Lesson 8 Part 2

Introduction to Hypothesis Testing and Hypothesis Tests of the Mean

Lesson 8 Part 2 Overview

- Lesson 8 part 1 covered estimation of the population mean: both point estimates and confidence intervals of the population mean.

- Part 2 of Lesson 8 extends the topic of inferential statistics about the population mean to hypothesis tests of the mean.

What is Hypothesis Testing?

- **Hypothesis testing is used when we wish to test whether the data provide significant evidence against some statement about the population from which the data were sampled.**

- **The statement under question is called the *null hypothesis*, often denoted by H_0.**

- **We are concerned here with statistical significance, not scientific or medical or economic or philosophical significance.**

- **A statistically significant result means that the sample statistic we observed is very unlikely to have occurred if the null hypothesis were true, assuming our model of the data is correct (e.g., assumption of normality, unbiased sampling).**

Steps in Hypothesis Testing

- Step1: The null and alternative hypothesis
- Step2: The test statistic
- Step3: The p-value
- Step4: The conclusion

Hypothesis Testing – Detailed Example

Clinical Example: Recurrent ear infections may play a role in the delay of language development. Typically children routinely begin to use sentences at the age of 36 months.

To test the hypothesis above, an investigator takes a random sample of 10 children who have recurrent ear infections. Assume the population distribution is normal.

Random Sample Statistics:

Average: 47 months

Sample size=10

Standard deviation = 12 months

Hypothesis Testing – Detailed Example

- Is this sample mean compatible with a population mean of 36 months?

- Is it unusual in a sample of 10 from a population centered at 36 months to get a sample average of 47 months?

- Do we have enough evidence to show that the mean age for routine sentence use for children with recurrent ear infections is greater than 36 months?

Hypothesis Testing: Introduction

- One approach is to construct a confidence interval for the population parameter;

- another is to construct a statistical hypothesis test.

The Null and Alternative Hypothesis- Step 1

- With statistical tests we assume that the mean of the population is equal to some postulated value μ_o which is called the null hypothesis (H_0).

- The alternative hypothesis (Ha) is a second statement that contradicts H_0.

Hypothesis Testing: -Step 1

- Write H_0 down using population parameters (not sample statistics).

- We are not trying to prove H_0; rather, we formulate H_0 such that we can *disprove* it.

- If there is "significant" evidence against H_0, we reject the null hypothesis, otherwise we fail to reject the null hypothesis.

- We choose the alternative that is most appropriate for the scientific question of interest.

Hypothesis Testing: Detailed Example- Step 1

$$H_0: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

$$H_A: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

- Null: The children in the sample come from a population centered at 36 months.
- Alternative: The children in the sample come from a population centered above 36 months.

Hypothesis Testing:
Test Statistic- Step 2

- To conduct a test of hypothesis we use our knowledge of the sampling distribution of the mean.

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \qquad \textbf{or} \qquad t_{n-1} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

For a given sample we can calculate the *test statistic.* Use of these distributions will depend on sample size.

Hypothesis Testing: Test Statistic- Step 2

- Use Z when σ is known and
  - The population distribution is normal (for any size n), or
  - The population distribution is not normal and $n > 30$.

- Use t when σ is unknown and
  - The population distribution is normal (for any size n), or
  - The population distribution is not too far from normal and $n > 15$, or
  - The population distribution is unimodal and $n > 40$.

Hypothesis Testing: Detailed Example-Step 2

- To test the hypothesis we have to compute the test statistic (σ is unknown and n=10, but the population is assumed to be normal.)

