"

   For example, how much data do I need to estimate the proportion of people who support a candidate to within 3%?

2. We remember that:

$$Margin\ of\ Error = E = z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Solving for the sample size, n:

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \bar{p}(1-\bar{p})$$

Thus, given:

    a. Margin of error, E
    b. Confidence level, $1-\alpha$, we can calculate $z_{\alpha/2}$
    c. An estimate of: $\bar{p}$

We can calculate the required sample size.

3. Problem: What do we use for $\bar{p}$? After all, we haven't collected any data yet.

    1. Previous and/or similar study.
    2. Educated guess, prior belief
    3. If no information is available use $\bar{p} = 0.5$, this is most conservative.

4. Note in the graph below that the closer the proportion is to 50%, the more data that is required to estimate it to within a fixed margin of error. For example, on the blue curve, where we are requiring a margin of error of 2%, if the true proportion is 50%, we would require about 2400 data values. On the other hand, if the true portion was 75% (or 25%) we would only need about 1800 data values to estimate to with 2%. Finally, note that if we can tolerate a 3% margin of error (the red curve), and the true proportion is 50%, we only need about 1067 data values, as opposed to 2400 for a 2% margin of error.
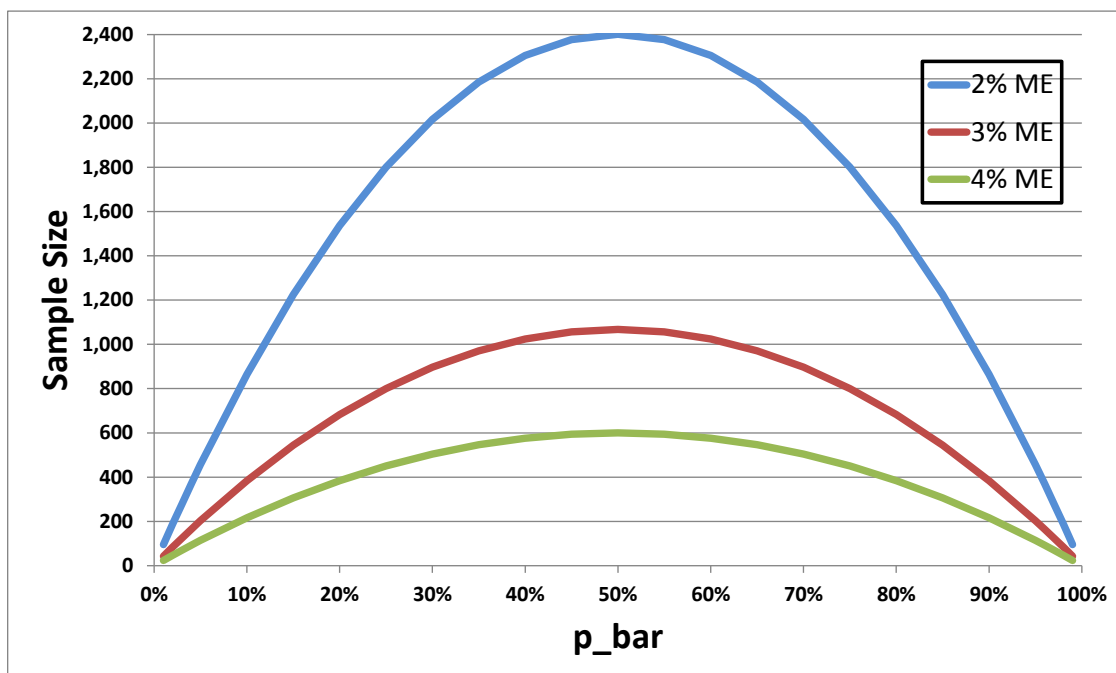
5. Most major, national opinion polls and surveys are conducted at the 95% confidence level with a 2-5% margin of error. A 2% margin of error, according to our formula would translate to a required sample size of 2401 people. It is interesting to note that by taking a simple random sample of about 2400 people you can infer what the beliefs of 200+ million adults in the US are to within 2%. Note also that our concept of large and small samples is different when we are talking about proportions. A sample size of 30 would have a margin of error of about 18% when dealing with proportions, hardly useful at all.

6. Example – Suppose that you want to estimate the proportion of voters who favor a particular candidate.

   a. What size SRS would be required in order to estimate the true proportion who favor the candidate with a margin of error of 5% with 90% confidence?

$$n = \left(\frac{1.645}{0.05}\right)^2 0.5(1 - 0.5) = 270.6 \implies 271$$

   b. What if you suspect that the candidate has about an 65% favorable rating. How much data would be required?

$$n = \left(\frac{1.645}{0.05}\right)^2 0.65(1 - 0.65) = 246.2 \implies 247$$

6.9 Confidence Interval Calculator Summary
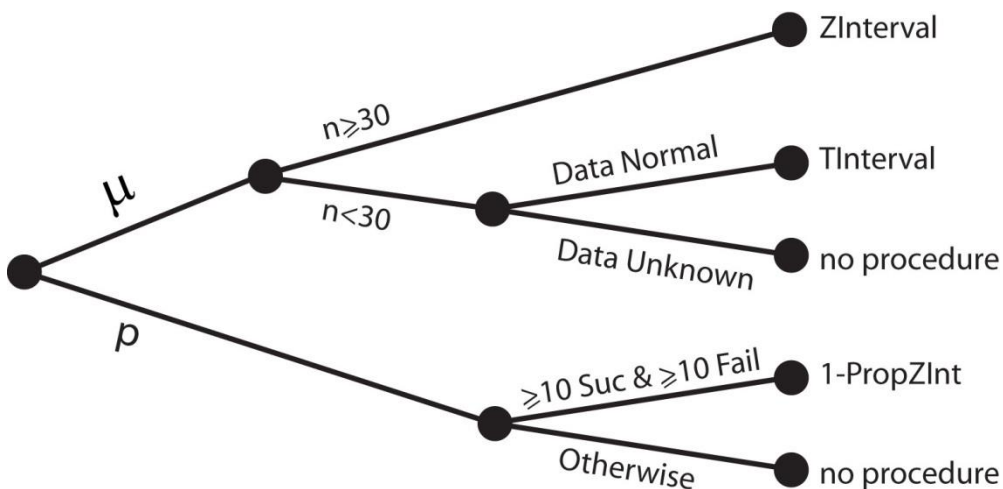
Homework 6.5

1. Suppose that you want to estimate the proportion of people that will use a potential new rail line that will be extended into the Northwest suburbs of Atlanta. How much data should you collect so that you are 90% confident and your margin of error is less than 5%?

2. A plant manager knows that in general about 10% of his products do not meet a customer's specifications. The customer has requested a 95% confidence interval on the true proportion of products that do not meet specifications with a margin of error of no more than 3%. How much data should the plant manager collect?

(a) $\left(\dfrac{0.05}{0.10}\right)^2 (0.03)(0.97)$

(b) $\left(\dfrac{0.05}{0.10}\right)^2 (0.10)(0.90)$

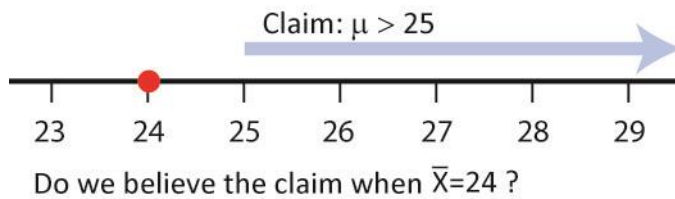(c) $\left(\dfrac{1.96}{0.03}\right)^2 (0.10)(0.90)$

(d) $\left(\dfrac{1.96}{0.10}\right)^2 (0.03)(0.97)$
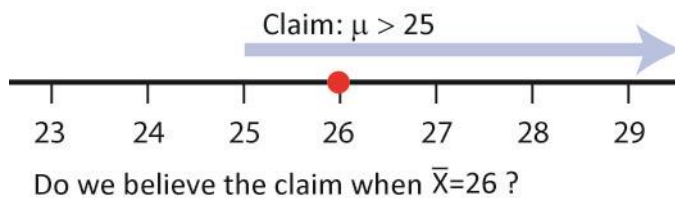
Chapter 7 – Hypothesis Testing

In this chapter we consider another type of statistical inference, hypothesis testing.

