http://wiki.stat.ucla.edu/socr/index.php/SOCR_Courses_2008_Thomson_ECON261

Lecture 7

Introduction to Hypothesis Testing

By Grace Thomson

INTRODUCTION TO HYPOTHESIS TESTING

What is Hypothesis Testing?

It's an inferential technique that allows managers and decision makers to identify and control the level of uncertainty. Through a hypothesis test you can draw conclusions as to the validity of your sample as an estimator of your population parameters.

There are basically 5 steps that we need to follow to perform a hypothesis test:

1. Specify Population of interest μ, and Formulate Null and Alternative Hypothesis

Alternative Hypothesis (Ha)

Also called research hypothesis, it includes the statement of what you wish to show. It's the statement to be accepted as true when the Null hypothesis is rejected. It never contains the equality sign and always contains opposite signs to the null hypothesis

Null Hypothesis (Ho)

It's a statement about the population that will be tested. A null hypothesis will be rejected only if sample provides evidence in contrary. It contains the equality sign "=". Represents the status quo (if things didn't change). This is the hypothesis to be REJECTED or NOT REJECTED.

A hypothesis test One-tailed or Two-tailed. This will define special characteristics to the test.

A two-tailed test is formulated as follows:

Ho: $\mu = \mu i$

Ha: $\mu \neq \mu i$

One tailed-tests may have the rejection area on the lower end of the distribution

Ho: $\mu > \mu i$

Ha: $\mu \leq \mu i$

Or on the upper end of the distribution

Ho: $\mu < \mu i$

Ha: $\mu \geq \mu i$

2. Determine the test to be used: Z test, t test or p-value

The type of test to be used will depend on factors such as sample size and information about the standard deviation.

If n $\geq$ 30, and σ is known → Use Z

In most cases we use Z test, whether σ is known or unknown, as long as n>30.

$$\text{Z test} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

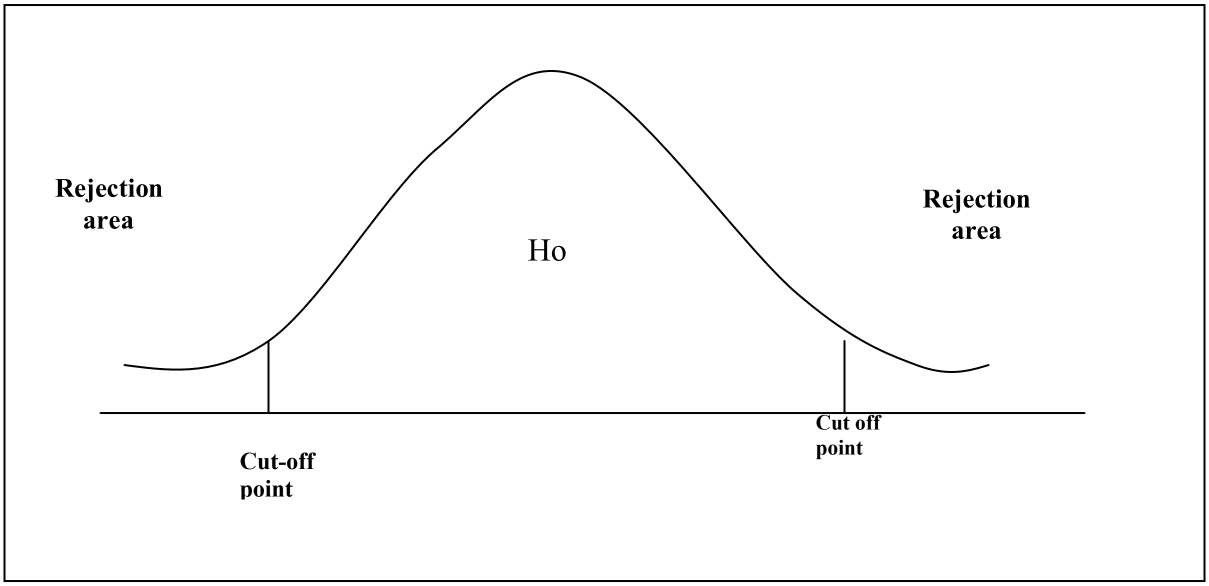
If σ is unknown and n< 30 → Use t

$$\text{t test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

If you are working with proportions, → Use p-value: p (Z < Zi)