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} =$$

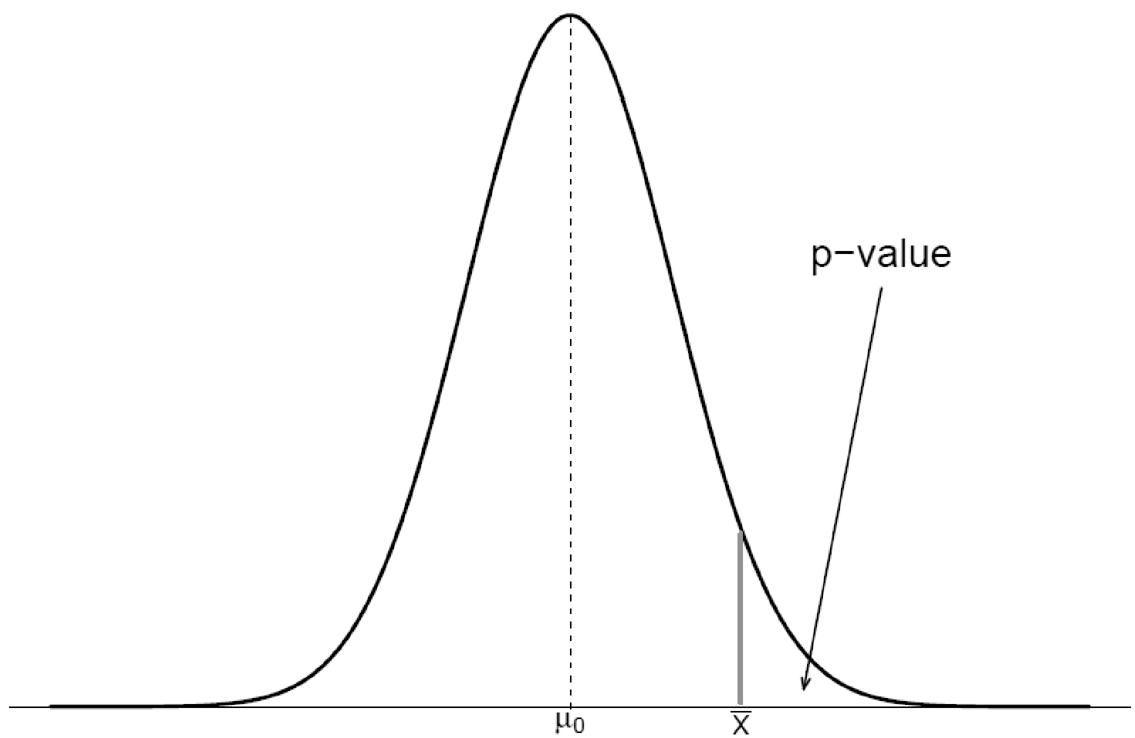
Hypothesis Testing: P-values – Step 3

- ***p-value*** : The probability of obtaining a mean (from our sample) as extreme or more extreme then the observed sample mean, given that the null hypothesis is true.

- If the p-value is "sufficiently small" we reject the null hypothesis. We have evidence against the null hypothesis.