7.1 – Introduction

1. Hypothesis testing involves evaluating a *claim* about μ.

2. **Example** – A claim is made that on average the Ford Mustang gets more than 25 *mpg.* A sample of 30 Mustangs is tested and it is found that the sample average is 24 *mpg*.

Claim: $\mu > 25$

23   24   25   26   27   28   29

Do we believe the claim when $\overline{X}=24$?

How about:

Claim: $\mu > 25$

23   24   25   26   27   28   29

Do we believe the claim when $\overline{X}=26$?

From here, it gets more interesting. Does the fact $\bar{x} = 26$ *prove* that $: \mu > 25$? The answer is, "it might," but we would have to have more information. Specifically, how much variability is there in the data and what confidence do we require. For example, the true average may be 24 and we just got "lucky" with $\bar{x} = 26$.

3. **Example** – A manufacturer of small parts may claim to a customer that "on average, my products meet your specifications." Then, the customer may choose a sample and conduct a hypothesis test to evaluate the claim that the products meet the specifications. Depending on the outcome of the test, the customer will either buy the products or refuse to buy them.

4. Steps in hypothesis testing:

    1. Specify hypotheses (considered next)
    2. Collect data
    3. Calculate a p-value (later)
    4. Use a p-value to make a decision about hypotheses (later)

7.2 – Specifying Hypotheses

1. The idea is to make two *competing* statements about μ, the true, unknown mean of some population.

   Examples:

   1. The *Ph* level of effluent meets EPA standards   *vs.*
      The *Ph* level of effluent exceeds EPA standards

   2. Shirts from a company have 18" neck length   vs.
      Shirts from a company do not have 18" neck length

   3. Light bulbs last on average, less than 1000 hours   *vs.*
      Light bulbs last on average, more than 1000 hours

2. **Hypotheses** - We call the two competing statements *hypotheses*. One is called the *null hypothesis* and the other is called the *alternate hypothesis*. Hypotheses are always statements about the *true* value of μ. In general, hypotheses are statements about a population parameter. The process we go through to evaluate a set of hypotheses is called *hypothesis testing*.

3. **Null hypothesis** - The null hypothesis is denoted by the symbol, H_o. It is the statement being *tested*. Hypothesis tests are designed to assess the strength of the evidence *against* the null hypothesis. It is *usually* a statement of "no effect", "no change", "no difference", "what has been true in the past."

4. **Alternate hypothesis** - The alternate hypothesis is denoted by the symbol, H_a. This is *usually* a statement of what we hope or suspect is true, a statement of what we want to "prove", or show to be true. When we look at a problem, we usually focus on the alternative hypothesis first. We adopt the point of view of the person (group) doing the hypothesis test. What do they want to *prove?* Put that in the alternative.

5. **Forms of hypotheses** – There are exactly three forms that a set of hypotheses may take.

   1. *Not equal to alternative*, two-sided hypotheses, two-tailed hypothesis

      $$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu \neq \mu_o$$

      Note that μ_o is called the *hypothesized value*. It is the *number* we are making a claim about.

      Example: Shirts are manufactured in Mexico and a company in Houston wants to purchase 10,000. The company requires that the shirts *have* a neck length of 18 inches. A quality control person is sent to Mexico to test 100 shirts' neck lengths to see if the average neck length is *close enough* to 18 inches.

      $$H_o: \mu = 18 \ \ vs. \ H_a: \mu \neq 18$$

   2. *Less than alternative*, one-sided hypothesis, one-tailed hypothesis

      $$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu < \mu_o$$

      Example: A company has developed a new drug, Advil. In preliminary testing, the drug appears to start fighting pain in less than 10 minutes. The company would now like to do a formal study to see if there is strong evidence that this is true.

$$H_o: \mu = 10 \ \ vs. \ \ H_a: \mu < 10$$

Who is doing the test? What do they want to show is true?

Example: A citizens tax watchdog group has made the claim that the average tax bill on motor vehicles is in excess of $150. State regulators feel that this value is inflated and would like to test to see if the claim is not valid.

$$H_o: \mu = 150 \ \ vs. \ \ H_a: \mu < 150$$

Who is doing the test? What do they want to show is true?

3. *Greater than alternative*, one-sided hypothesis, one-tailed hypothesis.

$$H_o: \mu = \mu_o \ \ vs. \ \ H_a: \mu > \mu_o$$

Example: Honda would like to be able to say that the Accord gets more than 28 *mpg* on average, on the highway. Remember, who is doing the test? What do they want to show is true?

$$H_o: \mu = 28 \ \ vs. \ \ H_a: \mu > 28$$

Example: Advil claims that on average it starts fighting pain in no more than 10 minutes. A consumer group would like to see if there is evidence that the claim is not true. Remember, who is doing the test? What do they want to show is true?

$$H_o: \mu = 10 \ \ vs. \ \ H_a: \mu > 10$$

6. **Examples** – *Research Hypothesis* - express what you want to prove (the research hypothesis) as the *alternate hypothesis.* The assumption is that some new product, process, *etc.,* has been developed. We would like to be able to make a claim about some characteristic that we are interested in, to show that it is better in some way. The proposed (suspected, hoped for) claim is expressed as the alternate hypothesis. In this way, if we reject H_o, (*i.e. accept H_a*) what we would like to happen, our conclusion will be a strong one.

1. example - A bakery equipment manufacturer has developed a new leavening process for preparing commercial bread loaves which he thinks will reduce the calorie content in an 24 *oz.* loaf of bread. The old equipment produced loaves that had on average 1500 calories. Loaves of bread produced by the new process were randomly selected and analyzed for calorie content.

   a. State the appropriate hypothesis. $H_o: \mu = 1500 \ \ vs. \ \ H_a: \mu < 1500$
   b. Suppose that in preliminary testing, it appears that the calorie content is less than 1400 calories. State the appropriate hypothesis. $H_o: \mu = 1400 \ \ vs. \ \ H_a: \mu < 1400$

2. example - A major oil company has developed a new gasoline additive that is designed to increase average gas mileage in subcompact cars. Before marketing the new additive, the company conducts the following experiment: 100 subcompact cars are randomly selected, the gasoline additive is dispensed into the tanks of the cars, and the miles-per-gallon obtained by each car in the study is then recorded. Is there sufficient evidence for the oil company to claim that the average gas mileage obtained by subcompact cars with the additive is greater than 30 *mpg*? State the appropriate hypothesis.
   $H_o: \mu = 30 \ \ vs. \ \ H_a: \mu > 30$

7. **Examples** – *Testing the validity of a claim* - This is in some ways the opposite of testing a research hypothesis.

In this case, a claim has been made about some product or process in the past. It is an established claim, it has existed for some time, it is perhaps publicly known through advertisements or other news media. We would like to evaluate this claim. We usually adopt the point of view, "innocent until proven guilty," which means that we express the *established claim* as the null hypothesis. In other words, if we are going to reject the claim, H_o, we want that to be a strong conclusion, we want to have solid statistical evidence that the claim is false.