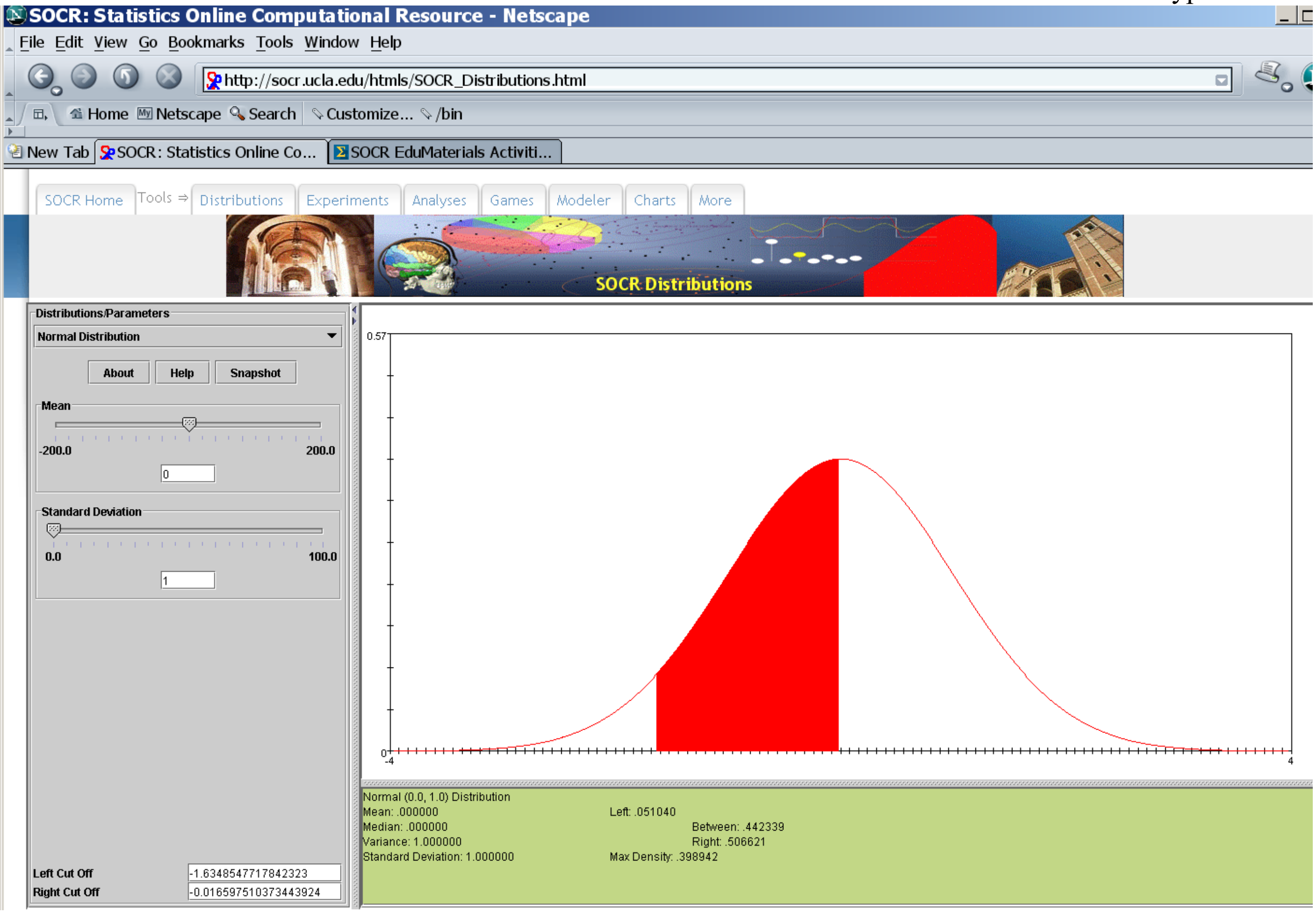
3. Define the rejection area, compute the critical values (cut-off points) of the distribution, Zcritic, tcritic and define the decision rule.

The most important part in a hypothesis test is to define what area under the curve is considered the area of rejection of the null hypothesis. This area is defined by the critical values of the distribution, which in turn are determined by the significance level that the researcher wants to give to their estimation.



The significance level is called alpha (α), and it will be usually given to you or you may choose it by experience. It is usually set in a level of 0.05, 001 or 0.10.

