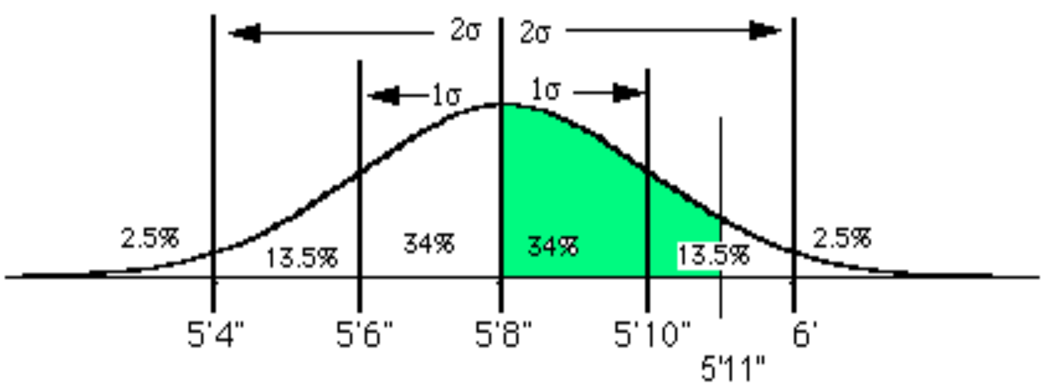
# CHAPTER 7

## Part a: Z-Scores

In previous chapters, we learned how to find the percentages and probabilities from a normal distribution curve when we are looking for a value that is EXACTLY between one or two standard deviations. But what if we are looking for information that is not exactly one or two (or even three) standard deviations from the mean? That is, information that is not directly between the lines we usually draw on a normal distribution curve. This is where we must use a set of values called Z-scores.

Let's go back for a minute to our example of the heights of men in the USA. As we saw, we found that the probability of finding a man taller than 6 feet (in a random sample) was .025. This is the area under the normal distribution curve beyond 6'. Now, let's say I wanted to find out the chance of finding (in a random sample of course) a man between 5'8" and 6'. Simply add the probabilities under the curve between these two heights! We find that 34% would be between 5'8" and 5'10" and 13.5% between 5'10" and 6'. Thus, adding these percentages together, we have 47.5% or there is a .475 chance of finding a man between 5'8" and 6' tall in a random sample.

Now, once again, let's assume one more thing. Let's say I wish to find the chance of finding a man who is between 5'8" and 5'11". Looking at our **normal distribution curve**, we don't have that exact percentage from 5'10" to 5'11" on the horizontal axis. We need to know the 'area' under the curve that corresponds to this little piece from 5'10" to 5'11". Fortunately, statisticians have perfected a way of calculating these numbers and they are called **'z-scores.'** These are simply a way to find the areas under the curve **from the mean** to the height we wish to find. (Areas under the curve **are** the probabilities!) Look at the diagram below.




The area from 5'8" to 5'11" is shaded for us. We know the areas for one standard deviation from the mean and for two standard deviations from the mean. But, 5'11" is more than 1 standard deviation away from the mean and not quite two standard deviations away. It is 1½ standard deviations away. So what is the area for 1½ standard deviations from the mean? This piece can be found by looking in the **table of z-scores** on the last page. We can estimate the answer since we know it must be less than 0.475. Looking at the table under **Z**, we see that the area for 1.5 standard deviations is .4332. This seems about right as it is more than 0.34 and less than 0.475.

We can calculate just about any probability between the mean and any height by finding the number of standard deviations our suspect height lies from the mean by using the following formula for '**z**'.

$$\boxed{z = \frac{x - \bar{x}}{\sigma}}$$

**Here's an example from the Chapter on Normal Curves, but this time we ask one more question:**

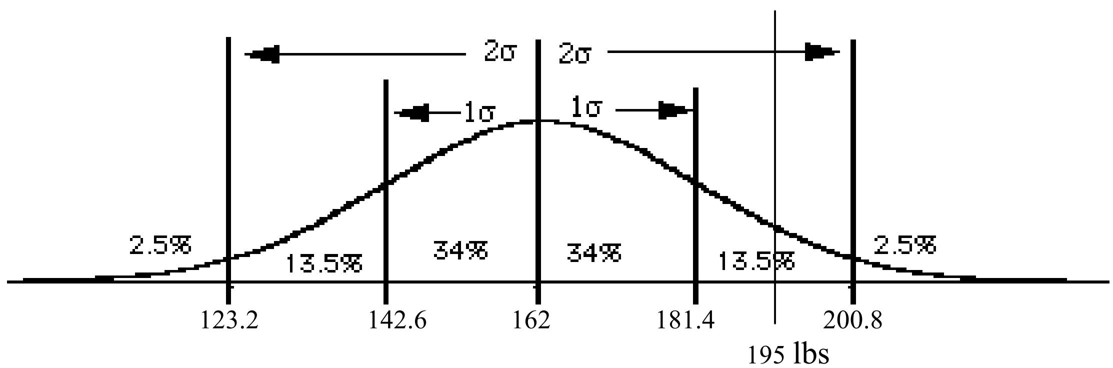
**ex. 1**) In a given population of males it is found that the mean weight is 162 lbs. Furthermore, the calculated standard deviation is 19.4 lbs. Assume that the distribution of weights among males is normal.