1. <u>example</u> - A manufacturer of fishing line wants to show that the mean breaking strength of a competitor's 22-pound line is really less than 22 pounds. State the appropriate hypothesis. $H_o: \mu = 22$ *vs.* $H_a: \mu < 22$

2. <u>example</u> - Advil claims that on average it starts fighting pain in no more than 10 minutes. A consumer group would like to see if there is evidence that the claim is not true. State the appropriate hypothesis. $H_o: \mu = 10$ *vs.* $H_a: \mu > 10$

3. **Examples** – *Testing in Decision Making Situations* - The previous two cases, testing research hypotheses and testing the validity of a claim usually involve taking action only when the null hypothesis is rejected (*i.e.* when a strong conclusion has been reached). In some instances, action must be taken in either case: rejecting or failing to reject the null hypothesis.

   1. <u>example</u> - Shirts are manufactured in Mexico and a company in Houston wants to purchase 10,000. The company requires that the shirts have a neck length of 18 inches. A quality control person is sent to Mexico to test 100 shirts' neck lengths to see if the average neck length is *close enough* to 18 inches. Should the company in Houston buy the shirts? State the appropriate hypothesis. $H_o: \mu = 18$ *vs.* $H_a: \mu \neq 18$

   2. <u>example</u> - A machine is set to put 12 *oz* of soda in a can. Periodically, a sample is tested to see if the true mean is really 12 *oz*. If there is evidence that the mean is not 12 *oz*. then the machine is taken off-line and a repair and calibration team is sent in to work on the machine. State the appropriate hypothesis. $H_o: \mu = 12$ *vs.* $H_a: \mu \neq 12$

Homework 7.1

1. An automobile assembly-line operation has a scheduled mean completion time of 2.2 minutes. Because of the effect of completion time on both earlier and later assembly operations, it is important to maintain the 2.2-minute standard. Which set of hypotheses is appropriate to test whether or not the operation is meeting its 2.2-minute standard?

   a. $H_o: \mu = 2.2$ vs. $H_a: \mu > 2.2$
   b. $H_o: \mu = 2.2$ vs. $H_a: \mu \neq 2.2$
   c. $H_o: \bar{x} = 2.2$ vs. $H_a: \bar{x} > 2.2$
   d. $H_o: \bar{x} = 2.2$ vs. $H_a: \bar{x} \neq 2.2$
   e. $H_o: \mu \leq 2.2$ vs. $H_a: \mu > 2.2$
   f. $H_o: \mu \geq 2.2$ vs. $H_a: \mu < 2.2$

2. The manager of an automobile dealership is considering a new bonus plan that is designed to increase sales volume. Currently, the mean sales volume is 14 automobiles per month. The manager would like to conduct a study to see if there is evidence that the new bonus plan increases sales volume. Which is the appropriate set of hypotheses?

   a. $H_o: \mu = 14$ vs. $H_a: \mu > 14$
   b. $H_o: \mu = 14$ vs. $H_a: \mu \neq 14$
   c. $H_o: \bar{x} = 14$ vs. $H_a: \bar{x} > 14$
   d. $H_o: \bar{x} = 14$ vs. $H_a: \bar{x} \neq 14$
   e. $H_o: \mu \leq 14$ vs. $H_a: \mu > 14$

3. The manager of a hotel has stated that the average guest bill for a weekend is less than $400. A member of the hotel's accounting staff has noticed that the total charges for guest bills have been increasing in recent months and suspects that the average is more than $400. Is there strong evidence to support the accountant's claim?

   a. $H_o: \mu \geq 400$ vs. $H_a: \mu < 400$
   b. $H_o: \mu = 400$ vs. $H_a: \mu < 400$
   c. $H_o: \bar{x} = 400$ vs. $H_a: \bar{x} < 400$
   d. $H_o: \mu = 400$ vs. $H_a: \mu > 400$
   e. $H_o: \mu \geq 400$ vs. $H_a: \mu < 400$

4. A restaurant averages $3,000 in net sales per night without any advertising. The manager of the restaurant is considering an advertising campaign. In talks with the advertising company the manager learns that on average total net sales increases by at least 10% when a particular advertising campaign is undertaken. The manager of the restaurant undertakes the advertising campaign and would now like to see if there is statistical evidence that net sales have increased by the claimed amount. What is the appropriate set of hypotheses:

   a. $H_o: \bar{x} = 3300$ vs. $H_a: \bar{x} < 3300$
   b. $H_o: \mu \geq 3300$ vs. $H_a: \mu < 3300$
   c. $H_o: \bar{x} = 3300$ vs. $H_a: \bar{x} > 3300$
   d. $H_o: \mu = 3000$ vs. $H_a: \mu > 3000$
   e. $H_o: \mu = 3300$ vs. $H_a: \mu > 3300$

5. In 2007, *The Motor Vehicle Manufacturers Association* reported statistics on the average number of years passenger cars were being used. The data were obtained from a sample of the owners of 100 passenger cars and showed a sample mean of 7.8 years and a sample standard deviation of 2.2 years. In 2000, the population mean was reported to be 6.5 years. If a researcher is looking for evidence to show that individuals are driving cars longer in 2007, then which is the correct set of hypotheses?

   a. $H_o: \mu_{2007} \geq 7.8$ vs. $H_a: \mu_{2007} < 7.8$
   b. $H_o: \mu_{2000} = 7.8$ vs. $H_a: \mu_{2000} > 7.8$
   c. $H_o: \mu_{2007} = 7.8$ vs. $H_a: \mu_{2007} < 7.8$
   d. $H_o: \mu_{2007} \geq 7.8$ vs. $H_a: \mu_{2007} < 7.8$
   e. $H_o: \mu_{2007} = 7.8$ vs. $H_a: \mu_{2007} > 7.8$
   f. $H_o: \mu_{2007} \leq 6.5$ vs. $H_a: \mu_{2007} > 6.5$
   g. $H_o: \mu_{2007} = 6.5$ vs. $H_a: \mu_{2007} > 6.5$
   h. $H_o: \mu_{2007} = 6.5$ vs. $H_a: \mu_{2007} \neq 6.5$
   i. $H_o: \mu_{2000} = 6.5$ vs. $H_a: \mu_{2000} > 6.5$

6. The label on a can of sardines says that on average there are at least 6 grams of protein in each serving. The Food and Drug Administration (FDA) has reason to suspect that filler is being added to the sardines and that there is less protein than stated. If this is the case, then the FDA will require the company to take the statement off the label. The FDA tests a sample of 50 cans in a foods laboratory. To determine whether the FDA should take action against the company, the appropriate set of hypotheses is...

    a. $H_o: \mu = 6$ vs. $H_a: \mu > 6$
    b. $H_o: \mu = 6$ vs. $H_a: \mu < 6$
    c. $H_o: \mu \neq 6$ vs. $H_a: \mu = 6$
    d. $H_o: \mu > 6$ vs. $H_a: \mu \leq 6$
    e. $H_o: \mu \leq 6$ vs. $H_a: \mu > 6$

7.3 – Decision Making

There are *two* possible *decisions* when doing a hypothesis test. Later, we will learn *how* decisions are made. Now, we study *what* those two decisions are. The first thing to note is that decisions are always made about the null hypothesis, H_o.

Reject H_o – This means that the evidence supplied by the data is sufficient to reject the null hypothesis. Rejecting the null hypothesis is called a *strong conclusion* (to be explained later). Thus, rejecting H_o says that we have strong reason to believe that H_o is not true. Thus, we have strong evidence that the alternate hypothesis is true. This is why, when possible, we put what we want to prove in the alternate hypothesis and in this case, we *want* to reject the null hypothesis so that a strong conclusion can be reached to believe the alternate hypothesis.

Fail to reject H_o – This means that there is *not* sufficient evidence to reject the null hypothesis. This is called a *weak conclusion*. This *does not* mean that we have *proved* that H_o is true. It simply means that there is not sufficient evidence to reject it. Neither does this mean that H_a is necessarily false. It just means that we do not have evidence to believe it. Note that the alternative hypothesis my actually be true, but the evidence in this case is not sufficient to believe it. Sometimes I call this a *wishy-washy* conclusion, we are not certain of anything. Sometimes I use the expression, when we fail to reject H_o that we are *stuck with H_o;* it might be true and it might not.

