Section 8.1
Inference for the Mean of a Population

In chapter 7, we discussed one type of inference about a population – the confidence interval. In this chapter we will begin discussion the **significance test**, the other type of statistical inference. The goal of a significance test is to assess a claim made about a population parameter given certain sample data. This assumption may or may not be true. **Hypothesis testing** refers to the formal procedures used by statisticians to reject or fail to reject statistical hypotheses.

Statistical Hypotheses

The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected. There are two types of statistical hypotheses.

H_0: is the **null hypothesis.** The **null hypothesis** states that there is no effect or change in the population.

H_a: is the **alternate hypothesis**. The **alternative hypothesis** describes the effect we suspect is true, in other words, it is the alternative to the "no effect" of the null hypothesis.

Since there are only two hypotheses, there are only two possible decisions: *reject the null hypothesis in favor of the alternative* or *don't reject the null hypothesis*. We will never say that we accept the null hypothesis. Here's an example why: *Let's say we're testing something that may be detrimental to someone's health, and we find that we should "fail to reject" H_0 (the null hypothesis). That doesn't mean all is ok with it, just that it may need further testing. Hence, we don't ever want to say, "yes, we'll accept the null hypothesis".*

For inference about a population mean: $H_0 : \mu = \mu_0$ where μ_0 represents the given (or claimed) population mean.

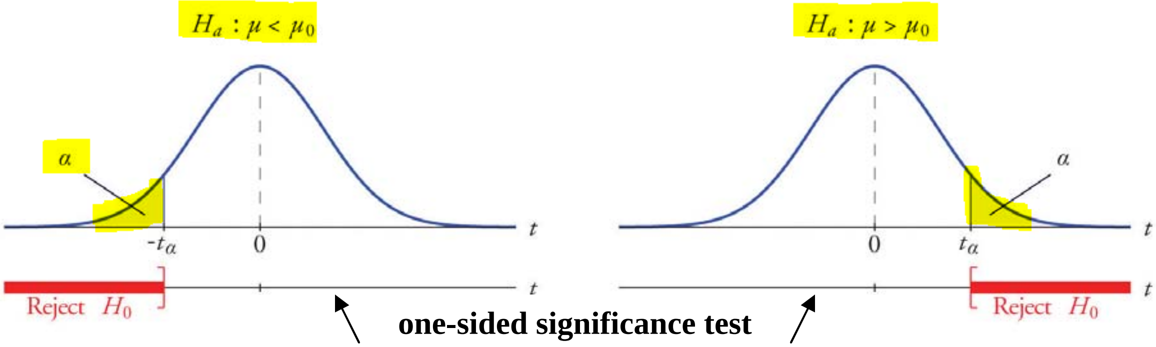
Examples:

- Suppose we claim that the average course grade in Business Calculus is 78%. We think it's more than that, and so we take a sample of this population and find that it's 81%. The hypotheses would be expressed as: $H_0 : \mu_0 = 78\%$; $H_a : \mu > \mu_0$. This is known as a **one-sided significance test**.

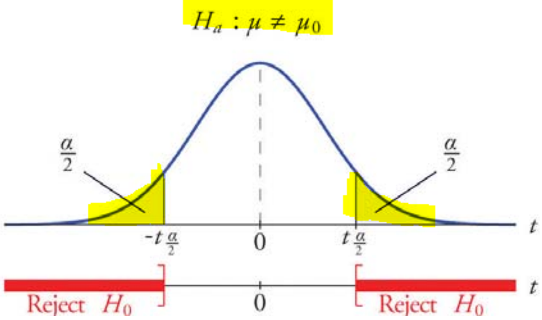
- Suppose we wanted to determine whether a coin was fair and balanced. A null hypothesis might be that half the flips would result in Heads and half, in Tails. The alternate hypothesis might be that the number of Heads and Tails would be very different. The hypotheses would be expressed as: $H_0 : \mu_0 = 0.5$; $H_a : \mu \neq \mu_0 \neq 0.5$. This is known as a **two-sided significance test**.