You can compute this critical values or cut-off using SOCR (see image below) or using Excel. Here are all the choices.



Using Excel

Z_{critic} → =NORMSINV (α) lower -tail test

= NORMSINV(1-α) upper-tail test

= NORMSINV (α/2) two-tailed test

t_{critic}→ =TINV(α*2, n-1) one-tail test (*write it in negative for lower-tail test, and in positive for upper-tail test)*

=TINV(α , n-1) two-tailed test

Decision rule: Make the statement of your decision comparing the statistic and the critical value.

Here are all the choices for the 3 types of tests:

If $Z_{statistic} > Z_{critic}$ or $Z_{statistic} < -Z_{critic}$ Reject Ho

If $t_{statistic} > t_{critic}$ or $t_{statistic} < -t_{critic}$ Reject Ho

If p-value < α Reject Ho

Not rejecting Ho means that *the difference between sample mean and is not large enough to attribute difference to anything but sampling error.*

4. Compute the statistics of the problem: sample mean, sample standard deviation, Z test, t test or p-value.

Z test = $\dfrac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ or t test = $\dfrac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ or p-value: p (Z < Zi)

5. Draw a conclusion and make a decision

Based on the decision rule stated for each case compare statistics vs critical values:

If Ho is rejected what is the interpretation for Ha?

If Ho is not rejected what is the interpretation for Ha?

These steps are constant in any type of hypothesis you may need to formulate.

CHEAT SHEET

Steps to perform a hypothesis test:

1. Specify Population of interest μ, and Formulate Null and Alternative Hypothesis
2. Determine the test to be used: Z test, t test or p-value
3. Define the rejection area, compute the critical values (cut-off points) of the distribution, Zcritic, tcritic and define the decision rule.
4. Compute the statistics of the problem: sample mean, sample standard deviation, Z test, t test or p-value.
5. Draw a conclusion and make a decision

Application: one-tailed test using z

Case 1 Efficiency in a Hospital

Let's say that you need to prove the efficiency of your area in the hospital, and you are interested in testing if the average waiting time per patient in your section is less than 35 minutes, you know that the standard deviation is 4.15 for the industry. You have taken a sample of 40 patients and their average waiting time is 32. Test this hypothesis using a level of significance of 0.05.

1. **Specify Population of interest μ, and Formulate Null and Alternative Hypothesis**

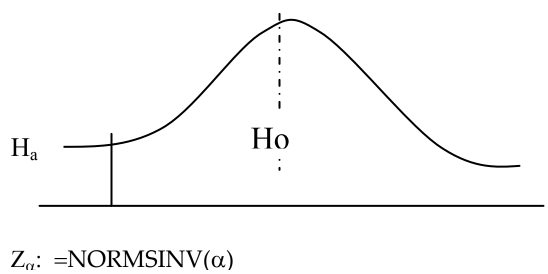
Average waiting time per patient is less than 35 minutes (this is Ha)

Population of interest: patients in the hospital

Ho: $\mu \geq 35$

Ha: $\mu < 35$

$\alpha = 0.05$



2. **Determine the test to be used: Z test, t test or p-value**

Since σ is known, n=40 , Use Z test

3. **Define the rejection area, compute the critical values (cut-off points) of the distribution, Z_α, t_α and define the decision rule.**

We will use Zcritic at a level of significance $\alpha = 0.05$.

Use excel → =NORMSINV(0.05)= 1.65

<u>Decision rule</u>

If $Z_{test} < -Z\alpha$, Reject Ho otherwise Don't reject it

Notice that the direction of the sign in the alternative Hypothesis (<) hints you that the rejection area is to the left.