7.5 – P-Value Approach to Hypothesis Testing

1. **The P-value Approach** – We can compute a **p-value** for any hypothesis test using a calculator or computer. A p-value is a probability. Using the p-value to make a decision is easy:

P-value	Decision
small (*e.g.* 0.05 or less)	Reject the null hypothesis (strong conclusion)
large	Fail to reject the null hypothesis (weak conclusion)

2. The p-value tells us how likely our results ($\bar{x}$) were, **if** the null hypothesis is true. Thus, a small p-value says that there is a small probability of obtaining a sample average as extreme as what was obtained when the null hypothesis is true. Thus, the smaller the p-value, the more reason to reject the null hypothesis.

3. Suppose that we are testing to see if a new drug starts fighting pain in less than 10 minutes, and we find that the sample average is 9.2 minutes, standard deviation is 3 minutes, and the sample size is 32. We test, $H_o: \mu = 10$ vs. $H_a: \mu < 10$ and find that the p value $= 0.07$

$H_a : \mu < 10$

$\bar{X} \sim N(\mu = 10, 3/\sqrt{32})$

p-value=0.07

9.2   μ=10
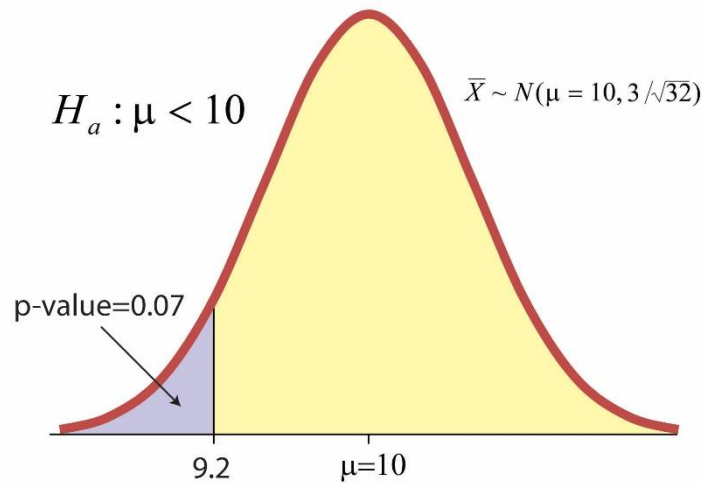
If we assume that the null hypothesis is true, then we are assuming that the sample average has the distribution, $\bar{X} \sim N(\mu = 10, \sigma = 3/\sqrt{32} = 0.53)$. To calculate the p-value, we ask how *unlikely* our sample results ($\bar{x} = 9.2$) were. Since we are testing a *less than* alternative, *unlikely* refers to values of 9.2 or less:

$$p - value = P(\bar{X} < 9.2 | \bar{X} \sim N(\mu = 10, \sigma = 0.53)$$

$$= normalcdf(-1E99, 9.2, 10, 0.53) = 0.0657 \approx 0.07$$

Thus, for this problem, a p-value of 7% means that if the null hypothesis is true, then there is about a 7/100 chance of obtaining a sample average of 9.2 or smaller. Because this probability is so small, it might lead us to suspect that the null hypothesis is not true. In some sense, there is, at most a 7% chance that we are wrong if we reject the null hypothesis. So what is our decision? Can we tolerate a 7% chance of error? Probably, in this case we would say, "Yes.". Thus, we would reject the null hypothesis and *conclude* by saying that there is strong evidence that the new drug starts fighting pain in less than 10 minutes.

4. Sometimes we are more formal about hypothesis testing. In this case, the *maximum tolerance for error* is stated *in advance* by the researcher or other person. We call this value, α. Common values for α are 1%, 5%, 10%. α is referred to as the *significance level* of a hypothesis test. Thus, since α is the maximum error we can tolerate, then we are going to reject the null hypothesis when our *actual error* (*p-value*) is *less than* α. When we are given a value for α, we call this the *fixed significance* approach to hypothesis testing. In this case, we use this rule:

   - If $p \ value < \alpha$ then Reject H_0

   - If $p \ value \geq \alpha$ then Fail to Reject H_0

This rule applies for *any* type of hypothesis testing, whether we learn it in this class or not.

5. Suppose in the previous example, we had chosen a significance level of 10% ($\alpha = 0.1$) for the hypothesis test. Then we would reject the null hypothesis because $p \ value = 0.07 < 0.1 = \alpha$.

6. Calculation of P-values – On your calculator, choose: Stats/Tests/Z-Test. The screen will show:

Z-Test
Inpt: Data   Stat

If you choose: <mark>Stat</mark>, the screen will look like this:

u_o:
σ:
$\bar{x}$:
n:
μ: $\neq \mu_o$, $< \mu_o$, $> \mu_o$:
Calculate Draw

You will type in or choose the appropriate values. Note that you specify the *alternate* hypothesis choosing:

Calculator Choice	Meaning
$\neq \mu_o$	$H_a: \mu \neq \mu_o$
$< \mu_o$	$H_a: \mu < \mu_o$
$> \mu_o$	$H_a: \mu > \mu_o$

When you choose *Calculate*, the p-value will be displayed along with some other values:

Calculator Display	Meaning
Z-test	
$\mu \neq 10$	Alternate hypothesis
z = 2.1082	test statistic
p = .0350	p-value
$\bar{x} = 11$	
$n = 40$	

7. Example – The mean cost of a heart by-bypass operation, nation-wide is $48,400 with several hundred thousand of these performed every year. A sample of 36 bypass operations in Atlanta revealed that the sample mean was $47,010 with a standard deviation of $8500.   Is there evidence at the 5% significance level that the mean cost of a heart-bypass operation in Atlanta is less than the national average.

Calculator:
u_o: 48400
σ: 8500
$\bar{x}$: 47010
n: 36
μ: $\neq \mu_o$, <mark>$< \mu_o$</mark>, $> \mu_o$:
<mark>Calculate</mark> Draw

Problem Type	Large sample Z test
Hypothesis	$H_o: \mu = \$48,400$ *vs.* $H_a: \mu < \$48,400$
P-value	0.1632
Decision	Since 0.1632 > 0.05, fail to reject the null hypothesis
Conclusion	There is not strong evidence that Atlanta's average cost is less than the national average.

8.   A machine has historically stamped CD's with an average 0.4 seconds per CD. Lately, it seems that the machine is taking slightly longer. Management would like to do a hypothesis test to determine whether the machine is averaging a time longer than the historical rate. If so, the machine will be taken off-line for maintenance and repair. A sample of size 40 is taken. The resulting average was 0.45 with a standard deviation of 0.06. Is there evidence at the 1% significance level that the mean length of time that it takes to stamp a CD has increased?

Calculator: u_o: 0.4
              σ: 0.06
              $\bar{x}$: 0.45
              n: 40
              μ: $\neq \mu_o$, $< \mu_o$, $> \mu_o$:
              Calculate Draw

Problem Type	Large sample Z test
Hypothesis	$H_o: \mu = 0.4$ vs. $H_a: \mu > 0.4$
P-value	$6.817836\,E - 8\ (calculator) \cong 6.82 * 10^{-8} = 0.0000000682 \approx 0$
Decision	Since $0 < 0.01$, reject the null hypothesis
Conclusion	The machine should be taken off-line for maintenance and repair. There is sufficient evidence to say that the average time to stamp a CD exceeds 0.4 seconds.

9.   Thread-count is used to measure the quality of a rug. A machine is set to produce rugs with a thread-count of 200. According to quality standards required by customers, if the thread-count is too far above or below the 200 value, it can't be sold as 200-thread. A sample of size 50 is taken from analyzing the thread count in small, randomly chosen blocks in a large rug. The average is found to be 192 with a standard deviation of 12.4. Test the appropriate hypothesis to determine if quality standards for 200-thread rugs have been met at the 5% significance.