Alternate Hypothesis and Rejection Region

Note: The value α is called the **significance level** of a test (In some problems we'll be given the significance level α. If we are not given α, we say it's 0.05.). It helps determine the **critical value** (either be t^* or z^*) and the **rejection region** (shaded region(s)) for the test.



one-sided significance test
Deviations in one direction from the null.



two-sided significance test
Deviations in any direction from the null.

The ***p*-value** of the test is a probability value that is either within the significance level α or not. It is the probability that the sample data is what it is assuming the null hypothesis is true. The *p*-value is said to be **statistically significant** when it is as small as, or smaller than, the given significance level, α.

When performing a significance test, we follow these steps:

1. Check assumptions.
2. State the null and alternate hypotheses.
3. Graph the rejection region, labeling the critical values.

To find critical values for a *z*-distribution, z^*, the R command is: $\text{qnorm}(\text{area to the left})$

To find critical values for a *t*-distribution, t^*, the R command is: $\text{qt}(\text{area to the left, df})$

4. Calculate the test statistic.
5. Find the *p*-value. If this answer is less than the significance level, α, we can reject the null hypothesis in favor of the alternate.

To find the *p*-value (probability) for a z-distribution, the R command is: pnorm(z) = area to the left. If you need area to the right of the z-value, then use 1 – pnorm(z).

To find the *p*-value (probability) for a *t*-distribution, the R command is: pt(*t*, df) = area to the left. If you need area to the right of the *t*-value, then use 1 – pt(*t*, df).

6. Give your conclusion using the context of the problem. When stating the conclusion you can give results with a confidence of $(1-\alpha)$(100)%.

***z* – test Assumptions:**

1. An SRS of size *n* from the population.
2. Known population standard deviation, σ.
3. Either a normal population or a large sample ($n \geq 30$). Otherwise, we assume normality.

To compute the *z* – test statistic, we use the formula: $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$.

***t* – test Assumptions:**

1. An SRS of size *n* from the population.
2. Unknown population standard deviation, but the sample standard deviation is given.
3. Either a normal population or large sample ($n \geq 30$). Otherwise, we assume normality.

To compute the *t* – test statistic, we use the formula: $t = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n}}$, where *s* is the sample standard deviation. The *t* – test will use *n* – 1 degrees of freedom.

Example 1: Mr. Murphy is an avid golfer. Suppose he has been using the same golf clubs for quite some time. Based on this experience, he knows that his average distance when hitting a ball with his current driver (the longest-hitting club) under ideal conditions is 200 yards with a standard deviation of 9. After some preliminary swings with a new driver, he obtained the following sample of driving distances:

205	198	220	210	194	201	213	191	211	203

He feels that the new club does a better job. Do you agree?

First of all, this will be a z-test since the population standard deviation is given.

We will also assume normality.

1. State the null and alternate hypotheses.

$H_0 : \mu_0 =$ 200

Since he believes the new clubs would do a ***better job*** then $H_a : \mu > \mu_0$.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \qquad n = 10 \qquad \mu_0 = 200 \qquad \sigma = 9$$

We need the mean to calculate the z-test statistic.

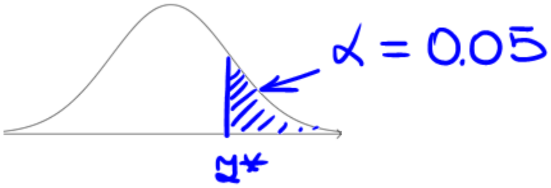
Commands: Answer:

```
>distances=c(205,198,220,210,194,201,213,191,211,203)
>mean(distances)
```

$\bar{x} = 204.6$

2. Graph the rejection region, labeling the critical values.

Since α is not given, then $\alpha = 0.05$ (one-sided tail; rejection region). *Recall: Since $\mu > \mu_0$, we shade to the right of the critical value.*



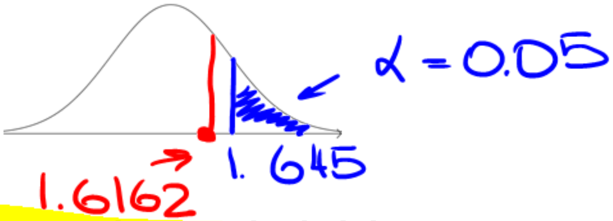
Now find z^*.