**a)** What percentage of males have weights between 162 lbs and 195 lbs?

**Step 1**: Draw the normal distribution curve with the percentages written INSIDE the curve.

**Step 2**: On the horizontal axis below the curve, put the mean in the center of the line.

**Step 3**: Using the mean and the given standard deviation, calculate one standard deviation above the mean simply by adding the standard deviation (in this case 19.4) to the mean. Put this result under the line to the right of the mean. Similarly, subtract the standard deviation from the mean and put that figure on the line to the left of center. Do the same for two standard deviations. Your result looks like this::




We see the question is asking for weights between 162 and 195 lbs. We must use z-scores since 195 lbs. is not in whole standard deviations from the mean.

Step 1: Draw a line on the graph about where 195 lbs might be. (see diagram above)

Step 2:..Find how many standard deviations 195 lbs is away from the mean of 162 lbs by calculating z.

$$z = \frac{195 - 162}{19.4} = 1.701 \ldots = 1.7 \text{ standard deviations}$$

Step 3:..Check the Z-score table. Look down the first column until you get to 1.7. Now look across the table to find the percentage associated with **1.7 s.d.** It is .4554 = **45.5%**

This is about right since we know the answer has to be more than 34% and less than 47.5%.
It's always a good idea to *estimate* your answer and see if your final calculation makes sense.

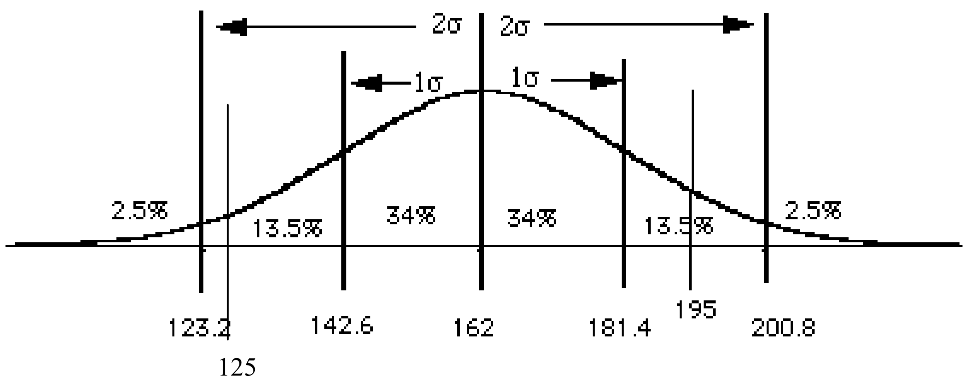
**Percentages to the Left of the Mean:** If we want a number to the left of the mean, we do the exact same thing. Do not worry about negative z's. Just use the absolute value to look up the associated percentage. We therefore re-write the formula for z:

$$\mathbf{z} = \left|\frac{\mathbf{x}-\bar{\mathbf{x}}}{\acute{\mathbf{o}}}\right|$$

ex.2) What percentage of males have weights between 125lbs and 162lbs? Look at the diagram below:

$$z = \left|\frac{125-162}{19.4}\right| = \left|\frac{-37}{19.4}\right| = 1.9$$

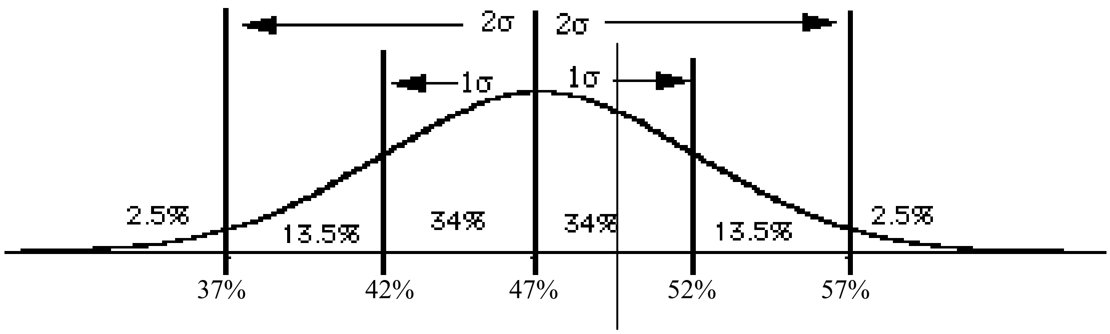
Associated percentage with z=1.9 is $\boxed{\mathbf{47.1\%}}$ Again this appears to make sense. It must be more than 34% and less than 47.5%.




**Reminder:** The z-table gives percentage between the mean only and the value you are working with. If you want a percentage beyond the mean either way, you must use the **B** column in the Z table. This will be demonstrated in the next example.