Calculator: u_o: 200
              σ: 12.4
              $\bar{x}$: 192
              n: 50
              μ: $\neq \mu_o$, $< \mu_o$, $> \mu_o$:
              Calculate Draw

Problem Type	Large sample Z test
Hypothesis	$H_o: \mu = 200$ vs. $H_a: \mu \neq 200$
P-value	$5.07 * 10^{-6} = 0.000000507 \approx 0$
Decision	Since $0 < 0.05$, reject the null hypothesis
Conclusion	There is strong evidence that the rug does not meet the 200-thread requirement.

Homework 7.2

1. An automobile assembly-line operation has a scheduled mean completion time of 2.2 minutes. Because of the effect of completion time on both earlier and later assembly operations, it is important to maintain the 2.2-minute standard. A simple random sample of size 40 is taken and the resulting sample mean is 2.0 minutes and standard deviation is 0.4 minutes. Is there evidence at the 5% significance level that the mean completion time differs from 2.2 minutes?

2. It is suspected that the average amount of garbage that households produce has increased beyond 50 lbs. per week. On the local rounds the garbage men randomly weigh the garbage at 36 homes. (a) The results showed an average of 53 lbs. and a standard deviation of 15 lbs. At the 10% significance level, test the appropriate hypothesis. (b) Suppose the average was 57 lbs., re-test the hypothesis.

3. Purchase requests in a large corporation have to be routed to a number of people for sign-offs. Traditionally, it has taken, on average 9 working days for the final signoff. A new process has been implemented and it is hoped that it reduces the average time for final sign-off. A sample of 30 purchase requests is analyzed and it is found that the average was 6.6 days with a standard deviation of 2.3 days. Can the company claim that the new process significantly reduces the average time for final sign-off? Interpret the p-value for this problem

and reach a decision and conclusion.

4. A bank's credit card customers spend on average $1987 per month on their credit cards. The bank considers a proposal to increase the amount that its credit card customers charge on their cards. If there is strong evidence that the proposal does increase credit card spending, then the proposal will be offered to all customers. To study the proposal more carefully, the bank offers the proposal to a simple random sample of 40 customers. The statistical analysis revealed that the null hypothesis should be rejected. You recommend that the company ...

   (a) ...reach the strong conclusion to offer the proposal to all of its customers
   (b) ...reach the weak conclusion to not offer the proposal to its customers.
   (c) ...reach the strong conclusion to not offer the proposal to its customers.
   (d) ...reach the weak conclusion to offer the proposal to all of its customers.

5. The life of a AA-cell flashlight battery is under study. Researchers would like to show that on average the battery lasts more than 10 hours, thus, the null hypothesis is that batteries, on average last less than 10 hours. After collecting data it was found that the p-value was 0.1276. Using a significance level of 0.05, the decision is to ...
   (a) ...reject the null hypothesis.
   (b) ...reject the alternate hypothesis.
   (c) ...fail to reject the null hypothesis.
   (d) ...fail to reject the alternate hypothesis.

7.6 – Decision Making Consequences

1. **Types of errors in hypothesis testing** – In hypothesis testing we either *reject* or *fail to reject* the null hypothesis, but the question that lingers is, "What if the decision we reached is wrong?" That possibility *always* exists. Consider the model used in the judicial system for defendants in criminal court:

$$H_o: Innocent \quad \text{vs.} \quad H_a: Guilty$$

   Someone is judged *guilty* if there is *overwhelming* (juries generally have to be unanimous) evidence of guilt. In this case, we *strongly reject innocence*. On the other hand, if there is *not* overwhelming evidence of guilt, a defendant is judged, *not guilty*. Notice the words do not say that the jury *proved innocence*, the jury simply finds the defendant *not guilty*. This is analogous to *failing to reject the null hypothesis*. However, in either case, whether the jury finds a defendant *guilty* or *not guilty*, there is the possibility that a wrong decision was reached. Suppose a jury convicts a person who is later proven innocent. Thus, the decision was in error. It is important to see that there is some ultimate truth to a set of hypotheses, which is usually unknowable and in light of that, we reach a decision based on a set of data and a p-value. Thus, it is important to consider the consequences of making an error. As shown in the table below, there are two possible errors in a jury trial (or hypothesis test): (a) a guilty person goes free, (b) an innocent person is wrongfully convicted. In statistics, we need to be aware of these two possible errors.

	Truth	
Decision	**Not Guilty**	**Guilty**
Not Guilty	Correct Decision	Error
Guilty	Error	Correct Decision

2. Now, let's be more specific about the types of errors that can occur in hypothesis testing. As you see in the table below, we give names to the two types of error that can be present in hypothesis testing: Type I and II errors.

Decision	Truth H_o true	H_a true	Conclusion
Fail to Reject H_o	Correct Decision	Type II Error, β	Weak
Reject H_o	Type I Error, α	Correct Decision	Strong

3. **Type I and II errors**

a. $P(Type\ I\ Error) = P(Reject\ H_0|H_0\ true) = P(H_a|H_0) = \alpha$

In hypothesis testing, *you select* α. It is usually a value like 0.1, 0.05, 0.01. In other words, the experimenter has control over the probability of a Type I error occurring. α is called the *significance level* of the hypothesis test.

A *Type I error* is when you...

   ...reject the null hypothesis when you shouldn't
   ...incorrectly reject a true null hypothesis

b. $P(Type\ II\ Error) = P(Fail\ to\ Reject\ H_a|H_a\ true) = P(H_0|H_a) = \beta$

For our purposes, the probability of a Type II error is *unknown*.

A *Type II error* is when you...

   ...fail to reject when you should have rejected
   ...incorrectly accept a false null hypothesis

4. Interpretation: A decision to reject the null hypothesis is a strong one because you know what the probability of making an error is, α. A decision to fail to reject the null hypothesis is a weak one because you don't know what the probability of making an error is, β is unknown.

5. **Consequences of Making Type I and II errors** – Even though we may never know if a Type I or II error has occurred, it is important, before doing a hypothesis test to consider and understand the risk of an incorrect decision. In the examples below, we present a problem and then ask you to state the possible errors, and then the consequences of such errors.

6. Example – Suppose that a new JIT (just-in-time) production method will be implemented if a hypothesis test supports the conclusion that the new method reduces the mean operating cost per hour.

a. State the appropriate hypothesis if the mean cost for the current production method is $350 per hour.
$$H_o: \mu_{New\ Method} = \$350 \quad vs. \quad H_a: \mu_{New\ Method} < \$350$$

b. What is the Type I error in this situation? (A Type I error is when you reject the null hypothesis when you shouldn't.)

c. What are the consequences of making a Type I error?

d. What is the Type II error in this situation? (A Type II error is when you fail to reject the null hypothesis when you should have rejected it.)

e. What are the consequences of making a Type II error?

7. Example - The National Institute for Drug Abuse (NIDA) has suggested a level of 60 ppm (parts per million) of THC (the active ingredient in marijuana) to detect marijuana use. Suppose that we want to test to see if there is strong evidence of marijuana use.

a. State the *appropriate* hypothesis.

b. What is the Type I error in this situation?

c. What are the consequences of making a Type I error?

d. What is the Type II error in this situation?

