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. 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.

We begin a significance test with stating the two claims we want to compare.

- The null hypothesis: $H_0$ - this is the claim we seek evidence against. Usually it is a statement of “no difference”.
- The alternate hypothesis: $H_a$ - this is the claim we suspect to be true instead of the null hypothesis.

Hypothesis testing refers to the formal procedures used by statisticians to reject or fail to reject statistical hypotheses.

Since there are only two hypotheses, there are only two possible decisions:

- reject the null hypothesis in favor of the alternative.
- 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”.
Remember we said we want to seek evidence against the null hypothesis? Why would we want to do that? Think of “innocent until proven guilty”.

In a trial in the United States the defendant is innocent until proven guilty; and the jury must evaluate the truth of two competing hypothesis:

- \( H_0 \): The defendant is not guilty. (The jury must assume this to be true until proven otherwise.)
- \( H_a \): The defendant is guilty. (This must be demonstrated to be true beyond a reasonable doubt.)

The judgement here is based on evidence and here we can find another example as to why we never want to say, “Accept \( H_0 \)”:

The jury then returns with one of two verdicts:
- “Not Guilty” i.e. Fail to reject \( H_0 \)
  or
- “Guilty” i.e. Reject \( H_0 \)

If the jury says, “Not Guilty” this does not mean innocence, just that there was not convincing evidence of guilt!!

In statistics we deal with evidence provided by a sample and use a probability to say how strong the evidence is.

The probability that measures the strength of the evidence against \( H_0 \) and in favor of \( H_a \) is called the p-value.

- Small p-values are evidence against \( H_0 \) since this indicates that the observed result is unlikely to occur when \( H_0 \) is true. In this case, reject \( H_0 \), there is convincing evidence for \( H_a \).
- Large p-values fail to give convincing evidence against \( H_0 \) since this indicates that the observed result is likely to occur by chance alone when \( H_0 \) is true. In this case, fail to reject \( H_0 \), there is not convincing evidence for \( H_a \).
How do we determine if the p-value is “small” or “large” so that we know if to reject $H_0$ or fail to reject $H_0$?

The value $\alpha$ is called the **significance level** (threshold) of a test. It is the probability of rejecting $H_0$ when it is true. Setting this value small makes it more difficult to wrongfully reject $H_0$ just by chance.

In some problems we’ll be given the significance level $\alpha$. If we are not given $\alpha$, we say it’s 5%.

Since $\alpha$ is our threshold then we’ll compare the p-value to it. The p-value of the test is a probability value that is either within the significance level $\alpha$ or not.

The p-value is said to be **statistically significant** when it is as small as, or smaller than, the given significance level $\alpha$.

**Statistically significant** means the result is not likely to occur just by chance.
The significance test will result in one of three tests:

- One-sided alternate hypothesis if it states that a parameter is larger than null hypothesis value.
- One-sided alternate hypothesis if it states that a parameter is smaller than null hypothesis value.
- Two-sided alternate hypothesis if it states that a parameter is different from the null hypothesis value.

**Example:** Suppose we claim that the average course grade in Business Calculus is 78%. An instructor thinks it’s more than that, and so we take a sample of this population and find that it’s 81%.

This is represented in the following standard normal curve for \( H_0 : \mu = 78\% \); \( H_a : \mu > 78\% \)

The critical value is either be \( t^* \) or \( z^* \) and the rejection region is \( \alpha \). If \( H_a \) was \( \mu < 78\% \) then the left tail would be shaded.

**Two-Sided Significance Test**

**Example:** 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: \( H_0 : \mu = 0.5 \); \( H_a : \mu \neq 0.5 \).
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.
   - Critical value, $z^*$: $\text{qnorm}(\text{area to the left})$
   - Critical value, $t^*$: $\text{qt}(\text{area to the left, df})$
4. Calculate the test statistic.
   - $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$
   - $t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$
5. Find the $p$-value.
6. $p$-value (probability): $\text{pnorm}(z)$ or $1 - \text{pnorm}(z)$.
   - $p$-value (probability): $\text{pt}(t, df)$ or $1 - \text{pt}(t, df)$.
7. Give your conclusion using the context of the problem.

---

### $z$ – test Assumptions:
1. An SRS of size $n$ from the population.
2. Known population standard deviation, $\sigma$.
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: $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$.

### $t$ – 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: $t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$, 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, is this a z-test or a t-test?

State the null and alternate hypotheses.

\[ H_0 : \mu = L \quad H_a : \mu \neq L \]

Graph the rejection region, labeling the critical value.

Since \( \alpha \) is not given, then \( \alpha = 0.05 \).

Find \( z^* \).

Command: Answer:

Calculate the test statistic:

\[ z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \]

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

Commands: Answer:

Find the \( p \)-value.

Command: Answer:
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.

First of all, is this a z-test or a t-test?

State the null and alternate hypotheses.

$H_0 : \mu$  \hspace{1cm} $H_a : \mu$

Graph the rejection region, labeling the critical values.

$\alpha =$

Find $t^*$.  
Command: \hspace{1cm} Answer:

Find the test statistic: 
\[ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \]

Find the $p$-value.

Command: \hspace{1cm} Answer:
**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?

<table>
<thead>
<tr>
<th>Neighborhood 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>18</td>
<td>35</td>
<td>44</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>After</td>
<td>21</td>
<td>23</td>
<td>30</td>
<td>19</td>
<td>24</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Neighborhood 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>18</td>
<td>35</td>
<td>44</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>After</td>
<td>21</td>
<td>23</td>
<td>30</td>
<td>19</td>
<td>24</td>
</tr>
</tbody>
</table>

**State the null and alternate hypotheses.**

$p_0 : \mu \quad p_a : \mu$

**Graph the rejection region, labeling the critical value.**

Since $\alpha$ is not given, then $\alpha = 0.05$.

Now find $t^*$.  

**Command:** Answer:
Calculate the test statistic:  \[ t = \frac{\bar{X} - \mu}{s / \sqrt{n}} \]

We need the mean and standard deviation to calculate the t-test statistic.

Commands:  

\[
\text{Answers:} \\
> \text{differences=\{c(-3,12,14,9,-2,8)\}} \\
> \text{mean(differences)} \\
> \text{sd(differences)} \\
\]

Find the \( p \)-value.

Command:  

\[
\text{Answer:} \\
\]