Picture of the P-value
Step 3

Distribution of $\overline{X}$ under H_0: $\mu = \mu_0$



for testing H_a: $\mu > \mu_0$

One or Two Sided p-values

- Typically represents what you are trying to prove

H_a: $\mu \neq \mu_o$ or

H_a: $\mu > \mu_o$ or

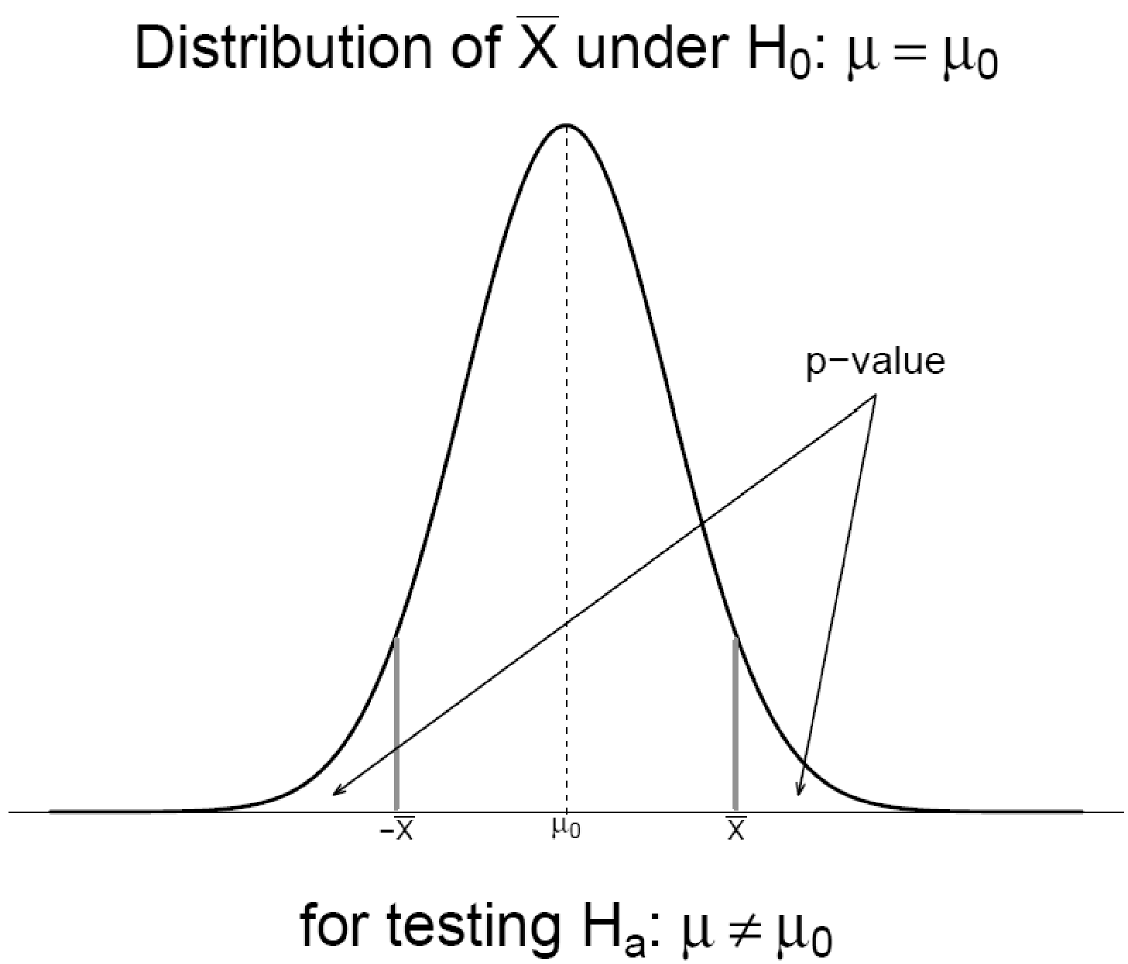
H_a: $\mu < \mu_o$

The children in our sample come from a population with a different mean (2-sided p-value)

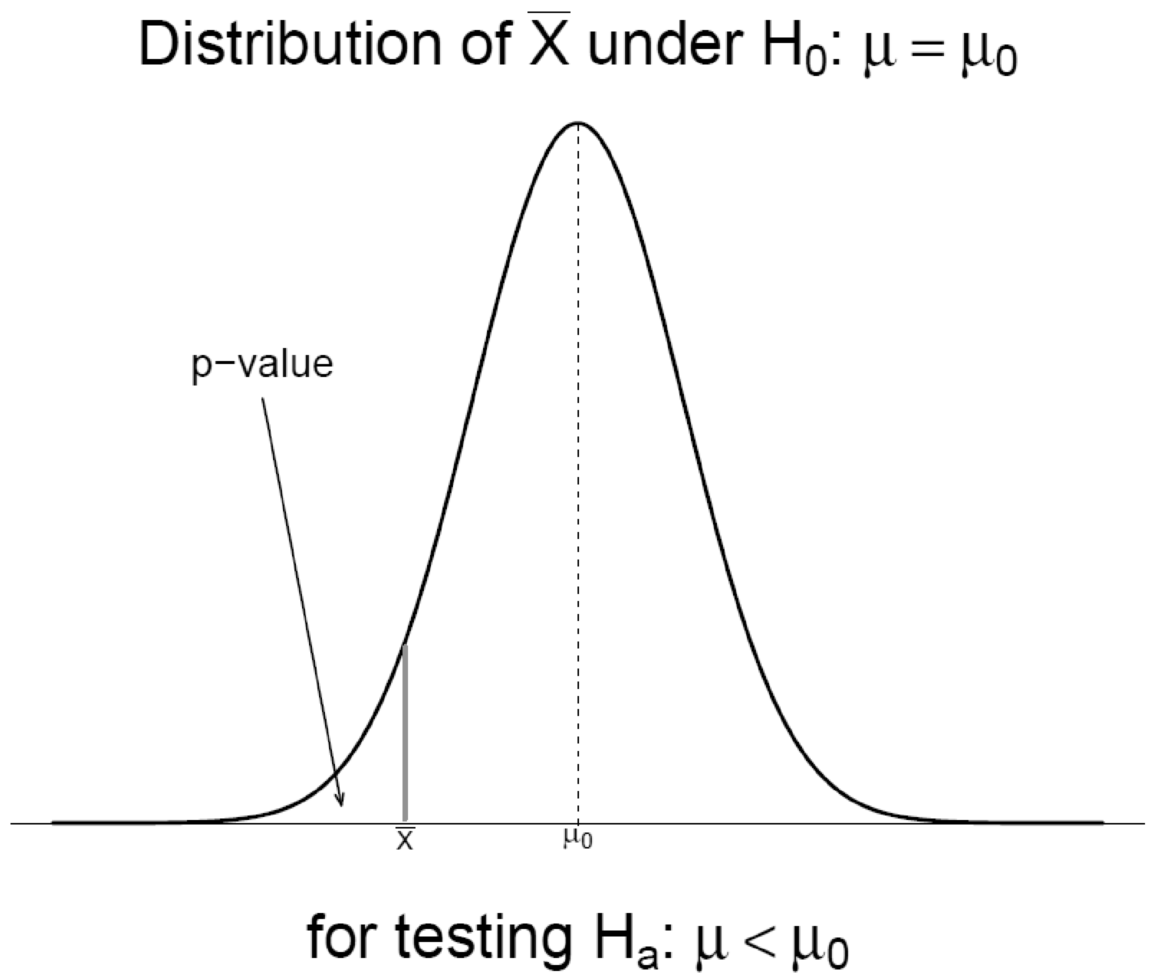
The children in our sample come from a population with a higher mean (1-sided p-value)