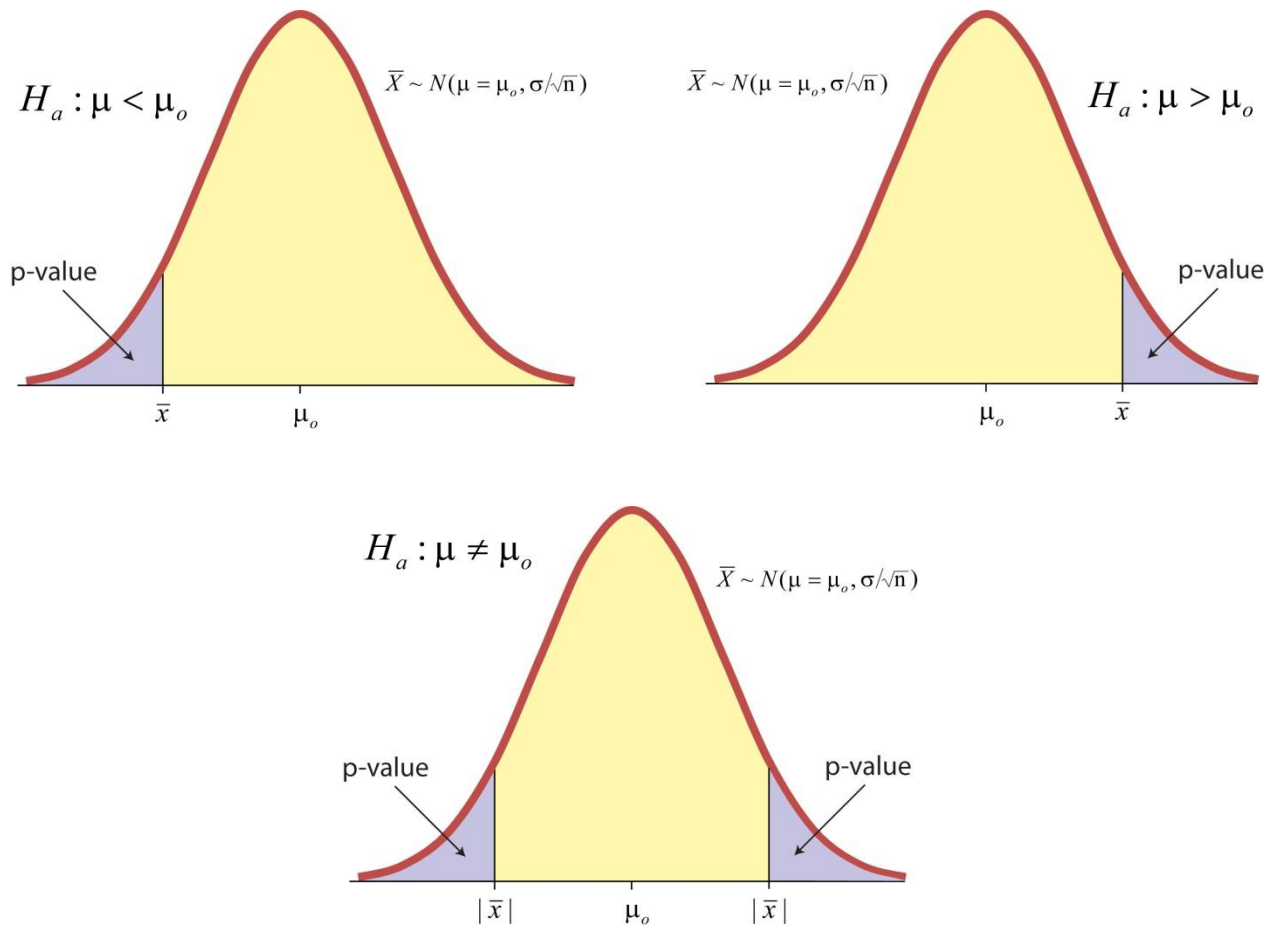e. What are the consequences of making a Type II error?

Homework 7.3

1. Consider homework 7.1, problem 4 to answer these questions:

    a. What is the Type I error in this situation?
    b. What are the consequences of making a Type I error?
    c. What is the Type II error in this situation?
    d. What are the consequences of making a Type II error?

2. Consider homework 7.1, problem 6 to answer these questions:

    a. What is the Type I error in this situation?
    b. What are the consequences of making a Type I error?
    c. What is the Type II error in this situation?
    d. What are the consequences of making a Type II error?

7.7 – More on P-Values

1. In some sense, a *p-value* tells you the *actual probability of a Type I error* for a *particular* problem. A given α is the *maximum* probability of Type I error that you can tolerate. Thus, as we saw earlier, $p-value < \alpha$ means reject the null hypothesis. Another approach to hypothesis testing is to simply report the p-value and let the reader interpret it according to their tolerance for risk. We saw earlier that a p-value assumes that the null hypothesis is true (in the *worst* case), *i.e.* that the true mean is $\mu = \mu_o$ and is the probability that a sample mean is *at least as unlikely* as what was observed. The phrase, *at least as unlikely,* is interpreted by considering the alternative hypothesis.

$H_a : \mu < \mu_o$ $\overline{X} \sim N(\mu = \mu_o, \sigma/\sqrt{n})$ $\overline{X} \sim N(\mu = \mu_o, \sigma/\sqrt{n})$ $H_a : \mu > \mu_o$

p-value $\overline{x}$ μ_o μ_o $\overline{x}$ p-value

$H_a : \mu \neq \mu_o$ $\overline{X} \sim N(\mu = \mu_o, \sigma/\sqrt{n})$

p-value   p-value $|\overline{x}|$ μ_o $|\overline{x}|$

2. Using a p-value to evaluate a hypothesis is very easy. There are two approaches:

   a. Reject H_o if $p-value < \alpha$ otherwise Fail to Reject H_o. Note that this rule is *always* true for any hypothesis testing situation, whether we learn it in this class or not. This is called the ***fixed significance*** approach.

   b. The smaller the p-value of a hypothesis test ...

      a. ...the more support for the alternate hypothesis.
      b. ...the more reason to reject the null hypothesis.
      c. ...the smaller the probability of a type I error.
      d. ...the smaller the probability that null hypothesis is true.

3. Example – Suppose that we are testing to see if a new drug starts fighting pain in less than 10 minutes, $H_o: \mu = 10$ vs. $H_a: \mu < 10$. A sample of size 40 is taken and it is found that the sample mean is 8.9 minutes and the standard deviation is 3 minutes. Using your calculator, the p-value for this problem would be: $p - value = 0.0102 \approx 1\%$.

    This number means that if the null hypothesis is true and we reject the null hypothesis, there is, at most a 1% chance that we are wrong. So what is our decision? Can you tolerate a 1% chance of error? If so, we would conclude by saying that there is strong evidence that the new drug starts fighting pain in less than 10 minutes on average.

4. Example – The average salary of clerical workers in the U.S. is $9.70 per hour. The state of Georgia took a sample of 30 clerical workers and found that the sample average was $10.05 and standard deviation was $2.06. Can Georgia make the claim that its clerical workers make more, on average than the national average?

$$H_o: \mu_{Georgia} = \$9.70 \quad vs. \quad H_a: \mu_{Georgia} > \$9.70$$

$$p - value = 0.1762 \approx 18\%$$

    This number means that if the null hypothesis is really true, then there is about an 18% chance of observing a sample average of $10.05 or higher. Thus, in this case, *if* we reject the null hypothesis, there is about an 18% chance that we are wrong. So what is our decision? Can you tolerate an 18% chance of error? If not, we would fail to reject the null hypothesis and we would conclude by saying that there is *not* strong evidence that the mean salary for clerical works in Georgia exceeds the national average. Or, we could say that there is *not* statistical evidence that the mean rate in Georgia differs from the national average.

Homework 7.4

1. It is claimed that the average entry-level salary for a Wall Street stockbroker is $250,000. Researchers would like to see if there is any strong evidence that the mean differs from this value. A sample of 50 stockbrokers reveals an average of $238,459 with a standard deviation of $50,223. (a) Calculate a p-value for this problem, (b) interpret it, (c) reach a decision, (d) state a conclusion.