4. Compute the statistics of the problem:

Sample mean is given: 32; Population standard deviation σ: 4.15

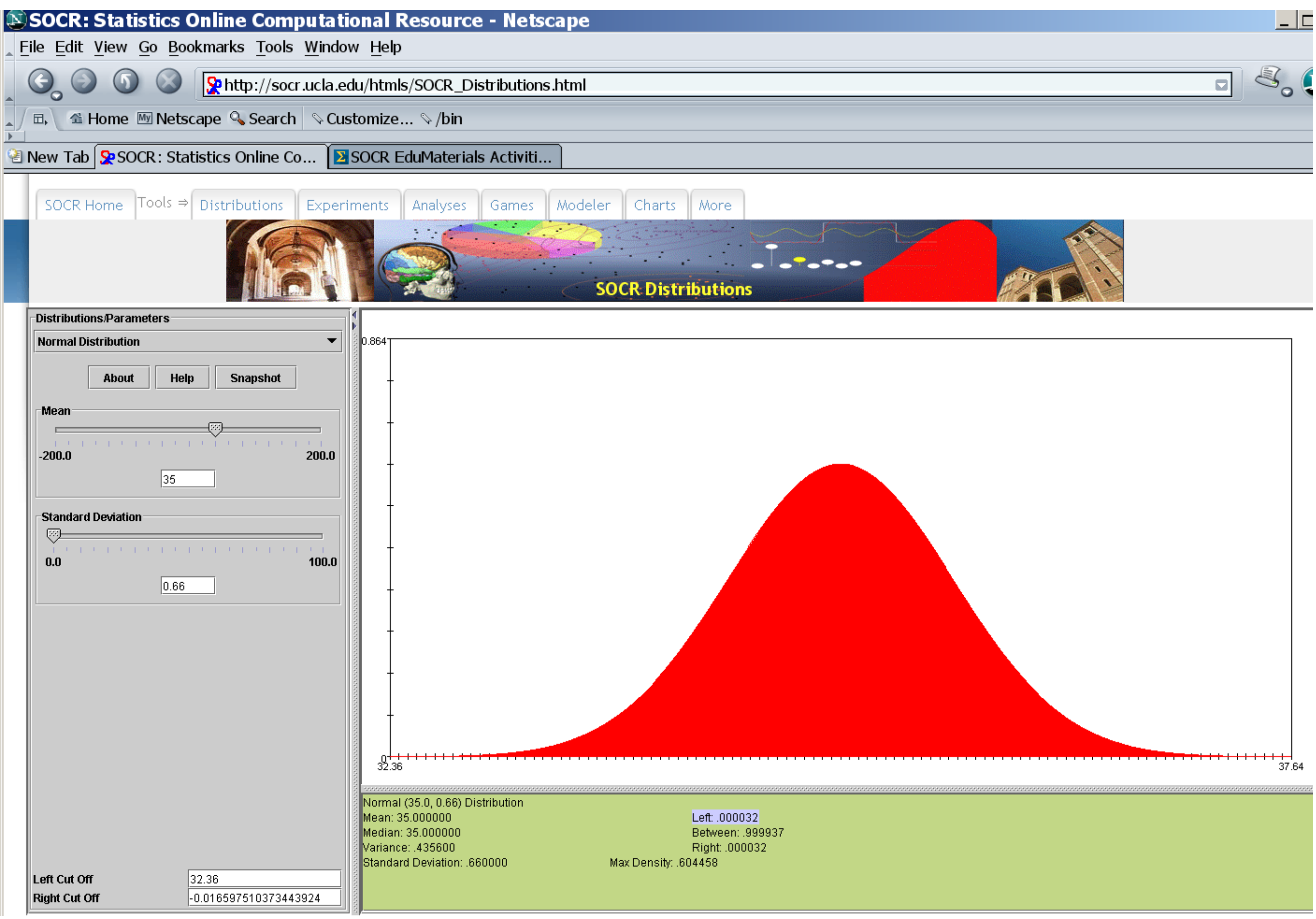
$$Z_{\text{test}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 35}{\frac{4.15}{\sqrt{40}}} = \frac{-3}{0.66} = -4.54$$

5. Draw a conclusion and make a decision

Because Z_{test} -4.54 < -$Z\alpha$ -1.65 , Reject Ho

There is statistical foundation to say that the average waiting time for patients in your area is less than 35 minutes.

Use SOCR Normal Distribution (http://socr.ucla.edu/htmls/SOCR_Distributions.html):



If you want to use p-value, you will need to find p(Z< -4.54), using excel formula =NORMSDIST(-4.54)

If p-value < α Reject Ho, otherwise don't reject it

p(Z< -4.54) = 0.000002

Since p-value 0.000002 < α 0.05 Reject Ho

We reach the same conclusion

Application: two-tailed test using z

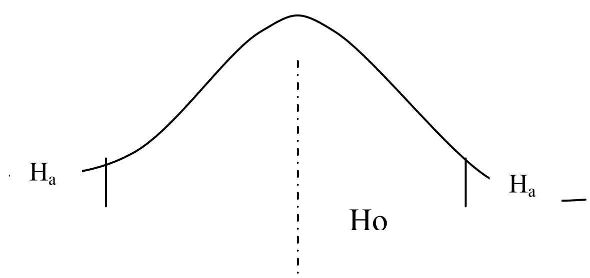
Using the same example about the efficiency in the hospital, test now if the average waiting time per patient is **equal** to the standard of 35 minutes. The standard deviation of the industry is 4.15 and you have taken a sample of 40 patients and their average waiting time is 38. Test this hypothesis using a level of significance of 0.05.