Command: qnorm (1-0.05) Answer: 1.645

qnorm (0.95)

3. Calculate the test statistic.

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{(204.6 - 200)}{(9 / \sqrt{10})} = 1.6162$$



Note that 1.616275 is to the left of 1.645 and is not in the shaded area (rejection region), so it's not in the region we will reject.

Hence, we'll FAIL TO REJECT this.

4. Find the *p*-value to verify this. If this answer is less than the significance level, α , we can reject the null hypothesis in favor of the alternate, otherwise we fail to reject the null hypothesis.

So we want to do: P(Z > 1.616) = 1 - pnorm(1.6162) = 0.0531 > α

$\alpha = 0.05$

So, we have verified that is it correct to FAIL TO REJECT the null hypothesis.

The p-value is the probability that we would get our observed result if the null hypothesis were true.

$(1 - 0.05) \times 100\%$

5. Conclusion: Based on 95% (because α = 0.05 i.e. a 5% significance level) certainty, we fail to reject the null hypothesis, i.e. don't favor the alternate which states that the new club does a better job.

Summary

- If the test statistic is not in the rejection region and the *p*-value is greater than α , then fail to reject H_0 .
- If the test statistic is in the rejection region and the *p*-value is less than α , then reject H_0 in favor of H_a .

Example 2: An association of college bookstores reported that the average amount of money spent by students on textbooks for the Fall 2010 semester was $325.16. A random sample of 75 students at the local campus of the state university indicated an average bill for textbooks for the semester in question to be $312.34 with a standard deviation of $76.42. Do these data provide significant evidence that the actual average bill is different from the $325.16 reported? Test at the 1% significance level.

$n = 75$

$\bar{x} = 312.34$

$s = 76.42$

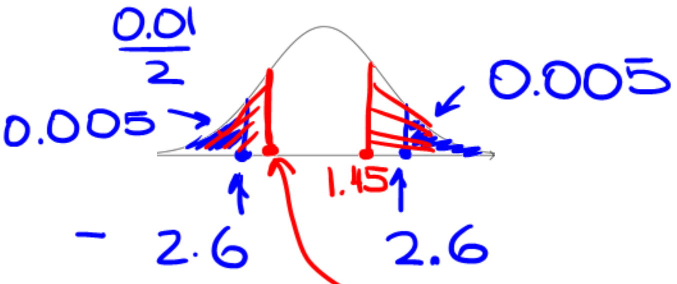
So are we going to do a [✓ *t-test*] or *z-test?* *Why?* sample s.d. is known

$\alpha = 0.01$

State the null and alternate hypotheses.

$H_0 : \mu_0 = 325.16$

$H_a : \mu \neq \mu_0 .$



Now find t^*.

Command: qt(0.005,75-1)

Answer: $t^* = -2.6439$

t-test statistic: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{312.34 - 325.16}{76.42/\sqrt{75}} = -1.4528$

Is it in the rejection region? NO

Fail to reject H_0?

Let's verify by finding the *p*-value:

Note: Since we cannot find P($t \neq$ number), the $\neq$ comes from H_a, then when finding the *p*-value, we will use:

- < if the test statistic is negative; that is, we'll find 2*P(t < negative number) = Command: 2*pt(negative number, df)
- > if the test statistic is positive; that is we'll find 2*P(t > positive number) = Command: 2*(1 – pt(positive number, df))

Command: 2*pt(-1.4528,74)

Answer: 0.1505 > α

Fail to reject H_0?