7.8 – Confidence Interval Approach to Hypothesis Testing

1. Instead of using a p-value, we can use a confidence interval to evaluate a hypothesis when it is a 2-sided hypothesis: $H_o: \mu = \mu_o$ vs. $H_a: \mu \neq \mu_o$. The rule we use is:

    - If μ_o is **not** in confidence interval, *Reject H_o*

    - If μ_o is in confidence interval, *Fail to Reject H_o*

2. Example – Thread-count is used to measure the quality of a rug. A machine is set to produce rugs with a thread-count of 200. According to quality standards required by customers, if the thread-count is too far above or below the 200 value, it can't be sold as 200-thread. A sample of size 50 is taken from analyzing the thread count in small, randomly chosen blocks in a large rug. The average is found to be 192 with a standard deviation of 12.4. Suppose quality standards for 200-thread rugs to pass a hypothesis test using 5% significance. Test the appropriate hypothesis using a 95% confidence interval

Problem Type	Large sample Z test
Hypothesis	$H_o: \mu = 200$ vs. $H_a: \mu \neq 200$
Confidence Interval	[188.56, 195.44]
Decision	Since 200 is not in the confidence interval, reject the null hypothesis
Conclusion	There is strong evidence that the rug does not meet the 200-thread requirement.

3. Note that statistical software such as Minitab will, in addition to the p-value, compute a 1-sided confidence interval when you do a hypothesis test. When you specify a hypothesis like this:

$$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu > \mu_o$$

in addition to the p-value, you will get a *lower bound* 1-sided confidence interval. This means that we are, say 95% sure that the true mean is *greater than* the lower bound.

Similarly, for this hypothesis:

$$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu < \mu_o$$

you will get an *upper bound* 1-sided confidence interval. This means that we are, say 95% sure that the true mean is *less than* the upper bound.

4. The same rule holds true for one-sided hypothesis tests and confidence intervals: reject when μ_o is not in the confidence interval (either Lower or Upper).

	1-sided	
Hypothesis	**Confidence Interval**	**Decision**
$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu > \mu_o$	Lower Bound	Reject if $\mu_o < Lower\ Bound$
$H_o: \mu = \mu_o \ \ vs. \ H_a: \mu < \mu_o$	Upper Bound	Reject if $\mu_o > Upper\ Bound$

5. Example – In one of the previous examples we evaluated:

$$H_o: \mu = 48{,}400 \ \ vs. \ H_a: \mu < 48{,}400$$

$\alpha = 0.05$ level. A 95% Upper Bound: 49,340. Thus, $\mu < 49{,}340$. This means that we are 95% sure that the true mean is less than 49,340. Thus, is there evidence that the true mean is less than 48,400? It *might* be, it could be 47,000. However, it might *not* be; the true mean could be 49,000 according to the confidence interval. Thus, there is not strong evidence so we fail to reject the null hypothesis. Using the rule above, we reach the same conclusion: μ_o *is* in the 1-sided confidence interval, so fail to reject.

6. Example – $H_o: \mu = 48{,}400$ vs. $H_a: \mu < 48{,}400$, 95% Upper Bound: 48,000. This means that we are 95% sure that the true mean is less than 48,000. Thus, is there strong evidence that the true mean is less than 48,400? Yes. We reject the null hypothesis. Using the rule above, we reach the same conclusion: μ_o is not in the 1-sided confidence interval, so reject.

7. Example – $H_o: \mu = 0.4$ vs. $H_a: \mu > 0.4$, 95% Lower Bound: 0.42793. This means that we are 95% sure that the true mean is greater than about 0.43. Thus, is there strong evidence that the true mean is greater than 0.4? Yes, definitely, it is most likely larger than 0.43. We reject the null hypothesis. Using the rule above, we reach the same conclusion: μ_o is not in the 1-sided confidence interval, so reject.

8. Example – $H_o: \mu = 0.4$ vs. $H_a: \mu > 0.4$, 95% Lower Bound: 0.38. This means that we are 95% sure that the true mean is greater than about 0.38. Thus, is there strong evidence that the true mean is greater than 0.4? It might be and it might not. We fail to reject the null hypothesis. Using the rule above, we reach the same conclusion: μ_o is in the 1-sided confidence interval, so fail to reject.

Homework 7.5

1. It is claimed that the average entry-level salary for a Wall Street stockbroker is $250,000. Researchers would like to see if there is any strong evidence that the mean differs from this value. A sample of 50 stockbrokers reveals an average of $238,459 with a standard deviation of $50,223. Use a 95% confidence interval to evaluate the appropriate hypothesis.

2. Suppose we are testing: $H_o: \mu = 0.005$, then we would reject the null hypothesis based on 95% CI, [-0.002, 0.010]. True *or* False.

3. Suppose we are testing to see if there is strong evidence that the average speed on a road is less than 72 *mph*. The 90% Upper Bound is 71.1 *mph*. Evaluate the appropriate hypothesis.

4. Suppose we are testing the null hypothesis that the average cost to fill up a passenger car is no more than $40. The 90% Lower Bound is 38.6 *mph*. Evaluate the appropriate hypothesis.

7.9 – Small Sample Hypothesis Testing

1. Hypothesis testing with small samples (n<30) is about the same as with large samples. However, we must assume the data is normally distributed. Statistical testing is based on the t-distribution which is similar to the standard normal distribution, but with fatter tails (more spread out). On TI-83 calculator, use Stats/Tests/T-Test.

2. Example – A writer believes that the average length of time that people drive a new car before selling it is at least 10 years. A simple random sample of size 20 was taken and it was found that the average was 11.21 years with a standard deviation of 5.48 years. If we assume that the time that people keep a car is normally distributed, then is there evidence at the 5% level that new cars are driven for at least 10 years, on average?

Problem Type	Small sample t-test
Hypothesis	$H_o: \mu = 10$ vs. $H_a: \mu > 10$
P-value	$0.1679 \approx 17\%$
Decision	Since 0.1679 > 0.05, fail to reject the null hypothesis
Conclusion	There is not strong evidence that the average length of time new cars are driven exceeds 10 years.

3. The average weight of a particular model of car is 2900 lbs. A new manufacturer for several engine components is being considered. One question of interest is whether this will change the average weight of the cars. A sample of 20 cars with the new components is taken and their weights are shown below. Has the mean weight of this model of car changed? (a) Evaluate the hypothesis at the 10% significance level (b) Use

a 90% confidence interval to evaluate this hypothesis.

3046	2970	2967	3060	3079	2948	2944	3043	2941	3015
2920	3036	3014	3000	2970	2986	2961	2916	2997	3049

a.

Problem Type	Small sample t-test
Hypothesis	$H_o: \mu = 2900$ vs. $H_a: \mu \neq 2900$
P-value	$5.47 * 10^{-8} = 0.0000000547 \approx 0 = 0\%$
Decision	Since 0 < 0.1, reject the null hypothesis
Conclusion	There is sufficient evidence that the average weight of cars with the new engine parts is not 2900 lbs.

b.

Problem Type	Small sample t-test
Hypothesis	$H_o: \mu = 2900$ vs. $H_a: \mu \neq 2900$
Confidence Interval	[2974, 3012]
Decision	Since 2900 is *not* in confidence interval, reject the null hypothesis
Conclusion	There is sufficient evidence that the average weight of cars with the new engine parts is not 2900 lbs.

Homework 7.6

1. The manager of an automobile dealership is considering a new bonus plan that is designed to increase sales volume. Currently, the mean sales volume is 14 automobiles per month. The manager adopts the new bonus plan and would now like to conduct a study to see if there is evidence that the new bonus plan increases sales volume. The previous 10 months data are taken as a sample (assumed to be normally distributed) and the mean is 15.4 cars with a standard deviation of 4 cars. Test the appropriate hypothesis at the 10% significance level.

