Math 2311
Bekki George – bekki@math.uh.edu
Office Hours: MW 11am to 12:45pm in 639 PGH
Online Thursdays 4-5:30pm
And by appointment

Class webpage: http://www.math.uh.edu/~bekki/Math2311.html
Ch7 - Confidence Intervals

Statistic ± margin of error

\[ \bar{x} \text{ or } \hat{p} \quad \pm \quad \pm \]

\[ z^* \text{ or } t^* \quad \text{times, st.dev.} \]

\[ \bar{x} \pm z^* \left( \frac{\sigma}{\sqrt{n}} \right) \quad \text{or} \quad \bar{x} \pm t^* \left( \frac{\sigma}{\sqrt{n}} \right) \]

\[ \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \]
1. What is the $z^*$ critical value for a 95% confidence interval?
   a. 1.96
   b. 1.645
   c. 2.33
   d. 2.06
   e. none of these
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.

\(H_0\) : is the null hypothesis. The null hypothesis states that there is no effect or change in the population. It is the statement being tested in a test of significance. \(\mu = \square\) stated population mean

\(H_a\) : 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. Changing to: \(<> \) or \(\neq\)

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.
For inference about a population mean:

\( H_0 : \mu = \mu_0 \) where \( \mu_0 \) represents the given population mean.

<table>
<thead>
<tr>
<th>Alternate Hypothesis</th>
<th>Rejection Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_a : \mu &lt; \mu_0 )</td>
<td>![One tailed test][1]</td>
</tr>
<tr>
<td>( H_a : \mu &gt; \mu_0 )</td>
<td>![One tailed test][2]</td>
</tr>
<tr>
<td>( H_a : \mu \neq \mu_0 )</td>
<td>![Two tailed test][3]</td>
</tr>
</tbody>
</table>

If I want to look at a critical value \( \alpha = .05 \)

\[ \text{one tailed tests} \]

\[ \text{two tailed test} \]
The probability, computed assuming that the null hypothesis is true, that the test statistic would take a value as extreme or more extreme than that actually observed is called the \textit{p-value} of the test. A result with a small \textit{p-value} is called \textit{statistically significant}. This means that chance alone would rarely produce so extreme a result. We say that a value is \textit{statistically significant} when the \textit{p-value} is as small as, or smaller than, the given significance level, \( \alpha \). If we are not given \( \alpha \), we can interpret the results like this:

- If the \textit{p-value} is less than 1\%, we say that there is overwhelming evidence to infer that the alternative hypothesis is true. (We also say that the test is highly significant)

- If the \textit{p-value} is between 1\% and 5\%, we say that there is strong evidence to infer that the alternative hypothesis is true. (We also say that the test is significant)

- If the \textit{p-value} is between 5\% and 10\%, we say that there is weak evidence to infer that the alternative hypothesis is true. (We also say that the test not statistically significant)

- If the \textit{p-value} is exceeds 10\%, we say that there is no evidence to infer that the alternative hypothesis is true.

\textit{If \( \alpha \) is not given, use \( \alpha = .05 \).}
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.
4. Calculate the test statistic.
5. Find the $p$-value. If this answer is less than the significance level, $\alpha$, we can reject the null hypothesis in favor of the alternate.
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, \( \sigma \).
3. Either a normal population or a large sample \((n \geq 30)\).

To compute the \( z \) – test statistic, we use the formula:

\[
 z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}
\]

\( \text{Test statistic} \)

\( \text{\( t \) – test} \)

Assumptions:
1. An SRS of size \( n \) from the population.
2. Unknown population standard deviation.
3. Either a normal population or large sample \((n \geq 30)\).

To compute the \( t \) – test statistic, we use the formula:

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

where \( s \) is the sample standard deviation. The \( t \) – test will use \( n - 1 \) degrees of freedom.
2. A \( t \) test is used instead of a \( z \) test when
   a. The population mean is unknown
   b. The population size is unknown
   c. The population variance is unknown
   d. None of these
Examples:

10. 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:

\[ 698 \quad 220 \quad 210 \quad 194 \quad 201 \quad 213 \quad 191 \quad 211 \quad 203 \]

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

\[ H_0: \mu = 200 \quad \bar{x} = 204.6 \]
\[ H_a: \mu > 200 \]
\[ \alpha = 0.05 \]

Test statistic:
\[
Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{204.6 - 200}{9 / \sqrt{10}} = 1.616
\]

P-value:
\[
P(Z > 1.616) = \text{normalcdf}(1.616, \text{big#}, 0, 1) = 0.053 > \alpha = 0.05
\]

Fail to reject the null hypothesis.
if test statistic is in rejection region means that the p-value < α
⇒ Reject H₀ in favor of Hₐ

if t.s. not in rejection region and p-value > α
⇒ fail to reject H₀
13. 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.

\[ n = 75 \quad \bar{x} = 312.34 \quad s = 76.42 \quad \alpha = .01 \]

\( H_0: \mu = 325.16 \)

\( H_a: \mu \neq 325.16 \)

\[ t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \]

\[ t = \frac{312.34 - 325.16}{76.42/\sqrt{75}} = -1.4528 \]

\[ \text{p-value: } 2 \cdot p(t < -1.4528) = .1505 > .01 \]

If \( H_a \neq \) then p-value use < if test stat neg > if pos and multiply 2

Fail to reject the null hypothesis.
Popper 19

3. Suppose a test of the hypothesis in a question gives a p-value of 0.02. The correct action based on $\alpha = 0.05$ would be to
   a. Reject $H_0$
   b. Reject $H_a$
   c. Accept $H_0$
   d. Fail to reject $H_0$
   x. None of these
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 ($n \geq 30$).
Example:

15. 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</th>
<th>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>
<td>37</td>
</tr>
<tr>
<td>After</td>
<td>21</td>
<td>23</td>
<td>30</td>
<td>19</td>
<td>24</td>
<td>29</td>
</tr>
</tbody>
</table>

Assume normal population

Before-After: -3 12 14 9 -2 8

H₀: no change: μₐ = 0  \( \alpha = .05 \)

Hₐ: \( \muₐ > 0 \)

\[ t = \frac{6.333 - 0}{7.174/\sqrt{6}} = 2.162 \]

\( t^* = 2.015 \)

\( \text{InvT}(0.95, 5) \)

\( 1 - pt(2.162, 5) < \alpha \)

p value: \( p(t > 2.162) = .0415 < .05 \)

Reject the null in favor of the alt. hypothesis.
4. A significance test was performed to test $H_0: \mu = 23$ versus $H_a: \mu < 23$. The test statistic is $z = -1.70$. What is the p-value for this test?
   
   a. 0.09
   b. 0.04
   c. 0.02
   d. 0.96
   e. none of these

5. A significance test was performed to test $H_0: \mu = 23$ versus $H_a: \mu \neq 23$. The test statistic is $z = -1.70$. What is the p-value for this test?
   
   a. 0.09
   b. 0.04
   c. 0.02
   d. 0.96
   e. none of these