Draw the normal distribution graph and define your rejection area.

Ho: $\mu = 35$

Ha: $\mu \neq 35$

α= 0.05



$Z_{\alpha/2}$: =NORMSINV(α/2) $Z_{\alpha/2}$: =NORMSINV(α/2)

We will use Z at α= 0.05, alpha has to be divided by the 2 tails of the distribution.

Use excel → =NORMSINV(0.05/2) =NORMSINV(0.25)= $\pm$1.96

If $Z_{test} > +Z_{\alpha/2}$

Or if $Z_{test} < -Z_{\alpha/2}$, Reject Ho otherwise Don't reject it

Notice that the equality sign in the alternative Hypothesis (≠) hints you that the rejection area is two-tailed.

Compute Z test

$$Z_{test} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{38 - 35}{\frac{4.15}{\sqrt{40}}} = \frac{3}{0.66} = 4.54$$

Make a decision and draw a conclusion

Because Z_{test} 4.54 > 1.96, Reject Ho

There is statistical foundation to say that the average waiting time for patients in not equal to 35 minutes.

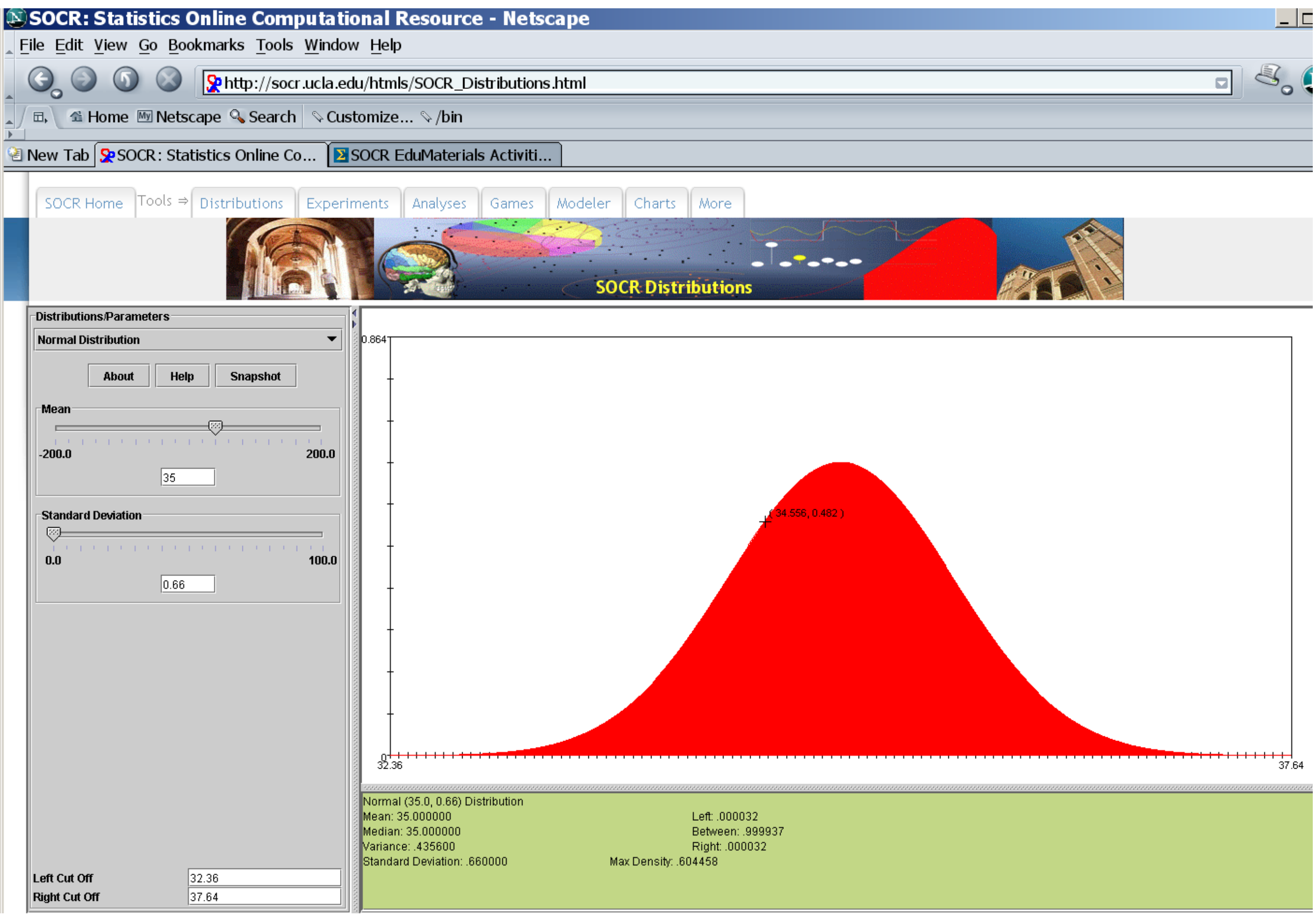
If you want to use p-value, you will need to find p(Z> 4.54) using excel formula =(0.5 - NORMSDIST(4.54)) =