2. A container of nuts says that ½ cup of nuts weight approximately 110 grams. A sample of ½ cup servings of nuts is taken and individually weighed. The results are shown below. A consumer's group would like to show that the average weight is really below 110 grams due to the nuts now being supplied from a different source. Is there evidence that the true average is less than 110 grams? Use a significance level of 5%.

98	103	109	112	108	102	114	105	96	112

7.10 – Hypothesis Testing for Proportions

1. Hypothesis testing for proportions is based on the normal distribution. On your calculator, choose: Stats/Tests/1-PropZTest (#5). The assumptions are the same as when calculating a confidence interval for a proportion. Some examples:

2. Example: A manufacturer claims to his customer that no more than 5% of his products will fail to meet the customer's specifications. Analysis of the first shipment of 1000 products revealed that 62 did not meet specifications. Is there strong statistical evidence that the manufacturer's claim is not valid? Use a significance level of 0.05.

Problem Type	Proportion z-test
Hypothesis	$H_o: p = 0.05$ vs. $H_a: p > 0.05$
P-value	$0.0408 \approx 4\%$
Decision	Since 0.0409 < 0.05, reject the null hypothesis
Conclusion	There is evidence at the 5% significance level that the manufacturer's claim is not true.

3. Example: Is drug use declining among teenagers between the ages of 13 and 17? Estimates from 2006 showed that about 11% of teenagers in this age bracket used some form of illegal drugs. An SRS of 500 teenagers, sampled from around the U.S. in 2007 revealed that 42 had used illegal drugs in the last year.

   a. Using $\alpha = 0.05$, can we make the claim that drug use in 2007 is declining from the 2006 level? (Use the fixed significance approach)

Problem Type	Proportion z-test
Hypothesis	$H_o : p_{2007} = 0.11$ vs. $H_a : p_{2007} < 0.11$
P-value	$0.0314 \approx 3\%$
Decision	Since 0.0314 < 0.05, reject the null hypothesis
Conclusion	There is sufficient evidence to say that the true proportion of drug use among teenagers in 2007 is less than the proportion in 2006.

   b. Suppose that 42 out of 450 had indicated that they had used drugs, can we make the claim that drug use is declining in 2007?

Problem Type	Proportion z-test
Hypothesis	$H_o : p_{2007} = 0.11$ vs. $H_a : p_{2007} < 0.11$
P-value	$0.1292 \approx 13\%$
Decision	Since 0.1292 > 0.05, fail to reject the null hypothesis
Conclusion	There is not evidence that drug use among teenagers has declined in 2007 compared to 2006.

4. Example: A large polling agency polls 1600 adult Americans at random and finds the following:

2007 Yearly Salary	Number
less than $20,000	46
between $20,000 and $50,000	734
between $50,000 and $80,000	500
between $80,000 and $110,000	224
between $110,000 and $140,000	64
more than $140,000	32

   Can we make the claim that about 5% of Americans have salaries less than $20,000 or more than $140,000?

   a. Use the fixed significance approach at the 10% level to evaluate the appropriate hypothesis.

Problem Type	Proportion z-test (1-PropZTest)
Hypothesis	$H_o : p_{<\$20K \; or > \$140K} = 0.05$ vs. $H_a : p_{<\$20K \; or > \$140K} \neq 0.05$
Point estimate of true proportion	$\bar{p} = \dfrac{46 + 32}{1600} = \dfrac{78}{1600} = 0.04875 = 4.875\%$
P-value	0.8185 = 81.85%
Decision	Since 0.8185 > 0.1, fail to reject the null hypothesis
Conclusion	There is not evidence that the claim is not valid, so we can claim the true percentage of Americans in this income category is 5%.

b. Use the (90%) confidence interval approach to evaluate this hypothesis.

Problem Type	Proportion CI (1-PropZInt)
Hypothesis	$H_o: p_{<\$20K \ or >\$140K} = 0.05 \quad vs. \quad H_a: p_{<\$20K \ or >\$140K} \neq 0.05$
Confidence Interval	$[0.03989, 0.0576] \approx [0.040, 0.058] = [4\%, 5.8\%]$
Decision	Since, 0.05 is in the confidence interval, fail to reject the null hypothesis
Conclusion	There is not evidence that the claim is not valid, so we can claim the true percentage of Americans in this income category is 5%.

7.11 Chapter 6 & 7 Calculator Summary
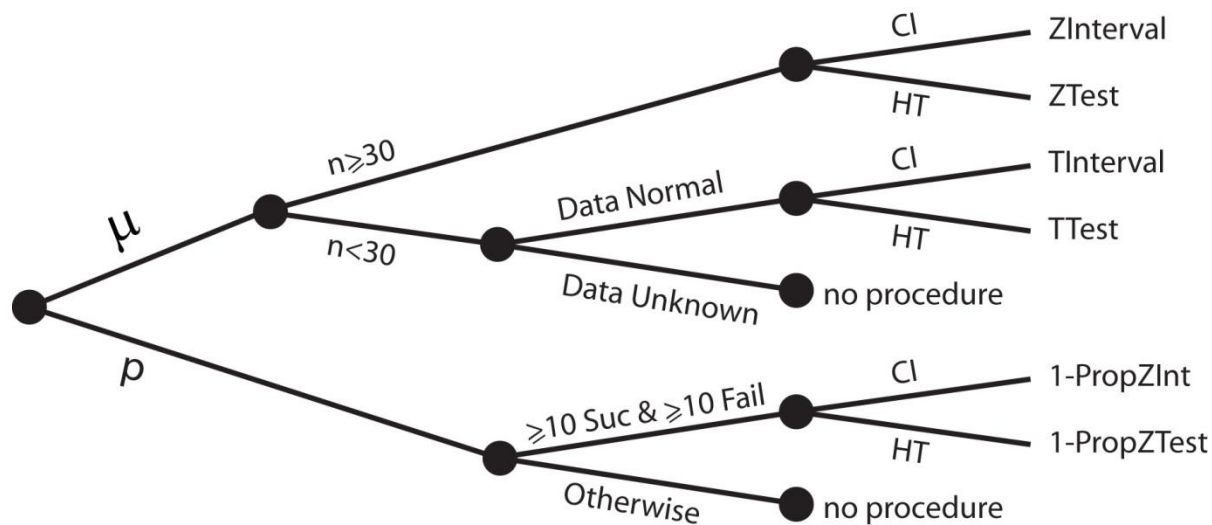
Homework 7.7

1. The length of a steel tube is required to be 10 *mm*. A manufacturer of these steel tubes is required by a customer to do a statistical test to demonstrate that more than 97% of his steel tubes will meet the specifications. What is the appropriate hypothesis to test?

   a. $H_o: p \geq 0.97$ vs. $H_a: p < 0.97$
   b. $H_o: p = 0.97$ vs. $H_a: p > 0.97$
   c. $H_o: \mu = 10$ vs. $H_a: \mu > 10$
   d. $H_o: p = 0.97$ vs. $H_a: p < 0.97$
   e. $H_o: \bar{p} = 0.97$ vs. $H_a: \bar{p} > 0.97$

2. An article in a New York City newspaper claimed that more than 30% of people in the city use the subway as their primary means of transportation in getting to work. A survey of 400 people shortly after the article was published, revealed that 140 people use the subway as their primary means of transportation to work. Test at the 5% significance level to see if there is evidence that the claim is true. Also find the point estimate of the true proportion that use the subway.

3. A survey of undergraduate students showed the following:

Rank in School	Number
Freshman	128
Sophomore	113
Junior	135
Senior	141

   a. Is there evidence at the 5% significance level that less than 50% of students are lower classmen (Freshmen and Sophomores)? Also find a point estimate of the true proportion that are lower classmen.
   b. Is there evidence at the 5% significance level that at least 45% of students are lower classmen (Freshmen and Sophomores)?