The children in our sample come from a population with a lower mean(1-sided p-value)

Picture of the P-value Two-Sided Alternative



Picture of the P-value
One-Sided Below Alternative



Calculation of the p-value - Step 3

- Use Excel
- Table C: t –table
- Table B: Z-table
- Use Applet

http://www.anu.edu.au/nceph/surfstat/surfstat-home/tables/t.php

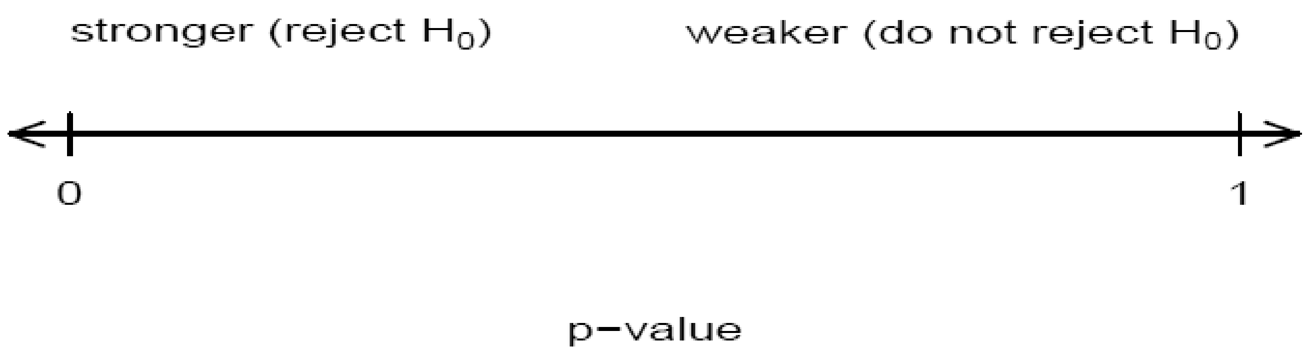
P-value:
Detailed Example – Step 3

If the null hypothesis is true, this statistic is the outcome of a student's t distribution and we can find the area to the right of 2.899 with 9 degrees of freedom.

=P(t>2.899) =TDIST(___,___,__) = 0.008812

Conclusion – Step 4

Remember the interpretation of a p-value: the probability of getting by chance a result as or more extreme than what we observed if H0 were true. The smaller the p-value (closer to 0), the stronger the evidence against H0 provided by the data. The larger the p-value (closer to 1), the weaker the evidence against H0.



Conclusion – Detailed Example- Step 4

p-value is small, Reject Ho

Our data provided sufficient evidence to suggest the average age at which children with recurrent ear aches use sentences is greater than 36 months. This may suggest a delay in language development.

How small is small?

Significance Level

- Definition

  We pre-specify a significance level , e.g.,α = 0.05. When $p < 0.05$ we reject H0 at level and say that the test result is statistically significant at level .
- Choosing $\alpha = 0.05$ is the convention. Sometimes exploratory analyses take a more 'liberal' level, e.g., α = 0.10; sometimes a more 'conservative' level is chosen, e.g., $\alpha = 0.01$.
- Hypothesis testing involves making a decision between two contradictory statements: the null hypothesis and the alternative hypothesis.