p-value = 0.000002

multiply by 2 = 0.000004

compare 2pvalue vs. alpha α

Use SOCR Normal Distribution (http://socr.ucla.edu/htmls/SOCR_Distributions.html):



APPLICATION: ONE TAILED TEST USING STUDENT- t

Now, you are interested in testing if the average waiting time per patient in your section is **less than 35** minutes? But you only got a small sample of 20 patients and don't have any reference about the population standard deviation. You know that your sample average waiting time is 32 and their standard deviation is 4.15. Test this hypothesis using a level of significance of 0.05.

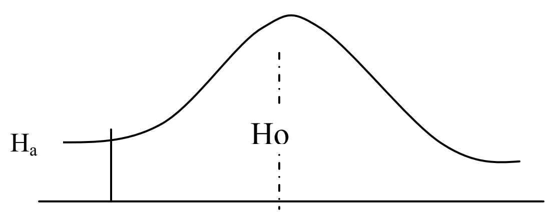
Notice that in this case the standard deviation is not a population number, is a sample number!!

Draw the normal distribution graph and define your rejection area.

Ho: $\mu \geq 35$

Ha: $\mu < 35$

$\alpha = 0.05$



Z_α: =TINV(α*2, n−1)

Use t value at $\alpha = 0.05$, degrees of freedom (n-1) = 19

Use =TINV(0.05*2 , 19) = -1.73.

Notice that we have to write it with a negative sign because it's a lower-tail test.

If $t_{test} < -t_\alpha$, Reject Ho otherwise Don't reject it

Notice that the direction of the sign in the alternative Hypothesis (<) hints you that the rejection area is to the left.

Compute t test

$$t_{test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{32 - 35}{\frac{4.15}{\sqrt{20}}} = \frac{-3}{0.93} = -3.23$$

Make a decision

Because t_{test} $-3.23 < \tau_{(\alpha,\, n-1)}$ -1.6848 , Reject Ho

Draw a conclusion

There is statistical foundation to say that the average waiting time for patients in your area is less than 35 minutes.

APPLICATION: TWO-TAILED TEST USING STUDENT- t

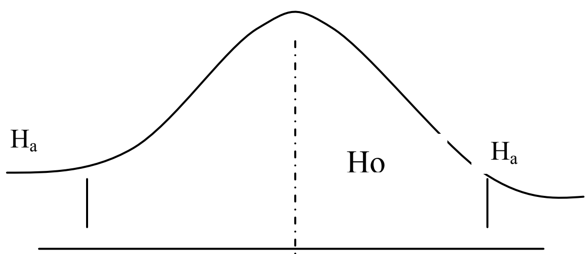
Now you are interested in testing if the average waiting time per patient is **equal to the standard** of 35 minutes? You only have a small sample of 20 patients and their average waiting time is 38 and their standard deviation is 4.15. Test this hypothesis using a level of significance of 0.05.

Draw the normal distribution graph and define your rejection area.

Ho: $\mu = 35$

Ha: $\mu \neq 35$

α= 0.05



Use t value at α= 0.05 and degree of freedom (n-1)= 19.