(1-.15)100%

α = 0.01

Assuming the null is true, we are resonably sure our observed data can be explained by chance along.
↳ 85%

Matched pairs is a special test when we are comparing corresponding values in data. This test is used only when our data samples are DEPENDENT upon one another (like before and after results).

Matched pairs *t* – test assumptions:

1. Each sample is an SRS of size *n* from the same population.
2. The test is conducted on paired data (the samples are NOT independent).
3. Unknown population standard deviation.
4. Either a normal population or large samples ($n \geq 30$). Otherwise, assume normality.

Example 3: A new law has been passed giving city police greater powers in apprehending suspected criminals. For six neighborhoods, the numbers of reported crimes one year before and one year after the new law are shown. Does this indicate that the number of reported crimes have dropped?

Neighborhood	**1**	**2**	**3**	**4**	**5**	**6**
Before	18	35	44	28	22	37
After	21	23	30	19	24	29

All conditions above are satisfied and we'll assume normality.

Next, let's calculate the differences. Since we're looking to see if the number of reported crimes have ***dropped*** then we'll need to do: Before – After. If the number of reported crimes dropped then this difference should be positive.

Neighborhood	**1**	**2**	**3**	**4**	**5**	**6**
Before	18	35	44	28	22	37
After	21	23	30	19	24	29
Differences:	-3	12	14	9	-2	8

State the null and alternate hypotheses.

$H_0 : \mu_0 = 0$ (What would this be if the new law makes "no difference"?)

$H_a : \mu > \mu_0$.

We need the mean to calculate the ~~z~~t-test statistic.

Commands:

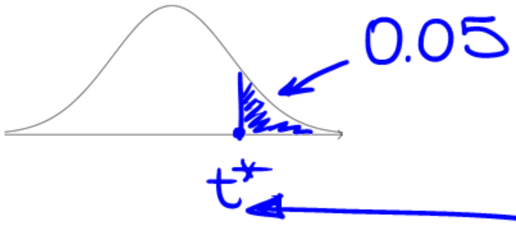
```
>differences=c(-3,12,14,9,-2,8)
>mean(differences)
>sd(differences)
```

Answers:

$\bar{x} = 6.3333$

$s = 7.174$

Graph the rejection region, labeling the critical values.
Since α is not given, then α = 0.05 (one-sided tail; rejection region).
Recall: Since $\mu > \mu_0$, we shade to the right of the critical value.



Now find t^*.
Command: qt (0.95, 6-1)

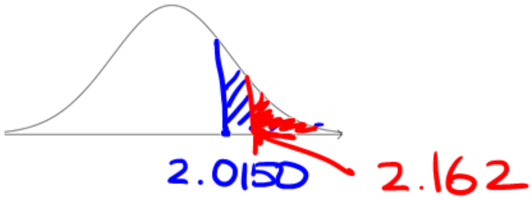
$1 - \alpha = 1 - 0.05$

Answer:

$t^* = 2.0150$

Calculate the test statistic.

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{6.333 - 0}{7.174/\sqrt{6}} = 2.162$$



Note that 2.162 is to the right of 2.015048 and is in the shaded area (rejection region).
Hence, we'll REJECT the null hypothesis.

Find the *p*-value to verify this. If this answer is less than the significance level, α, we can reject the null hypothesis in favor of the alternate, otherwise we fail to reject the null hypothesis.

So we want to do: P(t > 2.162) = 1- pt (2.162, 6-1) = 0.0415 $< \alpha$

So, we have verified that is it correct to REJECT the null hypothesis in favor of the alternate hypothesis.

Conclusion: We are rejecting the null hypothesis in favor of the alternate hypothesis with a confidence level of 95% (because α = 0.05). Or we can say, based on 95% certainty, we will reject that there is no change in the number of crimes reported in favor of saying there is a drop in the number of reported crimes.

Also, the *p*-value was 0.0415. This states that there is a 4.15% chance of observing a change (for the better; drop in crimes) assuming the null is true.