*Now let's go back once again to our hypothetical example of those who would vote for Barack Obama. This time we want to calculate Senator Obama's chances of winning i.e. of getting more than 50% of the vote! (Let's assume we have a two person election.)*

Our mean was 47% and the standard deviation was 5%. Draw the normal distribution curve for these figures. Put 47% as the mean in the center on the horizontal. Add 5% in one direction and subtract in the other, etc.




We ask the following question:

What is the probability that Senator Obama will get over 50% of the vote?

Step 1: Draw a line on the graph about where 50% is.

Step 2: Find how many standard deviations 50% is away from the mean of 47% by calculating z.

$$z = \frac{50\% - 47\%}{5\%}$$ Since this is all in percents, the % sign can be dropped.

$$z = \frac{3}{5} = .6$$ (If need be, round to one decimal.)

Step 3: Check the Z-score table. Look down the first column until you get to 0.6 Now look across to find the percentage associated with 0.6 s.d. It is 0.2257 in column A. This gives us the area between the mean and 50%. We need the area BEYOND the 50% or to the right of the line representing 50%. All we need to do is to look at the value in Column B. It is .2743

Therefore, Sen. Obama's chance of winning the presidency is only 27.4%.

Notice that .2743 + .2257 = .50. That is, they sum to 50% or half of the curve as it should be.

**Second method if you used column A.**

Step 4: We must subtract the amount of .2257 from the total to the right of the mean. Since this total is .50 or 50% we get

Prob = .50 – .2257 = .2743

So Senator Obama had a 27.4% chance of winning. (This figure may or maynot be enough for Senator Obama to decide to run????)

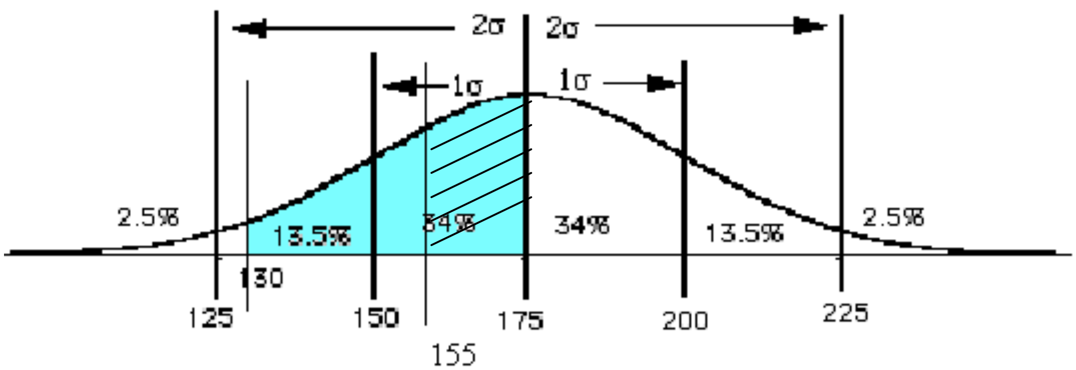
**Example of a slightly more involved Z-score Problem**

*The 911 emergency center in a certain city receives various numbers of calls on different days. A record of the center shows that the number of daily calls is normally distributed with mean of 175 and standard deviation of 25*. Calculate the following:

What percent of the days in this record have between 130 and 155 calls?

We need the percentage from 130 to 155. **But our z tables give us the percentages from the mean to x or beyond x.**

What we need to do is find two percentages and subtract to get the leftover. That is, find the percent between **130 and 175** and subtract the percent between **155 and 175.** What's left is the percent between 130 and 155. Look at the shaded area to see this more clearly. We need to subtract the area with the lines through it.




**Step 1:** From 130 to 175 we get:

$$z = \frac{130-175}{25} = -1.8$$

The negative in these cases just means that the 'x' value is to the left of the mean. Thus we only need the absolute value, or 1.8.

Looking up 1.8 in our table of z-scores, we find the percentage is .4641 or **46.4%. (you may leave it as the decimal to four places until you are done with the calculation.)**

**Step 2:** From 155 to 175 we get:

$$z = \left|\frac{155-175}{25}\right| = \frac{20}{25} = .8$$

Looking up the associated percentage we get: .2881

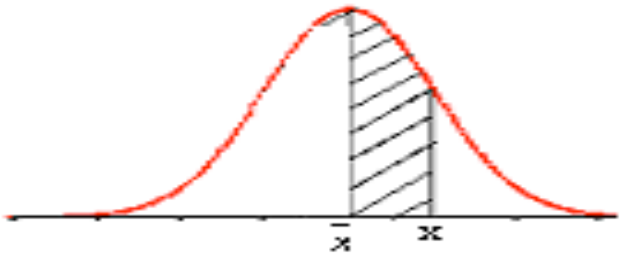
Finally, subtract the two percentages to get the space in between: **.4641 – .2881 = .176**

Thus the percent of days that get between 130 and 155 calls is $\boxed{\textbf{17.6\%}}$

Check to see if this makes sense. Between 130 and 150 we estimate less than 13.5%--maybe about 12% and between 150 and 155 we can estimate about 6%. Thus we should have an answer less than 18% which we do!