Use excel =TINV(α, n-1) =TINV(0.05, 19) = $\pm$2.022,

Notice that for two-tailed tests we directly type 0.05 in the formula.

If $t_{test} > +t_{\alpha/2}$

Or if $t_{test} < -t_{\alpha/2}$, Reject Ho otherwise Don't reject it

Notice that the equality sign in the alternative Hypothesis (≠) hints you that the rejection area is two-tailed.

Compute t test

$$t_{test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{38 - 35}{\frac{4.15}{\sqrt{20}}} = \frac{3}{0.93} = 3.23$$

Make a decision and draw a conclusion

Because t_{test} 3.23 > 2.022, Reject Ho

There is statistical foundation to say that the average waiting time for patients is not equal to 35 minutes.

Hypothesis Test for Proportions

When you use proportions you will state your Null and Alternate Hypothesis in terms of population parameters (p), using your sample statistic ($\overline{p}$) as an estimator.

Null Hypothesis will be a statement about the parameter that will include the equality

Alpha (α) determines the size of the rejection region

Test can be one or two-tailed, depending on how the alternative hypothesis is formulated.

The most important requirement to perform a hypothesis test for proportion is to assume that the distribution is normal, and in this case, n has to be sufficiently large such that $np \geq 5$ and $n(1-p) \geq 5$.

These are steps for a hypothesis test:

1. Formulate Hypothesis (one-tailed or two-tailed)

Ho: p = po

Ha: p ≠ po

2. Compute $Z_{test} = \dfrac{\overline{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$

Where:
$\overline{p}$ = sample proportion (x/n)
P = population proportion
n= sample size

3. Compute Z_α using the alpha provided by the problem
4. State the Decision rule:

If $Z_{test} > Z_\alpha$ Reject Ho (*use the decision rule that best fits to your needs)*

Type II Errors

Remember when we learned that alpha (α) measures the level of significance of a hypothesis test? Well, at the same time, alpha measures Error Type I, which is the probability of rejecting a Null Hypothesis when this is indeed true.   We can figure that this possibility is always present, considering that you might take samples that might be away from the true mean, or close to the true mean.

Now, we will learn how to measure the probability of accepting a Null Hypothesis when this is indeed false. This is a very important concern of researchers.   This probability is called Beta (β), or probability of committing Type II errors.

β is computed before the sample is taken. Beta is calculated based on the statement of the alternate hypothesis.

Remember the following key elements:

1. Beta is defined as the probability of every possible value of μ stated by Ha.
2. When computing Beta, we need to add a “What if” phrase to our statement.
3. Difference between μ stated in Ho and the μ that could be contained in Ha is what determines β

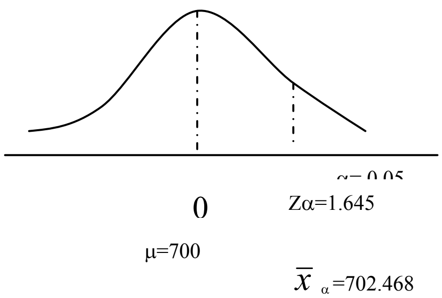
If μ_{Ho} is far from μ_{HA}
Chances of committing β is low

If μ_{Ho} is close to μ_{HA}
Chances of confusing them both and committing β is high

Let’s say for example that you want to test the hypothesis that □= 700, at an □= 0.05 with n=100 and σ=15.

Ho: μ ≤ 700

Ha: μ > 700



But you also want to compute the probability of committing Type II error (β) in this research. You will need to follow the steps below:

1. Find critical Z for an upper-tail hypothesis at a 0.05 alpha → =NORMSINV(1-0.05)= 1.645

2. Compute critical $\bar{x}_{\alpha}$ corresponding to that level of Z, using the formula of interval of confidence: $\overline{X}_a = \mu + Z_\alpha \frac{\sigma}{\sqrt{n}} = 700 + 1.645\frac{15}{\sqrt{100}} = 702.468$ (See graph above)

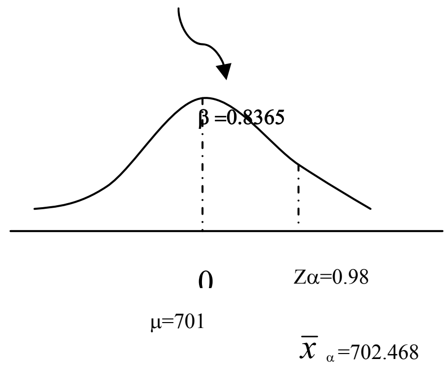
3. Calculate Z_{test} specifying a μ different to the critical value; what if the true mean μ is "701", for example:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{702.47 - 701}{\frac{15}{\sqrt{100}}} = 0.98$$