Hypothesis Testing: Types of Errors – Final Consideration

- Two possible ways to commit an error:
  1. Type I error: Reject Ho when it is true (α)
  2. Type II error: Fail to reject Ho when it is false (β)
- The goal in hypothesis testing is to keep α and β (the probabilities of type I and II errors) as small as possible.
- Usually α is fixed at some specific level - say 0.05 - significance level of the test.
- 1- β is called the power of the test.

Hypothesis Testing Outcomes

Hypothesis Test Decision	Population	
	Null Hypothesis is True	Null Hypothesis is False
Do Not Reject Null		
Reject Null		

Controlling Errors

- The ideal is to have low probability of both Type I and Type II errors BUT, for a given sample:
  - If α decreases, β increases
  - If β decreases, α increases

- The probability of Type I error is controlled by the significance level.

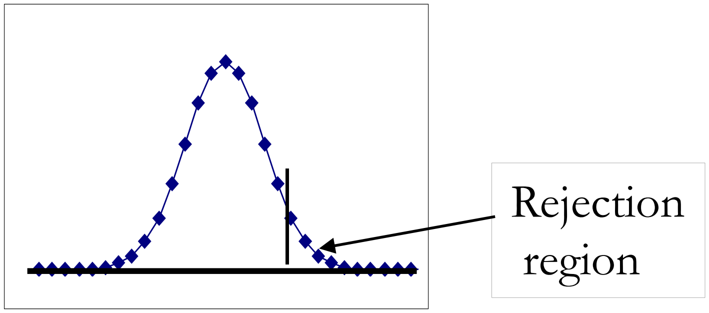
- For a given alpha level, the probability of Type II error can be decreased by increasing the sample size.

Step 3 Revisited: Rejection Regions OR Critical Values

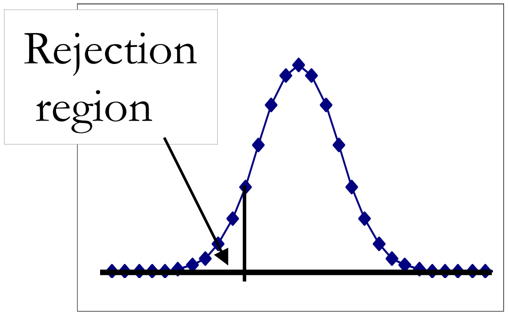
- There are two methods of making a decision about whether or not to reject the null hypothesis. Either method will result in the same decision.
  - Rejection region method: if the test statistic is in the ‘rejection region’ of the sampling distribution, the null hypothesis is rejected.
  - P-value method: if the p-value of the test statistic is less than the significance level (α) of the test, the null hypothesis is rejected.

Critical values for a z-test with significance level = 0.05

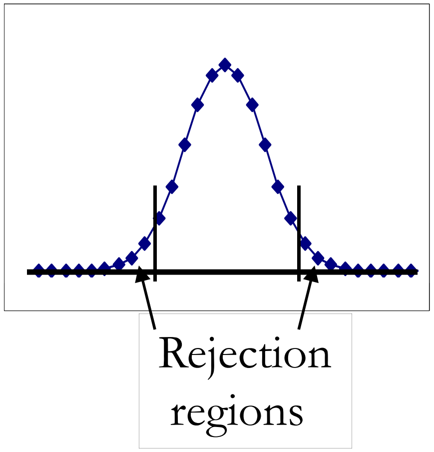
$H_{a:}\ \mu > \mu_0$
Critical value: 1.645
Area > 1.645 = 0.05



$H_{a:}\ \mu < \mu_0$
Critical value: - 1.645
Area < -1.645 = 0.05



$H_{a:}\ \mu \neq \mu_0$
Critical values: ±1.96
Area < -1.96 = 0.025
Area > 1.96 = 0.025



Using Excel to find Critical Values when $\alpha = 0.05$

- If the test statistic is a z-statistic, the critical value(s) are from the standard normal distribution
  - NORMSINV(0.025) = -1.96 and NORMSINV(1-0.025) = 1.96 for two-tailed test
  - NORMSINV(0.05) = -1.645 for one-tailed test of <
  - NORMSINV(0.95) = 1.645 for one-tailed test of >
- If the test statistic is a t-statistic, the critical value is from the t-distribution with n-1 df
  - TINV(0.05, n-1) for two-tailed test
  - TINV(0.10, n-1) for one-tailed test

Note that the TINV function assumes a two-sided test. To find the critical value for a one-sided test, you need to double the alpha level

Nutrition and Weight

- Typically, boys ages 10-14 years have a mean weight of 85 pounds. In a particular city neighborhood, kids were not getting proper nutrition and were eating high fat foods.