# Z-Scores*

A



B

| Z | Area between the Mean and x (curve A) | Area beyond x (curve B) |
|---|---|---|
| 0 | 0.0000 | 0.0000 |
| 0.1 | 0.0398 | 0.4602 |
| 0.2 | 0.0793 | 0.4207 |
| 0.3 | 0.1179 | 0.3821 |
| 0.4 | 0.1554 | 0.3446 |
| 0.5 | 0.1915 | 0.3085 |
| 0.6 | 0.2257 | 0.2743 |
| 0.7 | 0.2580 | 0.2420 |
| 0.8 | 0.2881 | 0.2119 |
| 0.9 | 0.3159 | 0.1841 |
| **1** | 0.3413 | 0.1587 |
| 1.1 | 0.3643 | 0.1357 |
| 1.2 | 0.3849 | 0.1151 |
| 1.3 | 0.4032 | 0.0968 |
| 1.4 | 0.4192 | 0.0808 |
| 1.5 | 0.4332 | 0.0668 |
| 1.6 | 0.4452 | 0.0548 |
| 1.7 | 0.4554 | 0.0446 |
| 1.8 | 0.4641 | 0.0359 |
| 1.9 | 0.4713 | 0.0287 |
| **2** | 0.4772 | 0.0228 |
| 2.1 | 0.4821 | 0.0179 |
| 2.2 | 0.4861 | 0.0139 |
| 2.3 | 0.4893 | 0.0107 |
| 2.4 | 0.4918 | 0.0082 |
| 2.5 | 0.4938 | 0.0062 |
| 2.6 | 0.4953 | 0.0047 |
| 2.7 | 0.4965 | 0.0035 |
| 2.8 | 0.4974 | 0.0026 |
| 2.9 | 0.4981 | 0.0019 |
| **3** | 0.4987 | 0.0013 |
| 3.1 | 0.4990 | 0.0010 |
| 3.2 | 0.4993 | 0.0007 |
| 3.3 | 0.4995 | 0.0005 |
| 3.4 | 0.4997 | 0.0003 |
| 3.5 | 0.4998 | 0.0002 |
| 3.6 | 0.4998 | 0.0002 |

* Adapted from "*Understanding Social Statistics,*" by Jane Fielding and Nigel Gilbert.

# Chapter 7b: Hypothesis Testing

Today we shall try and tie together the techniques of Statistics and Probability we have looked at and see if we can find ways of testing the various kinds of theories or **hypotheses** which people love to postulate, especially around election time. We shall need to keep in mind several important concepts from previous lessons:

a) The formula for calculating the standard deviation of sample fractions:

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

**p** is the *actual* (or in this case *hypothetical*) mean of the population - we hope either to **disprove it or show that we cannot disprove it;**

**n** is the **size** of the sample we are looking at.

b) Because the distribution of sample fractions is **normal**, we can use our "rough and ready" percentages:

- 68% lies within **one** standard deviation of the mean

- 95% lies within **two** standard deviations of the mean.

c) Remember that our samples must be **randomly** chosen.

**Hypothesis Testing:** This is a statistical way of testing a ***theory*** (hypothesis) about the population as a whole based on the ***observations*** of a sample drawn at random from the population. We use what is called "Hypothesis Testing." Key words in this method are:

a) The **null hypothesis** - this is a "straw man" hypothesis that we set up, only to try and disprove it or 'knock it down.' It's the opposite of what we really want to prove and, by disproving the null hypothesis, we end up supporting out own "hypothesis."

Notation indicating the Null Hypothesis: $\mathbf{H_0}$

b) A **5% level of significance** - this means that we are 95% sure that we are correct in rejecting the null hypothesis (more about that later), and also that there is a 5% (one in twenty, i.e. 1/20) chance that we are mistakenly rejecting it when it is in fact correct.

## *Hypothesis Testing – An Explanation*[1]

Suppose you are a grad student in a statistics class and you were assigned to do a study on President Obama's new $3.6 trillion budget. You took a random sample of 400 people and asked the question, "Do you agree with President Obama's new budget proposal?" Out of the 400 people you asked, 224 said they did not agree. In other words, out of 400 people, 224 gives a

$$\text{'sample fraction' of } \frac{224}{400} = 0.56 = 56\%$$

You would love to write your paper and say the results of your study showed that there is a clear majority of people in the country who are opposed to President Obama's budget proposal.

Your professor, who is trying to teach you how to do a hypothesis test, points out that although you have taken a random sample, any random sample is subject to fluctuations. We do not expect any one given sample to give a perfectly accurate figure. She reminds you that even if people's opinions are evenly divided, i.e. 50% opposed to the budget and 50% in agreement with the President, it's perfectly possible to get a random sample which has 56% of people against the budget. Your sample, she says, doesn't PROVE that the majority of people are against the budget. She says that opinion may be 50-50!