4. Calculate the probability of accepting 702.47 when 701 could be the true mean. P(0≤ Z ≤ 0.98)= 0.3365

5. Since you need to have the other half added to this probability.

β = 0.5 + 0.3365 = 0.8365



POWER OF THE TEST = 1-β

The power of the test measures how strong your estimators are to provide the more reliable hypothesis testing. If b is the probability of committing Error Type II, 1-β is of course the probability of not committing them, that is what gives power to the test. Easy, isn't it?

So for our example above the POWER OF THE TEST is = 1-0.8365 = 0.1635. There is a very weak probability of not committing this error, so you will need to do your best to decrease this error, by sampling more efficiently and for trying to have more accurate information.

LEARNING TEAM ACTIVITIES

HYPOTHESIS TESTING FOR MEANS AND PROPORTIONS

ANSWER TRUE OR FALSE TO THE FOLLOWING STATEMENTS.

1. A one-tailed hypothesis for a population mean with a significance level equal to .05 will have a critical value equal to z = .45.

2. Whenever possible, in establishing the null and alternative hypotheses, the research hypothesis should be made the alternative hypothesis.

3. If a hypothesis test is conducted for a population mean, a null and alternative hypothesis of the form:

Ho : $\mu = 100$

HA : $\mu \neq 100$

will result in a one-tailed hypothesis test since the sample result can fall in only one tail.

4. A local medical center has advertised that the mean wait for services will be less than 15 minutes. Given this claim, the hypothesis test for the population mean should be a one-tailed test with the rejection region in the lower (left-hand) tail of the sampling distribution.

5. A local medical center has advertised that the mean wait for services will be less than 15 minutes. In an effort to test whether this claim can be substantiated, a random sample of one-hundred customers was selected and their wait times were recorded. The mean wait time was 17.0 minutes. Based on this sample result, there is sufficient evidence to reject the medical center's claim.

6. The Adams Shoe Company believes that the mean size for men's shoes is now more than 10 inches. To test this, they have selected a random sample of n = 100 men. Assuming that the test is to be conducted using a .05 level of significance, a p-value of .07 would lead the company to conclude that their belief is correct.

7. A large tire manufacturing company has claimed that its top line tire will average more than 80,000 miles. If a consumer group wished to test this claim, they would formulate the following null and alternative hypotheses:

$Ho : \mu \geq 80{,}000$

$Ha : \mu < 80{,}000$

8. A large tire manufacturing company has claimed that its top line tire will average more than 80,000 miles. If a consumer group wished to test this claim, the research hypothesis would be

$Ha : \mu > 80{,}000\ miles.$

9. If a hypothesis test leads to incorrectly rejecting the null hypothesis, a Type II statistical error has been made.

10. The police chief in a local city claims that the average speed for cars and trucks on a stretch of road near a school is at least 45 mph. If this claim is to be tested, the null and alternative hypotheses are:

$Ho : \mu < 45 mph$

$Ha : \mu \geq 45 mph$

11. The loan manager for State Bank and Trust has claimed that the mean loan balance on outstanding loans at the bank is over $14,500. To test this at a significance level of 0.05, a random sample of n = 100 loan accounts is selected. Assuming that the population standard deviation is known to be $3,000, the null and alternative hypotheses to be tested are:

Ho : $\mu \leq \$14,500$

Ha : $\mu > \$14,500$

12. The director of the city Park and Recreation Department claims that the mean distance people travel to the city's greenbelt is more than 5.0 miles. Assume that the population standard deviation is known to be 1.2 miles and the significance level to be used to test the hypothesis is 0.05 when a sample size of n = 64 people are surveyed. Given this information, if the sample mean is 15.90 miles, the null hypothesis should be rejected.

13. The state insurance commissioner believes that the mean automobile insurance claim filed in her state exceeds $1,700. To test this claim, the agency has selected a random sample of 20 claims and found a sample mean equal to $1,733 and a sample standard deviation equal to $400. They plan to conduct the test using a 0.05 significance level. Based on this, the null hypothesis should be rejected if $x > \$1,854.66$ approximately.