- A random sample of 25 boys ages 10-14 from this neighborhood was taken. They were found to have an average weight of 89.06 pounds with a standard deviation s =11.60 pounds.

- Test the hypothesis that boys ages 10-14 in this neighborhood weigh more on average than typical boys. Assume approximate normality of weights.

[Example adapted from: Chap T. Le, Introductory Biostatistics, page 248]

WHAT IS H_0?

1. The average weight across all boys ages 10-14 in this neighborhood is equal to typical boys.
2. The average weight across all boys ages 10-14 in this neighborhood is different from typical boys.
3. Boys ages 10-14 in this neighborhood weigh more on average than typical boys.
4. The average weight across this sample of boys ages 10-14 in this neighborhood is equal to typical boys.

WHAT IS H_A?

1. The average weight of boys ages 10-14 in this neighborhood equals 85 lbs.
2. The average weight of boys ages 10-14 in this neighborhood is different from 85 lbs.
3. Boys ages 10-14 in this neighborhood on average weigh more than 85 lbs.
4. Boys ages 10-14 in this neighborhood on average weigh less than 85 lbs.

Should we use T-test or Z-test?

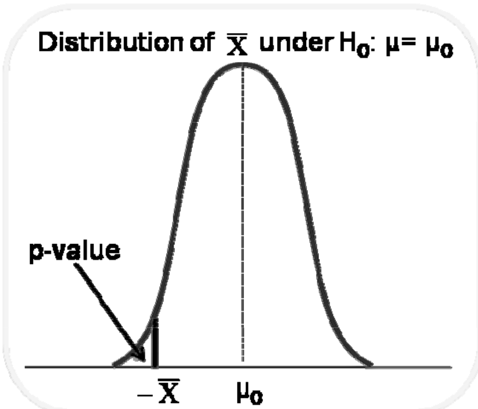
1. T-Test
2. Z-Test

Calculate the test statistic.

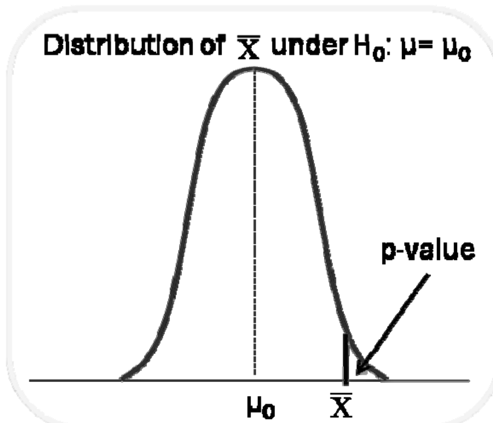
1. 38.39
2. 0.35
3. 1.75
4. -1.75

Select the p-value we need.

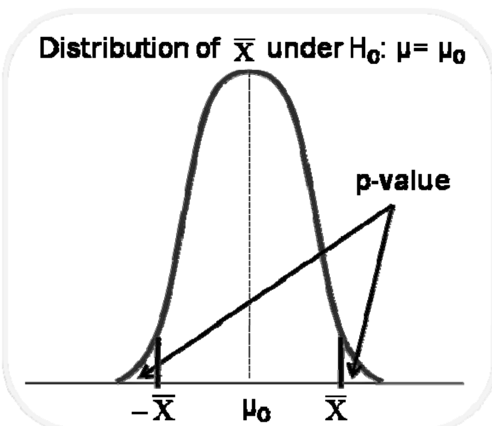
1.



2.



3.



p-value=0.046

What is the conclusion at α=0.05 ?

1. We reject the null hypothesis at level 0.05 and conclude that boys ages 10-14 in this neighborhood weigh more on average than typical boys.
2. We fail to reject the null at level 0.05 and conclude that boys ages 10-14 in this neighborhood don't weigh more on average than typical boys.

What is the conclusion at α=0.01?

1. We reject the null hypothesis at level 0.01 and conclude that boys ages 10-14 in this neighborhood weigh more on average than typical boys.
2. We fail to reject the null at level 0.01 and conclude that boys ages 10-14 in this neighborhood don't weigh more on average than typical boys.

Readings and Assignments

- Reading
  - Chapter 5 pgs. 102-110
- Complete the Lesson 8 Part 2 Practice Exercises
- Work through Excel Module 8 examples
- Complete Homework 5 and submit by due date