You now need to show the professor what you have learned about Hypothesis Testing. You know there is no sure-fire way to PROVE it unless you went and asked absolutely everyone, a daunting task. However, we should be willing to settle for less than absolute certainty because we very often make decisions every day without being "absolutely" certain. You just need to show that it is very unlikely that there is a 50-50 split in opinion.

Let's step aside for a minute from the student's thinking and see what we are talking about in regards to absolute certainty. For example, if I told you that the Boston Red Sox won a baseball game yesterday by a score of 14 to -5, you would tell me that it was impossible because you could not get a negative score. But, what if I told you they won by a score of 658 to 7? You would probably think I made a mistake, not because it was impossible, but because it is SO UNLIKELY to score that many runs in a game.

Let's look at yet another example. If I told you that every one of the 30 math instructors in the math department was born in March, you again would not believe me, not because you were "absolutely certain" it was untrue, but because it is so UNLIKELY that all 30 would be born in the same month.

Getting back to your study, although you cannot prove that people's opinion is evenly split about the new budget proposal, is there a way you can show the instructor that it is SO UNLIKELY that we may reasonably conclude that it is possible that the majority are against it? This type of thinking is called ***Hypothesis Testing***. We <u>assume</u> that opinions ARE evenly split, i.e. we assume the CLAIM of the skeptic, (in

---

[1] Adapted from "CORE CURRICULUM QUANTITATIVE REASONING REQUIREMNT." Gleason, Molay, Hallett, et. al., 1991. President and Fellows of Harvard College.

this case the instructor), and then say that if the claim (50-50) were true, then it would be EXTREMELY UNLIKELY to get the survey results we did, i.e.
56% against. Thus we doubt the assumption. (50-50). In statistical language, the skeptic's claim that opinions are evenly divided is called the ***null hypothesis***. Your plan is – for the sake of argument – to assume the null hypothesis is true (even though you believe otherwise) and show that this leads to an extremely unlikely result.

Okay, back to the 'skeptical' instructor. Let's see what happens if you go along with her and say that indeed 50% of the people are against the new budget and 50% of the people are for the new budget. You need to see how LIKELY it is to get a survey result at least as far from 50% as you did by CHANCE alone, i.e. just by the luck of the draw in selecting this random sample, if indeed opinion was really in fact evenly split. Here is what you do.

1) You assume the true value of those against the budget is 50%. This is 'p', or the mean in the population. That is, 50% of all people are against the budget and 50% are for it. So if true, in a random sample of 400 people, you would EXPECT to get 50% against. Our expected mean is 50% or

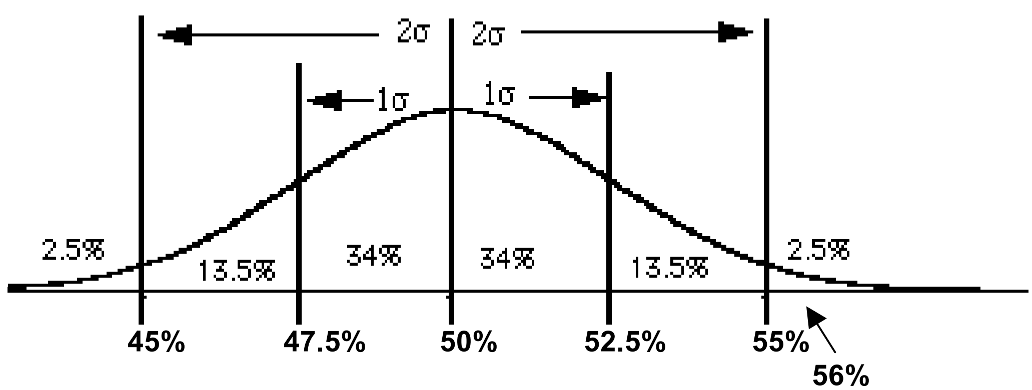
p = 50%    n = 400

Your histogram of sample fractions if it were possible to take many, many samples, would look normal, with the mean = 50% and standard deviation equal to the calculated sample standard deviation using p = 50%.

2) Calculate the standard deviation using the new formula, $\sigma = \sqrt{\dfrac{p(1-p)}{n}}$

$$\sigma = \sqrt{\frac{.50(1-.50)}{400}} = \sqrt{\frac{.50(.50)}{400}} = \sqrt{\frac{.025}{400}} = .025 = 2.5\%$$

3) Draw the histogram.




4) Now look to see where your OBSERVED sample fraction of 56% lands. It's way to the right of 55%. That is, it is two full standard deviations beyond the EXPECTED value of 50%. How likely is it to get a sample result at least this far from the expected value? Using our normal approximation methods we learned recently, we can compute the likelihood. Remember from lesson 6, we found out that 95% of all samples should yield results within two standard deviations of the mean, in this case within two standard deviations of 50%. So that leaves 5% further away from 50%. In other words, if opinion were really split 50-50, less than 5 out of 100 would give results at least as lopsided as ours.

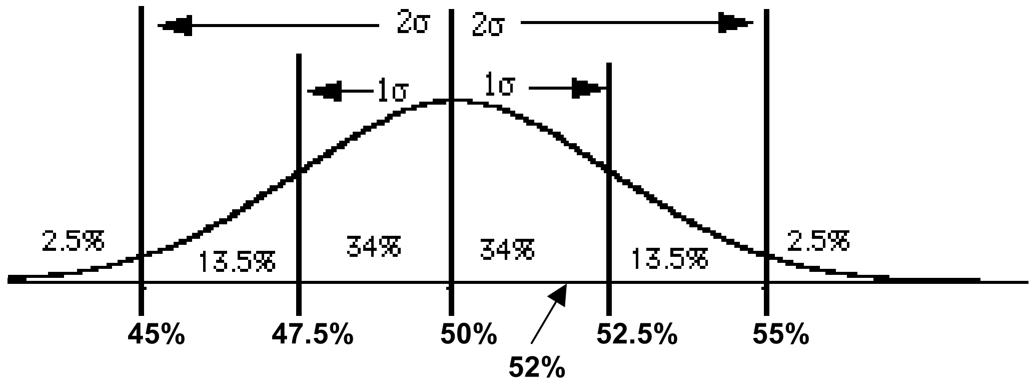
That means getting a sample result like 56% is very unlikely. (We could figure out the probability even more closely using z-scores.) In fact, looking at our diagram of the normal distribution curve, we see that the 56% lands in the tail to the right so we can say that there is less than a 2.5% chance of getting a sample like we did if opinions were evenly split. Thus it is more reasonable to conclude that the 50-50 split was wrong, and that our sample showed that a majority of people really did oppose the new budget proposal. Of course, it is POSSIBLE that opinions are evenly split, but our sample result would be SO UNLIKELY that we don't believe the assumption.

This is how hypothesis testing works. If we would like to show something (in our case, that the majority of people were opposed to the new budget), then we first assume the contrary. We then try to show that the observation we made (56%) is so unlikely that we must strongly doubt the assumptions and thus believe that the premise we originally wanted to show.

In essence, we set up a straw man and try to knock him down. If a doctor is trying to show that a drug is effective against a certain disease, he assumes as his null hypothesis that the drug has NO effect. (This is where the word 'null' comes from.) He then tries to show that the probability of his results occurring under this assumption is so small that he must REJECT the hypothesis that the treatment has no effect and conclude that it does make a difference. When we make a statistical test of a hypothesis like, "the majority of people are opposed to the new budget proposal" or "this drug is effective against a disease", we are making an argument against chance. We want to show that it is simply too unlikely to get evidence like ours as a result of variability in the sampling process alone.

When trying to decide whether our chances are sufficiently small to reject the null hypothesis, how small is small enough? The usual convention is that the probability be less than 5%. By our rules for normal approximation, we know that this means our sample result must fall at least two standard deviations from the result predicted by the null. But this 5% convention is an arbitrary choice, and in any given situation you may wish for stronger evidence that the null hypothesis is untenable (i.e. an even smaller probability, corresponding to a sample result even further than two standard deviations away) or you may be willing to settle for weaker proof (i.e. a larger probability, corresponding to a closer sample fraction). For example, in the case of the doctor testing a new drug, he may desire stronger evidence, especially if there are potentially harmful side effects.

Now suppose, for example, that we had found that only 52% of our sample was opposed to the budget proposal. It still looks like the majority of people are opposed since 52% > 50%. But let's again test to see whether this hypothesis holds up again. Once more, for the sake of argument, we'll play along with the skeptic and assume opinion is evenly split, 50-50. If that were true, we would expect the same distribution of sample results as we computed above, centered at the claim of 50% and with a standard deviation of 2.5%. But now our observed result is only 2 percentage points above the expected, so it's less than one standard deviation away. Look at the diagram below.




Notice that about 1/3 of samples (34%) would look like ours even if opinions were split evenly. This means that it is simply not UNLIKELY enough to feel confident rejecting the 50-50 theory.

On the other hand, if we had obtained an even larger percentage, say 60%, from our 400 people sample, our result would have fallen 10 percentage points away from the expected mean or four standard deviations away from the mean. The probability of getting a sample like that would be miniscule. In fact, 99.7% of the data should fall within three standard deviations of center. Four standard deviations would be about 99.994% of the data, so the probability of being at least this far away would only be about 0.006%. This means only 6 samples in 100,000 would give results as lopsided as ours. Now we can be extremely confident that the claim of 50-50 split was wrong. To sum up how to use hypothesis testing you need to do the following:

- Formulate the null hypothesis, i.e. the skeptic's claim that nothing more than the luck of the draw in the sampling process is responsible for producing our results.

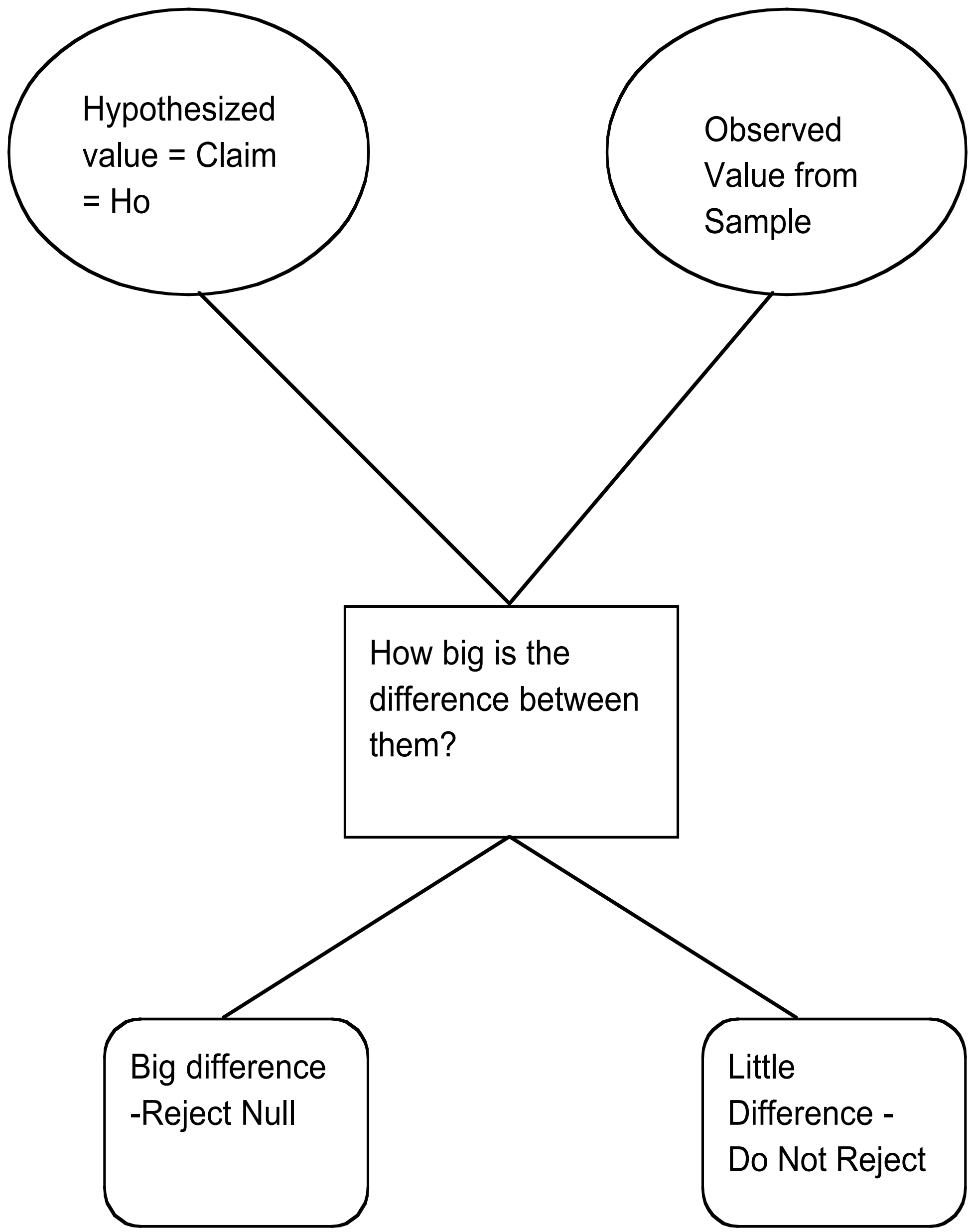
- Assume the null is true (even if you don't believe it) and estimate the probability of getting an outcome at least as far from the one expected by the null as yours is. To do this, look at how many standard deviations away your sample result falls from the expected.

- If the sample is too far from the expected, **REJECT** the null and conclude that there really is something more than chance variation at work. (If the probability is NOT small enough, you **CANNOT REJECT** the null. You have not proven the null hypothesis is true, but you don't have sufficiently strong reason to doubt it either.)

***Level of Significance***:

In *general,* statisticians use two standard deviations as a guide for when we are far enough away from the expected mean to feel comfortable rejecting the null. We should not get too hung up on such an artificial cut-off. The true odds range over a continuum of values. A probability of 4.9% does not 'prove' our hypothesis was true while odds of 5.1% 'disprove' it. In using our artificial cut-off, we are doing the best we can. We are working on probabilistic grounds so we are always going to be taking a risk that we made an error. Using the 5% 'level of significance' or the two standard deviations is the usual guide. In the next section, there is a diagram of the process as well as the <u>steps</u> outlined in detail using a specific example.

# How it works:




*Taken from Applied Statistics for the Behavioral Sciences, by Hinkle, Wiersma, and Jurs.

## <u>Comparison with Probability Type II - (Lectures 6 and 7)</u>

We now add another example of Probability Type II. Notice how you use the different percentages.

| | **Lecture 6 – The Normal Distribution Curve** | **Lecture 7 – Null Hypothesis Testing** |
|---|---|---|
| **Mean** | Could be a number = $\bar{x}$<br>or a Percent **= p** | Use the **CLAIM** percent<br>**Not** the Observed |
| **Standard Deviation** | If mean is a number, so is s.d.<br>If mean is a percent, so is s.d.--<br>usually given to you | Calculate the standard deviation based on the **CLAIM** percent **Not** the observed |
| **Diagram** | Set up Normal Distribution Curve with the mean in the center. One and Two standard Deviations to left and right. Fill inside of Curve with rough and ready rules. | Set up Normal Distribution Curve.<br>The Mean is in the center. This is the **CLAIM** percent **Not** the observed. Use the s.d. you calculated. |
| **Conclusion** | Check for Probability of something occurring or the percent of things occurring. | Look at your **OBSERVED** percent for the **first** time and see where is LANDS. If **INSIDE** the *likely region* you **CANNOT** reject.<br>If ***OUTSIDE – reject***. |

**On the following page, you will find an example of the steps you need in doing a Hypothesis Test. Follow these steps when doing your homework problems. You must show all work including the diagram.**

## Step Method for Solving Hypothesis Tests.

**Ex:** The manager of a radio station **claims** that **25%** of all college students in the area listen to his station. A random sample of 200 college students in the area showed that only 45 of them did listen to the station. **Test** the accuracy of the manager's statement at a 5% level of significance.

**Step 1: State the Null Hypothesis (You Must State the Hypothesis First for Every Problem**. This is similar to a scientific experiment. State it as a fact even though you do not believe it's true.)

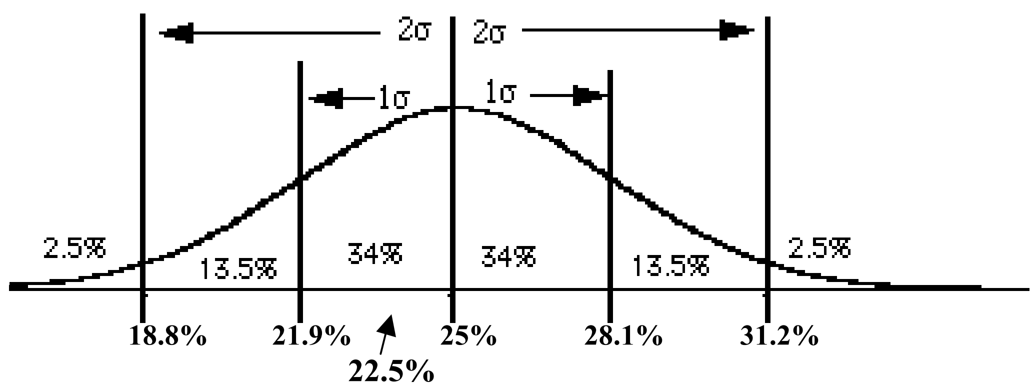
**Null Hypothesis:** 25% of all college students in the area listen to this radio station.
(You may use the symbol for Null Hypothesis, $\mathbf{H_o}$, or just write **NH**: or the whole phrase.)
e.g. $\mathbf{H_o}$**:** 25% of all college students in the area listen to this radio station.

**Step 2: Calculate the standard deviation** using the **claim % for p** and the number of people in the sample for **n.** so **p = 0.25 and n = 200**

$$\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.25(1-0.25)}{200}} = \sqrt{\frac{0.25(0.75)}{200}} = \sqrt{\frac{0.1875}{200}} = \sqrt{0.0009375} = 0.03061862 = 3.1\%$$

**Note:** we rounded the standard deviation to one decimal place.

**Step 3:** Draw the diagram **(calculate p + 2σ and p - 2σ)**




**Step 4: Compare** the percentage that you got from your **observed sample results** to this interval. If it falls inside, you **cannot** reject. If it falls outside you can reject. For this problem we must calculate the percentage from our sample observation since only the total number of students was given.

$\mathbf{\frac{45}{200} = 22.5\%}$ Now look to see where this percentage lands on your diagram. This one falls **within** the two standard deviation range, i.e. **inside** the likely region.

**Step 5: Construct** the proper sentence ***either rejecting or not rejecting the Null Hypothesis***.

***Since*** our observed percentage falls inside the likely region, we ***cannot reject*** the manager's claim that 25% of all students in the area listen to his station, at a 5% level of significance**.** (You must include three things in this statement: **1**. Where the observed percentage falls-inside or outside the 95% interval;
**2**. Whether or not we reject the NH; **3.** at what level of significance you did the calculation – it will generally be a 5% level of significance for our class.)

**Step 6: Write a short concluding** statement in your own words.

We conclude that it is possible (but not certain) that the manager is accurate in his assessment.